NEUROPSYCHOPHARMACOLOGY
The Fifth Generation of Progress

An Official Publication of the American College of Neuropsychopharmacology

5th Edition

2002

Lippincott Williams & Wilkins

Philadelphia

530 Walnut Street, Philadelphia, PA 19106 USA LWW.com

0-7817-2837-1

Printed in the USA

Library of Congress Cataloging-in-Publication Data

Neuropsychopharmacology: the fifth generation of progress: an official publication of the American College of Neuropsychopharmacology / editors, Kenneth L. Davis… [et al.].

p. ; cm.

Includes bibliographical references and index.

ISBN 0-7817-2837-1

1. Neuropsychopharmacology. 2. Neurobehavioral disorders. I. Davis, Kenneth L., 1947-II. American College of Neuropsychopharmacology.

[DNLM: 1. Psychopharmacology—methods. 2. Diagnostic Imaging—methods. 3. Drug Evaluation. 4. Neurotransmitters—physiology. QV 77 N4968 2002]

RM315 .N4825 2002

616.8′0461—dc21

2001038463

10 9 8 7 6 5 4 3 2 1

Editors

Kenneth L. Davis MD

Esther and Joseph Klingenstein Professor of Psychiatry and Chairman
Department of Psychiatry, Mount Sinai School of Medicine, New York, New York

Dennis Charney MD

Chief
Mood and Anxiety Disorder Research Program, National Institute of Mental Health Bethesda, Maryland

Joseph T. Coyle MD

Eben S. Draper Professor of Psychiatry and of Neuroscience; Chairman
Department of Psychiatry, Harvard Medical School, Boston, Massachusetts

Charles Nemeroff MD, PhD

Reunette W. Harris Professor and Chairman
Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia

SECTION EDITORS

Samuel H. Barondes MD

Jeanne and Sanford Robertson, MD, Director
Center for Neurobiology and Psychiatry, Department of Psychiatry, University of California, San Francisco, San Francisco, California

Dennis Charney MD

Chief
Mood and Anxiety Disorder Research Program, National Institute of Mental Health, Bethesda, Maryland

Joseph T. Coyle MD

Eben S. Draper Professor of Psychiatry and of Neuroscience; Chairman
Department of Psychiatry, Harvard Medical School, Boston, Massachusetts

Kenneth L. Davis MD

Esther and Joseph Klingenstein Professor of Psychiatry and Chairman
Department of Psychiatry, Mount Sinai School of Medicine, New York, New York

Robert Desimone PhD

Scientific Director
National Institute of Mental Health, Director, Intramural Research Program, National Institute of Mental Health, Bethesda, Maryland

Ronald Duman PhD
Departments of Psychiatry and Pharmacology, Laboratory of Molecular Psychiatry, Yale University School of Medicine, New Haven, Connecticut

Marian W. Fischman PhD

Professor of Behavioral Biology
Department of Psychiatry, College of Physicians and Surgeons of Columbia University and, The New York Psychiatric Institute, New York, New York

Eric Hollander MD

Professor of Psychiatry; Director of Clinical Psychopharmacology; Director of the Compulsive; Impulsive and Anxiety Disorders

Program Clinical Director
Seaver Autism Research Center, Department of Psychiatry, Mount Sinai School of Medicine, New York, New York

David Kupfer MD, PhD

Thomas Detre Professor and Chair
Department of Psychiatry, University of Pittsburgh Medical Center Western Psychiatric Institute and Clinic, Pittsburgh, Pennsylvania

Herbert Meltzer MD

Bixler Professor of Psychiatry and Pharmacology
Vanderbilt University School of Medicine, Nashville, Tennessee

Charles Nemeroff MD, PhD

Reunette W. Harris Professor and Chairman

Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia

Eric Nestler MD, PhD

Lou and Ellen McGinley Distinguished Professor and Chairman Department of Psychiatry
The University of Texas Southwestern Medical Center at Dallas, Dallas, Texas

Charles P. O'Brien MD, PhD

Chief of Psychiatry; Philadelphia VA Medical Center Kenneth Appel Professor and, Vice Chair of Psychiatry
University of Pennsylvania, Philadelphia, Pennsylvania

Carol Tamminga MD

Professor of Psychiatry and Pharmacology
University of Maryland School of Medicine; Deputy Director, Maryland Psychiatric Research Center, Baltimore, Maryland

Daniel Weinberger MD

Chief
Clinical Brain Disorders Branch, National Institutes of Health, National Institute of Mental Health, Bethesda, Maryland

Secondary Editors

Charley Mitchell

Acquisitions Editor

Ray Reter

Developmental Editor

Patrick Carr

Production Editor

Tim Reynolds

Manufacturing Manager

Mark Lerner

Cover Designer

Compositor: Maryland Composition

Printer: Courier Westford

CONTRIBUTORS

Henry David Abraham M.D.

Lecturer
Department of Psychiatry, Harvard Medical School, Boston, Massachusetts

Paul D. Acton Ph.D

Research Assistant Professor
Radiology, University of Pennsylvania, Philadelphia, Pennsylvania

George K. Aghajanian M.D
Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut

Huda Akil Ph.D.
Mental Health Research Institute, University of Michigan, Ann Arbor, Michigan

Meenakshi Alreja Ph.D.
Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut

James C. Anthony Ph.D

Professor
Johns Hopkins University, School of Hygeine and Public Health, Baltimore, Maryland

Paul S. Appelbaum M.D

Professor and Chair
Department of Psychiatry, University of Massachusetts Medical School, Worcester, Massachusetts

Jean-Michel Arrang Ph.D.

Research Director

Unité de Neurobiologie et Pharmacologie Moléculaire, INSERM, Paris, France

Gary Aston-Jones Ph.D.
Department of Psychiatry, University of Pennsylvania, VAMC, Philadelphia, Philadelphia

C. Bruce Baker M.D., J.D.

Assistant Professor
Department of Psychiatry, Yale University School of Medicine, Connecticut Mental Health Center, New Haven, Connecticut

Vaishali P. Bakshi Ph.D
Department of Psychiatry, University of Wisconsin at Madison, Madison, Wisconsin

Peter A. Bandettini Ph.D

Director
Functional Neuroimaging Facility, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland

Robert F. Barnes M.D
University of Washington School of Medicine, Department of Psychiatry and Behavioral Sciences, Seattle, Washington

Samuel H. Barondes M.D

Jeanne and Sanford Robertson M.D. Director
Center for Neurobiology and Psychiatry, Department of Psychiatry, University of California, San Francisco, San Francisco, California

Vincenzo S. Basile B.Sc., B.Ed.

Research Scientist
Psychiatric Neurogenetics, Centre for Addiction and Mental Health, Clarke Division, Toronto, Ontario, Canada

Russell Bauer Ph.D.
University of Florida, Health Science Center, Department of Clinical Psychology, Gainesville, Florida

Kevin L. Behar

Neal L. Benowitz M.D
Department of Psychiatry, University of California, San Francisco, San Francisco, California

Karen F. Berman M.D

Chief
Unit on Integrative Neuroimaging, Clinical Brain Disorders Branch, National Institute of Mental Health, National Institutes of Health, Intramural Research Program, Bethesda, Maryland

Wade Berrettini M.D., Ph.D.
University of Pennsylvania School of Medicine, Center for Neurobiology and Behavior, Philadelphia, Pennsylvania

Edward H. Bertram III M.D.

Associate Professor
Department of Neurology, University of Virginia, Charlottesville, Virginia

Joseph Biederman M.D.

Chief
Joint Program in Pediatric Psychopharmacology, Massachusetts General and McLean Hospitals; Professor, Department of Psychiatry, Massachusssets General Hospital, Boston, Massachusetts

Randy D. Blakely M.D., Ph.D
Department fo Pharmacology, Vanderbilt University, Nashville, Tennessee

Daryl E. Bohning Ph.D

Professor
Department of Radiology, Medical University of South Carolina, Charleston, South Carolina

Robert J. Boland M.D
Miriam Hospital, Department of Psychiatry, Providence, Rhode Island

David L. Braff M.D
Department of Psychiatry, University of California, San Diego, La Jolla, California

James Robert Brasic M.D., M.P.H.

Postdoctoral Fellow
Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, Maryland

Monique Breteler M.D., Ph.D.
Department of Epidemiology and Biostatistics, Erasmus Medical Center, Rotterdam, Netherlands

Michael Brownstein M.D.
National Institutes of Health, National Institute of Mental Health, Bethesda, Maryland

Alastair Buchan FRCPC
Department of Clinical Neuroscience, Foothills Hospital, Calgary, Alberta, Canada

Christian Büchel M.D.
Department of Neurology, Hamburg University, Hamburg, Germany

Robert E. Burke M.D

Professor of Neurology and Pathology
Departments of Neurology and Pathology, Columbia University, New York, New York

David J. Burn F.R.C.P., M.D., M.A., M.B.B.S

Consultant and Service Lecturer in Neurology
Department of Neurology, Regional Neurosciences Centre, Newcastle upon Tyne, United Kingdom

Alistair Burns M.D.
Department of Old Age Psychiatry, Withington Hospital, West Disbury, Manchester, United Kingdom

Joseph D. Buxbaum Ph.D

Assistant Professor of Psychiatry
Department of Psychiatry, Mount Sinai School of Medicine, New York, New York

Daniel Buysee M.D.

Associate Professor of Psychiatry
University of Pittsburgh Health System, WPIC, Pittsburgh, Pennsylvania

William A. Carlezon Jr. Ph.D

Assistant Professor of Psychiatry
Harvard Medical School, McLean Hospital, Belmont, Massachusetts

James V. Cassella Ph.D
Neurogen Corporation, Branford, Connecticut

V. S. Caviness Jr. CMA
MGH-East, Charlestown, Massachusetts

Britton Chance M.D.
Department of Biophysics and Physical Biochemistry, University of Pennsylvania, Johnson Research Foundation, Philadelphia, Pennsylvania

Dennis Charney M.D.

Chief
Mood and Anxiety Disorder Research Program, National Institute of Mental health, Bethesda, Maryland

A. R. Childress
University of Pennsylvania School of Medicine, Department of Psychiatry, Addiction Treatment, Philadelphia, Pennsylvania

Emil F. Coccaro M.D
The University of Chicago, The Pritzker School of Medicine, Department of Psychiatry, Chicago, Illinois

C. Keith Conners M.D.
Behavioral Neurology Department, Durham, North Carolina

Jeremy D. Coplan M.D
Columbia University, New York Psychiatric Institute, New York, New York

Nancy H. Covell Ph.D

Project Director
Research Division, Connecticut Department of Mental Health and Addiction Services, Hartford, Connecticut, and, Research Associate, Department of Psychology, University of Connecticut, Storrs, Connecticut

Joseph T. Coyle M.D

Eben S. Draper Professor of Psychiatry and Neuroscience; Chairman

Department of Psychiatry, Harvard Medical School, Boston, Massachusetts

Merit E. Cudkowicz

Anders M. Dale M.D
NMR Center, Charlestown, Massachusetts

Richard Davidson M.D.
Department of Psychology, University of Wisconsin, Madison, Wisconsin

Michael Davidson
Sheba Medical Center, Tel Aviv, Israel

Linda Davies M.Sc.
Health Economics, Reader, Health Economics, School of Psychiatry and Behavioural Sciences, Univesity of Manchester, Manchester, United Kingdom

Kenneth L. Davis M.D

Esther and Joseph Klingenstein Professor of Psychiatry and, Chairman
Department of Psychiatry, Mount Sinai School of Medicine, New York, New York

Michael Davis Ph.D.
Department of Psychiatry, Emory University School of Medicine, Atlanta, Georgia

Mahlon R. DeLong M.D
Emory University, Department of Neurology, Atlanta, Georgia

Robert Desimone Ph.D.

Scientific Director
National Institute of Mental Health; Director, Intramural Research Program/National Institute of Mental Health, Bethesda, Maryland

Errol B. DeSouza Ph.D
Research and Development, Neurocrine Bioscience, Inc., San Diego, California

William L. Dewey
Virginia Commonwealth University, Department of Pharmacology and Toxicology, Richmond, Virginian

Vincenzo M. Di Marzo Ph.D.
Primo Richercatore, Istituto Chimica Biomoleculare, Consiglio Nazionale Ricerche, Pozzuoli (NA), Italy

David F. Dinges Ph.D

Professor of Psychology in Psychiatry
Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania

Cynthia M. Dorsey Ph.D

Assistant Professor
Department of Psychiatry, Harvard Medical School, Boston, Massachusetts, and, Director, Sleep Research Laboratory, Behavioral Psychopharmacology Research Laboratory, McLean Hospital, Belmont, Massachusetts

Wayne C. Drevets M.D

Chief
Section on Mood and Anxiety Disorders Imaging, Molecular Imaging Branch, National Institute of Mental Health, Bethesda, Maryland

Karen Duff Ph.D.
Nathan Kline Institute, Dementia Research Program, Orangeburg, New York

Ronald S. Duman Ph.D

Professor of Psychiatry and Pharmacology, and, Director
Abraham Ribicoff Research Facilities, Departments of Psychiatry and Pharmacology, Yale University School of Medicine, New Haven, Connecticut

Gary E. Duncan Ph.D

Research Associate Professor
Department of Psychiatry, University of North Carolina, Chapel Hill, North Carolina

Jane L. Eisen M.D

Assistant Professor
Department of Psychiatry and Human Behavior, Brown Medical School, Providence, Rhode Island

Mary-Anne Enoch M.D.

Visiting Associate
Laboratory of Neurogenetics, NIAAA, National Institutes of Health, Bethesda, Maryland

Charles M. Epstein M.D
Emory University, Department of Neurology, Atlanta, Georgia

Susan M. Essock Ph.D
Mount Sinai School of Medicine, Department of Psychiatry, New York, New York

Alan Evans M.D.
Department of Neurology, McGill University, Montreal Neurological Institute, Montreal, Quebec, Canada

Dwight L. Evans M.D
Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania

StephenV. Faraone Ph.D.

Associate Professor
Department of Psychiatry, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts

Marian W. Fischman Ph.D

Professor of Behavioral Biology
Department of Psychiatry, College of Physicians and Surgeons of Columbia University and, The New York Psychiatric Institute, New York, New York

Joanna S. Fowler Ph.D

Senior Chemist
Chemistry Department, Brookhaven National Laboratory, Upton, New York

Teresa R. Franklin Ph.D

Substance Abuse Fellow
Department of Psychiatry, University of Pennsylvania, and, Philadelphia Veterans Affairs Medical Center, Philadelphia, Pennsylvania

Alan Frazer Ph.D.

Chair and Professor
Department of Pharmacology, University of Texas Health Science Center, San Antonio, Texas

Robert Freedman M.D.

Professor
Department of Psychiatry and Pharmacology, University of Colorado, Denver, Colorado

Nelson B. Freimer M.D
University of California, San Francisco, Department of Psychiatry, San Francisco, California

Linda K. Frisman Ph.D

Director of Research
Research Division, Connecticut Department of Mental Health and Addiction Services, Hartford, Connecticut, and, Research Professor, Department of Psychology, University of Connecticut, Storrs, Connecticut

Karl Friston
Wellcome, Department of Cognitive Neurology, London, United Kingdom

Paul J. Fudala Ph.D

Research Associate Professor of Pharmacology in Psychiatry
Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania, and, Clinical Toxicologist, Department of Behavioral Health Service, Veterans Affairs Medical Center, Philadelphia, Pennsylvania

Masahiro Fujita M.D., Ph.D.

Assistant Professor
Department of Psychiatry, Yale University School of Medicine, West Haven, Connecticut

Fred H. Gage Ph.D

Professor
Laboratory of Genetics, The Salk Institute, La Jolla, California

Stephen J. Garlow M.D., Ph.D

Assistant Professor
Department of Psychiatry and Behavioral Science, Emory University, Atlanta, Georgia

L. Garrido
Massachusetts General Hospital-East, Charlestown, Massachusetts

Nori Geary Ph.D.

Research Professor
Department of Psychiatry, Weill Medical College of Cornell University, New York, New York

Mark S. George M.D
Departments of Psychiatry, Radiology and Neurology, Medical University of South Carolina, Charleston, South Carolina

Mark A. Geyer Ph.D

Professor
Department of Psychiatry, University of California, San Diego, La Jolla, California

Patricia S. Godman-Rakic Ph.D.
Department of Neurobiology, Yale University School of Medicine, New Haven, Connecticut

Donald C. Goff M.D

Director
Schizophrenia Program, Massachusetts General Hospital, Boston, Massachusetts, and, Associate Professor, Department of Psychiatry, Harvard Medical School, Boston, Massachusetts

Terry E. Goldberg Ph.D
National Institutes of Mental Health, Bethesda, Maryland

David Goldman M.D.

Chief
Laboratory of Neurogenetics, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Rockville, Maryland

Patricia S. Goldman-Rakic Ph.D.
Department of Neurobiology, Yale University School of Medicine, New Haven, Connecticut

Jack M. Gorman M.D
Columbia University, Department of Psychiatry, New York, New York

Anthony A. Grace M.D
University of Pittsburgh, Department of Neuroscience and Psychiagtry, Pittsburgh, Pennsylvania

Jean-Michel Gracies M.D., Ph.D.

Assistant Professor
Department of Neurology, The Mount Sinai Medical Center, New York, New York

Steven H. Graham M.D., Ph.D

Associate Professor and Vice-Chair
Department of Neurology, University of Pittsburgh, Pittsburgh, Pennsylvania, and, Associate Director for Research, Department of Geriatric Research Education and Clinical Center, Veterans Affairs Pittsburgh Health System, Pittsburgh, Pennsylvania

Michael F. Green Ph.D

Professor
Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, California

Paul E. Greenberg M.A., M.S
Analysis Group, Inc., Cambridge, Massachusetts

David J. Greenblatt M.D
Dept of Pharmacology and Experimental Therapeutics, Tufts University School of Medicine, Boston, Massachusetts

Dimitri E. Grigoriadis Ph.D

Senior Director
Department of Pharmacology and Lead Discovery, Neurocrine Biosciences, Inc., San Diego, California

Amy Grogg PharM.D.

Associate Director
CNS Outcomes Research, Department of Outcomes Research, Janssen Pharmacetica, Titusville, New Jersey

Jerold S. Harmatz B.A

Assistant Professor of Medicine
Department of Pharmacology and Experimental Therapeutics, Tufts University School of Medicine, Boston, Masachusetts

Vahram Haroutunian Ph.D.

Associate Professor
Department of Psychiatry, Mount Sinai School of Medicine and, Bronx Veterans Affairs Medical Center, New York, New York

James C. Harris M.D

Director
Developmental Neuropsychiatry, Professor of Psychiatry and Behavioral Sciences and Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland

Philip D. Harvey Ph.D

Professor of Psychiatry
Mount Sinai School of Medicine, New York, New York

Robert W. Hickey M.D

Assistant Professor
Department of Pediatrics, Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania

Makoto Higuchi Ph.D.
University of Pennsylvania, School of Medicine, Department of Pathology and Laboratory Medicine, Philadelphia, Pennsylvania

Steven A. Hillyard M.D
University of California, San Diego, Department of Neuroscience, La Jolla, California

Robert M. A. Hirschfeld M.D.

Titus H. Harris Chair; Professor and Chair
Department of Psychiatry and Behavioral Science, University of Texas Medical Branch, Galveston, Texas

J. Allan Hobson M.D.

Professor of Psychaitry
Massachusetts Mental Health Center, Boston, Massachusetts

Eric Hollander M.D.

Professor of Psychiatry; Director of Clinical Psychopharmacology; Director of the Compulsive, Impulsive and Anxiety Disorders

Program; Clinical Director Seaver Autism Research Center
Department of Psychiatry, Mount Sinai School of Medicine, New York, New York

Fahmeed Hyder
Department of Diagnostic Radiology, Yale University, School of Medicine, New Haven, Connecticut

Steven E. Hyman M.D
National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland

Robert B. Innis M.D

Professsor
Department of Psychiatry and Pharmacology, Yale University School of Medicine, West Haven, Connecticut

Bruce G. Jenkins Ph.D

Director
Neurochemical Imaging, MGH-NMR Center/Radiology, Massachusetts General Hospital, Charlestown, Massachusetts

J. David Jentsch Ph.D.
Department of Neuroscience, University of Pittsburgh, Pittsburgh, Pennsylvania

David S. Johnson M.D., Ph.D
Nina Ireland Laboratory of Developmental Neurobiology, Department of Psychiatry, Langley Porter Psychiatric Institute, University of California, San Francisco, San Francisco, California

Reese T. Jones M.D
Langely Porter Psychiatric Institute, University of California, San Francisco, San Francisco, California

Paramjit Joshi M.D.

Chair
Department of Psychiatry, Children's National Medical Center, Washington, DC

Ned H. Kalin M.D

Professor and Chair
Department of Psychiatry, University of Wisconsin Medical School, Madison, Wisconsin

Peter W. Kalivas Ph.D
Department of Physiology and Neuroscience, Medical University of South Carolina, Charleston, South Carolina

John Kane M.D.
Hillside Hospital, Glen Oaks, New York

Shatij Kapur M.D., Ph.D.
Department of Psychiatry, Clark Institute; University of Toronto, Ontario, Canada

Walter Kaye M.D.
Western Psychiatric Institute & Clinic, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania

Paul E. Keck Jr. M.D.
c/o Husseini K. Manji, M.D Department of Psychiatry, Wayne State University School of Medicine, Detroit, Michigan

John H. Kehne Ph.D

Executive Director
Department of Behavioral Biology, Neurogen Corporation, Branford, Connecticut

Martin B. Keller M.D

Professor and Chairman
Brown University School of Medicine, Butler Hospital, Providence, Rhode Island

David Kennedy M.D.
Morphometric Analysis Laboratory, Neuroscience Center, Charlestown, Massachusetts

James L. Kennedy M.D., FRCP(C)

Head
Neurogenetics Section, Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

Justine M. Kent M.D
Columbia University, Department of Psychiatry, New York, New York

Matcheri S. Keshavan M.D

Professor of Psychiatry
Department of Psychiatry, Western Psychiatric Institute and Clinic, Pittsburgh, Pennsylvania

Ronald C. Kessler Ph.D
Harvard Medical School, Boston, Massachusetts

Clint Kilts M.D.
Emory University School of Medicine, Department of Psychiatry, Atlanta, Georgia

Joel E. Kleinman M.D., Ph.D
Clinical Brain Disorders Branch, National Institutes of Mental Health Neuroscience Center, Washington, DC

Violetta Klimek Ph.D.

Assistant Professor
Department of Psychiatry, University of Mississippi Medical Center, Jackson, Mississippi

Jacqueline D. Kloss Ph.D

Assistant Professor of Psychology
Department of Psychology, Sociology, Anthropology, Drexel University, and, Adjunct Assistant Professor of Psychology, Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania

George F. Koob Ph.D
Department of Neuropharmacology, The Scripps Research Institute, La Jolla, California

Kathy L. Kopnisky
National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland

Thomas R. Kosten M.D

Professor of Psychiatry

Department of Psychiatry, Yale University School of Medicine, and, Chief, Department of Psychiatry, Veterans Affairs Connecticut Healthcare System, West Haven, Connecticut

Chris M. Kozma Ph.D

Senior Research Director
Strategic Outcomes Services of Care Science, Inc., Philadelphia, Pennsylvania

Mark S. Kramer M.D., Ph.D

Senior Director
Department of Clinical Neuroscience, Merck & Co., West Point, Pennsylvania, and, University of Pennsylvania, Philadelphia, Pennsylvania

Mary Jeanne Kreek M.D.
The Rockefeller University, Laboratory of the Biology of Addictive Diseases, New York, New York

Suchitra Krishnan-Sarin Ph.D.
Yale University School of Medicine, Department of Psychiatry, New Haven, Connecticut

John H. Krystal M.D

Psychiatry Service
Veterans Affairs Connecticut Healthcare System, West Haven, Connecticut

David Kupfer M.D.

Thomas Detre Professor and Chair
Department of Psychiatry, University of Pittsburgh Medical Center, Western Psychiatric Institute and Clinic, Pittsburgh, Pennsylvania

Marta Kutas Ph.D.

Professor of Cognitive Science and Neurosciences
Department of Cognitive Science, University of California, San Diego, La Jolla, California

Ariel J. Lang Ph.D

Assistant Professor in Residence
Department of Psychiatry, University of California, San Diego, San Diego, California, and, Psychologist, Veterans Affairs San Diego Health System, San Diego, California

Marc Laruelle M.D.
New York Psychiatric Institute, Department of Psychiatry, New York, New York

Paul Leber M.D.

Director
Neuro-Pharm Group, LLC, Potomac, Maryland

James F. Leckman M.D

Neison Harris Professor of Child Psychiatry and Pediatrics
Child Study Center, Yale University School of Medicine, New Haven, Connecticut

Virginia M.-Y. Lee Ph.D.

The John H. Ware 3rd Professor in Alzheimer's Research and Co-Director
Center for Neurodegenerative Disease Research, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

Robert H. Lenox M.D

Vice President
Head CNS Group, Aventis Pharmaceuticals, Bridgewater, New Jersey

Jane Leserman Ph.D.

Research Associate Professor
Department of Psychiatry, University of North Carolina School of Medicine, Chapel Hill, North Carolina

Michael L. Leski Ph.D

Instructor
Department of Psychiatry, McLean Hospital, Belmont, Massahusetts

David A. Lewis M.D

Professor
Department of Psychiatry and Neuroscience, University of Pittsburgh, Pittsburgh, Pennsylvania

Alfred J. Lewy M.D
Oregon Health Sciences University, Sleep & Mood Disorders Laboratory, Portland, Oregon

Jeffrey A. Lieberman M.D
Department of Psychiatry, University of North Carolina Chapel Hill, Chapel Hill, North Carolina

J. Listerud
University of Pennsylvania School of Medicine, Department of Psychiatry, Addiction Treatment, Philadelphia, Pennsylvania

Shauna N. MacMillan B.Sc.

Research Assistant
Department of Psychiatry and Behavioral Neurosciences, Wayne University School of Medicine, Detroit, Michigan

Pierre Magistretti M.D., Ph.D.
Inst de Physiologie, Univ Lausanne, Lausanne, Switzerland

Ramin Mahmoud M.D.

Director
Outcome Research, Janssen Research Foundation, Titusville, New Jersey

Nikos Makris M.D., Ph.D.

Instructor
Department of Neurology, Harvard Medical School, Boston, Massachusetts

Robert C. Malenka M.D., Ph.D
Department of Psychiatry and Behavioral Science, Stanford University School of Medicine, Palo Alto, California

Andrea Manca M.Sc.

Research Fellow
Centre for Health Economics, University of York, York, United Kingdom

Husseini K. Manji M.D
Department of Psychiatry, Wayne University School of Medicine, Detroit, Michigan

John D. Mann M.D
Department of Neurology, University of North Carolina, Chapel Hill, School of Medicine, Chapel Hill, North Carolina

J. John Mann M.D.

Chief of Neuroscience
Department of Psychiatry, Columbia University, New York, New York

Russell L. Margolis M.D

Associate Professor and Attending Psychiatrist
Department of Psychiatry, Johns Hopkins University School of Medicine and Hospital, Baltimore, Maryland

Athina Markou
The Scripps Research Institute, Department of Neuropharmacology, La Jolla, California

Billy R. Martin Ph.D
Department of Pharmacology and Toxicology, Virginia Commonwealth University, Richmond, Virginia

Mario Masellis B.Sc., M.Sc.

Research Scientist
Department of Psychiatric Neurogenetics, Clarke Division, Centre for Addiction and Mental Health, Toronto, Ontario, Canada

Karen I. Mason Ph.D

Neuropsychology Fellow
Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, School of Medicine, Los Angeles, California

Graeme F. Mason

Helen S. Mayberg M.D
Baycrest Center, Rotman Research Institute, Toronto, Ontario, Canada

W. Vaughn McCall M.D.

Professor
Department of Psychiatry and Behavioral Medicine, Wake Forest University School of Medicine, Winston-Salem, North Carolina

Una D. McCann M.D
Mt. Auburn Hospital, Department of Psychiatry and Behavioral Sciences, Lexington, Massachusetts

Robert W. McCarley M.D
Harvard Medical School, Department of Psychiatry, Brockton, Massachusetts

William M. McDonald M.D
Department of Psychiatry, Wesley Woods Health Center, Atlanta, Georgia

Christopher J. McDougle M.D

Raymond E. Houk Professor of Psychiatry; Professor of Pediatrics and Neurobiology
Department of Psychiatry, Indiana University School of Medicine, Indianapolis, Indiana

L. Alison McInnes M.D., M.S.

Assistant Adjunct Professor
Department of Psychiatry and Neurogenetics, University of California, San Francisco, San Francisco, California

Ian G. McKeith M.D
Department of Old Age Psychiatry, Wolfson Research Centre, Newcastle General Hospital, Newcastle upon Tyne, United Kingdom

Thomas McLellan Ph.D.
Philadelphia Veterans Affairs Medical Center, Philadelphia, Pennsylvania

Gavan P. McNally

Research Fellow
Mental Health Research Institute, University of Michigan, Ann Arbor, Michigan

James H. Meador-Woodruff M.D.

Senior Associate Research Scientist
Mental Health Research Institute, and, Associate Professor, Department of Psychiatry, University of Michigan, Ann Arbor, Michigan

Brian Meissner PharM.D.

Janssen Outcomes Fellow
Department of Pharmacy, University of South Carolina, Columbia, South Carolina

Herbert Meltzer M.D.

Bixler Professor of Psychiatry and Pharmacology
Vanderbilt University School of Medicine, Nashville, Tennessee

Wallace B. Mendelson M.D

Professor
Departments of Psychiatry, Medicine, and Clinical Pharmacology, University of Chicago, Chicago, Illinois

Kathleen Merikangas Ph.D.

Professor of Epidemiology
Public Health and Psychiatry, Yale University School of Medicine, New Haven, Connecticut

Emmanuel Mignot M.D., Ph.D.

Associate Professor
Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

Mark A. Mintun M.D
Washington University School of Medicine, Department of Radiology, St. Louis, Missouri

Seiya Miyamoto M.D., Ph.D.

Visiting Research Fellow
Department of Psychiatry, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, and, Department of Neuropsychiatry, St. Marianna University School of Medicine, Kawasaki, Japan

Bita Moghaddam
Yale University School of Medicine, Department of Psychiatry, West Haven, Connecticut

Richard C. Mohs Ph.D
Research Service, Veterans Affairs Medical Center, Bronx, New York

Paul Morley M.D.
Receptors and Ion Channels Group, Institute of Biological Science, National Research Council of Canada, Ottawa, Ontario, Canada

John H. Morrison Ph.D
Mount Sinai Medical Center, New York, New York

Pierandrea Muglia M.D.

Postdoctoral Fellow
Neurogenetics Section, Centre for Addiction and Mental Health, Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

Emanuela Mundo M.D.

Assistant Professor
Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada

Charles Nemeroff M.D., Ph.D.

Reunette W. Harris Professor and Chairman
Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia

Eric J. Nestler M.D., Ph.D

Lou and Ellen McGinley Distinguished Professor and Chairman
Department of Psychiatry, The University of Texas Southwestern Medical Center at Dallas, Dallas, Texas

Rachael Neve M.D.
McLean Hospital, Belmont, Massachusetts

Seiji Nishino M.D., Ph.D.

Senior Research Scientist
Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California

Ralph A. Nixon M.D., Ph.D
Nathan Kline Institute, Dementia Research Program, Orangeburg, New York

Eric Nofzinger M.D.
University of Pittsburgh Health System, Department of Psychiatry, Pittsburgh, Pennsylvania

Charles P. O'Brien M.D., Ph.D.

Chief of Psychiatry
Department of Behavioral Health, Philadelphia Veterans Affairs Medical Center, and, Kenneth Appel Professor and Vice Chair, Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania

John T. O'Brien D.M.
Department of Old Age Psychiatry, Wolfson Research Centre, Newcastle General Hospital, Newcastle upon Tyne, United Kingdom

Stephanie S. O'Malley Ph.D.
Yale University School of Medicine, Department of Psychiatry, New Haven, Connecticut

C. Warren Olanow M.D., F.R.C.P.L.

Professor and Chairman
Department of Neurology, Mount Sinai Medical Center, New York, New York

Berend Olivier Ph.D.
Department of Psychopharmacology, Utrecht University, Utrecht, Netherlands, and, Department of Psychiatry, Yale University School of Medicine, Hawthorne, New York

Richard W. Olsen Ph.D

Professor of Neuroscience, Pharmacology, and Anesthesiology
Department of Molecular and Medical Pharmacology, University of California Los Angeles School of Medicine, Los Angeles, California

Gregory A. Ordway Ph.D

Division of Neurobiology and Behavior Research
Department of Psychiatry, University of Mississippi Medical Center, Jackson, Mississippi

Michael J. Owen Ph.D
University of Wales College of Medicine, Department of Psychological Medicine, Cardiff, United Kingdom

Vural Özdemir M.D., Ph.D.

Senior Scientist
Department of Pharmacogenetics, R. W. Johnson Pharmaceutical Research Institute, Raritan, New Jersey

Edward F. Pace-Schott M.S., M.A., L.M.H.C.

Instructor
Department of Psychiatry, Harvard Medical School, Boston, Massachusetts

Stefano Pallanti M.D.
Mount Sinai School of Medicine, Department of Psychiatry, New York, New York

G. M. Papadimitriou
Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts

S. Parvathy Ph.D.

Postdoctoral Fellow
Department of Psychiatry, Mount Sinai School of Medicine, New York, New York

David L. Pauls Ph.D
Child Study Center, Yale University School of Medicine, New Haven, Connecticut

Robert H. Perry M.B., Ch.B.
Department of Neuropathology, Newcastle General Hospital, Newcastle upon Tyne, United Kingdom

Elaine Perry M.D.
MRC Neurochemical Pathology Unit, Newcastle General Hospital, Newcastle upon Tyne, United Kingdom

John Petitto M.D.
Department of Psychiatry, University of Florida, College of Medicine, Gainesville, Florida

Ognen A. C. Petroff

Michael Phelps M.D.
Department of Molecular and Medical Pharmacology, Center Health Services, University of California, Los Angeles, School of Medicine, Los Angeles, California

Marina Picciotto M.D.
Department of Psychiatry, Yale University School of Medicne, New Haven, Connecticut

Debra A. Pinals M.D

Director; Fellowship in Law and Psychiatry
Department of Psychiatry, University of Massachusetts Medical School, Worcester, Massachusetts

Daniel S. Pine M.D

Associate Professor of Clinical Psychiatry
Department of Psychiatry, Division of Child and Adolescent Psychiatry, Columbia University / New York State Psychiatric Institute, New York, New York

Joseph Piven M.D.
Department of Psychiatry, University of North Carolina, Chapel Hill, Chapel Hill, North Carolina

Mark Potenza M.D., Ph.D.

Director
Problem Gambling Clinic and, Assistant Professor, Department of Psychiatry, Yale University School of Medicine and, Connecticut Mental Health Center, New Haven, Connecticut

William Z. Potter M.D., Ph.D

K. Ranga Rama Krishnan M.D.
Department of Psychiatry, Duke University Medical Center, Durham, North Carolina

Bruce R. Ransom M.D., Ph.D

Warren Magnuson Professor and Chair
Department of Neurology, University of Washington, Seattle, Washington

Murray A. Raskind M.D
Psychiatry Service, Seattle Veterans Affairs Medical Center, Seattle, Washington

Steven A. Rasmussen M.D
Brown University, Butler Hospital, Providence, Rhode Island

Scott L. Rausch M.D

Associate Professor
Massachusetts General Hospital, Charlestown, Massachusetts

Assistant Professor
Section of Neurobiology, University of California, San Diego, La Jolla, California

Helen E. Scharfman Ph.D

Director
Center for Neural Recovery and Rehabilitation Research, Helen Hayes Hospital, West Haverstraw, New York, and, Associate Professor, Departments of Pharmacology and Neurology, Columbia University, New York, New York

Alan F. Schatzberg M.D
Department of Psychiatry, Atlanta, Georgia

Peter J. Schmidt M.D

Chief
Unit on Reproductive Endocrinology, Behavioral Endocrinology Branch, National Institute of Mental Health, Bethesda, Maryland

Marc Alan Schuckit M.D.

Professor of Psychiatry
Department of Psychiatry, San Diego Veterans Affairs Medical Center and, University of California, San Diego Medical School, San Diego, California

Ann C. Schulte Ph.D

Associate Professor
Department of Psychology, North Carolina State University, Raleigh, North California

Robert Schwarcz Ph.D.

Director
Neuroscience Program, Maryland Psychiatric Research Center, Baltimore, Maryland

J. C. Schwartz M.D.
Unite de Neurobiologie et Pharmacol., Centre Paul Broca de l'inserm, Paris, France

Richard I. Shader M.D
Newton Centre, Massachusetts

Yvette Sheline M.D.
Department of Psychiatry, Washington University of St. Louis, St. Louis, Missouri

Jun Shen

Lisa M. Shin Ph.D

Assistant Professor
Department of Psychology, Tufts University, Medford, Massachusetts

Toni S. Shippenberg Ph.D
National Institutes of Health, National Institute on Drug Abuse, Integrative Neurosciences, Bethesda, Maryland

Robert G. Shulman

Nicola Sibson

Larry J. Siever M.D
Bronx Veterans Affairs Medical Center, Psychiatry Service, Bronx, New York

Daphne Simeon M.D.

Assistant Professor
Department of Psychiatry, Mt. Sinai School of Medicine, New York, New York

Daniel L. Small Ph.D

Receptors and Ion Channels Group
Institute for Biological Sciences, National Research Council of Canada, Ottawa, Ontario, Canada

Gary W. Small M.D

Director
Center on Aging, Parlow-Solomon Professor on Aging, Professor, Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Neuropsychiatric Institute, Los Angeles, California

Gerard P. Smith M.D

Professor of Psychiatry (Behavioral Neuroscience)
Department of Psychiatry, Weill Medical College of Cornell University, New York, New York

A. Gregory Sorensen M.D.

Assistant Radiologist
Department of Radiology / NMR Center, Massachusetts General Hospital, and, Assistant Professor in Radiology, Department of Radiology, Harvard Medical School, Boston, Massachusetts

Dan J. Stein M.B.

Director
MRC Research Unit on Anxiety Disorders, Department of Psychiatry, University of Stellenbosch, Cape Town, South Africa

Murray B. Stein M.D
Anxiety & Traumatic Stress Disorders, Department of Psychiatry, University of California, San Diego, La Jolla, California

William S. Stone Ph.D

Assistant Professor
Harvard University, Department of Psychiatry, Massachusetts Mental Health Center, Boston, Massachusetts

James S. Sutcliffe Ph.D

Assistant Professor
Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee

Neal R. Swerdlow M.D

Professor
Department of Psychiatry, University of California, San Diego, School of Medicine, La Jolla, California

Martin P. Szuba M.D

Assistant Professor
Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania

Boris Tabakoff Ph.D.
Department of Pharmacology, University of Colorado School of Medicine, Denver, Colorado

John F. Tallman Ph.D

Scientific Director
Neurogen Corporation, Branford, Connecticut

Carol A. Tamminga M.D

Professor of Psychiatry and Pharmacology
University of Maryland School of Medicine, and, Deputy Director, Maryland Psychiatric Research Center, Baltimore, Maryland

Laurence H. Tecott M.D., Ph.D
Department of Psychiatry, University of California, San Francisco, San Francisco, California

Kenneth E. Towbin M.D
The George Washington University, School of Medicine, Department of Psychiatry, Washington, DC

John Q. Trojanowski M.D., Ph.D
University of Pennsylvania School of Medicine, Department of Pathology and Laboratory Medicine, Philadelphia, Pennsylvania

Ming T. Tsuang M.D., Ph.D

Stanley Cobb Professor of Psychiatry; Director
Harvard Institute of Psychiatry Genetics, Head, Department of Psychiatry, Harvard Medical School, Massachusetts Mental Health Center, Boston, Massachusetts

David S. Tuch B.A

(Physics); Graduate Research Assistant
Massachusetts General Hospital, NMR Center, Charlestown, Massachusetts

Henriette van Praag M.D.

Research Associate
Laboratory of Genetics, The Salk Institute for Biological Studies, La Jolla, California

Gabriel A. Vargas M.D., Ph.D

Research Fellow
Department of Psychiatry, University of California at San Francisco, San Francisco, California

Veronica J. Vieland Ph.D

Professor
Department of Biostatistics and Psychiatry, University of Iowa, Iowa City, Iowa

Nora D. Volkow M.D

Professor
Department of Psychiatry, SUNY, Stony Brook, Stony Brook, NY, and, Medical Department, Brookhaven National Laboratory, Upton, New York

Joseph Volpicelli M.D.
Department of Psychiatry, University of Pennsylvania, Veterans Affairs Medical Center, Philadelphia, Pennsylvania

Lisa L. von Moltke M.D.
Tufts University School of Medicine, Department of Pharmacology and Experimental Therapeutics, Boston, Massachusetts

Mark Von Zastrow M.D.
University of California, San Francisco, Department of Psychiatry, San Francisco, California

B. Timothy Walsh M.D.

Ruane Professor of Pediatric Psychopharmacology
College of Physicians and Surgeons, Columbia University; Director, Eating Disorders Research Unit, New York State Psychiatric Institute, New York, New York

Thomas H. Wassink M.D

Assistant Professor
Department of Psychiatry, University of Iowa College of Medicine, Iowa City, Iowa

V. J. Wedeen
MGH-East, NMR Center, Charlestown, Massachusetts

Daniel R. Weinberger M.D

Chief
Clinical Brain Disorders Branch, National Institutes of Health, National Institute of Mental Health, Bethesda, Maryland

Myrna M. Weissman Ph.D

Professor of Psychiatry and Epidemiology
Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York

Thomas Wichmann M.D.

Assistant Professor
Department of Neurology, Emory University, Atlanta, Georgia

Michael Williams Ph.D., DSc

Divisional Vice President
Neurological and Urological Diseases Research, Pharmaceuticals Products Division, Abbott Laboratories, Abbott Park, Illinois

Margaret G. Woerner Ph.D

Associate Director
Psychiatry Research, Hillside Hospital, North Shore-Long Island Jewish Health System, Glen Oaks, New York, Associate Professor, Department of Psychiatry, Albert Einstein College of Medicine, Bronx, New York

Dean F. Wong M.D., Ph.D

Radiology Vice-Chair for Research Administration and Training; Professor of Radiology, Psychiatry, and Environmental Health

Sciences
Johns Hopkins Medical Institutions, Baltimore, Maryland

Scott W. Woods M.D
Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut

George Woody M.D.
Veterans Affairs Medical Center, Substance Abuse Treatment Unit, Philadelphia, Pennsylvania

Audrey J. Worth
CMA, MGH-East, Charlestown, Massachusetts

Ona Wu M.S.

Research Assistant
Department of Radiology, MGH NMR Center, Charlestown, Massachusetts

Larry J. Young Ph.D

Associate Professor
Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, Georgia

Michael J. Zigmond Ph.
Department of Neuroscience, University of Pittsburgh, Pittsburgh, Pennsylvania

Joseph Zohar

Associate Professor
Department of Psychiatry, Sheba Medical Center, Tel Hashomer, Israel

PREFACE

Neuropsychopharmacology: The Fifth Generation of Progress appears at an important moment in the history of psychopharmacology. We have recently ended the decade of the brain, a decade that witnessed enormous progress in understanding fundamental physiology of the central nervous system. The fruits of these basic science discoveries have already resulted in important progress in the treatment of mental illness. The importance of these fundamental discoveries has recently been acknowledged by the awarding of the Nobel Prize in Psychology or Medicine to three members of the College, Arvid Carlsson, Paul Greengard and Eric Kandel for their discoveries on neuronal signaling. This edition in the *Generation of Progress* series details advances in both the basic science and clinical application of recent research in psychopharmacology. It also demonstrates the prospects for even greater advances in the future.

Charles P. O'Brien M.D., Ph.D.

ACNP President, 2001

Joseph T. Coyle M.D.

ACNP President, 2002

Contents

Section I
Neurotransmitters and Signal Transduction

Eric J. Nestler

Ronald S. Duman

Neurotransmitters and Signal Transduction - Introduction

Neuropsychopharmacology continues to be organized primarily according to the neurotransmitters that are utilized by various populations of neurons for synaptic transmission. This is because the vast majority of psychotropic drugs presently used clinically to treat neuropsychiatric disorders still have as their initial targets proteins that regulate the availability of a particular neurotransmitter (e.g., presynaptic reuptake transporters, synthetic or degradative enzymes) or that serve as ligands for particular neurotransmitter receptors. It is entirely appropriate then that this edition of *Neuropsychopharmacology: A Generation of Progress* begins with a section devoted to the major neurotransmitter systems in the brain. Rather than provide a comprehensive review of the now vast literature on neurotransmitter systems, the goal of this section is to highlight recent advances in the field.

Glutamate, as described in the chapter by Joseph Coyle, Michael Leski, and John Morrison, is the major excitatory neurotransmitter in the brain. During the past decade, numerous subtypes of glutamate transporters and glutamate receptors have been identified and characterized. Each of these represents a potentially exciting target for new pharmacotherapeutic agents. Richard Olsen focuses on γ-aminobutyric acid (GABA), which serves as the major inhibitory neurotransmitter in the brain. GABA receptors and GABA transporters are important targets for commonly used antianxiety, anticonvulsant, and antimanic medications. Agents with improved specificity toward subtypes of these proteins may offer substantial benefit as future treatments.

The next several chapters focus on other small-molecule neurotransmitters, which are used by relatively small fractions of neurons and generally serve to modulate the efficacy of glutamatergic and GABAergic synapses through diffuse projections throughout the neuraxis. Such neurotransmitters include the catecholamines, norepinephrine, and dopamine. Norepinephrine, covered in a chapter by Gary Aston-Jones, regulates mood, attention, and alertness and is a substrate for many commonly used antidepressants. Dopamine, discussed in a chapter by Anthony Grace, plays a critical role in movement and reward. Accordingly, it is involved in movement disorders such as Parkinson's disease and is a common target for most drugs of abuse. The catecholamines, along with serotonin and histamine, are often referred to as *monoamine neurotransmitters* because they contain a single amine group. Serotonin is critically involved in many brain functions and is the target for many commonly used antidepressants. George Aghajanian and Elaine Sanders-Bush focus on new findings about serotonin, including the discovery and characterization of 14 distinct serotonin receptors and their physiologic functions. Histamine is discussed by Jean-Charles Schwartz and Jean-Michel Arrang. Although it has been known for some time that histamine regulates alertness and sleep, new advances in histamine pharmacology have been made possible by the cloning of three distinct histamine receptors. Finally, acetylcholine is often categorized along with the monoamines because it too is concentrated in discrete regions of the brain, many of which project diffusely to other parts of the brain. A major goal of neuropsychopharmacology research, as discussed by Marina Picciotto, Meenakashi Alreja, and J. David Jentsch, continues to be the development of drugs that are selective for the many subtypes of cholinergic receptors expressed in the central nervous system.

Many other types of molecules serve neurotransmitter functions. Michael Williams covers the so-called purinergic neurotransmitters, which include adenosine and adenosine triphosphate. The last few years have seen the cloning and characterization of a vast number of purinergic receptors,

with very different transmitter selectivities and functional properties. It is believed that selective ligands at these various receptors may serve as novel drugs in the treatment of Parkinson's disease, insomnia, anxiety, and pain, to name a few. Many types of polypeptides serve as neurotransmitters; these molecules are often termed *neuropeptides*. Significant recent progress has been made in understanding the physiologic role and pharmacology of certain neuropeptides, which are discussed in several chapters in this section. Gavan McNally and Huda Akil cover the opioid peptides, including a newly discovered opioid-like peptide, termed *orphanin-FQ* or *nociceptin*, that promotes nociception. Errol De Souza and Dimitri Grigoriadis review recent advances in the understanding of corticotropin-releasing factor, including the identification of two main types of receptors for corticotropin-releasing factor and other peptides (e.g., urocortin) that serve as endogenous ligands for the receptors. Nadia Rupniak and Mark Kramer focus on substance P and related neurokinins. Long known to be involved in the regulation of pain perception, recent evidence suggests that antagonists at certain neurokinin receptors may be effective antidepressants.

Despite the importance of neurotransmitter systems in neuropsychopharmacology, it must be emphasized that all the proteins that account for neurotransmitter synthesis and degradation, reuptake, and receptors, and for neuropeptide transmitters themselves, represent a small fraction of the perhaps hundreds of thousands of proteins expressed in the adult brain. A central promise of neuropsychopharmacology as we enter a new century is to evaluate these vast arrays of other proteins as targets for entirely new families of pharmacotherapeutic agents. Robert Malenka reviews what we know about synaptic plasticity, the processes by which the efficacy of transmission at particular synapses is altered as a consequence of synaptic activity. Mark von Zastrow covers the molecular and cellular mechanisms underlying receptor internalization, a process in which the numbers of many and perhaps most types of neurotransmitter receptors on the plasma membrane are regulated by synaptic activity. David Russell and Ronald Duman offer an overview of neurotrophic factors and their signaling pathways. Neurotrophic factors have long been recognized for their role in neural growth and differentiation during development, and we now know they are also important for regulating the survival and plasticity of adult neurons. Eric Nestler and Steven Hyman review the intracellular signaling pathways by which neurotransmitters, acting on plasma membrane receptors, regulate gene expression. Such regulation represents a prominent mechanism of long-term plasticity in the nervous system, including the actions of repeated exposure to psychotropic drugs (e.g., antidepressant action and drug addiction). Pierre Magistretti and Bruce Ransom discuss the role of glial cells in the central nervous system—in particular, their control of the energy metabolism in the brain. Finally, Fred Gage and Henriette van Praag summarize new knowledge of neurogenesis in the adult brain. The recent discovery that new neurons are born in certain regions of the brain each day, and may be incorporated into the existing circuitry within those regions, raises new hope for the treatment of neurodegenerative and other neuropsychiatric disorders. The subject matter of these last several chapters has not yet been exploited pharmacologically, but it is believed that the next generation of progress will see new pharmacologic agents directed at these nontraditional mechanisms.

1

Acetylcholine

Marina R. Picciotto

Meenakshi Alreja

J. David Jentsch

Marina R. Picciotto and Meenakshi Alreja: Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.

J. David Jentsch: Department of Neuroscience, University of Pittsburgh, Pittsburgh, Pennsylvania.

Acetylcholine (ACh) is critical for communication between neurons and muscle at the neuromuscular junction, is involved in direct neurotransmission in autonomic ganglia, and has been implicated in cognitive processing, arousal, and attention in the brain (1). Cholinergic transmission can occur through muscarinic (G protein-coupled) or nicotinic (ionotropic) receptors and is terminated by the action of cholinesterases. Seventeen different subunits of the nicotinic ACh receptor (nAChR) (2) and five different subtypes of the muscarinic receptor (3) have been cloned to date, and a majority of those are known to be expressed in the brain. Although the anatomic locations of cholinergic cell bodies and their projections have been known for some time (Fig. 1.1), recent studies using specific cholinotoxins, electrophysiology, or molecular genetics have altered our view of the functional role of the cholinergic system in the brain. The anatomic, pharmacologic, and biochemical complexity of the cholinergic system indicates an intricate involvement in nervous system function, and new advances in this field are discussed here.

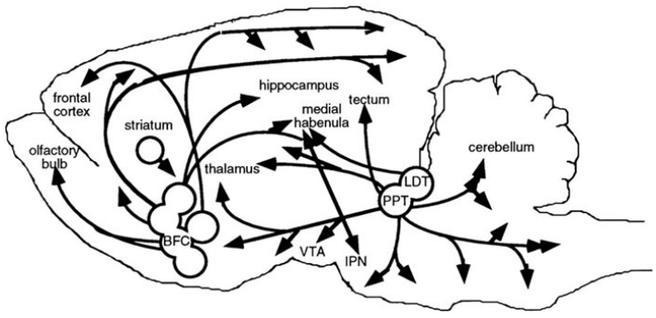


FIGURE 1.1. Anatomy of major cholinergic pathways in the brain. The principal source of cholinergic input to the cortex and hippocampus is the basal forebrain complex, whereas the pedunculopontine and laterodorsal tegmental areas innervate brain stem and midbrain targets preferentially. Cholinergic interneurons are found in the olfactory tubercle, striatum, nucleus accumbens, and islands of Calleja. BFC, basal forebrain complex; VTA, ventral tegmental area; IPN, interpeduncular nucleus; PPT, pedunculopontine tegmental nucleus; LDT, laterodorsal tegmental nucleus.

KNOCKOUT OF MUSCARINIC- AND NICOTINIC-RECEPTOR SUBUNITS

Part of "1 - Acetylcholine "

A particularly useful tool in identifying the role of individual molecules in the physiologic and behavioral functions of the cholinergic system are transgenic animals that lack specific subunits or subtypes of muscarinic receptors or nAChRs. These animals, termed *"knockout" mice*, can be generated by means of genetic engineering techniques and have been extremely useful in determining the functional role of many proteins that have been identified through molecular cloning (see refs. 4 and 5 for a review of this technology). Mice lacking the α3 (6), α4 (7), α5 (8), α7 (9), α9 (10), β2 (11), β3 (12), or β4 (13) subunit of the nAChR have been reported. In addition, mice lacking the M1 (14), M2 (15), and M4 (16) subtypes of the muscarinic receptor have been generated. These mice have already been used to demonstrate the role of particular receptor subtypes in the physiologic effects of ACh in muscle, the peripheral ganglia, and the central nervous system (Table 1.1).

TABLE 1.1. KNOCKOUTS OF MUSCARINIC AND NICOTINIC SUBUNITS

Subunit	Knockout	Knockout Phenotype	References
Muscarinic receptors			
M1	Viable	Disruption of M current and muscarinic seizures	(14,21)
M2	Viable	Disruption of muscarinic receptor-dependent movement and temperature control and antinociception	(15,21)
M4	Viable	Enhancement of D1 receptor-mediated locomotor stimulation	(16,21)
Nicotinic subunits			
α3	High mortality rate before and after weaning	Impaired growth, megalocystis (inflamed urinary bladder) and mydriasis (widely dilated ocular pupils)	(6)
α4	Viable	Reduced antinociception	(7)
α5	Viable	Not yet reported	(8)
α7	Viable	Largely normal; lack MLA-sensitive nicotine response in hippocampal interneurons; may have slightly decreased anxiety response	(9,22)
α9	Viable	Involved in cochlear efferent innervation development and function	(10)
β2	Viable	Lack nicotine-induced increases in passive avoidance, reinforcement, antinociception; show increased neurodegeneration during aging	(7,11,24,31,32)
β3	Viable	Altered locomotor activity	(12)
β4	Viable	Viable, but lethal when combined with β2 subunit knockout	(13)

MLA, methyl-lyaconitine, a β7 antagonist.

The function of ACh has been best studied at the neuromuscular junction, where signaling occurs through the muscle form of the nAChR. In the embryo, the nAChR at the neuromuscular junction is a pentamer made up of two α, one β, one γ, and one δ subunit. After birth, the γ subunit is replaced by the ε subunit, so that the physiologic properties of the receptor are altered. In mice in which the ε subunit has been knocked out, the neuromuscular junction nAChRs remain in the embryonic form; the consequence is survival past birth with progressive muscle degeneration and lethality by 2 to 3 months of age (17). These experiments demonstrate that maturation of the neuromuscular junction nAChR is necessary for muscle cell function and survival and imply that the kinetics of ACh neurotransmission are critical for the health of muscle fibers in adulthood.

Cholinergic neurotransmission within the sympathetic ganglia occurs through several receptor subtypes. In the peripheral nervous system, the issue of which ACh-receptor subtypes are involved in cholinergic neurotransmission has been addressed both by knocking out muscarinic and nicotinic subunits and by treating sympathetic neurons from isolated chick sympathetic ganglia with antisense oligonucleotides (short stretches of DNA that can inhibit the translation of a particular protein of interest) to decrease the expression of α3, α4, and α7 nAChR subunits. Antisense experiments have indicated that the α3 nAChR subunit plays a primary role in nicotinic transmission in sympathetic ganglia and that the α7 subunit also contributes to the observed currents (18 ,19). These data are in agreement with electrophysiologic and immunoprecipitation studies of nAChR subunits from ganglionic neurons (20). Although the results of studies using this powerful technique are compelling,

some problems have been noted with antisense approaches, including issues of specificity, so that it is useful to complement these studies with other techniques that can be used to manipulate levels of nAChR subunits.

In studies of knockout mice, disruptions of two nicotinic-receptor subunits expressed in sympathetic ganglia, α7 (9) and ß2 (11), do not grossly alter ganglionic function. In contrast, if the ß2 and the ß4 nAChR-subunit mutations are combined (13), or if the α3 nAChR subunit is knocked out (6), mutant mice die perinatally of severe autonomic failure. These experiments suggest that a nicotinic cholinergic receptor composed of the α3/ß4 or ß2 subunit, or both, is responsible for mediating direct neurotransmission by ACh between ganglionic neurons.

Muscarinic function has also been studied in the autonomic ganglia with knockout technology. Mutation of the M1 muscarinic receptor is sufficient to abolish the M current, a muscarine-mediated potassium current, in the sympathetic ganglia, but M1 mutation does not significantly perturb ganglionic function (14). In contrast, mice lacking

the M2 muscarinic-receptor subtype do not show carbachol-induced bradycardia, confirming that the effect of ACh on sympathetic control of heart rate is mediated through the M2 receptor (15). In addition, knockout animals have been very useful in determining which subtypes of muscarinic receptors are responsible for the effects of ACh on modulation of calcium channels in sympathetic neurons (21). A slow, voltage-independent modulation is mediated by M1 receptors, whereas a fast, voltage-dependent modulation is mediated through M2, and neither is affected in M4 knockout mice.

The function of ACh in the brain has also been examined in electrophysiologic experiments with mice lacking cholinergic-receptor subtypes. For example, a rapidly desensitizing nicotinic current in the hippocampus is mediated through an α7-containing receptor (9). Mice lacking the α7 subunit appear grossly normal in behavioral experiments (22), but future experiments should determine whether these currents contribute to nicotine-induced improvements in cognitive function or to nicotine-induced seizure activity. Antisense experiments have also demonstrated that the α5 nAChR subunit can alter the electrophysiologic properties of nAChRs containing the α4 and ß2 subunits *in vivo* (23). Mice lacking the ß2 subunit have been used to characterize four classes of nAChR in the brain by means of pharmacologic and electrophysiologic techniques (24) and to extend the existing pharmacologic characterization of nicotinic-receptor subtypes (Fig. 1.2). Future experiments using mice lacking individual nAChR α subunits should allow a finer definition of these receptor classes.

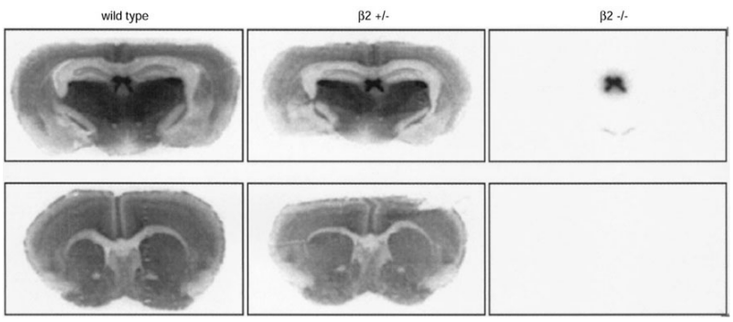


FIGURE 1.2. Nicotinic ligand binding in brain slices from wild-type and ß2-subunit knockout mice. Mice lacking individual subunits of the nicotinic acetylcholine receptor (nAChR) can be used to distinguish between subclasses of receptors. For example, although ß2-subunit knockout mice lack the highest-affinity subclass of nicotine binding sites, the frog toxin epibatidine, shown here, still reveals ß4 subunit-containing nAChRs in the medial habenula (remaining binding shown in panel at *top, far right*). Binding of epibatidine in brain slices through thalamus *(top)* or striatum *(bottom)* is shown in wild-type heterozygous (ß2 +/-) and homozygous (ß2 -/-) ß2-subunit knockout mice.

A significant development in thinking about nicotinic-receptor function has been the idea that nicotine exerts many of its functions in brain through the regulation of neurotransmitter release, at least partly through terminal and preterminal nAChRs (25 ,26). Experiments on synaptosomes (nerve terminals) isolated from mice lacking the ß2 subunit of the nAChR have shown that presynaptic regulation of GABA release by nicotine is mediated through ß2 subunit-containing receptors in most brain areas (27). This is also likely to be the case for other neurotransmitters because the efflux of rubidium, a radioactive tracer that serves as a marker of neurotransmitter vesicle fusion, is mediated through ß2 subunit-containing receptors in most brain areas (28). More recently, dopamine release from striatal synaptosomes has been shown to be disrupted in ß2-subunit knockout mice, while ACh release in the interpeduncular nucleus is preserved (29). This suggests that a distinct nAChR subtype, most likely containing the ß4 subunit, mediates nicotine-elicited ACh release in the interpeduncular nucleus.

Systems-level function and behavioral effects of ACh have also been examined in knockout mice. Muscarinic agonist-induced seizures are dependent on the presence of the M1 receptor because M1 knockout mice are resistant to pilocarpine-induced seizure activity (14). Interestingly, although the M1 receptor has been implicated in the modulation

of potassium channels in the hippocampus in pharmacologic experiments, muscarinic modulation of potassium channels is unchanged in the hippocampus of M1 knockout mice (30). In contrast, the pharmacologic effects of muscarinic agonists on movement, temperature control, and antinociception appear to be mediated through the M2 receptor because these responses are absent in M2 knockout mice (15). M4 receptors are also involved in locomotion; these knockout animals exhibit increased basal locomotor activity and a potentiated locomotor response to D1-selective dopaminergic agonists (16).

Like the M2 receptors, the α4/ß2 subtype of nAChR is implicated in antinociceptive cholinergic pathways. Mice lacking either of these subunits show decreased nicotine-induced analgesia (7). In behavioral experiments, the ß2 nicotinic subunit mediates the ability of nicotine to improve avoidance learning and may also be involved in the circuitry underlying this form of associative learning in wild-type mice (11). In addition, this subunit appears to be necessary for the mouse to experience the reinforcing properties of nicotine because animals without the ß2 subunit will not self-administer nicotine (31). Extensions of these experiments to mice lacking other subunits of the nicotinic receptor should allow identification of the receptor subtypes that are activated by smoking in humans and result in tobacco addiction. An interesting effect of ACh on neuronal survival was demonstrated in mice lacking the ß2 nAChR subunit (32). Mice that lack this cholinergic-receptor subtype show progressive neuronal loss with age in cortical and hippocampal brain areas, which appears to lead to age-related impairments in spatial learning. These experiments demonstrate that the effects of ACh on cognition, antinociception, locomotion, and overall neuronal activity are differentially mediated through the various subtypes of muscarinic and nicotinic receptors, and that the various roles of ACh may be separated pharmacologically, suggesting new targets for rational drug design.

ROLE FOR CHOLINERGIC NEURONS IN AROUSAL AND SLEEP

Part of "1 - Acetylcholine "

Traditionally, the basal forebrain complex, the primary source of cholinergic innervation to the telencephalon (Fig. 1.1), was thought to be involved in arousal or sleep regulation. Either lesions or electric stimulation of subregions of the basal forebrain can facilitate sleep and synchronize the EEG, and cholinergic drugs regulate EEG synchrony (33). Moreover, a correlation between cortical ACh release and the state of behavioral activation or sleep has been observed in rodents. Thus, it was hypothesized that cholinergic input to the neocortex from the basal forebrain is critical for regulating arousal (see ref. 34 for review).

The pontomesencephalic tegmentum is also critical for the sleep–wake cycle. These neurons largely do not innervate the neocortex but project to the diencephalon (thalamus) and the basal forebrain complex. Stimulation of tegmental brainstem cholinergic neurons can evoke cortical ACh release and EEG desynchrony, and these effects are blocked by reversibly decreasing the activity of the basal forebrain (35). Moreover, application of cholinergic agonists to the basal forebrain produces behavioral activation and EEG desynchrony (33). Although the brainstem cholinergic projections to the thalamus undoubtedly also contribute to EEG regulation (36), these findings suggest that cholinergic projections to the basal forebrain from the pontomesencephalic tegmentum regulate behavioral arousal.

It was subsequently noted that cholinergic tegmental projections largely formed connections with noncholinergic neurons within the basal forebrain (37). This finding is critical because it could explain why stimulation of the horizontal diagonal band, preoptic area, and substantia innominata, but not of the septal nucleus and nucleus basalis, produces sleep in the cat (33). The ratio of cholinergic to noncholinergic neurons in the horizontal diagonal band, preoptic area, and substantia innominata is significantly lower than in the septum and nucleus basalis. This observation has led to the hypothesis that activation of primarily noncholinergic neurons is responsible for producing sleep after basal forebrain stimulation (33). These noncholinergic neurons are believed to be GABAergic and achieve their effects through inhibition of cholinergic basal forebrain neurons and neurons within the brainstem reticular formation. In contrast, stimulation of the nucleus basalis or septal nucleus produces behavioral activation and cortical ACh release, and this is consistent with the notion that basal forebrain cholinergic neurons are involved in behavioral arousal (activation), whereas noncholinergic basal forebrain neurons are involved in regulating the sleep state. These two effects are related (sleep vs. arousal), but the qualitative contributions of the GABA and cholinergic systems to sleep and arousal are opposed.

ROLE FOR CHOLINERGIC NEURONS IN MOTIVATION AND REWARD

Part of "1 - Acetylcholine "

Cholinergic neurons have also been implicated in motivation and reward. The strongest evidence for the hypothesis that nAChRs are involved in motivation and reward is that nicotine is abused by humans and is reinforcing in animals (see ref. 38 for review). The effects of nicotine on tests of reinforcement and behavioral sensitization are primarily mediated through the mesolimbic dopamine system (39). Indeed, the ventral tegmental area (VTA) may be sufficient to mediate the reinforcing properties of nicotine, as local injection of nicotine or nicotinic agonists into the VTA can result in increased locomotion (40) or conditioned place preference (41).

Basal forebrain cholinergic neurons may also be involved in modulating cortical processing of stimuli with conditioned or unconditioned rewarding properties because these neurons are more responsive to stimuli with a high incentive value. Novel stimuli that typically elicit orienting responses and attention in animals increase cortical ACh release, but this effect is diminished with repeated exposure if the stimulus has no contingent incentive valence. In contrast, if the stimulus is repeatedly paired with an incentive stimulus (e.g., food or foot shock), the (now-conditioned) stimulus can evoke ACh release even after multiple exposures (42). These sorts of changes are reflected in the firing of “reinforcement-related” neurons within the primate basal forebrain (43). Pontomesencephalic cholinergic neurons are also involved in motivation and reward, although these effects are likely mediated, in part, by projections to the dopamine neurons within the VTA (44 ,45).

Stimulation of the VTA by the pedunculopontine tegmental nucleus (PPT) enhances mesostriatal dopamine transmission (45 ,46). While a significant proportion of the PPT neurons that project to the tegmental dopamine neurons are noncholinergic (44), the cholinergic input *per se* appears to stimulate dopamine neurons (47). Thus, ascending projections from the PPT to the dopamine cells may regulate the ability of mesostriatal dopamine neurons to affect incentive/motivational processes.

This innervation of dopamine cells by cholinergic neurons may explain the finding that lesions of the PPT can modulate the rewarding qualities of addictive drugs. Lesions of the PPT reduce the self-administration of nicotine (48) and opiates (49). Moreover, conditioned place preference for food, opiates (50), morphine (51), and amphetamine (52) is blocked or reduced by PPT lesions, whereas cocaine-induced reward is unaffected (53). Although the mesolimbic dopamine pathway is known to be involved in drug reward (see ref. 54 for review), it is not yet known whether the influence of the PPT on drug reinforcement is through cholinergic projections. It is also not known whether the effect of PPT lesions on these processes is mediated through projections to areas other than the dopamine cell groups within the VTA.

The PPT may have another, more critical, role in motivation and reward via afferent inputs from the striatum (55). Excitotoxic lesions of the PPT (that equivalently destroy both cholinergic and noncholinergic neurons) disrupt responding for conditioned reinforcement and augment stimulant-induced orofacial stereotypy, yet no difference is observed in stimulant-induced locomotion or other measures of food consumption (42 ,56). These data may implicate the PPT (and its innervation from the striatum) in response selection when discrimination is involved because the disruption of responding for conditioned reinforcement resulted from decreased discrimination of response between a lever associated with reinforcement and an inactive lever (56). However, a recent study found that although PPT lesions increased sucrose consumption, similar lesions did not affect discrimination or contrast effects (57). Nevertheless, the hypothesis of Winn (58) is that lesions of the PPT affect responding for rewarding stimuli similarly to lesions of the frontal cortex, so that the role of the PPT, like that of the basal forebrain, is expanded into higher-order cognitive processes.

ROLE FOR CHOLINERGIC NEURONS IN COGNITIVE PROCESSES

Part of "1 - Acetylcholine "

Lesion Studies

The hypothesis of cholinergic involvement in learning and memory processes arose from several findings. Both destruction of the basal forebrain complex and the administration of cholinergic antagonists produce profound deficits in a variety of forms of cognition, including learning and memory (59 ,60).

The original finding that lesions of the basal forebrain could produce deficits in a variety of cognitive tasks suggested a role for ACh in cognitive function. Electrolytic, radiofrequency, or nonspecific excitotoxic lesions of cholinergic subnuclei within the basal forebrain (particularly the medial septum/diagonal band) profoundly impair performance on a variety of tests of learning, memory, and attention, particularly the Morris water maze and passive avoidance learning (see ref. 59 for review). These deficits appeared to be reversed following regeneration of cholinergic projections across a bridging graft (61) or after grafting of ACh-producing cells in the hippocampus (62). These findings have been interpreted as support for the hypothesis of cholinergic involvement in cognitive functions; however (as with arousal and sleep), noncholinergic neurons within the basal forebrain may likewise be involved in these effects, and more specific approaches must be employed to address these issues.

Novel approaches for selectively destroying cholinergic neurons depend on the differential sensitivity of basal forebrain neurons to excitotoxins and new types of immunotoxins. Systematic studies have demonstrated that cholinergic and noncholinergic neurons within the basal forebrain are differentially sensitive to excitotoxic amino acids such as quisqualate, α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) (Fig. 1.3), kainate, and *N*-methyl-D-aspartate (NMDA) (59). Based on the results of these studies, new methods for preferentially destroying cholinergic neurons have been described (63). Moreover, an IgG–saporin toxin has been developed that takes advantage of the fact that basal forebrain cholinergic neurons are particularly enriched with low-affinity receptors for nerve growth factor (64). The toxin selectively binds to the receptor for nerve growth factor and then kills the neuron expressing the receptor. More excitingly, recent studies suggest that IgG–saporin can be used to destroy the cholinergic innervation of

terminal regions into which the toxin is injected (65). These methods have been applied to studies of learning and memory in an attempt to qualify earlier findings.

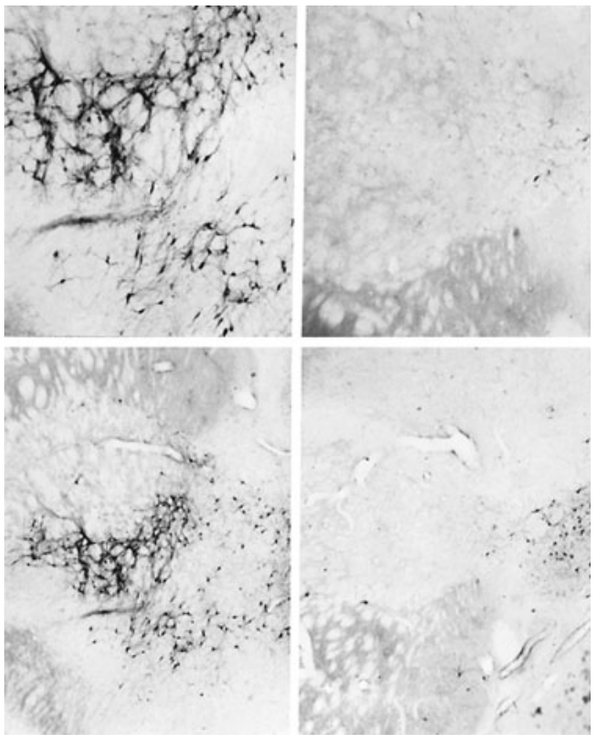


FIGURE 1.3. Acetylcholinesterase staining of the nucleus basalis magnocellularis after infusion of saline solution or AMPA to destroy cholinergic neurons preferentially. Low concentrations of the glutamatergic agonist AMPA selectively destroy cholinergic neurons (measured by acetylcholinesterase staining) and spare γ-aminobutyric acid (GABA) neurons *(left)*. In contrast, control sections show robust acetylcholinesterase staining after infusion of saline solution *(right)*. This process allows more specific cholinergic lesions to be generated, so that the function of the neurons in behavioral processes can be clarified. (Courtesy of Professor Barry J. Everitt, University of Cambridge.)

Based on either excitotoxic or saporin lesions of the basal forebrain, the hypothesis for cholinergic function has been revised considerably. Essentially, selective damage to cholinergic neurons of the basal forebrain has failed to produce the retrograde or anterograde amnesia or deficits in learning that have been reported to result from nonspecific lesions of the basal forebrain (59 ,66). Previously, the medial septal/diagonal band nuclei and their projections to posterior cortical regions were thought to be critical for spatial learning and contextual conditioning (59). By means of saporin lesions, however, cholinergic depletion within the hippocampus or posterior parietal cortex has been shown to result in impairments in latent inhibition or unblocking (65), with sparing of spatial learning (67) and spatial working memory (68). Moreover, selective excitotoxic lesions of the medial septum/diagonal band produce enhancements in contextual conditioning but impairments in discrete cue (trace) conditioning (69). Both sets of data may suggest that the attentional processing of discrete stimuli is disrupted following cholinergic depletion from posterior cortical regions. It is possible, however, that the depletion of ACh from caudal or rostral cortical regions alone may be insufficient to impair performance of some tasks, whereas combined depletions may have more than additive effects (70).

Other investigators have further argued that the cholinergic innervation of rostral (e.g., frontal) cortex from the nucleus basalis is also involved in attentional functions, such as vigilance or sustained, divided attention (59 ,71). Direct pharmacologic manipulation of basal forebrain neurons has been used to alter activated cholinergic efflux in the frontal cortex and performance of tasks related to stimulus processing or detection (72). Selective excitotoxic lesions or pharmacologic manipulation of the nucleus basalis has also been reported to impair performance in a five-choice serial reaction task that requires animals to detect and respond to brief visual stimuli (73). Interestingly, the observation that appetitive pavlovian learning for a discrete cue is enhanced after nucleus basalis lesions (74) suggests that attentional processing of discrete cues may not be affected by depletion of ACh from the rostral neocortex except when divided attention is required. The findings of these latter studies are also bolstered by advances in the measurement of ACh transmission *in vivo*, which allows investigators to quantify directly the extent of the lesions produced by the toxins for the first time (75). Taken together, the available data seem to suggest that basal forebrain cholinergic neurons are capable of regulating the cortical processing of sensory stimuli within a variety of domains, which may be explained by a role for basal forebrain ACh in the regulation of cortical processing.

Tegmental cholinergic neurons have also been implicated in cognitive processes (58 ,76). Although some of the effects of PPT lesions on learning and memory may be related to generalized anxiety (76), PPT lesions also produce a set of behavioral deficits that are consistent with executive dysfunction and impairments in frontal lobe functioning (58). In particular, PPT lesions result in deficits of behavioral inhibition and motor perseveration. Notably, working memory performance does not seem to be affected by destruction of the PPT (77). The position of the PPT as a modulator of dopaminergic systems (which affect frontal cortex function), in addition to the influence of the frontal cortex on the PPT (mediated through the striatum), suggests that this nucleus is in an excellent position to affect the functions of the frontostriatal system. Further research that attempts to control for the extent and selectivity of PPT lesions is necessary.

Muscarinic Mechanisms

Although lesions of cholinergic nuclei have implicated ACh in various behavioral processes, it is also of interest to determine

which cholinergic-receptor subtypes mediate these responses to ACh. Systemic infusions of the muscarinic-receptor antagonists atropine and scopolamine produce an amnesic syndrome in humans (78), monkeys (79), and rats (80). Several lines of evidence suggest that multiple central nervous system structures, including the medial septum/diagonal band region, are critical in mediating the effects of muscarinic drugs on mnemonic functions (80). Infusions of muscarinic-receptor antagonists into a variety of cortical regions, including the hippocampus, prefrontal cortex, and amygdala, can impair the cognitive functions associated with these respective regions (81). Similarly, the effects of systemic muscarinic antagonists are attenuated by intraseptal injections of muscarinic agonists, and intraseptal applications of muscarinic antagonists mimic the amnesic effects of systemic treatment with muscarinic antagonists in experimental animals (82). These results suggest that activation of muscarinic receptors by ACh at multiple forebrain sites, including within the somatodendritic regions of the cholinergic neurons, may be involved in the behavioral dysfunction produced by muscarinic cholinergic antagonists.

Figure 1.4 presents the results of an experiment aimed at determining the relationship between *in vivo* cortical cholinergic transmission and the cognitive effects of muscarinic-receptor antagonists. Scopolamine was administered systemically to rats performing a test of working memory, the spatial delayed alternation task, both alone and in combination with FG7142, an anxiogenic ß-carboline that acts as an inverse agonist of the benzodiazepine site of the $GABA_A$ receptor. Consistent with previous findings, scopolamine produced dose-dependent performance impairments when administered 45 minutes before testing on the delayed alternation task, suggesting that decrements in cholinergic stimulation of muscarinic receptors result in cognitive dysfunction. FG7142 (20 mg/kg) significantly elevated prefrontal cortical ACh release *in vivo* (measured in parallel studies), and FG7142 on its own impaired delayed alternation performance. Interestingly, the fact that coadministration of FG7142 and scopolamine did not affect the slope of the dose–response curve for scopolamine suggests that these two drugs act on different mechanisms to impair delayed alternation performance. The additivity of these effects indicates that supranormal ACh transmission produced by FG7142 likely does not contribute to the working memory deficits produced by this drug; moreover, the data indicate that the impairments produced by scopolamine are independent of the level of ongoing cortical cholinergic transmission. Thus, it is possible that the cognitive effects of muscarinic antagonists may not be solely the consequence of changes in cortical cholinergic transmission.

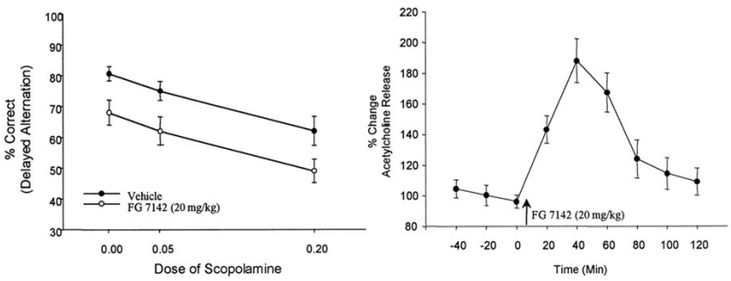


FIGURE 1.4. The cognitive effects of scopolamine administration are insensitive to phasic changes in cortical acetylcholine (ACh) release. Scopolamine dose-dependently impairs performance on a test of spatial working memory, the delayed alternation task, in control rats and rats treated with FG7142, an inverse agonist of the benzodiazepine site of the γ-aminobutyric acid subtype A ($GABA_A$) receptor *(left)*. Although FG7142 increases prefrontal cortical ACh release *in vivo (right)* and produces performance deficits on its own *(left)*, it does not alter the slope of the dose– response curve for scopolamine.

The septohippocampal pathway was first believed to convey only cholinergic fibers to the hippocampus (83); the noncholinergic, GABAergic component was discovered almost two decades later (84). Work focusing on the GABA limb of the septohippocampal GABA pathway has suggested that the septohippocampal GABA and cholinergic pathways may both be critical for the effects of septal efferents on cognitive functioning (85). In support of this hypothesis, agents that increase impulse flow in the septohippocampal GABA pathway, including muscarinic agonists, augment

learning and memory (86), whereas agents that impair learning and memory decrease impulse flow in the GABA pathway (87). Interestingly, impulse flow in the septohippocampal GABA pathway is maintained by ACh released via the *tonic* firing activity of septohippocampal cholinergic neurons. This release occurs via local axon collaterals of septohippocampal neurons, which then synapse on septohippocampal GABA neurons within the medial septum/diagonal band (Fig. 1.5). Thus, interaction between the septohippocampal GABA pathway and muscarinic mechanisms within the medial septum/diagonal band may be crucial for learning and memory (86).

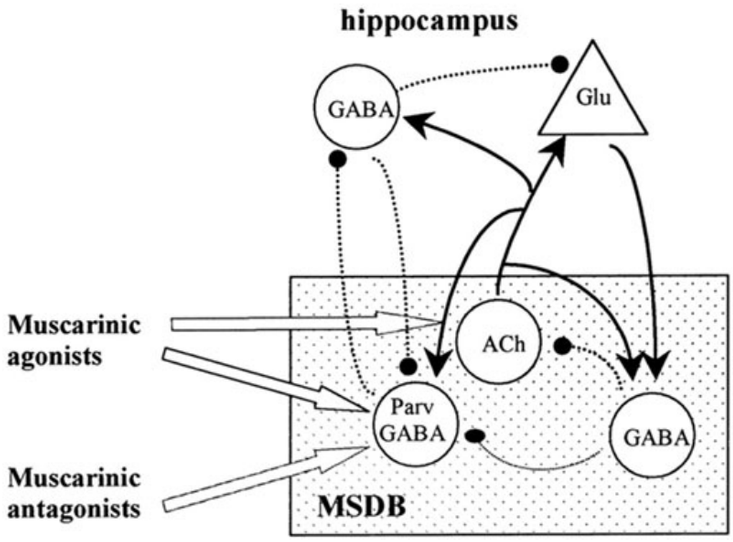


FIGURE 1.5. Schematic representation of the septohippocampal pathway. The medial/septum diagonal band region is composed primarily of cholinergic and GABAergic neurons, and the activity of both neuronal populations is regulated by locally released γ-aminobutyric acid (GABA). The cholinergic neurons and a subpopulation of GABA neurons, containing the calcium-binding protein parvalbumin (parv), project to the hippocampus via the fimbria/fornix. Muscarinic agonists may not increase hippocampal acetylcholine release directly, but rather activate septohippocampal GABA neurons via M3 (and possibly M5) receptors. Similarly, muscarinic antagonists disrupt impulse flow in the septohippocampal GABA pathway.

Cholinergic neurons, the primary source of ACh input to the hippocampus, innervate both the hippocampal pyramidal neurons and subpopulations of GABAergic interneurons (88). In contrast, septohippocampal GABA neurons are very selective in their innervation pattern, do not innervate the pyramidal cells at all, but innervate almost every type of hippocampal interneuron (89). Septohippocampal GABA neurons are able to produce a powerful disinhibitory effect on pyramidal cells via this connectivity and so enhance their excitability (90). Loss of cholinergic neurons severely disables the septohippocampal pathway by reducing both the direct excitatory cholinergic drive and the indirect disinhibitory GABA drive to the hippocampus via locally released ACh. A restoration of cholinergic function within the medial septum/diagonal band, not just in the hippocampus, could therefore be critical for the treatment of cognitive deficits associated with the septohippocampal pathway.

At a molecular level, the excitatory effects of ACh on hippocampal pyramidal cells were at first thought to be mediated via the M1 subtype of muscarinic receptor, partly as a result of closing of M-type potassium channels, so that specific M1-receptor agonists were developed. However, M1-receptor agonists were found to be of limited use in improving cognition. This might not be surprising because studies of knockout mice lacking M1 receptors show no change in muscarinic enhancement of potassium currents in the hippocampus (30). The finding that non-M1 receptors (M3 and possibly M5) mediate the effects of ACh in the medial septum/diagonal band may further explain the limited effectiveness of M1 agonists in improving learning and memory functions and supports the need for M3-selective agonists (85).

Nicotinic Mechanisms

Nicotinic systems are also involved in several important aspects of cognitive function, including attention, learning, and memory (60). Nicotinic ACh receptors are expressed throughout the brain, including areas involved in cognitive function, such as the hippocampus and frontal cortex (91). Nicotinic agonists improve performance on a variety of memory tasks, particularly following lesions or aging, whereas nicotinic antagonists such as mecamylamine impair working memory function (60). The nAChR subtypes involved in cognitive function are under investigation, and different subtypes may be involved in the performance of different cognitive tasks. As mentioned above, experiments on knockout mice have implicated nAChRs containing the β2 subunit in both passive avoidance learning (11) and maintenance of spatial learning during aging (32). Although the cellular basis for the effects of nicotine are likely to be diverse, one site of action for nicotine, excitation of hippocampal GABAergic interneurons through both α7 and non-α7 subtypes of the nAChR, has been demonstrated by several groups (see ref. 92 for review). Further, although theta rhythm in the hippocampus, a mechanism that appears to facilitate the induction of synaptic plasticity, is abolished by atropine (93), it is converted to burst-mode activity by nicotinic antagonists (94).

Another major contributor to the cholinergic hypothesis of cognitive functioning was the discovery in the early 1980s that cholinergic neurons in the basal forebrain degenerate in Alzheimer's disease (95). Since then, loss/atrophy of cholinergic neurons has been reported not only in Alzheimer's disease but also in Parkinson's disease, Lewy body dementia, progressive supranuclear palsy, and several other disorders (96), although not all studies have reported losses in cholinergic neurons (97). In those that have reported losses, the greatest reduction in numbers, of the order of 50% to 65%, has been observed in cholinergic neurons of the nucleus basalis and the medial septal/diagonal band regions of patients pathologically verified as having Alzheimer's disease (96). Loss of high-affinity nAChRs has also been seen in the brains of patients with Alzheimer's disease (98), and

nicotinic agonists have been proposed as potential therapeutic agents to treat the disease (60).

ROLE FOR CHOLINERGIC NEURONS IN STIMULUS PROCESSING

Part of "1 - Acetylcholine "

Several lines of evidence suggest that cholinergic neurotransmission through nAChRs can affect stimulus processing. In support of this notion, nicotine has been reported to alleviate some sensory gating deficits in schizophrenic patients (99), and animal studies also suggest that nicotine may act to facilitate sensory inhibition, such as prepulse inhibition of startle in mice (100) and rats (101). In another animal model of sensory processing, latent inhibition, nicotine can either enhance or disrupt sensory habituation, depending on the preexposure parameters (102). Lesions of the nucleus accumbens or the pedunculopontine nucleus have been shown to block prepulse inhibition (103), whereas lesions of the hippocampus, septum, medial raphe, and nucleus accumbens disrupt latent inhibition (104), observations suggesting that nicotine may act in one or more of these brain areas to affect sensory processing. Another brain area that is likely to mediate the effect of nicotine on sensory gating in schizophrenia is the hippocampus. Postmortem studies have shown a reduced number of α-bungarotoxin-sensitive nAChRs (α7-containing nAChRs) in the hippocampus in schizophrenic patients (105). Further, pharmacologic (106) and genetic (107) studies have suggested a role for the α7 nAChR in prepulse inhibition in rodents.

A series of physiologic studies also supports the concept that ACh, acting on muscarinic receptors within the cerebral cortex, promotes cortical responses to exogenous stimuli. ACh can produce a biphasic effect on membrane polarization in cortical neurons: a rapid hyperpolarization followed by a prolonged depolarization (108). The inhibitory component was mediated through ACh-induced activation of GABAergic interneurons that inhibited the pyramidal cells in a feed-forward manner. In contrast, the long-lasting depolarization was mediated through direct effects of ACh on the cortical neuron. Subsequent studies suggested that this effect is mediated by blockade of I_m, a voltage-sensitive rectifying K^+ channel (14). In addition, ACh reduced spike frequency adaptation by blocking the after-hyperpolarization effect.

The net physiologic effect of these changes in cortical cell physiology may be to render pyramidal cells more responsive to afferent input. Because the membrane is more depolarized, neurons are more likely to fire in response to a given excitatory stimulus; also, the response to that stimulus may be prolonged because the after-hyperpolarization has been blocked. Thus, it seems plausible that muscarinic cholinergic effects on cortical pyramidal cells may indeed promote stimulus access to the cortical circuit. Inasmuch as attentional processing may represent the ability of stimuli to be processed actively within the neocortex, these physiologic actions of ACh may be consistent with the reported behavioral effects of cholinergic lesions.

CONCLUSIONS

Part of "1 - Acetylcholine "

Recent studies using new physiologic techniques, cholinergic-selective toxins, and molecular genetic techniques have refined our ideas about the role of ACh in the brain. In particular, it is clear that cholinergic and GABAergic pathways are intimately connected in the hippocampus and basal forebrain complex and may combine to exert their effects on cognition, attention, and arousal. Further, the subtypes of cholinergic receptors that mediate these effects of ACh are beginning to be elucidated with the use of knockout mice that lack specific receptor subunits. These techniques have contributed to a minirevolution in our views of how ACh contributes to cognitive processes. Research in this area is moving very quickly, and it is likely that these ideas will continue to be refined as the new techniques are applied to previously described behavioral paradigms. Improvements in existing techniques—for example, through the development of inducible and site-specific mutations in cholinergic-receptor subtypes—will also contribute to further refinements in our view of cholinergic functions in the brain.

SUMMARY

Part of "1 - Acetylcholine "

Acetylcholine is critical for communication between neurons and muscle at the neuromuscular junction, is involved in direct neurotransmission in autonomic ganglia, and has been implicated in cognitive processing, arousal, and attention in the brain. The results of recent studies in which specific cholinotoxins, electrophysiology, or molecular genetic techniques were used have altered our view of the functional role of the cholinergic system in the brain. Mice that lack specific subunits or subtypes of muscarinic or nAChRs have recently been generated and used to demonstrate the role of particular receptor subtypes in physiologic effects of ACh in muscle, peripheral ganglia, and the central nervous system. Roles for cholinergic neurons have been found in arousal and sleep, motivation and reward, cognitive processes, and stimulus processing. The evaluation of these functions by means of novel cholinotoxins and new electrophysiologic techniques have refined our ideas about the role of ACh in the brain. The evidence that cholinergic and GABAergic pathways are intimately connected in the hippocampus and basal forebrain complex and may combine to affect cognition, attention, and arousal is reviewed. In addition, the subtypes of cholinergic receptors that mediate these effects of ACh are discussed based on studies of knockout mice that lack specific receptor subunits. Improvements in existing techniques—for example, through the development of inducible and site-specific mutations in

cholinergic-receptor subtypes—will contribute to further refinements in our view of cholinergic functions in the brain.

ACKNOWLEDGMENTS

Part of "1 - Acetylcholine "

M.R.P. is supported by grants DA10455, DA00167, and DA07290 from the National Institutes of Health and the Christiane Brooks Johnson Foundation. M.A. is supported by grant DA09797 from the National Institutes of Health. We thank Professor B. J. Everitt for the use of his unpublished figure.

REFERENCES

1. Karczmar AG. Brief presentation of the story and present status of studies of the vertebrate cholinergic system. *Neuropsychopharmacology* 1993;9:181–199.

2. Picciotto M, Caldarone BJ, King SL, et al. Nicotinic receptors in the brain: links between molecular biology and behavior. *Neuropsychopharmacology* 2000;22:451–465.

3. Nadler LS, Rosoff ML, Hamilton SE, et al. Molecular analysis of the regulation of muscarinic receptor expression and function. *Life Sci* 1999;64:375–379.

4. Picciotto MR. Knock-out mouse models used to study neurobiological systems. *Crit Rev Neurobiol* 1999;13:103–149.

5. Picciotto MR, Wickman K. Using knockout and transgenic mice to study neurophysiology and behavior. Physiol Rev 1998;78:1131–1163.

6. Xu W, Gelber S, Orr-Urtreger A, et al. Megacystis, mydriasis, and ion channel defect in mice lacking the alpha3 neuronal nicotinic acetylcholine receptor. *Proc Natl Acad Sci U S A* 1999;96:5746–5751.

7. Marubio L, del Mar Arroyo-Jimenez M, Cordero-Erausquin M, et al. Reduced antinociception in mice lacking neuronal nicotinic receptor subunits. *Nature* 1999;398:805–810.

8. Xu W, Sutcliffe CB, Lorenzo I, et al. Gene targeting of the alpha7 and beta2 subunits and the beta4, alpha3 and alpha5 cluster of neuronal nicotinic acetylcholine receptors. *Soc Neurosci Abst* 1997;23:391.

9. Orr-Urtreger A, Goldner FM, Saeki M, et al. Mice deficient in the alpha7 neuronal nicotinic acetylcholine receptor lack alpha-bungarotoxin binding sites and hippocampal fast nicotinic currents. *J Neurosci* 1997;17:9165–9171.

10. Vetter DE, Liberman MC, Mann J, et al. Role of alpha9 nicotinic ACh receptor subunits in the development and function of cochlear efferent innervation. *Neuron* 1999;23:93–103.

11. Picciotto MR, Zoli M, Léna C, et al. Abnormal avoidance learning in mice lacking functional high-affinity nicotine receptor in the brain. *Nature* 1995;374:65–67.

12. Allen RS, Cui C, Heinemann SF. Gene-targeted knockout of the beta3 neuronal nicotinic acetylcholine receptor subunit. *Soc Neurosci Abst* 1998;24:1341.

13. Xu W, Orr-Urtreger A, Nigro F, et al. Multiorgan autonomic dysfunction in mice lacking the beta2 and the beta4 subunits of neuronal nicotinic acetylcholine receptors. *J Neurosci* 1999;19:9298–9305.

14. Hamilton SE, Loose MD, Qi M, et al. Disruption of the M1 receptor gene ablates muscarinic receptor-dependent M current regulation and seizure activity in mice. *Proc Natl Acad Sci U S A* 1997;94:13311–13316.

15. Gomeza J, Shannon H, Kostenis E, et al. Pronounced pharmacologic deficits in M2 muscarinic acetylcholine receptor knockout mice. *Proc Natl Acad Sci U S A* 1999;96:1692–1697.

16. Gomeza J, Zhang L, Kostenis E, et al. Enhancement of D1 dopamine receptor-mediated locomotor stimulation in M(4) muscarinic acetylcholine receptor knockout mice. *Proc Natl Acad Sci U S A* 1999;96:10483–10488.

17. Witzemann V, Schwarz H, Koenen M, et al., Acetylcholine receptor epsilon-subunit deletion causes muscle weakness and atrophy in juvenile and adult mice. *Proc Natl Acad Sci U S A* 1996;93:13286–13291.

18. Yu CR, Role LW. Functional contribution of the alpha-7 subunit to multiple subtypes of nicotinic receptors in embryonic chick sympathetic neurones. *J Physiol (Lond)* 1998;509:651–665.

19. Listerud M, Brussaard AB, Devay P, et al. Functional contribution of neuronal AChR subunits revealed by antisense oligonucleotides. *Science* 1991;254:1518–1521.

20. Sargent PB. The diversity of neuronal nicotinic acetylcholine receptors. *Annu Rev Neurosci* 1993;16:403–443.

21. Shapiro MS, Loose MD, Hamilton SE, et al. Assignment of muscarinic receptor subtypes mediating G-protein modulation of Ca(2+) channels by using knockout mice. *Proc Natl Acad Sci U S A* 1999;96:10899–10904.

22. Paylor R, Nguyen M, Crawley JN, et al. Alpha-7 nicotinic receptor subunits are not necessary for hippocampal-dependent learning or sensorimotor gating—a behavioral characterization of ACRA7-deficient mice. *Learn Mem* 1998;5:302–316.

23. Ramirez-Latorre J, Yu CR, Qu X, et al. Functional contributions of alpha5 subunit to neuronal acetylcholine receptor channels. *Nature* 1996;380:347–351.

24. Zoli M, Lééna C, Picciotto MR, et al. Identification of four classes of brain nicotinic receptors using b2-mutant mice. *J Neurosci* 1998;18:4461–4472.

25. Wonnacott S, Irons J, Rapier C, et al. Presynaptic modulation of transmitter release by nicotinic receptors. *Prog Brain Res* 1989;79:157–163.

26. McGehee DS, Role LW. Presynaptic ionotropic receptors. *Curr Opin Neurobiol* 1996;6:342–349.

27. Lu Y, Grady S, Marks MJ, et al. Pharmacological characterization of nicotinic receptor-stimulated GABA release from mouse brain synaptosomes. *J Pharmacol Exp Ther* 1998;287:648–657.

28. Marks MJ, Whiteaker P, Calcaterra J, et al. Two pharmacologically distinct components of nicotinic receptor-mediated rubidium efflux in mouse brain require the beta 2 subunit. *J Pharmacol Exp Ther* 1999;289:1090–1103.

29. Grady SR, Meinerz NM, Cao J, et al. Nicotinic agonists stimulate acetylcholine release from mouse interpeduncular nucleus: a function mediated by a different nAChR than dopamine release from striatum. *J Neurochem* 2001;76:258–268.

30. Rouse ST, Hamilton SE, Potter LT, et al. Muscarinic-induced modulation of potassium conductances is unchanged in mouse hippocampal pyramidal cells that lack functional M-1 receptors. *Neurosci Lett* 2000;278:61–64.

31. Picciotto MR, Zoli M, Rimondini R, et al. Acetylcholine receptors containing the beta-2 subunit are involved in the reinforcing properties of nicotine. *Nature* 1998;391:173–177.

32. Zoli M, Picciotto MR, Ferrari R, et al. Increased neurodegeneration during ageing in mice lacking high-affinity nicotine receptors. *EMBO J* 1999;18:1235–1244.

33. Szymusiak R. Magnocellular nuclei of the basal forebrain: substrates of sleep and arousal regulation. *Sleep* 1995;18:478–500.

34. Sarter M, Bruno JP. Cortical cholinergic inputs mediating arousal, attentional processing and dreaming: differential afferent regulation of the basal forebrain by telencephalic and brainstem afferents. *Neuroscience* 2000;95:933–952.

35. Rasmusson DD, Clow K, Szerb JC. Modification of neocortical acetylcholine release and electroencephalogram desynchronization due to brainstem stimulation by drugs applied to the basal forebrain. *Neuroscience* 1994;60:665–677.

36. Steriade M. Arousal: revisiting the reticular activating system. *Science* 1996;272:225–226.

37. Semba K, Reiner P, McGeer E, et al. Brainstem afferents to the magnocellular basal forebrain studied by axonal transport, immunohistochemistry, and electrophysiology in the rat. *J Comp Neurol* 1988;267:433–453.

38. Picciotto MR. Common aspects of the action of nicotine and other drugs of abuse. *Drug Alcohol Depend* 1998;51:165–172.

39. Corrigall WA, Franklin KB, Coen KM, et al. The mesolimbic dopaminergic system is implicated in the reinforcing effects of nicotine. *Psychopharmacology* 1992;107:285–289.

40. Leikola-Pelho T, Jackson DM. Preferential stimulation of locomotor activity by ventral tegmental microinjections of (-)-nicotine. *Pharmacol Toxicol* 1992;70:50–52.

41. Museo E, Wise RA. Place preference conditioning with ventral tegmental injections of cytisine. *Life Sci* 1994;55:1179–1186.

42. Inglis FM, Day JC, Fibiger HC. Enhanced acetylcholine release in hippocampus and cortex during the anticipation and consumption of a palatable meal. *Neuroscience* 1994;62:1049–1056.

43. Wilson FA, Rolls ET. Learning and memory is reflected in the responses of reinforcement-related neurons in the primate basal forebrain. *J Neurosci* 1990;10:1254–1267.

44. Kitai S, Shepard P, Callaway J, et al. Afferent modulation of dopamine neuron firing pattern. *Curr Opin Neurobiol* 1999;9:690–697.

45. Lokwan S, Overton P, Berry M, et al. Stimulation of the pedunculopontine tegmental nucleus in the rat produces burst firing in A9 dopaminergic neurons. *Neuroscience* 1999;92:245–254.

46. Blaha CD, Winn P. Modulation of dopamine efflux in the striatum following cholinergic stimulation of the substantia nigra in intact and pedunculopontine tegmental nucleuslesioned rats. *J Neurosci* 1993;13:1035–1044.

47. Blaha CD, Allen LF, Das S, et al. Modulation of dopamine efflux in the nucleus accumbens after cholinergic stimulation of the ventral tegmental area in intact, pedunculopontine tegmental nucleus-lesioned and laterodorsal tegmental nucleuslesioned rats. *J Neurosci* 1996;16:714–722.

48. Lanca A, Adamson K, Coen K, et al. The pedunculopontine tegmental nucleus and the role of cholinergic neurons in nicotine self-administration in the rat: a correlative neuroanatomical and behavioral study. *Neuroscience* 2000;96:735–742.

49. Olmstead MC, Munn EM, Franklin KB, et al. Effects of pedunculopontine tegmental nucleus lesions on responding for intravenous heroin under different schedules of reinforcement. *J Neurosci* 1998;18:5035–5044.

50. Bechara A, van der Kooy D. A single brain stem substrate mediates the motivational effects of both opiates and food in nondeprived rats but not in deprived rats. *Behav Neurosci* 1992;106:351–363.

51. Olmstead MC, Franklin KB. Effects of pedunculopontine tegmental nucleus lesions on morphine-induced conditioned place preference and analgesia in the formalin test. *Neuroscience* 1993;57:411–418.

52. Olmstead MC, Franklin KB. Lesions of the pedunculopontine tegmental nucleus block drug-induced reinforcement but not amphetamine-induced locomotion. *Brain Res* 1994;638:29–35.

53. Parker JL, van der Kooy D. Tegmental pedunculopontine nucleus lesions do not block cocaine reward. *Pharmacol Biochem Behav* 1995;52:77–83.

54. Koob GF. Drugs of abuse: anatomy, pharmacology and function of reward pathways. *Trends Pharmacol Sci* 1992;13:177–184.

55. Winn P, Brown VJ, Inglis WL. On the relationships between the striatum and the pedunculopontine tegmental nucleus. (Review; 115 refs.) *Crit Rev Neurobiol* 1997;11:241–261.

56. Inglis WL, Dunbar JS, Winn P. Outflow from the nucleus accumbens to the pedunculopontine tegmental nucleus: a dissociation between locomotor activity and the acquisition of responding for conditioned reinforcement stimulated by d-amphetamine. *Neuroscience* 1994;62:51–64.

57. Olmstead MC, Inglis WL, Bordeaux CP, et al. Lesions of the pedunculopontine tegmental nucleus increase sucrose consumption but do not affect discrimination or contrast effects. *Behav Neurosci* 1999;113:732–743.

58. Winn P. Frontal syndrome as a consequence of lesions in the pedunculopontine tegmental nucleus: a short theoretical review. (Review; 141 refs.) *Brain Res Bull* 1998;47:551–563.

59. Everitt BJ, Robbins TW. Central cholinergic systems and cognition. (Review; 210 refs.) *Annu Rev Psychol* 1997;48:649–684.

60. Levin ED, Simon BB. Nicotinic acetylcholine involvement in cognitive function in animals. *Psychopharmacology* 1998;138:217–230.

61. Eagle KS, Chalmers GR, Clary DO, et al. Axonal regeneration and limited functional recovery following hippocampal deafferentation. *J Comp Neurol* 1995;363:377–388.

62. Dunnett SB, Low WC, Iversen SD, et al. Septal transplants restore maze learning in rats with fornix-fimbria lesions. *Brain Res* 1982;251:335–348.

63. Page KJ, Sirinathsinghji DJ, Everitt BJ. AMPA-induced lesions of the basal forebrain differentially affect cholinergic and non-cholinergic neurons: lesion assessment using quantitative *in situ* hybridization histochemistry. *Eur J Neurosci* 1995;7:1012–1021.

64. Waite JJ, Chen AD, Wardlow ML, et al. 192 immunoglobulin g-saporin produces graded behavioral and biochemical changes accompanying the loss of cholinergic neurons of the basal forebrain and cerebellar Purkinje cells. *Neuroscience* 1995;65:463–476.

65. Bucci DJ, Holland PC, Gallagher M. Removal of cholinergic input to rat posterior parietal cortex disrupts incremental processing of conditioned stimuli. *J Neurosci* 1998;18:8038–8046.

66. Voytko ML, Olton DS, Richardson RT, et al. Basal forebrain lesions in monkeys disrupt attention but not learning and memory. *J Neurosci* 1994;14:167–186. [Published erratum appears in *J Neurosci* 1995;15(3 Pt 2) following table of contents.]

67. Baxter MG, Bucci DJ, Gorman LK, et al. Selective immunotoxic lesions of basal forebrain cholinergic cells: effects on learning and memory in rats. *Behav Neurosci* 1995;109:714–722.

68. Chappell J, McMahan R, Chiba A, et al. A re-examination of the role of basal forebrain cholinergic neurons in spatial working memory. *Neuropharmacology* 1998;37:481–487.

69. McAlonan GM, Wilkinson LS, Robbins TW, et al. The effects of AMPA-induced lesions of the septo-hippocampal cholinergic projection on aversive conditioning to explicit and contextual cues and spatial learning in the water maze. *Eur J Neurosci* 1995;7:281–292.

70. Baxter MG, Bucci DJ, Holland PC, et al. Impairments in conditioned stimulus processing and conditioned responding after combined selective removal of hippocampal and neocortical cholinergic input. *Behav Neurosci* 1999;113:486–495.

71. Voytko ML. Cognitive functions of the basal forebrain cholinergic system in monkeys: memory or attention? (Review; (123 refs.) *Behav Brain Res* 1996;75:13–25.

72. McGaughy J, Decker MW, Sarter M. Enhancement of sustained attention performance by the nicotinic acetylcholine receptor agonist ABT-418 in intact but not basal forebrain-lesioned rats. *Psychopharmacology* 1999;144:175–182.

73. Muir JL, Everitt BJ, Robbins TW. AMPA-induced excitotoxic lesions of the basal forebrain: a significant role for the cortical cholinergic system in attentional function. *J Neurosci* 1994;14:2313–2326.

74. Olmstead MC, Robbins TW, Everitt BJ. Basal forebrain cholinergic lesions enhance conditioned approach responses to stimuli predictive of food. *Behav Neurosci* 1998;112:611–629.

75. McAlonan GM, Dawson GR, Wilkinson LO, et al. The effects of AMPA-induced lesions of the medial septum and vertical limb nucleus of the diagonal band of Broca on spatial delayed non-matching to sample and spatial learning in the water maze. *Eur J Neurosci* 1995;7:1034–1049.

76. Leri F, Franklin KB. Learning impairments caused by lesions to the pedunculopontine tegmental nucleus: an artifact of anxiety? *Brain Res* 1998;807:187–192.

77. Steckler T, Keith AB, Sahgal A. Lesions of the pedunculopontine tegmental nucleus do not alter delayed non-matching to position accuracy. *Behav Brain Res* 1994;61:107–112.

78. Ghoneim MM, Mewaldt SP. Studies on human memory: the interactions of diazepam, scopolamine, and physostigmine. *Psychopharmacology* 1977;52:1–6.

79. Rupniak NM, Steventon MJ, Field MJ, et al. Comparison of the effects of four cholinomimetic agents on cognition in primates following disruption by scopolamine or by lists of objects. *Psychopharmacology* 1989;99:189–195.

80. Givens B, Olton DS. Bidirectional modulation of scopolamine-induced working memory impairments by muscarinic activation of the medial septal area. *Neurobiol Learn Mem* 1995;63:269–276.

81. Broersen LM, Heinsbroek RPW, de Bruin JPC, et al. The role of the medial prefrontal cortex of rats in short-term memory functioning: further support for involvement of cholinergic rather than dopaminergic mechanisms. *Brain Res* 1994;674:221–229.

82. Markowska AL, Olton DS, Givens B. Cholinergic manipulations in the medial septal area: age-related effects on working memory and hippocampal electrophysiology. *J Neurosci* 1995;15(3 Pt 1):2063–2073.

83. Raisman G. The connexions of the septum. *Brain* 1966;89:317–348.

84. Kohler C, Chan-Palay V, Wu JY. Septal neurons containing glutamic acid decarboxylase immunoreactivity project to the hippocampal region in the rat brain. *Anat Embryol (Berl)* 1984;169:41–44.

85. Wu M, Shanabrough M, Leranth C, et al. Cholinergic excitation of septohippocampal GABA but not cholinergic neurons: implications for learning and memory. *J Neurosci* 2000;20:8103–8110.

86. Wu M, Shanabrough M, Leranth C, et al. Cholinergic excitation of septohippocampal GABA but not cholinergic neurons: Implications for learning and memory. *J Neurosci* 2000;20:3900–3908.

87. Alreja M, Shanabrough M, Liu W, et al. Opioids suppress IPSCs in neurons of the rat medial septum/diagonal band: involvement of septohippocampal GABAergic neurons. J Neurosci 2000;20:1179–1189.

88. Frotscher M, Leranth C. The cholinergic innervation of the rat fascia dentata: identification of target structures on granule cells by combining choline acetyltransferase immunocytochemistry and Golgi impregnation. *J Comp Neurol* 1986;243:58–70.

89. Miettinen R, Freund TF. Convergence and segregation of septal and median raphe inputs onto different subsets of hippocampal inhibitory interneurons. *Brain Res* 1992;594:263–272.

90. Toth K, Freund TF, Miles R. Disinhibition of rat hippocampal pyramidal cells by GABAergic afferents from the septum. *J Physiol* 1997;500(Pt 2):463–474.

91. Wada E, Wada K, Boulter J, et al. Distribution of alpha2, alpha3, alpha4, and beta2 neuronal nicotinic subunit mRNAs in the central nervous system: a hybridization histochemical study in rat. *J Comp Neurol* 1989;284:314–335.

92. Jones S, Sudweeks S, Yakel JL. Nicotinic receptors in the brain: correlating physiology with function. *Trends Neurosci* 1999;22:555–561.

93. Buzsaki, G., Two-stage model of memory trace formation: a role for "noisy" brain states. *Neuroscience* 1989;31:551–570.

94. Cobb SR, Bulters DO, Suchak S, et al. Activation of nicotinic acetylcholine receptors patterns network activity in the rodent hippocampus. *J Physiol* 1999;518(Pt 1):131–140.

95. Whitehouse PJ, Price DL, Clark AW, et al. Alzheimer disease: evidence for selective loss of cholinergic neurons in the nucleus basalis. *Ann Neurol* 1981;10:122–126.

96. Arendt T, Bigl V, Arendt A. Neurone loss in the nucleus basalis of Meynert in Creutzfeldt-Jakob disease. *Acta Neuropathol* 1984;65:85–88.

97. Mufson EJ, Bothwell M, Kordower JH. Loss of nerve growth factor receptor-containing neurons in Alzheimer's disease: a quantitative analysis across subregions of the basal forebrain. *Exp Neurol* 1989;105:221–232.

98. Nordberg A, Winblad B. Reduced number of (^{3}H)nicotine and (^{3}H)acetylcholine binding sites in the frontal cortex of Alzheimer brains. *Neurosci Lett* 1986;72:115–119.

99. Adler LE, Olincy A, Waldo M, et al. Schizophrenia, sensory gating, and nicotinic receptors. *Schizophr Bull* 1998;24:189–202.

100. Stevens KE, Wear KD. Normalizing effects of nicotine and a novel nicotinic agonist on hippocampal auditory gating in two animal models. *Pharmacol Biochem Behav* 1997;57:869–874.

101. Acri JB, Morse DE, Popke EJ, et al. Nicotine increases sensory gating measured as inhibition of the acoustic startle reflex in rats. *Psychopharmacology* 1994;114:369–374.

102. Rochford J, Sen AP, Quirion R. Effect of nicotine and nicotinic receptor agonists on latent inhibition in the rat. *J Pharmacol Exp Ther* 1996;277:1267–1275.

103. Swerdlow N, Caine S, Braff D, et al. The neural substrates of sensorimotor gating of the startle reflex: a review of recent findings and their implications. *J Psychopharmacol* 1992;6:176–190.

104. Weiner I, Feldon J. The switching model of latent inhibition: an update of neural substrates. *Behav Brain Res* 1997;88:11–25.

105. Freedman R, Hall M, Adler LE, et al. Evidence in postmortem brain tissue for decreased numbers of hippocampal nicotinic receptors in schizophrenia. *Biol Psychiatry* 1995;38:22–33.

106. Caine SB, Geyer MA, Swerdlow NR. Carbachol infusion into the dentate gyrus disrupts sensorimotor gating of startle in the rat. *Psychopharmacology* 1991;105:347–354.

107. Stevens KE, Freedman R, Collins AC, et al. Genetic correlation of inhibitory gating of hippocampal auditory evoked response and alpha-bungarotoxin-binding nicotinic cholinergic receptors in inbred mouse strains. *Neuropsychopharmacology* 1996;15:152–162.

108. McCormick DA, Prince DA. Mechanisms of action of acetylcholine in the guinea-pig cerebral cortex *in vitro. J Physiol* 1986;375:169–194.

2
Serotonin

George K. Aghajanian

Elaine Sanders-Bush

George K. Aghajanian: Departments of Psychiatry and Pharmacology, Yale University School of Medicine, New Haven, Connecticut.

Elaine Sanders-Bush: Department of Pharmacology, Vanderbilt University School of Medicine, Nashville, Tennessee.

Serotonin, or 5-hydroxytryptamine (5-HT), has been implicated in almost every conceivable physiologic or behavioral function—affect, aggression, appetite, cognition, emesis, endocrine function, gastrointestinal function, motor function, neurotrophism, perception, sensory function, sex, sleep, and vascular function (1). Moreover, most drugs that are currently used for the treatment of psychiatric disorders (e.g., depression, mania, schizophrenia, autism, obsessive-compulsive disorder, anxiety disorders) are thought to act, at least partially, through serotoninergic mechanisms (see elsewhere, *this volume*). How is it possible for 5-HT to be involved in so many different processes? One answer lies in the anatomy of the serotoninergic system, in which 5-HT cell bodies clustered in the brainstem raphe nuclei are positioned through their vast projections to influence all regions of the neuraxis. Another answer lies in the molecular diversity and differential cellular distribution of the many 5-HT receptor subtypes that are expressed in brain and other tissues.

During the past decade, molecular cloning techniques have confirmed that putative 5-HT receptor subtypes, predicted from radioligand binding and functional studies (e.g., 5-HT_1, 5-HT_2, 5-HT_3, 5-HT_4), represent separate and distinct gene products. This knowledge has revolutionized contemporary research on the serotoninergic system. Through the use of *in situ* messenger RNA (mRNA) hybridization and immunocytochemical maps, studies of previously recognized 5-HT receptors could be directed more precisely toward neurons and model cell lines that express these specific 5-HT receptor subtypes. Moreover, by the use of cloning techniques, investigations could be initiated to determine the functional role of previously unrecognized 5-HT receptors (e.g., 5-HT_5, 5-HT_6, 5-HT_7). Concurrently, much progress has been made in delineating the signal transduction pathways of the various 5-HT-receptor subtypes. The focus of this review is on the molecular and cellular aspects of individual 5-HT receptor subtypes and their transduction mechanism, in addition to interactions between different receptor subtypes within a single neuron or region. The implications of this work in understanding the global functions of the 5-HT system are discussed.

5-HT RECEPTOR SUBTYPES: MOLECULAR AND CELLULAR ASPECTS

Part of "2 - Serotonin "

Molecular Biology

In the first half of the last decade, the cloning of the major known families of 5-HT receptors was accomplished. More recently, attention has turned to issues of transcriptional and post-transcriptional regulation.

RNA Processing

The 5′-flanking region of several 5-HT-receptor genes has been cloned, and consensus sequences for transcription factors have been identified in the promoter region (2 ,3 and 4). The identification of these potential regulatory sites sets the stage for investigations on possible functionally significant regulation of gene transcription *in vivo* (5). A prominent form of post-transcriptional regulation is alternative RNA splicing, in which the splicing out of intronic sequence varies. Alternative splicing is common and occurs for a number of 5-HT receptors, including the 5-HT_{2C}, 5-HT_4, and 5-HT_7 receptors. The two splice variants of the 5-HT_{2C} receptor described in the literature encode severely truncated proteins with no obvious function (6 ,7 and 8). In contrast, the splice variants of the 5-HT_4 receptor (5-$HT_{4(a)}$-5-$HT_{4(f)}$) and 5-HT_7 receptor (5-$HT_{7(a)}$-5-$HT_{7(d)}$) differ in length and composition in the carboxyl terminus (see refs. 9 and 10 for review). Marked species differences and perhaps regional differences lead to different patterns of splicing. Recently, Claeysen et al. (11) showed that the shortest 5-HT_4 receptor variants have the highest degree of constitutive activity, suggesting that the long tail provides structural stability to the molecule. Splice variants of the 5-HT_7 receptor have no known functional differences. In contrast, a second form of post-transcriptional regulation, RNA editing, tends to

have marked effects on the functional properties of proteins. For example, RNA editing changes a single amino acid in the ß subunit of the AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) receptor, which dictates the gating properties of this ligand-gated ion channel (see ref. 12 for review).

RNA editing in mammalian systems was discovered about a decade ago and is defined as any modification, other than alternative splicing, that occurs at the level of mRNA. Several mechanisms of RNA editing exist, but mammalian editing generally involves the conversion of adenosine residues to inosines by the action of a family of adenosine deaminases (13). Such editing events have the potential to alter the genetic code at the level of RNA; the resulting is the formation of multiple protein isoforms with altered function. The discovery of RNA editing of the 5-HT_{2C} receptor provided the first, and so far only, example of editing of a G protein-coupled receptor (14). Editing of the human 5-HT_{2C} receptor mRNA involves five sites, *A* through *E*, where adenosine is converted to inosine; inosine substitutes for guanosine in the genetic code, thus generating different protein isoforms. Multiple RNA isoforms have been found for the 5-HT_{2C} receptor in human brain, predicting the formation of protein isoforms with up to three amino acids changed in the second intracellular loop of the receptor (15 ,16). Editing at the *A, B, C*, and *D* adenosine residues of human 5-HT_{2C}-receptor mRNA leads to predicted changes in all three amino acids to yield valine, serine, valine (VSV) at positions 156, 158, and 160 rather than isoleucine, asparagine, isoleucine (INI) at these positions in the unedited receptor isoform (Fig. 2.1). Editing at all five sites predicts the formation of the valine, glycine, valine (VGV) isoform. Because the second intracellular loop has been implicated in receptor–G protein coupling, initial functional studies have focused on the intracellular signaling properties of 5-HT_{2C}-receptor isoforms. These studies have shown that edited receptor isoforms couple less efficiently to G_q proteins, evidenced by lowered agonist potencies to activate phospholipase C (7 ,14 ,15) and reduced receptor constitutive activity (16 ,17). The discovery that the 5-HT_{2C} receptor is regulated by RNA editing presents a challenge for pharmacologists because multiple isoforms with potentially different pharmacologic properties and functions are predicted to exist in brain. It is not clear, for example, which receptor isoform should be used for *in vitro* modeling of the receptor and to characterize newly developed drugs. The unedited INI isoform is predicted to represent less than 10% of the total population of receptors in human brain; the principal isoform is VSV (15 ,16). To date, all studies of function have involved recombinant cells expressing a single receptor isoform. Evaluation of the *in vivo* functional consequences of RNA editing of the 5-HT_{2C} receptor awaits the development of experimental methods for isolating the function of a single, specific isoform in brain. Strategies such as the generation of blocking antibodies that target specific amino acid combinations in the second intracellular loop or the development of transgenic mice that express a single isoform may be successful, although they are experimentally challenging and time-consuming. It is not known whether other 5-HT receptors, or for that matter other G protein-coupled receptors, are subject to RNA editing. It seems likely that the 5-HT_{2C} receptor would not be unique. However, screening methods for reliably detecting RNA editing are not available; instead, the discovery of edited substrates depends on comparing genomic and cyclic DNA sequences. Consequently, new edited substrates are slow to emerge.

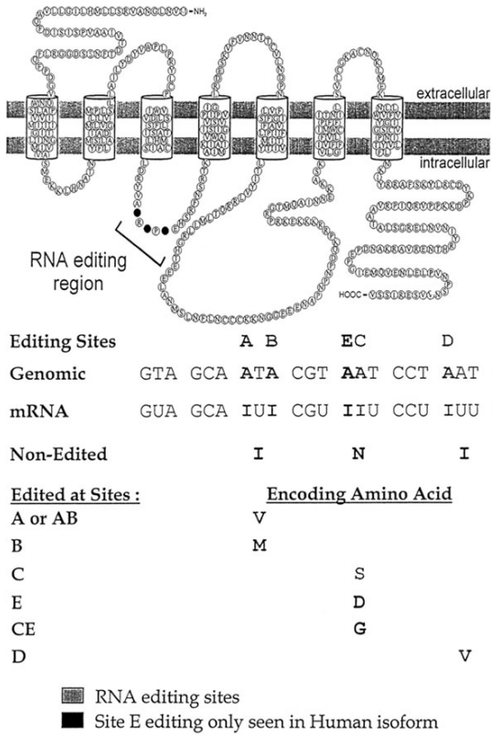


FIGURE 2.1. RNA editing of the 5-hydroxytryptamine subtype 2C (5-HT_{2C}) receptor. Editing of the 5-HT_{2C}-receptor messenger RNA transcript generates multiple receptor isoforms that differ in one to three amino acids in the second intracellular loop.

Post-translational Regulation

Receptor desensitization and down-regulation are common adaptive responses to sustained agonist exposure. The most widely accepted model of desensitization of G protein-coupled receptors is based on extensive studies of the ß-adrenergic receptor, a G_s-linked receptor. In a simplified rendition of the model, agonist binding to a cell surface receptor leads to receptor phosphorylation, arrestin binding, receptor internalization into endosomes, dephosphorylation of the receptor, and recycling back to the cell surface. Receptor phosphorylation is thought to mediate desensitization by uncoupling the receptor from G protein. For many receptors, this phosphorylation event is promoted by a family of G protein-coupled receptor kinases (GRKs). However, second messenger-dependent kinases and protein kinases C and A, in addition to GRKs, have all been implicated in the desensitization of 5-HT_{1A} receptor (18). Abundant *in vivo* studies have documented a blunting of 5-HT_{1A} receptor-mediated behavioral responses after long-term treatment with agonists or serotonin uptake inhibitors that indirectly promote receptor activation. Indeed, desensitization of raphe 5-HT_{1A} autoreceptors has been proposed to play a role in the delayed therapeutic onset of antidepressant drugs (see ref. 19 for review).

Protein kinase C has also been implicated in 5-HT_{2A}-receptor desensitization (20). Subsequent steps in the desensitization-resensitization cycle have been demonstrated for the 5-HT_{2A} receptor, including arrestin binding to the third intracellular loop of the receptor (21) and internalization into endocytotic vesicles (22). Surprisingly, 5-HT_{2A}-receptor antagonists also cause receptor internalization, which may be related to the earlier findings of antagonist down-regulation of 5-HT_{2A} receptors (see ref. 23 for review). Importantly, antagonist-mediated 5-HT_{2A}-receptor internalization has been confirmed in cortical pyramidal cells and is accompanied by an apparent redistribution from dendrites to cell bodies (24). The fact that atypical antipsychotic drugs such as clozapine and olanzapine, but not haloperidol, promote 5-HT_{2A}-receptor internalization has led to speculation that this novel antagonist property may be related to therapeutic action in schizophrenia.

Receptor Constitutive Activity and Inverse Agonism

In *Psychopharmacology: The Fourth Generation of Progress*, the concept of inverse agonist properties of serotonin-receptor antagonists was novel and unique to the 5-HT_{2C} receptor (25). *Inverse agonism* is the ability of certain antagonists to block the spontaneous (also referred to as *constitutive*) activity of a G protein-coupled receptor, in addition to blocking the binding of an agonist. These antagonists are referred to as *inverse agonists* because their effects are opposite to those of agonists. In contrast, other antagonists, referred to as *neutral antagonists*, have no apparent activity when added alone, even though they block the action of an agonist. To explain receptor constitutive activity and inverse agonism of antagonists, the model of receptor–G protein coupling was modified to propose spontaneous receptor isomerization to an active form (*R**) in the absence of an agonist (26). Antagonists with inverse agonist activity bind to and stabilize the inactive receptor conformation (*R*), whereas neutral antagonists were proposed to bind equally well to the active and inactive forms of the receptor. Since 1995, 5-HT_{1A} receptors (27, 28), 5-HT_{1B} receptors (29), 5-HT_{1D} receptors (30), mutant 5-HT_{2A} receptors (31), 5-HT_4 receptors (11 ,32), and 5-HT_7 receptors (33) have been shown to exhibit constitutive activity; both inverse agonists and neutral antagonists have been described for these receptors.

All these studies demonstrating inverse agonist properties of 5-HT antagonists have been performed *in vitro* in cells expressing recombinant receptors. It is not known whether inverse agonism is relevant to the *in vivo* actions of receptor antagonists, including their therapeutic properties. Recent *in vivo* studies of a simple motor reflex have produced convincing evidence for constitutive activity of the 5-$HT_{2A/2C}$ receptor and differential effects of inverse agonists versus neutral antagonists (34 ,35). However, Millan et al. (36) were unable to show differential effects of inverse agonists versus neutral antagonists at the 5-HT_{1B} receptor and concluded that *in vitro* demonstrations of inverse agonist activity cannot be extrapolated to the *in vivo* situation. Recent studies by Berg et al. (37) suggest that even in the absence of measurable effects on basal activity, prolonged treatment with inverse agonists at the 5-HT_{2C} receptor produces enhanced phospholipase C activation, likely because of increased expression of G_q proteins.

Electrophysiology

Although the electrophysiologic actions of 5-HT may seem quite varied, considerable uniformity is found within each of the major receptor families. For example, all members of the 5-HT_1 family tend to have inhibitory actions either presynaptically or postsynaptically. Similarly, all members of the 5-HT_2 family tend to have excitatory actions. Therefore, the discussion of 5-HT electrophysiology is organized according to receptor family subtypes.

5-HT_1 Receptors

Dense concentrations of 5-HT_{1A} binding sites and high levels of 5-HT_{1A} mRNA expression are found in a number of regions, including the dorsal raphe nucleus, hippocampal pyramidal cell layer, and cerebral cortex (38 ,39 and 40). Studies in these regions have been useful in delineating the physiologic role of this receptor.

Raphe Nuclei

Serotoninergic neurons of the raphe nuclei are inhibited by the local (microiontophoretic) application of 5-HT to their cell body region. Thus, the receptor mediating this effect has been termed a *somatodendritic autoreceptor* (as opposed to the prejunctional autoreceptor). Early studies in the dorsal raphe nucleus showed that lysergic acid diethylamide (LSD) and other indolamine hallucinogens are powerful agonists at the somatodendritic 5-HT autoreceptor (41 ,42). Functionally, the somatodendritic 5-HT autoreceptor has been shown to mediate collateral inhibition (43). The ionic basis for the autoreceptor-mediated inhibition, either by 5-HT or LSD, is an opening of K^+ channels to produce a hyperpolarization (44); these channels are characterized by their inwardly rectifying properties (45). As in the dorsal raphe nucleus, serotoninergic neurons of the lower brainstem are also hyperpolarized by 5-HT via the opening of inwardly rectifying K^+ channels (46 ,47). Similar findings in acutely isolated (48) and individually microcultured (49) dorsal raphe neurons underscore the fact that autoreceptor inhibition is independent of any inputs to the raphe nucleus. Patch-clamp recordings in the cell-attached and outside-out configuration from such acutely isolated dorsal raphe neurons show that the increase in K^+ current results from a greater probability of opening of unitary K^+ channel activity (50).

The somatodendritic autoreceptors of serotoninergic neurons in both the dorsal raphe nucleus and the nucleus raphe magnus appear to be predominantly of the 5-HT_{1A} subtype; a variety of drugs with 5-HT_{1A} selectivity (e.g., 8-OH-DPAT and the anxiolytic drugs buspirone and ipsapirone) share the ability to inhibit raphe cell firing potently in a dose-dependent manner (47 ,50a ,50b). Recently, a highly selective 5-HT_{1A} antagonist (WAY 100635) has been found that potently blocks the direct inhibition of dorsal raphe serotoninergic neurons both by 5-HT and selective 5-HT_{1A} agonists (51 ,52). WAY 100635 also blocks the indirect inhibition of dorsal raphe neurons induced by selective 5-HT reuptake inhibitors (53). Complementing this autoinhibitory role of local 5-HT_{1A} receptors is a long-loop negative-feedback system activated by postsynaptic 5-HT_{1A} receptors in the medial prefrontal cortex (54 ,55).

In addition to long-loop feedback systems, short-loop regulatory circuits are found within the dorsal raphe nucleus and the adjacent periaqueductal gray (PAG). These short-loop circuits involve interactions between 5-HT and local inhibitory GABAergic (γ-aminobutyric acid) and excitatory glutamatergic neurons (56). Interestingly, both the local excitatory and inhibitory inputs to 5-HT cells are negatively modulated by μ opiate receptors. Local GABAergic neurons are activated by 5-HT via 5-$HT_{2A/2C}$ receptors in a local, negative feedback loop that complements 5-HT_{1A}-mediated autoinhibition (57). Neurokinins such as substance P and neurokinin B, via NK_1 and NK_3 receptors, respectively, activate mostly local glutamatergic excitatory inputs to 5-HT cells (58). Some of these local circuits are depicted schematically in Fig. 2.2 .

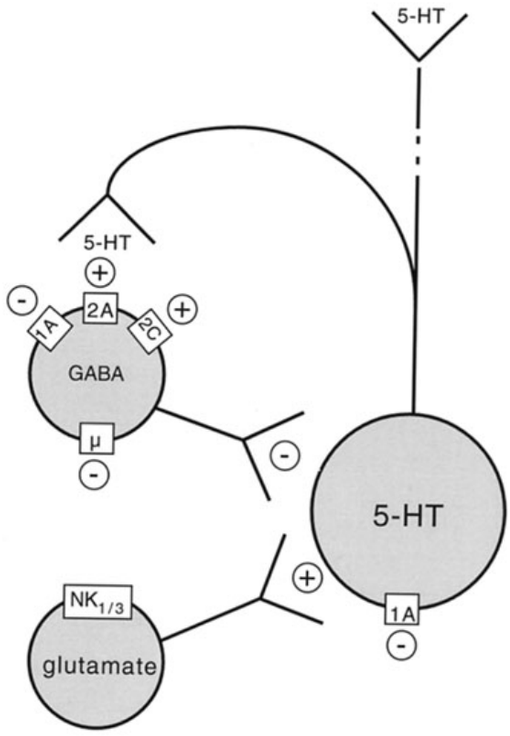


FIGURE 2.2. Schematic representation of local regulatory circuitry within the dorsal raphe nucleus (DRN). In addition to somatodendritic 5-hydroxytryptamine subtype 1A (5-HT_{1A}) autoreceptors on the 5-HT neurons *per se*, local GABAergic (γ-aminobutyric acid) and glutamatergic neurons in the DRN/ventral periaqueductal gray (PAG) region modulate the activity of serotoninergic neurons. Note the location of inhibitory μ opiate receptors on both categories of local neurons. Also depicted are excitatory 5-$HT_{2A/2C}$ and inhibitory 5-HT_{1A} receptors on GABAergic neurons and excitatory NK_1 (substance P) and NK_3 (neurokinin B) receptors on glutamate neurons in the DRN/PAG.

Other Subcortical Regions

Inhibitory or hyperpolarizing responses to 5-HT have been reported in a wide variety of neurons in the spinal cord, brainstem, and diencephalon. In general, such responses have been attributed to mediation by 5-HT_1 receptors. In sensory neurons of dorsal root ganglia, a 5-HT_1-like receptor has been reported to reduce the calcium component of action potentials and to produce hyperpolarizations that can be mimicked by 5-HT_{1A} agonists such as 8-OH-DPAT (59). In cerebellar Purkinje cells, 5-HT-induced inhibition, but not excitation, is mediated through 5-HT_{1A} receptors (60). In brain slices of the nucleus prepositus hypoglossi, focal electric stimulation evokes inhibitory postsynaptic potentials (IPSPs) that are mediated by 5-HT_{1A} receptors to activate an inwardly rectifying K^+ conductance (61) and a novel outwardly rectifying K^+ conductance (62). In the midbrain PAG, a region known to be involved in pain modulation and fear responses, approximately half the cells are inhibited/hyperpolarized by 8-OH-DPAT, suggesting mediation by 5-HT_{1A} receptors (63). In the ventromedial hypothalamus (64) and lateral septum (65 ,66), 5-HT and 5-HT_{1A} agonists produce inhibitory effects, also by activating a K^+ conductance. In addition to these postsynaptic effects, 5-HT has been shown to suppress glutamatergic synaptic transmission via presynaptic 5-HT_{1B} receptors in various regions, including the hypoglossal nucleus (67) and the nucleus accumbens (68).

In the rat laterodorsal tegmental nucleus, bursting cholinergic neurons are hyperpolarized by 5-HT via 5-HT_1 receptors (69). In freely behaving rats, the direct injection of 5-HT into the laterodorsal tegmental nucleus has been found to suppress rapid-eye-movement (REM) sleep (70). In unanesthetized cats, a corresponding population of neurons that are active selectively during REM states (REM-on neurons) in the laterodorsal tegmental nucleus has been shown to be inhibited by direct application of the 5-HT_{1A} agonist 8-OH-DPAT (71). It has been proposed that during REM sleep, the removal of a tonic inhibitory 5-HT influence from these cholinergic neurons may be responsible for the emergence of an activated EEG during this behavioral state.

Hippocampus

Pyramidal cells of the CA1 region express high levels of 5-HT_{1A}-receptor mRNA and 5-HT_{1A}-receptor binding (72). Early on, intracellular recordings in brain slices showed that the 5-HT-induced inhibition was caused by hyperpolarization resulting from an opening of K^+ channels (73). Subsequent work, in which various pharmacologic approaches have been used in brain slices, has shown that the 5-HT-induced inhibition in both CA1 and CA3 pyramidal cells is mediated by the activation of receptors of the 5-HT_{1A} subtype (74 ,75 ,76 and 77). After long-term but not short-term administration of various antidepressant treatments (selective 5-HT reuptake inhibitors, monoamine oxidase inhibitors,

tricyclic drugs, electroconvulsive therapy), disinhibitory responses are seen with the selective 5-HT_{1A} antagonist WAY 100635, which suggests increased 5-HT_{1A}-mediated inhibitory tone on CA3 hippocampal pyramidal cells (78). Interestingly, this increase in 5-HT_{1A} tone after long-term antidepressant treatment is potentiated by short-term treatment with lithium (79).

In addition to the above-mentioned direct effects on pyramidal cells, 5-HT has been shown to depress both excitatory and inhibitory synaptic potentials in the hippocampus. Relatively high concentrations of 5-HT cause a reduction in electrically evoked excitatory postsynaptic potentials (EPSPs) in CA1 pyramidal cells (80), an effect that is mimicked by 8-OH-DPAT, which suggests mediation by 5-HT_{1A} receptors. Indirect measures indicate that 5-HT acts presynaptically to reduce Ca^{2+} entry and thereby glutamatergic synaptic transmission. In addition, a 5-HT_{1A}-mediated inhibitory effect on putative inhibitory interneurons of the hippocampus has been observed (81 ,82). Consistent with an opening of K^+ channels, the inhibitory effects of 5-HT on interneurons result from a hyperpolarization associated with a reduction in input resistance. Functionally, the 5-HT_{1A}-mediated inhibition of GABAergic interneurons in the hippocampus leads to a disinhibition of pyramidal cells in CA1. Clearly, the effects of 5-HT in the hippocampus are highly complex, involving both presynaptic and postsynaptic actions that may, to varying degrees, be inhibitory or disinhibitory, facilitative or disfacilitative.

Cerebral Cortex

Hyperpolarizing/inhibitory responses in pyramidal cells of the cerebral cortex induced by 5-HT_{1A} have been described in a number of studies. In entorhinal cortex, where the density of 5-HT_{1A} receptors is especially high (and the density of 5-HT_{2A} receptors low), unopposed 5-HT_{1A} receptor-mediated hyperpolarizing responses are seen (83). However, cortical neurons in most other regions typically display mixed inhibitory and excitatory responses to 5-HT because of expression by the same pyramidal cells of multiple 5-HT receptor subtypes (e.g., 5-HT_{1A} and 5-$HT_{2/2C}$) (84 ,85 ,86 and 87). Hyperpolarizing responses mediated by 5-HT_{1A} receptors are often unmasked or enhanced in the presence of 5-HT_2 antagonists, consistent with the idea that an interaction occurs between 5-HT_{1A} and 5-HT_{2A} receptors at an individual neuronal level (84 ,88 ,89). A similar suggestion of a shift in the balance between 5-HT-mediated excitation and inhibition comes from another *in vivo* study, in which both systemic and local application of 5-HT_2 antagonists was shown to prevent an enhancement of the unit activity (and cortical desynchronization) that normally occurs in response to noxious stimuli (tail compression) in anesthetized rats (90).

In addition to the above-mentioned postsynaptic effects, various presynaptic effects are mediated by 5-HT_1 receptors in the cerebral cortex. In cingulate cortex, 5-HT, acting on presynaptic 5-HT_{1B} receptors, reduces the amplitude of electrically evoked EPSPs, including both *N*-methyl-D-aspartate (NMDA) and non-NMDA components (87). Similar modulations of EPSPs, mediated by 5-HT_{1A} or 5-HT_{1B} receptors, have been reported for several cortical regions, including medial prefrontal (91) and entorhinal cortex (92).

5-HT_2 Receptors

Quantitative autoradiographic studies show high concentrations of 5-HT_2 binding sites and mRNA expression in certain regions of the forebrain, such as the neocortex (layers IV/V), piriform cortex, claustrum, and olfactory tubercle (93). With few notable exceptions (e.g., motor nuclei and the nucleus tractus solitarius), relatively low concentrations of 5-HT_2 receptors or mRNA expression are found in the brainstem and spinal cord. Studies aimed at examining the physiologic role of 5-HT_2 receptors in several of these regions are discussed in the following sections.

Motoneurons

In the facial and other cranial motor nuclei, motoneurons have a high density of 5-HT_2-receptor binding sites. Early studies *in vivo* showed that 5-HT applied microiontophoretically does not by itself induce firing in the normally quiescent facial motoneurons, but does facilitate the subthreshold and threshold excitatory effects of glutamate (94). Intracellular recordings from facial motoneurons *in vivo* or in brain slices *in vitro* (95 ,96) show that 5-HT induces a slow, subthreshold depolarization associated with an increase in input resistance, indicating a decrease in resting K^+ conductance. Ritanserin and other 5-HT_2 antagonists are able to block the excitatory effects of 5-HT in facial motoneurons selectively (97). Indolamine (e.g., LSD and psilocin) and phenethylamine (e.g., mescaline and DOI) hallucinogens act as 5-HT_2 agonists at facial motoneurons. Iontophoretically administered LSD, mescaline, and psilocin, although having relatively little effect by themselves, produce a prolonged facilitation of facial motoneuron excitability (98). Intracellular studies in brain slices show that the enhancement is in part caused by a small but persistent depolarizing effect of the hallucinogens (97 ,99).

Other Subcortical Regions

In brain slices of the medial pontine reticular formation, 5-HT induces a hyperpolarization in some cells and a depolarization in other cells (100). The hyperpolarizing responses are associated with an increase in membrane conductance and have a 5-HT_1 pharmacologic profile. The depolarizing responses have a 5-HT_2 pharmacology and are associated with a decrease in membrane conductance resulting from a decrease in an outward K^+ current. These two actions of 5-HT do not appear to coexist in the same neurons because none of the cells display dual responses to selective agonists.

In brain slices of the substantia nigra pars reticulata, a majority of neurons are excited by 5-HT via 5-HT_2 receptors (101), possibly of the 5-HT_{2C} rather than 5-HT_{2A} subtype (102). Neurons in the inferior olivary nucleus are excited by 5-HT via 5-HT_{2A} receptors, so that the oscillatory frequency of input to cerebellar Purkinje cells is altered (103). In the nucleus accumbens, the great majority of neurons are depolarized by 5-HT, and they are induced to fire (104). This depolarization is associated with an increase in input resistance secondary to a reduction in an inward rectifier K^+ conductance. Pharmacologic analysis shows that the depolarization is mediated by a 5-HT_2- rather than a 5-HT_1- or 5-HT_3-type receptor.

GABAergic neurons of the nucleus reticularis thalami show marked depolarizing responses to 5-HT, associated with a decrease in a resting or "leak" K^+ conductance; these excitatory responses are blocked by the 5-HT_2 antagonists ketanserin and ritanserin (105). The 5-HT-induced slow depolarization potently inhibits burst firing in these cells and promotes single-spike activity. It has been suggested that the 5-HT-induced switch in firing mode from rhythmic oscillation to single-spike activity, which occurs during states of arousal and attentiveness, contributes to the enhancement of information transfer through the thalamus during these states. GABAergic neurons within the medial septal nucleus are also excited by 5-HT via 5-HT_2 receptors (106). In the dentate gyrus of the hippocampus, a subpopulation of GABAergic neurons is activated via 5-HT_{2A} receptors, evidenced by an increase in IPSP frequency in granule cells in the dentate gyrus (107). Recently, similar activation of GABAergic neurons via 5-HT_{2A} receptors has been reported in the CA1 region of the hippocampus (108). These observations closely resemble findings in the piriform cortex, where a subpopulation of GABAergic interneurons is excited by 5-HT via 5-HT_{2A} receptors (see below). Also, indirect evidence suggests that 5-HT-induced inhibition of dentate/interpositus neurons of the deep cerebellar nuclei is mediated indirectly by the activation of GABAergic interneurons through 5-HT_2 receptors (109). Taken together, these findings suggest that in multiple locations within the central nervous system, excitation of subpopulations of interneurons by 5-HT via 5-HT_2 receptors gives rise to indirect inhibitory effects.

Cerebral Cortex

The electrophysiologic effects of 5-HT have been studied in several cortical regions. *In vitro* studies in brain slice preparations have shown that pyramidal cells in various cortical regions respond to 5-HT by either a small hyperpolarization, depolarization, or no change in potential (84 ,85 and 86 ,110). Depending on the region of cortex under study, as described below, the depolarizations appear to be mediated by 5-HT_{2A} or 5-HT_{2C} receptors.

In addition to these postsynaptic effects, 5-HT induces an increase in "spontaneous" (not electrically evoked) postsynaptic potentials or currents (PSPs/PSCs) in brain slices from various cortical regions. Originally, recordings were made from pyramidal cells in a paleocortical region, the piriform cortex. In that region, as in the hippocampus (see above), 5-HT, acting through 5-HT_{2A} receptors, induces an increase in spontaneous IPSPs (86 ,111 ,112 ,113 ,114 and 115). *In vivo* studies have also provided evidence for a 5-HT_{2A} receptor-mediated activation of GABAergic neurons in piriform cortex (116). As in piriform cortex, 5-HT can increase IPSCs in pyramidal cells in various layers of the neocortex (117 ,118). The IPSCs result from the activation of cortical interneurons via 5-$HT_{2A/2C}$ or 5-HT_3 receptors (117). Immunocytochemical evidence has been found in primate cerebral cortex for a segregation of 5-HT_{2A}- and 5-HT_3-expressing interneurons; the former project to somatobasilar and the latter to distal apical dendritic regions of pyramidal cells (119).

Quantitatively, in layer V pyramidal cells, synaptic events induced by 5-HT consist largely of EPSPs/EPSCs (118). Thus, approximately 85% of all PSCs are blocked by AMPA/kainate glutamate-receptor antagonists (e.g., LY293558) but not by the $GABA_A$ antagonist bicuculline (118). The 5-HT-induced increase in EPSCs is most pronounced in frontal regions, including the medial prefrontal cortex (118). In that region, 5-HT_{2A} receptors are denser than in more posterior regions (40 ,120). Recent studies, in which intracellular labeling with biocytin was used, have confirmed that 5-HT-induced increases in EPSCs occur predominantly in layer V pyramidal cells, whereas responses are minimal in layer II/III cells and lacking in layer VI cells (121).

The 5-HT-induced EPSCs are antagonized competitively by low concentrations of the highly selective 5-HT_{2A} antagonist MDL 100,907 (pA_2, 8.8), which indicates mediation by 5-HT_{2A} receptors (118 ,122). Norepinephrine, via α_1 adrenoceptors, also induces an increase in EPSPs in layer V pyramidal cells, but (at least in the rat) the increase is only a fraction of that produced by 5-HT (122). Changes in the frequency of synaptic currents or potentials are generally regarded as indicative of a modulation of presynaptic function. Accordingly, the nonspecific group II/III metabotropic glutamate receptor agonist (1*S*,3*S*)-ACPD (118) and the selective mGlu II/III metabotropic agonist LY354740 (123), which act at inhibitory autoreceptors and are located on glutamatergic nerve terminals, suppress the 5-HT-induced increase in the frequency of EPSCs. In addition, the activation of μ receptors, located presynaptically on thalamocortical inputs, also suppresses 5-HT-induced EPSCs, particularly in the medial prefrontal cortex (124). These results are consistent with the idea that activation of 5-HT_{2A} receptors increases the release of glutamate onto layer V pyramidal cells through a presynaptic mechanism. 5-HT also produces a small but significant increase in the *amplitude* of spontaneous EPSCs, an effect that may involve postsynaptic amplification mechanisms (118). Such postsynaptic

effects are consistent with the finding of a high level of 5-HT_{2A}-receptor immunoreactivity in the apical dendrites of cortical pyramidal cells (125 ,126 and 127).

The 5-HT-induced EPSCs are blocked by bath application of the slice with the fast sodium channel blocker tetrodotoxin or perfusion with a solution containing no added calcium (“zero” calcium) (118). Ordinarily, tetrodotoxin sensitivity and Ca^{2+} dependence would suggest that activation of glutamatergic cells within the slice by 5-HT had led to an impulse flow-dependent release of glutamate. However, several lines of evidence argue against this conventional interpretation. First, rarely were neurons within the confines of the brain slice induced to fire by bath application of 5-HT. Second, none of the pyramidal cells (a potential source of intracortical excitatory inputs) was depolarized sufficiently by 5-HT as recorded under our conditions to reach the threshold for firing. Third, EPSCs can be induced by the microiontophoresis of 5-HT onto “hot spots” along the trajectory of apical dendrites of layer V pyramidal cells (118). Together, these experiments suggest that the increase in spontaneous EPSCs induced by 5-HT in neocortical pyramidal cells occurs through a focal action involving a Ca^{2+}-dependent mechanism that is not based on an increase in impulse flow in excitatory afferents.

As an alternative to a conventional impulse flow-related mechanism, it has been hypothesized that the 5-HT-induced EPSCs result from an activation of the “asynchronous” release pathway (128). One of several distinguishing characteristics of this alternative mechanism of transmitter release is that Sr^{2+} can substitute for Ca^{2+} in the asynchronous, but not synchronous, release (129). This feature appears to be the result of the differential involvement of two isoforms of the calcium-sensing protein synaptotagmin in the two alternative release mechanisms (130). Consistent with this idea, Sr^{2+} is highly effective in enabling 5-HT to induce an increase in the frequency of EPSCs in the absence of Ca^{2+} (128).

Recently, it has been found that LSD and other hallucinogenic drugs, acting as partial agonists at 5-HT_{2A} receptors, promote a late component of electrically evoked EPSPs (131). It is possible that this late component, rather than representing conventional polysynaptic transmission, is mediated through the mechanism of asynchronous transmitter release, possibly involving a release of intraterminal Ca^{2+} stores via the phospholipase C, inositol triphosphate (IP_3) pathway. An enhancement of asynchronous evoked EPSPs via 5-HT_{2A} receptors would provide a possible synaptic mechanism for the hallucinogenic effects of these drugs. In contrast, 5-HT itself does not promote the late component of *electrically evoked* release except during the washout phase, presumably because of opposing actions at 5-HT_1 or other non-5-HT_{2A} receptors (132). Figure 2.3 depicts the proposed location of various 5-HT-receptor subtypes and their interactions with other neurotransmitter receptors within cortical circuitry.

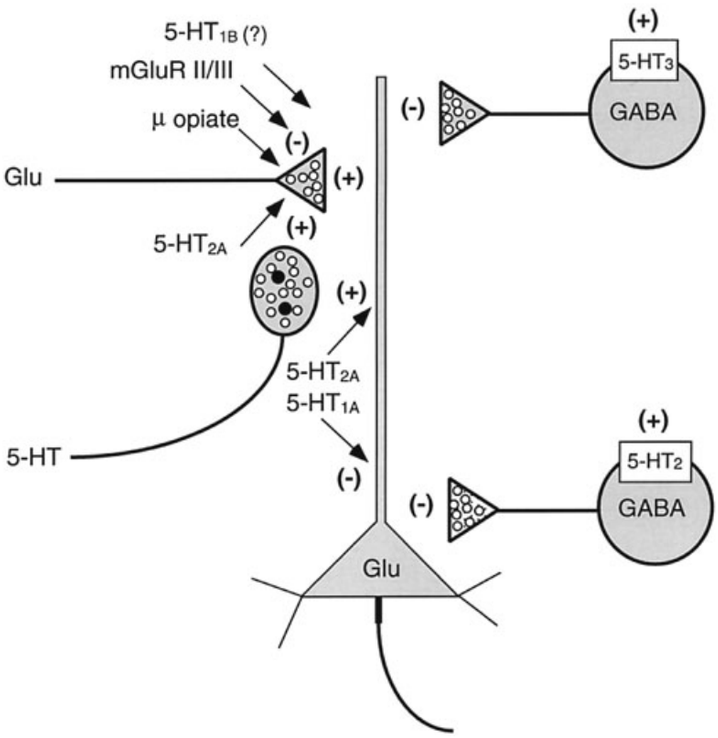


FIGURE 2.3. Modulation of excitatory and inhibitory transmission by multiple 5-hydroxytryptamine (5-HT) receptors in the cerebral cortex. 5-HT_{2A} receptors are depicted as enhancing glutamate release from a glutamatergic terminal onto a layer V pyramidal cell; the same terminal is seen to be negatively modulated by various G_i/G_o-coupled receptors (e.g., μ-opiate, 5-HT_{1B}, and mGluR II/III). In addition, 5-HT_{2A} receptors are shown to have a direct postsynaptic excitatory effect that is opposed by postsynaptic 5-HT_{1A} receptors. Finally, 5-HT_2 and 5-HT_3 receptors are shown on anatomically distinct GABAergic inputs to the somatobasilar and apical regions, respectively, of the pyramidal cell.

5-HT_3 Receptors

Excitatory responses to 5-HT have been found in various central neurons that have many of the characteristics of peripheral 5-HT_3 responses, including rapid onset and rapid desensitization, features typical of ligand-gated ion channels rather than G protein-coupled receptor responses (133 ,134). In cultured NG108-15 cells, the permeation properties of the 5-HT_3 channel are indicative of a cation channel with relatively high permeability to Na^+ and K^+ and low permeability to Ca^{2+} (134). A 5-HT-gated ion channel has been cloned that has physiologic and pharmacologic properties appropriate for a 5-HT_3 receptor (135). In the oocyte expression system, this receptor shows rapid desensitization and is blocked by 5-HT_3 antagonists (e.g., ICS 205–930 and MDL 72222). Its sequence homology with the nicotinic acetylcholine receptor (27%) and the $ß_1$ subunit of the $GABA_A$ receptor (22%) indicates that this 5-HT_3-receptor clone is a member of the ligand-gated ion channel superfamily. Typically, members of this superfamily are comprised of multiple subunits; however, only one 5-HT_3-receptor subunit and an alternatively spliced variant have been cloned to date (136).

In hippocampus slices, 5-HT has been reported to increase spontaneous GABAergic IPSPs, most likely through a 5-HT_3 receptor-mediated excitation of inhibitory interneurons; these responses also show fading with time (137 ,138). A similar 5-HT_3 receptor-mediated induction of IPSCs has been reported in the neocortex (117). Whole-cell patch-clamp recordings have confirmed a direct 5-HT_3 receptor-mediated excitatory effect on hippocampal interneurons independent of G-protein activation (139). Although fast, rapidly inactivating excitation has generally become accepted as characteristic of 5-HT_3 receptors, nondesensitizing responses have also been reported. In dorsal root ganglion cells, a relatively rapid but noninactivating depolarizing response has been described that has a 5-HT_3 pharmacologic profile (140). In neurons of nucleus tractus solitarius brain slices, a postsynaptic depolarizing response to 5-HT_3 agonists has been observed that is not rapidly desensitizing (141). In addition to these postsynaptic effects, a 5-HT_3 receptor-mediated increase in Ca^{2+} influx has been described in a subpopulation of striatal nerve terminals (142).

5-HT_4, 5-HT_6, and 5-HT_7 Receptors

The first known protein G_s-coupled 5-HT receptor, the 5-HT_4 receptor, was identified on the basis of pharmacologic and biochemical criteria (e.g., positive coupling of 5-HT responses to adenylyl cyclase) (9). Subsequently, a receptor with matching pharmacologic and other properties was cloned and found to be expressed in various regions of the brain (143). Two other 5-HT receptors positively coupled to adenylyl cyclase have been cloned. Because their pharmacology differed from that of the previously described 5-HT_4 site, they were designated as 5-HT_6 and 5-HT_7 receptors (144 ,145 and 146). At this time, electrophysiologic studies are available only for the 5-HT_4 and 5-HT_7 receptors and are described below.

5-HT_4 Receptors

Binding studies using a selective 5-HT_4 ligand indicate that 5-HT_4 receptors are present in several discrete regions of the mammalian brain, including the striatum, substantia nigra, olfactory tubercle, and hippocampus (147). Because these regions also express 5-HT_4-receptor mRNA, it appears likely that the receptors function postsynaptically to mediate certain actions of 5-HT. The best studied of these regions is the hippocampus, in which both biochemical and electrophysiologic studies have provided a detailed picture of the actions of 5-HT at 5-HT_4 receptors. Electrophysiologic studies show that 5-HT_4 receptors mediate an inhibition of a calcium-activated potassium current that is responsible for the generation of a slow after-hyperpolarization in hippocampal pyramidal cells of the CA1 region (74 ,148 ,149). A suppression of the after-hyperpolarization would enhance the ability of these cells to respond to excitatory inputs with robust spike activity.

5-HT_7 Receptors

The circadian rhythm in mammals is set by a pacemaker located primarily in the suprachiasmatic nucleus of the hypothalamus. This pacemaker activity can be maintained in hypothalamic slices, in which suprachiasmatic neurons display diurnal changes in neuronal firing rate. Administration of 5-HT appears to produce a phase shift in this activity (150) by acting on a receptor that may be of the 5-HT_7 subtype (144). This shift is mediated by stimulation of adenylate cyclase because it is mimicked by increasing intracellular cyclic adenosine monophosphate (cAMP) and blocked by inhibiting protein kinase A (151). However, the precise mechanism by which 5-HT_7 receptors act is not presently known because it is unclear whether suprachiasmatic neurons themselves express the 5-HT_7 receptors (144). Furthermore, the effect of 5-HT on the membrane properties of these cells has not been examined. 5-HT_7 receptor activation has been reported to inhibit $GABA_A$ currents on suprachiasmatic neurons in culture (152), but the relationship, if any, between these observations and 5-HT changes in circadian activity remains to be determined.

Another electrophysiologic effect that may be mediated through 5-HT receptors that are positively coupled to adenylate cyclase is the enhancement of the hyperpolarizing-activated nonselective cationic current I_h. The I_h channels, which are homologous to cyclic nucleotide-gated channels in specialized sensory neurons, are positively modulated by cAMP (153 ,154). An increase in I_h tends to prevent excessive hyperpolarization and increase neuronal excitability. In a number of regions of the brain, including the thalamus (155), prepositus hypoglossi (156), substantia nigra zona compacta (157), and hippocampus (158), 5-HT has been shown to enhance I_h through a cAMP-dependent mechanism. Results of a pharmacologic analysis with multiple nonselective drugs suggested that the increase in I_h induced by 5-HT in dorsal root ganglion cells is mediated by 5-HT_7 receptors (159). Recently, the first drug with selectivity toward the 5-HT_7 receptor was shown to block activation of adenylyl cyclase by 5-HT agonists in guinea pig hippocampus (33). The increasing availability of such selective drugs should greatly enhance the electrophysiologic evaluation of G_s-coupled 5-HT receptors.

INTRACELLULAR SIGNAL TRANSDUCTION PATHWAYS

Part of "2 - Serotonin "

Multiple Signaling Pathways: G Proteins and Second Messengers

Multiple intracellular signaling pathways constitute a common theme for G protein-coupled receptors, and the 5-HT receptor family is not unique. Inhibition of adenylate cyclase

was the first intracellular pathway to be described for $G_{i/o}$ protein-coupled receptors, such as the 5-HT_{1A} receptor. However, it is now clear that these receptors regulate multiple signaling pathways and effector molecules (Fig. 2.2), including activation of G protein-gated inwardly rectifying K^+ (GIRK) channels, inhibition of voltage-sensitive Ca^{2+} channels, activation of phospholipase Cß, and activation of mitogen-activated protein kinase (see ref. 18 for review). Although all these signals are sensitive to pertussis toxin, so that $G_{i/o}$ proteins are implicated, they may be mediated by distinct G protein complexes. For example, coupling to GIRK channels is mediated by ßγ subunits released from G_i (and possibly G_o) proteins, whereas inhibition of Ca^{2+} channels is mediated by ßγ subunits released from G_o proteins. The profile of signaling molecules varies from cell to cell, offering diverse signaling possibilities and contributing additional complexity. For example, 5-HT_{1A} receptor activation of phospholipase C is cell-type dependent; this signal is mediated by G protein ßγ subunits and thus requires the presence of a ßγ-regulated phospholipase C isoform. The ßγ subunits, generated by dissociation of the heterotrimeric G_i protein, also activate the type 2 isoform of adenylate cyclase. This activation is conditional, dependent on the coactivation by $G_{\alpha s}$ (i.e., $G_{\alpha i}$ potentiates the action of $G_{\alpha s}$). The obvious question is why the opposing actions of $G_{\alpha i}$ and $G_{\beta\gamma}$ do not offset each other. The answer may lie in the details. In addition to the large family of G proteins (21 α subunits, 5 ß subunits, and 11 γ subunits), the adenylate cyclase family comprises at least nine members, each regulated by distinct inputs. Most of these molecules are found in the central nervous system. The G protein that contributes ßγ activation of type 2 adenylate cyclase is $G_{\alpha i1}$ or $G_{\alpha i2}$ heterotrimer (160), whereas all three $G_{\alpha i}$ subunits ($\alpha_{i3} > \alpha_{i2} > \alpha_{i1}$) have the ability to inhibit adenylate cyclase types 5 and 6 (161). Thus, in cells in the brain in which $G_{\alpha i1}$ or $G_{\alpha i2}$ heterotrimer is coexpressed with type 2 adenylate cyclase, 5-HT_{1A}-receptor activation may potentiate G_s-mediated increases in cAMP. This type of interaction has been shown to occur in brain, in which G_i-linked receptors enhance ß-adrenergic responses (162); a similar interaction may take place in cells that coexpress a 5-HT_{1A} receptor family member with one of the 5-HT receptors (5-HT_4, 5-HT_6, or 5-HT_7) linked to activation of adenylate cyclase.

Although a 5-HT receptor-mediated increase in cAMP formation in superior colliculus was one of the earliest second messenger pathways defined in brain, the 5-HT_4 receptor was one of the last 5-HT receptors to be cloned (143). This receptor and the 5-HT_6 and 5-HT_7 receptors have in common the ability to activate adenylate cyclase via $G_{\alpha s}$ (Fig. 2.2). In transfected cells, the 5-HT_6 receptor couples to adenylate cyclase type 5, the typical $G_{\alpha s}$-sensitive isoform (163). In contrast, the 5-HT_7 receptor increases intracellular calcium, which activates calmodulin-stimulated adenylate cyclase type 1 or 8. A recent characterization of rat hippocampal homogenates suggests that both the 5-HT_4 and 5-HT_7 receptors are involved in cAMP formation (adenylate cyclase isoform unknown) in the hippocampus (164). Interestingly, the 5-HT_{1A} receptor produces a slight increase in cAMP formation, perhaps reflecting $G_{\beta\gamma}$ potentiation of G_s activation of adenylate cyclase type 2 mediated by the 5-HT_4 or 5-HT_7 receptor.

Receptors that couple to the G_q family members (G_q, G_{11}, G_{14}, and $G_{15/16}$) activate phospholipase C in a pertussis toxin-insensitive manner. Activation of phospholipase C was the first signal transduction mechanism identified for the 5-HT_2-receptor family and is essentially universal. This probably reflects the wide distribution of $G_{q/11}$ and the functional redundancy of these two G proteins. The 5-HT_{2C} receptor has been shown to couple in a pertussis toxin-sensitive manner to $G_{i/o}$ in *Xenopus* oocytes (e.g., see ref. 165) and in some transfected cell lines (166). In contrast, recent evidence suggests that phospholipase C activation in a native setting (choroid plexus) is mediated entirely by $G_{q/11}$ coupling (167). Coupling of the 5-HT_{2C} receptor to G_{13} with subsequent cytoskeletal rearrangement has been recently described in a transfected cell line (168). Extensive evidence suggests that 5-HT_{2A} and 5-HT_{2C} receptors couple to other effector pathways, in addition to phospholipase C (Fig. 2.4). Phospholipase A_2 is a well-characterized independent signal transduction pathway that leads to arachidonic acid, with subsequent prostaglandin and leukotriene formation (169). 5-HT_{2A}-receptor activation of mitogen-activated protein kinase has been extensively characterized in vascular smooth muscle and is also thought to be independent of phospholipase C activation (170 ,171). The 5-HT_{2A} receptor increases phospholipase D activity via a small G-protein ARF (adenosine diphosphate ribosylation factor) pathway, with protein kinase C activation being the principal consequence (172). 5-HT_{2A} receptors differentially regulate brain-derived neurotrophic factor in hippocampus and cortex and play a role in stress-induced down-regulation of brain-derived neurotrophic factor expression in hippocampus (173 ,174). In addition, a 5-HT_{2A} receptor-mediated increase in transforming growth factor-α_1, secondary to protein kinase C activation, has been described (175). The 5-HT_{2A} and 5-HT_{2C} receptors elicit region-specific increases in immediate early genes c-*fos* and *Arc* in rat brain (176), which are likely downstream of phospholipase C activation. Extensive, complex cross-talk between the 5-HT_{2A} and 5-HT_{2B} receptor and the 5-$HT_{1B/D}$ receptor has been demonstrated in immortalized serotoninergic cells, in which the 5-HT_{2B} receptor, via a phospholipase A_2 product, attenuates 5-$HT_{1B/D}$ receptor-mediated adenylate cyclase inhibition (177). Coactivation of the 5-HT_{2A} receptor blocks this interaction by an unknown mechanism. These examples of parallel, interacting, and converging intracellular signaling pathways illustrate the complexity of receptor signaling, even within a single receptor subclass.

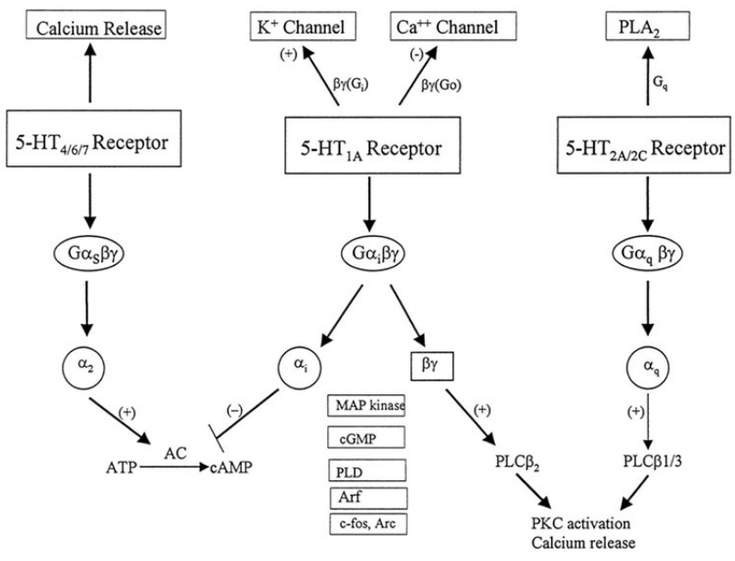


FIGURE 2.4. Examples of potential converging and interacting signaling pathways for 5-hydroxytryptamine-receptor subtypes. This figure illustrates only a few of the nearly unlimited possibilities, depending on the cell phenotype. Also listed are additional effectors activated by one or another of these receptors with pathways of activation that have not yet been determined.

Physiologic Correlates

In general, the electrophysiologic effects of 5-HT correspond well to the G-protein and second messenger coupling of the various receptor subtypes. The G_i/G_o-coupled 5-HT_1 receptors generally mediate inhibitory effects on neuronal firing through an opening of inwardly rectifying K^+ channels or a closing of voltage-gated Ca^{2+} channels. Inhibitions mediated by 5-HT_1 receptors have been observed in neurons located in diverse regions of the central nervous system, ranging from pyramidal cells of the cerebral cortex and hippocampus to serotoninergic neurons of the brainstem raphe nuclei. The $G_{q/11}$-coupled 5-HT_2 family of receptors generally mediates slow excitatory effects through a decrease in K^+ conductance or an increase in nonselective cation conductance. Slow excitatory effects mediated by 5-HT_2 receptors have been observed in a number of regions, including the spinal cord and brainstem (e.g., motoneurons), subcortical regions (e.g., nucleus accumbens), and cerebral cortex, where these receptors are most concentrated. The 5-HT_3 receptors, which are ligand-gated channels with structural homology to nicotinic cholinergic receptors, mediate fast excitatory effects of 5-HT. Specific examples are given below for 5-HT_1, 5-HT_2, and 5-HT_4 receptors, for which intracellular transduction pathways have been studied most intensively.

5-HT_1 Receptors

The opening of K^+ channels via 5-HT_{1A} receptors in dorsal raphe neurons is mediated by pertussis toxin-sensitive G proteins (178 ,179). The molecular mechanisms underlying the opening of K^+ channels are most likely common to all neurotransmitter receptors that couple through the G_i/G_o family of G proteins. As in the dorsal raphe, these receptors activate a pertussis toxin-sensitive G protein that couples to the opening of inwardly rectifying K^+ channels through a membrane-delimited pathway (74 ,180). It is widely accepted that the βγ rather than α subunits regulate the channels (181 ,182 and 183). The effector mechanism that ultimately mediates the inhibitory effect signaled by 5-HT_{1A} receptors is the inwardly rectifying K^+ channel. Interestingly, at least one of the potassium K^+ subunits identified in heart, GIRK-1, is expressed at high levels in hippocampus (184), which suggests that it might be involved in mediating the 5-HT_{1A} receptor-induced hyperpolarization in this region.

Consistent with this possibility, the K^+ current activated by 5-HT_{1A} receptors in the CA1 region does show the characteristic signature of this potassium channel family-namely, inward rectification (74).

5-HT_2 Receptors

The role of G proteins in mediating the 5-HT_2-induced slow inward current that results from K^+ channel closure has been evaluated in facial motoneurons by using the hydrolysis-resistant guanine nucleotide analogues GTPγS and GDPβS (185). The 5-HT-induced inward current becomes largely irreversible in the presence of intracellular GTPγS. Mediation by G proteins is also suggested by the fact that the inward current is reduced by intracellular GDPβS, which prevents G-protein activation. Although the identity of the G protein(s) mediating the electrophysiologic responses has not yet been determined directly, the 5-HT_2 family of receptors is known to be coupled to phospholipase C. Thus, a member of the $G_{q/11}$ family may be involved because the latter can directly activate phospholipase C (186).

5-HT_4 Receptors

Initially, it was shown that 5-HT suppresses a calcium-activated potassium current that is responsible for the generation of a slow after-hyperpolarization in hippocampus pyramidal cells of the CA1 region (see above). Subsequent studies have implicated 5-HT_4 receptors, acting via cAMP and protein kinase A, in mediating this action (187). A similar activation of a cAMP-dependent protein kinase has been implicated in the suppression of a voltage-activated K^+ current in cultured neurons from the superior colliculus (188). More recently, it has been shown that 5-HT_4 receptors reduce after-hyperpolarization in hippocampus pyramidal cells by inhibiting calcium-induced calcium release from intracellular stores (189).

Pharmacologic Significance

The pharmacologic significance of a single receptor regulating multiple signaling pathways is only just beginning to be defined. The most explicit studies of promiscuous coupling of receptors to multiple G-protein signaling pathways have involved transfection of a recombinant receptor into various cell models that do not normally express the receptor. Powerful genetic strategies involving antisense techniques, overexpression of signaling molecules, and expression of constitutively active and dominant negative mutants have exposed a multitude of fascinating possibilities for a single receptor to sculpt multiple signals depending on the properties of the cell. In addition, theoretical arguments (190) and experimental evidence (191 ,192 and 193) have appeared in support of the novel concept of agonist-directed trafficking of the intracellular signal. This model proposes that when a single receptor interacts with multiple signaling pathways, the pattern of intracellular signaling may differ depending on the agonist. Although the mechanism of agonist-specified signaling is not known, one possibility is that different agonists promote distinct receptor conformations, thereby exposing interfacial domains with altered protein–protein interaction properties. All these studies in artificial conditions tell us only what can occur, not what does occur *in vivo*. Techniques for studying the role of multiple signaling pathways in native preparations are needed to tease out the significance of the various signaling molecules in normal physiology and in pathologic states. Transgenic and knockout strategies have some utility; however, targeting signaling molecules will have a multitude of unwanted consequences because of their universal role in cell physiology. Another strategy was recently described that has significant potential for teasing out signaling pathways downstream of receptor activation (167). Synthetic blocking peptides targeting specific protein–protein interactions in a signaling pathway are rendered membrane-permeable by a novel conjugation reaction, so that the function of a particular signaling step in native systems can be defined.

BEHAVIORAL CORRELATES

Part of "2 - Serotonin "

5-HT Neuronal Activity and Behavioral State

In a variety of mammalian species, serotoninergic neurons of the raphe nuclei have been found to have a slow, tonic pattern of firing (approximately one to two spikes per second). The maintenance of rhythmic firing under a wide variety of conditions has suggested that serotoninergic neurons possess intrinsic tonic pacemaker mechanisms. Intracellular recordings from dorsal raphe neurons show that spikes arise from gradual depolarizing ramps (pacemaker potentials) rather than synaptic potentials. The pacemaker rhythm of serotoninergic neurons results from a complex interplay of intrinsic ionic currents (e.g., a voltage-dependent transient outward potassium current, a low-threshold inward calcium current, and a calcium-activated outward potassium current) (194). Also modulating the activity of serotoninergic neurons are various neurotransmitters, including norepinephrine and 5-HT itself. Norepinephrine, acting via α_1 adrenoceptors, accelerates pacemaker activity of serotoninergic neurons by closing potassium channels. Conversely, 5-HT itself, acting on 5-HT_{1A} autoreceptors, opposes excessive activity of serotoninergic neurons.

The highly regulated pacemaker activity of serotoninergic neurons suggests that the 5-HT system serves an important homeostatic function. Through its effects on neuronal excitability in diverse regions of the brain and spinal cord, the serotoninergic system is in a strategic position to coordinate complex sensory and motor patterns during different

behavioral states. Recordings from serotoninergic neurons in unanesthetized animals have shown that activity is highest during periods of waking arousal, reduced in quiet waking, reduced further in slow-wave sleep, and absent during REM (dream) sleep (195). It can be hypothesized that the function of the 5-HT system, by its coordinated fluctuations in activity, is to promote a given behavioral state. This concept is illustrated in the following scenario. When serotoninergic neurons are in a tonic firing mode, the following conditions would prevail: (a) Motoneurons would be in a relatively depolarized, excitable state (via 5-HT_2 receptors) and thus receptive to the initiation of movement; (b) neurons of the nucleus reticularis thalami would be in a depolarized, single-spike mode (via 5-HT_2 receptors) and thus conducive to thalamocortical sensory information transfer (105 ,155); (c) GABAergic neurons of the septohippocampal pathway would be activated (in part via 5-HT_{2A} receptors), potentially enhancing long-term potentiation by inhibiting GABAergic neurons of the hippocampus (106 ,196); (d) neurons of the laterodorsal tegmental nucleus would be hyperpolarized (via 5-HT_1 receptors) and therefore not able to generate the bursting activity of REM sleep (69 ,70 and 71). Conversely, with a reduction in serotoninergic activity during various stages of sleep, the above conditions would switch such that motoneurons would become less excitable, thalamocortical sensory information transfer would be diminished, hippocampal function would be reduced, and sleep spindles and pontogeniculo–occipital (PGO) waves would emerge.

Molecular Genetics (Including Genetic Polymorphisms)

5-HT-Receptor/Transporter Knockouts

New drugs are beginning to appear that show considerable selectivity for a particular serotonin receptor subtype; however, many are not yet readily available to the general scientific community. Genetically modified mice that fail to express a specific receptor provide a powerful means to complement pharmacologic tools for evaluating the behavioral consequences of a particular serotonin-receptor protein (see ref. 197 for review). The first 5-HT-receptor knockout mouse was described in 1994 (198), in which the 5-HT_{1B} receptor was eliminated by homologous recombination technique. These original studies showed markedly enhanced aggression in 5-HT_{1B}-receptor knockout mice. Since then, altered responses to drugs of abuse, including enhanced alcohol consumption (199) and sensitization to cocaine (200), in addition to impaired paired-pulse inhibition (201) and paradoxical sleep (202), have been shown to be prominent phenotypic traits. In 1995, a “knockout” mouse line expressing a mutant, nonfunctional 5-HT_{2C} receptor was described (203). Subsequently, enhanced seizure susceptibility (204), obesity and late-onset diabetes (205), and a specific deficit in dentate gyrus long-term potentiation (206) have been reported. Mouse lines have recently been generated that are null for other important 5-HT-related molecules; these including the 5-HT_{1A} receptor, which is associated with enhanced anxiety (207 ,208 and 209), the serotonin transporter, with enhanced cocaine sensitivity (210 ,211), and the 5-HT_{5A} receptor, with reduced sensitivity to LSD (212). Although monoamine oxidase A-null mice have general alterations in biogenic amine dynamics, evidence suggests that the enhanced levels of 5-HT found in these mice are associated with neurodevelopmental abnormalities (213 ,214). Innovative technologies such as inducible, conditional knockouts, which have the potential for temporally and spatially controlling gene manipulation, hold great promise for the future. This is illustrated in a recent study in which localized rescue of knocked-out genes was used to study the differential sorting of the 5-HT_{1A} and 5-HT_{1B} receptor in striatal neurons (215). In these transgenic mice, but not transfected neurons in culture, reproduction of the normal targeting of the 5-HT_{1B} receptor to axon terminals set the stage for mutagenesis studies of molecular determinants of receptor targeting to axon terminals *in vivo*.

Genetic Polymorphisms

Molecules involved in brain 5-HT pathways have been favorite targets for candidate gene studies, and the number of publications dealing with genetic variations in 5-HT systems has increased dramatically during the past few years. Recent population studies have probed for single nucleotide polymorphisms in synthetic enzymes, inactivation molecules, and receptors for 5-HT. The list of human diseases studied is extensive and includes obsessive-compulsive disorder, major depression, bipolar depression, schizophrenia, Alzheimer’s disease, eating disorders, anxiety, neuroticism, fibromyalgia, alcoholism, suicide, homicide, substance abuse, pathologic gambling, and responses to psychotherapeutic agents. Despite the abundance of publications, no definitive, reproducible links between allelic variants of 5-HT-related molecules have been found in human populations with behavioral disorders or brain diseases. More often than not, results are not reproducible from study to study, in large part because of the heterogeneous nature of psychiatric diseases, the absence of a specific diagnostic laboratory test, and the modest numbers of patients in many studies. The most extensively studied genetic polymorphism in a 5-HT-related molecule is the insertion/deletion polymorphism in the promoter region of the 5-HT transporter gene (216). These variable-length polymorphisms are biologically significant because *in vitro* studies have shown that the short form reduces the expression of transporter mRNA, with subsequently reduced uptake capacity (217). Although many studies suggest that the short form is associated with affective disorders, others have failed to replicate these findings (218).

Some commonly studied polymorphisms, such as the C103T variant in the 5-HT_{2A} receptor, are silent (i.e., do not change the genetic code), whereas other polymorphisms, such as the 5-HT_{2C} receptor Cys23Ser allele (219), produce mutant proteins with no apparent alterations in functional properties. The clinical importance of such a subtle genetic variant may require analysis of other related genes in tandem. Methods for detecting genetic polymorphisms are advancing rapidly and now allow simultaneous genotyping of several nucleotide polymorphisms; for example, a method was recently described to detect multiple single-nucleotide polymorphisms of 5-HT-related genes (220).

OVERVIEW AND CONCLUSIONS

Part of "2 - Serotonin "

This review has emphasized recent developments in molecular, transductional, and cellular aspects of the 5-HT system. Molecular topics that were hardly mentioned in the previous edition of this book include RNA editing, post-translational processing, genetic polymorphisms, and the use of selective 5-HT-receptor and transporter gene knockouts. Notable developments in cellular physiology since the previous edition include growing numbers of studies on the more recently cloned 5-HT-receptor subtypes (e.g., 5-HT_{3-7}) and refinements in the analysis of 5-HT_1- and 5-HT_2-receptor function. Examples of the latter include the following: (a) the recognition that the well-known 5-HT_{1A} autoreceptor feedback regulation of 5-HT neurons occurs within the context of a complex set of local and long-loop regulatory circuits; (b) the finding that 5-HT_2 receptors have a dramatic influence on cortical information processing, which has allowed new insights into the mechanism of action of hallucinogenic and atypical antipsychotic drugs. In turn, advances in molecular and cellular research on individual receptor subtypes have provided new experimental tools for the behavioral analysis of the 5-HT system (e.g., pharmacologic agents with more precisely defined actions and gene knockouts). The question remains of whether the diverse cellular and molecular actions of 5-HT mediated by the various receptor subtypes can be incorporated into a holistic scheme that can define the overall function of the 5-HT system. Selected examples have been given of how the 5-HT system can be seen as modulating, in a complex but coordinated fashion, a number of motor, sensory, and other systems to promote a given behavioral state or function. The recent molecular and cellular advances, by enabling a more comprehensive analysis of the elementary actions of individual 5-HT-receptor subtypes, have set the stage for a more precise analysis of overall function.

ACKNOWLEDGMENTS

Part of "2 - Serotonin "

This work was supported by grants from the National Institute of Mental Health, the National Institute on Drug Abuse, and the state of Connecticut.

REFERENCES

1. Bloom FE, Kupfer DJ. *Psychopharmcology: the fourth generation of progress*. New York: Raven Press, 1995:407–471.

2. Albert PR, Lembo P, Storring JM, et al. The 5-HT_{1A} receptor: signaling, desensitization, and gene transcription [see Comments]. *Neuropsychopharmacology* 1996;14:19–25.

3. Bedford FK, Julius D, Ingraham HA. Neuronal expression of the 5-HT_3 serotonin receptor gene requires nuclear factor 1 complexes. *J Neurosci* 1998;18:6186–6194.

4. Zhu QS, Chen K, Shih JC. Characterization of the human 5-HT_{2A} receptor gene promoter. *J Neurosci* 1995;15:4885–4895.

5. Lesch KP, Heils A. Serotoninergic gene transcriptional control regions: targets for antidepressant drug development? *Int J Neuropsychopharmacol* 2000;3:67–79.

6. Canton H, Emeson RB, Barker EL, et al. Identification, molecular cloning, and distribution of a short variant of the 5-hydroxytryptamine$_{2C}$ receptor produced by alternative splicing. *Mol Pharmacol* 1996;50:799–807.

7. Wang Q, O'Brien PJ, Chen CX, et al. Altered G protein-coupling functions of RNA editing isoform and splicing variant serotonin$_{2C}$ receptors. *J Neurochem* 2000;74:1290–1300.

8. Xie E, Zhu L, Zhao L, et al. The human serotonin 5-HT_{2C} receptor: complete cDNA, genomic structure, and alternatively spliced variant. *Genomics* 1996;35:551–561.

9. Bockaert J, Fozard JR, Dumuis A, et al. The 5-HT_4 receptor: a place in the sun. *Trends Pharmacol Sci* 1992;13:141–145.

10. Hamblin MW, Guthrie CR, Kohen R, et al. G_s protein-coupled serotonin receptors: receptor isoforms and functional differences. *Ann N Y Acad Sci* 1998;861:31–37.

11. Claeysen S, Sebben M, Becamel C, et al. Novel brain-specific 5-HT_4 receptor splice variants show marked constitutive activity: role of the C-terminal intracellular domain [In Process Citation]. *Mol Pharmacol* 1999;55:910–920.

12. Seeburg PH. The role of RNA editing in controlling glutamate receptor channel properties. *J Neurochem* 1996;66:1–5.

13. Reuter H, Porzig H. Localization and functional significance of the Na^+/Ca^{2+} exchanger in presynaptic boutons of hippocampal cells in culture. *Neuron* 1995;15:1077–1084.

14. Burns CM, Chu H, Rueter SM, et al. Regulation of serotonin-2C receptor G-protein coupling by RNA editing [see Comments]. *Nature* 1997;387:303–308.

15. Fitzgerald LW, Iyer G, Conklin DS, et al. Messenger RNA editing of the human serotonin 5-HT_{2C} receptor. *Neuropsychopharmacology* 1999;21:82S–90S.

16. Niswender CM, Copeland SC, Herrick-Davis K, et al. RNA editing of the human serotonin 5-hydroxytryptamine 2C receptor silences constitutive activity. *J Biol Chem* 1999;274:9472–9478.

17. Herrick-Davis K, Grinde E, Niswender CM. Serotonin 5-HT_{2C} receptor RNA editing alters receptor basal activity: implications for serotoninergic signal transduction. *Neurochemistry* 1999;73:1711–1717.

18. Raymond JR, Mukhin YV, Gettys TW, et al. The recombinant 5-HT_{1A} receptor: G protein coupling and signalling pathways. *Br J Pharmacol* 1999;127:1751–1764.

19. Pineyro G, Blier P. Autoregulation of serotonin neurons: role in antidepressant drug action. *Pharmacol Rev* 1999;51:533–591.

20. Roth BL, Palvimaki EP, Berry S, et al. 5-Hydroxytryptamine$_{2A}$ (5-HT_{2A}) receptor desensitization can occur without down-regulation. *J Pharmacol Exp Ther* 1995;275:1638–1646.

21. Gelber EI, Kroeze WK, Willins DL, et al. Structure and function of the third intracellular loop of the 5-hydroxytryptamine$_{2A}$ receptor: the third intracellular loop is alpha-helical and binds purified arrestins. *J Neurochem* 1999;72:2206–2214.

22. Berry SA, Shah MC, Khan N, et al. Rapid agonist-induced internalization of the 5-hydroxytryptamine$_{2A}$ receptor occurs via the endosome pathway *in vitro. Mol Pharmacol* 1996;50:306–313.

23. Roth BL, Berry SA, Kroeze WK, et al. Serotonin 5-HT$_{2A}$ receptors: molecular biology and mechanisms of regulation. *Crit Rev Neurobiol* 1998;12:319–338.

24. Willins DL, Berry SA, Alsayegh L, et al. Clozapine and other 5-hydroxytryptamine-2A receptor antagonists alter the subcellular distribution of 5-hydroxytryptamine-2A receptors *in vitro* and *in vivo. Neuroscience* 1999;91:599–606.

25. Sanders-Bush E, Canton H. *Serotonin receptors: signal transduction pathways*. New York: Raven Press, 1995.

26. Samama P, Cotecchia S, Costa T, et al. A mutation-induced activated state of the beta 2-adrenergic receptor. Extending the ternary complex model. *J Biol Chem* 1993;268:4625–4636.

27. Newman-Tancredi A, Conte C, Chaput C, et al. Inhibition of the constitutive activity of human 5-HT$_{1A}$ receptors by the inverse agonist spiperone but not the neutral antagonist WAY 100,635. *Br J Pharmacol* 1997;120:737–739.

28. Newman-Tancredi A, Verriele L, Chaput C, et al. Labelling of recombinant human and native rat serotonin 5-HT$_{1A}$ receptors by a novel, selective radioligand, [^{3}H]-S 15535: definition of its binding profile using agonists, antagonists and inverse agonists. *Naunyn Schmiedebergs Arch Pharmacol* 1998;357:205–217.

29. Selkirk JV, Scott C, Ho M, et al. SB-224289—a novel selective (human) 5-HT$_{1B}$ receptor antagonist with negative intrinsic activity. *Br J Pharmacol* 1998;125:202–208.

30. Thomas DR, Faruq SA, Balcarek JM, et al. Pharmacological characterisation of [^{35}S]-GTPγS binding to Chinese hamster ovary cell membranes stably expressing cloned human 5-HT$_{1D}$ receptor subtypes. *J Recept Signal Transduct Res* 1995;15:199–211.

31. Egan CT, Herrick-Davis K, Teitler M. Creation of a constitutively activated state of the 5-hydroxytryptamine$_{2A}$ receptor by site-directed mutagenesis: inverse agonist activity of antipsychotic drugs. *J Pharmacol Exp Ther* 1998;286:85–90.

32. Blondel O, Gastineau M, Langlois M, et al. The 5-HT$_{4}$ receptor antagonist ML10375 inhibits the constitutive activity of human 5-HT$_{4(c)}$ receptor. *Br J Pharmacol* 1998;125:595–597.

33. Thomas DR, Middlemiss DN, Taylor SG, et al. 5-HT stimulation of adenylyl cyclase activity in guinea pig hippocampus: evidence for involvement of 5-HT$_{7}$ and 5-HT$_{1A}$ receptors. *Br J Pharmacol* 1999;128:158–164.

34. Harvey JA, Welsh SE, Hood H, et al. Effect of 5-HT$_{2}$ receptor antagonists on a cranial nerve reflex in the rabbit: evidence for inverse agonism. *Psychopharmacology (Berl)* 1999;141:162–168.

35. Welsh SE, Romano AG, Harvey JA. Effects of serotonin 5-HT$_{(2A/2C)}$ antagonists on associative learning in the rabbit. *Psychopharmacology (Berl)* 1998;137:157–163.

36. Millan MJ, Gobert A, Audinot V, et al. Inverse agonists and serotoninergic transmission: from recombinant, human serotonin (5-HT)$_{1B}$ receptors to G-protein coupling and function in corticolimbic structures *in vivo* [In Process Citation]. *Neuropsychopharmacology* 1999;21:61S–67S.

37. Berg KA, Stout BD, Cropper JD, et al. Novel actions of inverse agonists on 5-HT$_{2C}$ receptor systems. *Mol Pharmacol* 1999;55:863–872.

38. Chalmers DT, Watson SJ. Comparative anatomical distribution of 5-HT$_{1A}$ receptor mRNA and 5-HT$_{1A}$ binding in rat brain—a combined *in situ* hybridisation/*in vitro* receptor autoradiographic study. *Brain Res* 1991;561:51–60.

39. Miquel MC, Doucet E, Boni C, et al. Central serotonin$_{1A}$ receptors: respective distributions of encoding mRNA, receptor protein and binding sites by *in situ* hybridization histochemistry, radioimmunohistochemistry and autoradiographic mapping in the rat brain. *Neurochem Int* 1991;19:453–465.

40. Pazos A, Palacios JM. Quantitative autoradiographic mapping of serotonin receptors in the rat brain. I. Serotonin-1 receptors. *Brain Res* 1985;346:205–230.

41. Aghajanian GK, Foote WE, Sheard MH. Lysergic acid diethylamide: sensitive neuronal units in the midbrain raphe. *Science* 1968;161:706–708.

42. Aghajanian GK, Haigler HJ, Bloom FE. Lysergic acid diethylamide and serotonin: direct actions on serotonin-containing neurons in rat brain. *Life Sci* 1972;11:615–622.

43. Wang RY, Aghajanian GK. Antidromically identified serotoninergic neurons in the rat midbrain raphe: evidence for collateral inhibition. *Brain Res* 1977;132:186–193.

44. Aghajanian GK, Lakoski JM. Hyperpolarization of serotoninergic neurons by serotonin and LSD: studies in brain slices showing increased K^{+} conductance. *Brain Res* 1984;305:181–185.

45. Williams JT, Colmers WF, Pan ZZ. Voltage- and ligand-activated inwardly rectifying currents in dorsal raphe neurons *in vitro. J Neurosci* 1988;8:3499–3506.

46. Bayliss DA, Li Y-W, Talley M. Effects of serotonin on caudal raphe neurons: activation of an inwardly rectifying potassium current. *J Neurophysiol* 1997;77:1349–1361.

47. Pan ZZ, Wessendorf MW, Williams JT. Modulation by serotonin of the neurons in rat nucleus raphe magnus *in vitro. Neuroscience* 1993;54:421–429.

48. Penington NJ, Kelly JS, Fox AP. Whole-cell recordings of inwardly rectifying K^{+} currents activated by 5-HT$_{1A}$ receptors on dorsal raphe neurones in the adult rat. *J Physiol* 1993;469:387–405.

49. Johnson MD. Electrophysiological and histochemical properties of postnatal rat serotoninergic neurons in dissociated cell culture. *Neuroscience* 1994;63:775–787.

50. Penington NJ, Kelly JS, Fox AP. Unitary properties of potassium channels activated by 5-HT in acutely isolated rat dorsal raphe neurones. *J Physiol* 1993;469:407–426.

50a. Sprouse JS, Aghajanian GK. Electrophysiological responses of serotoninergic dorsal raphe neurons to 5-HT1A and 5-HT1B agonists. *Synapse* 1987;1:3–9.

50b. Sprouse JS, Aghajanian GK. Responses of hippocampal pyramidal cells to putative serotonin 5-HT1A and 5-HT1B agonists: a comparative study with dorsal raphe neurons. *Neuropsychopharmacology* 1988;27:707–715.

51. Craven R, Graham-Smith D, Newberry N. Way-100635 and GR127935: effects on 5-hydroxytryptamine-containing neurons. *Eur J Pharmacol* 1994;271:R1–R3.

52. Forster EA, Cliffe IA, Bill DJ, et al. A pharmacological profile of the selective silent 5-HT$_{1A}$ receptor antagonist, WAY-100635. *Eur J Pharmacol* 1995;281:81–88.

53. Gartside SE, Umbers V, Hajos M, et al. Interaction between a selective 5-HT$_{1A}$ receptor antagonist and an SSRI *in vivo*: effects of 5-HT cell firing and extracellular 5-HT. *Br J Pharmacol* 1995;115:1064–1070.

54. Ceci A, Baschirotto A, Borsini F. The inhibitory effect of 8-OH-DPAT on the firing activity of dorsal raphe serotoninergic neurons in rats is attenuated by lesion of the frontal cortex. *Neuropsychopharmacology* 1994;33:709–713.

55. Hajos M, Hajos-Korcsok E, Sharp T. Role of the medial prefrontal cortex in 5-HT$_{1A}$ receptor-induced inhibition of 5-HT neuronal activity in the rat. *Br J Pharmacol* 1999;126:1741–1750.

56. Jolas T, Aghajanian GK. Opioids suppress spontaneous and NMDA-induced inhibitory postsynaptic currents in the dorsal raphe nucleus of the rat *in vitro. Brain Res* 1997;755:229–245.

57. Liu R, Jolas T, Aghajanian GK. Serotonin 5-HT$_{2}$ receptors activate local GABA inhibitory inputs to serotoninergic neurons of the dorsal raphe nucleus. *Brain Res* 2000;873:34–45.

58. Liu RJ, Aghajanian GK. Neurokinins activate local glutamatergic inputs to serotoninergic neurons of the dorsal raphe nucleus. *Soc Neurosci Abst* 2001 *(in press)*.

59. Scroggs RS, Anderson EG. 5-HT_1 receptor agonists reduce the Ca^+ component of sensory neuron action potentials. *Eur J Pharmacol* 1990;178:229–232.

60. Darrow EJ, Strahlendorf HK, Strahlendorf JC. Response of cerebellar Purkinje cells to serotonin and the 5-HT_{1A} agonists 8-OH-DPAT and ipsapirone *in vitro. Eur J Pharmacol* 1990;175:145–153.

61. Bobker DH, Williams JT. Serotonin-mediated inhibitory postsynaptic potential in guinea-pig prepositus hypoglossi and feedback inhibition by serotonin. *J Physiol (Lond)* 1990;422:447–462.

62. Bobker DH, Williams JT. The serotoninergic inhibitory postsynaptic potential in prepositus hypoglossi is mediated by two potassium currents. *J Neurosci* 1995;15:223–229.

63. Behbehani MM, Liu H, Jiang M, et al. Activation of $serotonin_{1A}$ receptors inhibits midbrain periaqueductal gray neurons of the rat. *Brain Res* 1993;612:56–60.

64. Newberry NR. 5-HT_{1A} receptors activate a potassium conductance in rat ventromedial hypothalamic neurones. *Eur J Pharmacol* 1992;210:209–212.

65. Joels M, Shinnick-Gallagher P, Gallagher JP. Effect of serotonin and serotonin analogues on passive membrane properties of lateral septal neurons *in vitro. Brain Res* 1987;417:99–107.

66. Van den Hooff P, Galvan M. Actions of 5-hydroxytryptamine and 5-HT_{1A} receptor ligands on rat dorsolateral septal neurones *in vitro. Br J Pharmacol* 1992;106:893–899.

67. Singer JH, Bellingham MC, Berger AJ. Presynaptic inhibition of glutamatergic synaptic transmission to rat motoneurons by serotonin. *J Neurophysiol* 1996;76:799–807.

68. Muramatsu M, Danet M, Lapiz S, et al. Serotonin inhibits synaptic glutamate currents in rat nucleus accumbens neurons via presynaptic 5-HT_{1B} receptors. *Eur J Neurosci* 1998;10:2371–2379.

69. Luebke JI, Greene RW, Semba K, et al. Serotonin hyperpolarizes cholinergic low-threshold burst neurons in the rat laterodorsal tegmental nucleus *in vitro. Proc Natl Acad Sci U S A* 1992;89:743–747.

70. Horner R, Sanford LD, Annis D, et al. Serotonin at the laterodorsal tegmental nucleus suppresses rapid-eye-movement sleep in freely behaving rats. *J Neurosci* 1997;17:7541–7552.

71. Thakkar MM, Strecker RE, McCarley RW. Behavioral state control through differential serotoninergic inhibition in the mesopontine cholinergic nuclei: a simultaneous unit recording and microdialysis. *J Neurosci* 1998;18:5490–5497.

72. Pompeiano M, Palacios JM, Mengod G. Distribution and cellular localization of mRNA coding for 5-HT_{1A} receptor in the rat: correlation with receptor binding. *J Neurosci* 1992;12:440–453.

73. Segal M. The action of serotonin in the rat hippocampal slice preparation. *J Physiol (Lond)* 1980;303:423–439.

74. Andrade R, Nicoll RA. Pharmacologically distinct actions of serotonin on single pyramidal neurones of the rat hippocampus recorded *in vitro. J Physiol (Lond)* 1987;394:99–124.

75. Okuhara DY, Beck SG. 5-HT_{1A} receptor linked to inward-rectifying potassium current in hippocampal CA3 pyramidal cells. *J Neurophysiol* 1994;71:2161–2167.

76. Segal M, Azmitia EC, Whitaker-Azmitia PM. Physiological effects of selective 5-HT_{1A} and 5-HT_{1B} ligands in rat hippocampus: comparison to 5-HT. *Brain Res* 1989;502:67–74.

77. Zgombick JM, Beck SG, Mahle CD, et al.. Pertussis toxin-sensitive guanine nucleotide-binding protein(s) couple adenosine A_1 and 5-$hydroxtryptamine_{1A}$ receptors to the same effector system in rat hippocampus: biochemical and electrophysiological studies. *Mol Pharmacol* 1989;35:484–494.

78. Haddjeri N, de Montigny C, Blier P. Long-term antidepressant treatments result in a tonic activation of forebrain 5-HT_{1A} receptors. *J Neurosci* 1998;19:10150–10156.

79. Haddjeri N, Szabo ST, de Montigny C, et al. Increased tonic activation of rat forebrain 5-HT_{1A} receptors by lithium addition to antidepressant treatments. *Neuropsychopharmacology* 2000;22:346–356.

80. Schmitz D, Empson RM, Heinemann U. Serotonin and 8-OH-DPAT reduce excitatory transmission in rat hippocampal area CA1 via reduction in presumed presynaptic Ca^{2+} entry. *Brain Res* 1995;701:249–254.

81. Schmitz D, Empson RM, Heinemann U. Serotonin reduces inhibition via 5-HT_{1A} receptors in area CA1 of rat hippocampal slices *in vitro. J Neurosci* 1995;15:7217–7225.

82. Segal M. Serotonin attenuates a slow inhibitory postsynaptic potential in rat hippocampal neurons. *Neuroscience* 1990;36:631–641.

83. Grunschlag CR, Haas HL, Stevens DR. 5-HT inhibits lateral entorhinal cortical neurons of the rat *in vitro* by activation of potassium channel-coupled 5-HT_{1A} receptors. *Brain Res* 1997;770:10–17.

84. Araneda R, Andrade R. 5-$Hydroxytryptamine_2$ and 5-$Hydroxytryptamine_{1A}$ receptors mediate opposing responses on membrane excitability in rat association cortex. *Neuroscience* 1991;40:399–412.

85. Davies MF, Deisz RA, Prince DA, et al. Two distinct effects of 5-hydroxytryptamine on single cortical neurons. *Brain Res* 1987;423:347–352.

86. Sheldon PW, Aghajanian GK. Serotonin (5-HT) induces IPSPs in pyramidal layer cells of rat piriform cortex: evidence for the involvement of a 5-HT_2-activated interneuron. *Brain Res* 1990;506:62–69.

87. Tanaka E, North RA. Actions of 5-hydroxytryptamine on neurons of the rat cingulate cortex. *J Neurophysiol* 1993;69:1749–1757.

88. Ashby CRJ, Edwards E, Wang RY. Electrophysiological evidence for a functional interaction between 5-HT_{1A} and 5-HT_{2A} receptors in the rat medial prefrontal cortex: an iontophoretic study. *Synapse* 1994;17:173–181.

89. Lakoski JM, Aghajanian GK. Effects of ketanserin on neuronal responses to serotonin in the prefrontal cortex, lateral geniculate and dorsal raphe nucleus. *Neuropharmacology* 1985;24:265–273.

90. Neumann RS, Zebrowska G. Serotonin (5-HT_2) receptor-mediated enhancement of cortical unit activity. *Can J Physiol Pharmacol* 1992;70:1604–1609.

91. Read HL, Beck SG, Dun NJ. Serotoninergic suppression of interhemispheric cortical synaptic potentials. *Brain Res* 1994;643:17–28.

92. Schmitz D, Gloveli T, Empson RM, et al. Serotonin reduces synaptic excitation in the superficial medial entorhinal cortex of the rat via a presynaptic mechanism. *J Physiol* 1998;508.1:119–129.

93. Mengod G, Pompeiano M, Martinez-Mir MI, et al. Localization of the mRNA for the 5-HT_2 receptor by *in situ* hybridization histochemistry: correlation with the distribution of receptor binding sites. *Brain Res* 1990;524:139–143.

94. McCall RB, Aghajanian GK. Serotoninergic facilitation of facial motoneuron excitation. *Brain Res* 1979;169:11–27.

95. Aghajanian GK, Rasmussen K. Intracellular studies in the facial nucleus illustrating a simple new method for obtaining viable motoneurons in adult rat brain slices. *Synapse* 1989;3:331–338.

96. Larkman PM, Penington NJ, Kelly JS. Electrophysiology of adult rat facial motoneurones: the effect of serotonin (5-HT) in novel *in vitro* brainstem slice. *J Neurosci Methods* 1989;28:133–146.

97. Rasmussen K, Aghajanian GK. Serotonin excitation of facial motoneurons: receptor subtype characterization. *Synapse* 1990;5:324–332.

98. McCall RB, Aghajanian GK. Hallucinogens potentiate responses to serotonin and norepinephrine in the facial motor nucleus. *Life Sci* 1980;26:1149–1156.

99. Garratt JC, Alreja M, Aghajanian GK. LSD has high efficacy relative to serotonin in enhancing the cationic current I_h: intracellular studies in rat facial motoneurons. *Synapse* 1993;13:123–134.

100. Stevens DR, McCarley RW, Greene RW. Serotonin$_1$ and serotonin$_2$ receptors hyperpolarize and depolarize separate populations of medial pontine reticular formation neurons *in vitro. Neuroscience* 1992;47:545–553.

101. Pessia M, Jiang ZG, North RA, et al. Actions of 5-hydroxytryptamine on ventral tegmental area neurons of the rat *in vitro. Brain Res* 1994;654:324–330.

102. Rick CE, Stanford IM, Lacey MG. Excitation of rat substantia nigra pars reticulata mediated by 5-hydroxytryptamine$_{2C}$ receptors. *Neuroscience* 1995;69:903–913.

103. Sugihara I, Lang EJ, Linas R. Serotonin modulation of inferior olivary oscillations and synchronicity: a multiple-electrode study in the rat cerebellum. *Eur J Neurosci* 1995;7:521–534.

104. North RA, Uchimura N. 5-Hydroxytryptamine acts at 5-HT$_2$ receptors to decrease potassium conductance in rat nucleus accumbens neurones. *J Physiol (Lond)* 1989;417:1–12.

105. McCormick DA, Wang Z. Serotonin and noradrenaline excite GABAergic neurones of the guinea-pig and cat nucleus reticularis thalami. *J Physiol (Lond)* 1991;442:235–255.

106. Alreja M. Excitatory actions of serotonin on GABAergic neurons of the medial septum and diagonal band of Broca. *Synapse* 1996;22:15–27.

107. Piguet P, Galvan M. Transient and long-lasting actions of 5-HT on rat dentate gyrus neurones *in vitro. J Physiol (Lond)* 1994;481:629–639.

108. Shen R-Y, Andrade R. 5-Hydroxytryptamine$_2$ receptor facilitates GABAergic neurotransmission in rat hippocampus. *J Pharmacol Exp Ther* 1998;285:805–812.

109. Cumming-Hood PA, Strahlendorf HK, Strahlendorf JC. Effects of serotonin and 5-HT$_{2/1C}$ receptor agonist DOI on neurons of the cerebral dentate/interpositus nuclei: possible involvement of a GABAergic interneuron. *Eur J Pharmacol* 1994;236:457–465.

110. McCormick DA, Williamson A. Convergence and divergence of neurotransmitter action in human cerebral cortex. *Proc Natl Acad Sci U S A* 1989;86:8098–8102.

111. Gellman RL, Aghajanian GK. Pyramidal cells in piriform cortex receive a convergence of inputs from monoamine activated GABAergic interneurons. *Brain Res* 1993;600:63–73.

112. Gellman RL, Aghajanian GK. Serotonin$_2$ receptor-mediated excitation of interneurons in piriform cortex: antagonism by atypical antipsychotic drugs. *Neuroscience* 1994;58:515–525.

113. Marek GJ, Aghajanian GK. Excitation of interneurons in piriform cortex by 5-hydroxytryptamine: blockade by MDL 100,907, a highly selective 5-HT$_{2A}$ receptor antagonist. *Eur J Pharmacol* 1994;259:137–141.

114. Marek GJ, Aghajanian GK. LSD and the phenethylamine hallucinogen DOI are potent partial agonists at 5-HT$_{2A}$ receptors on neurons in the rat piriform cortex. *J Pharmacol Exp Ther* 1996;278:1373–1382.

115. Sheldon PW, Aghajanian GK. Excitatory responses to serotonin (5-HT) in neurons of the rat piriform cortex: evidence for mediation by 5-HT$_{1C}$ receptors in pyramidal cells and 5-HT$_2$ receptors in interneurons. *Synapse* 1991;9:208–218.

116. Bloms-Funke P, Gernert M, Ebert U, et al. Extracellular single-unit recordings of piriform cortex neurons in rats: influence of different types of anesthesia and characterization of neurons by pharmacological manipulation of serotonin receptors. *J Neurosci Res* 1999;55:608–619.

117. Zhou F-M, Hablitz JJ. Activation of serotonin receptors modulates synaptic transmission in rat cerebral cortex. *J Neurophysiol* 1999;82:2989–2999.

118. Aghajanian GK, Marek GJ. Serotonin induces excitatory postsynaptic potentials in apical dendrites of neocortical pyramidal cells. *Neuropharmacology* 1997;36:589–599.

119. Jakab RL, Goldman-Rakic PS. Segregation of serotonin 5-HT$_{2A}$ and 5-HT$_3$ receptors in inhibitory circuits of the primate cerebral cortex. *J Comp Neurol* 2000;417:337–348.

120. Lopez-Gimenez JF, Mengod G, Palacios JM, et al. Selective visualization of rat brain 5-HT$_{2A}$ receptors by autoradiography with [^{3}H]MDL 100,907. *Naunyn Schmiedebergs Arch Pharmacol* 1997;356:446–454.

121. Lambe EK, Goldman-Rakic PS, Aghajanian GK. Serotonin induces EPSCs preferentially in layer V pyramidal neurons of the frontal cortex in rat. *Cereb Cortex* 2000;10:974–980.

122. Marek GJ, Aghajanian GK. 5-HT$_{2A}$ receptor or α_1 adrenoceptor activation induces excitatory postsynaptic currents in layer V pyramidal cells of the medial prefrontal cortex. *Eur J Pharmacol* 1999;367:197–206.

123. Marek GJ, Wright RA, Schoepp DD, et al. Physiological antagonism between 5-hydroxytryptamine$_{2A}$ and group II metabotropic glutamate receptors in prefrontal cortex. *J Pharmacol Exp Ther* 2000;292:76–87.

124. Marek GJ, Aghajanian GK. 5-Hydroxytryptamine-induced EPSCs in neocortical layer V pyramidal cell of prefrontal cortex: suppression by mu opiate receptor activation. *Neuroscience* 1998;86:485–497.

125. Hamada S, Senzaki K, Hamaguchi-Hamada K, et al. Localization of 5-HT$_{2A}$ receptor in rat cerebral cortex and olfactory system revealed by immunohistochemistry using two antibodies raised in rabbit and chicken. *Mol Brain Res* 1998;54:199–211.

126. Jakab RL, Goldman-Rakic PS. 5-Hydroxytryptamine$_{2A}$ serotonin receptors in the primate cerebral cortex: possible site of action of hallucinogenic and antipsychotic drugs in pyramidal cell apical dendrites. *Proc Natl Acad Sci U S A* 1998;95:735–740.

127. Willins DL, Deutch AY, Roth BL. Serotonin 5-HT$_{2A}$ receptors are expressed on pyramidal cells and interneurons in the rat cortex. *Synapse* 1997;27:79–82.

128. Aghajanian GK, Marek GJ. Serotonin, by 5-HT$_{2A}$ receptors, increase EPSCs in layer V pyramidal cells of prefrontal cortex by an asynchronous mode of glutamate release. *Brain Res* 1999;825:161–171.

129. Goda Y, Stevens CF. Two components of transmitter release at a central synapse. *Proc Natl Acad Sci U S A* 1994;91:12942–12946.

130. Li C, Bazbek CL, Davletov A, et al. Distinct Ca^{2+} and Sr^{2+} binding properties of synaptotagmins. *J Biol Chem* 1995;270:24898–24902.

131. Aghajanian GK, Marek GJ. Serotonin and hallucinogens. *Neuropsychopharmacology* 1999;21:165–235.

132. Aghajanian GK, Marek GJ. Serotonin 5-HT$_{2A}$ receptors enhance asynchronous excitatory transmission in pyramidal cells (layer V) of prefrontal cortex. *Soc Neurosci Abst* 1998;24:1366.

133. Yakel JL, Jackson MB. 5-HT$_3$ receptors mediate rapid responses in cultured hippocampus and a clonal cell line. *Neuron* 1988;1:615–621.

134. Yakel JL, Shao XM, Jackson MB. The selectivity of the channel coupled to the 5-HT$_3$ receptor. *Brain Res* 1990;533:46–52.

135. Maricq AV, Peterson AS, Brake AJ, et al. Primary source and functional expression of the 5-HT$_3$ receptor, a serotonin-gated ion channel. *Science* 1991;254:432–437.

136. Fletcher S, Barnes NM. Desperately seeking subunits: are native 5-HT_3 receptors really homomeric complexes? *Trends Pharmacol Sci* 1998;19:212–215.

137. Ropert N. Inhibitory action of serotonin in CA1 hippocampal neurons *in vitro. Neuroscience* 1988;26:69–81.

138. Ropert N, Guy N. Serotonin facilitates GABAergic transmission in the CA1 region of rat hippocampus *in vitro. J Physiol (Lond)* 1991;441:121–136.

139. McMahon LL, Kauer JA. Hippocampal interneurons are excited via serotonin-gated ion channels. *J Neurophysiol* 1997;78:2493–2502.

140. Todorovic S, Anderson EG. 5-HT_2 and 5-HT_3 receptors mediate two distinct depolarizing responses in rat dorsal root ganglion neurons. *Brain Res* 1990;511:71–79.

141. Glaum SR, Brooks PA, Spyer KM, et al. 5-Hydroxytryptamine-3 receptors modulate synaptic activity in the rat nucleus tractus solitarius *in vitro. Brain Res* 1992;589:62–68.

142. Ronde P, Nichols RA. High calcium permeability of serotonin 5-HT_3 receptors on presynaptic nerve terminals from rat striatum. *J Neurochem* 1998;70:1094–1103.

143. Gerald C, Adham N, Kao H-T, et al. The 5-HT_4 receptor: molecular cloning and pharmacological characterization of two splice variants. *EMBO J* 1995;14:2806–2815.

144. Lovenberg TW, Baron BM, deLecea L, et al. A novel adenylyl cyclase-activating serotonin receptor (5-HT_7) implicated in the regulation of mammalian circadian rhythms. *Neuron* 1993;11:449–458.

145. Monsma FJ Jr, Shen Y, Ward RP, et al. Cloning and expression of a novel serotonin receptor with high affinity for tricyclic psychotropic drugs. *Mol Pharmacol* 1992;43:320–327.

146. Shen Y, Monsma FJJ, Metcalf MA, et al. Molecular cloning and expression of 5-hydroxytryptamine$_7$ serotonin receptor subtype. *J Biol Chem* 1993;268:18200–18204.

147. Grossman CJ, Killpatrick GJ, Bunce KT. Development of a radioligand binding assay for 5-HT_4 receptors in guinea-pig and rat brain. *Br J Pharmacol* 1993;109:618–624.

148. Andrade R, Chaput Y. 5-HT_4-like receptors mediate the slow excitatory response to serotonin in the rat hippocampus. *J Pharmacol Exp Ther* 1991;257:930–937.

149. Beck S. G. 5-Carboxyamidotryptamine mimics only the 5-HT-elicited hyperpolarization of hippocampal pyramidal cells via 5-HT_{1A} receptor. *Neurosci Lett* 1989;99:101–106.

150. Medanic M, Gillette MU. Serotonin regulates the phase of the rat suprachiasmatic circadian pacemaker *in vitro* only during the subjective day. *J Physiol (Lond)* 1995;450:629–642.

151. Prosser RA, Heller HC, Miller JD. Serotoninergic phase advances of the mammalian circadian clock involve protein kinase A and K^+ channel opening. *Brain Res* 1994;644:67–73.

152. Kawahara F, Saito H, Katsuki H. Inhibition of 5-HT_7 receptor stimulation of $GABA_A$ receptor-activated current in cultured rat suprachiasmatic neurons. *J Physiol (Lond)* 1995;478:67–73.

153. Ludwig A, Zong X, Jeglitsch M, et al. A family of hyperpolarization-activated mammalian cation channels. *Nature* 1998;393:587–591.

154. Santora B, Liu DT, Yao H, et al. Identification of a gene encoding a hyperpolarization-activated pacemaker channel of brain. *Cell* 1998;83:717–729.

155. Pape HC, McCormick DA. Noradrenaline and serotonin selectively modulate thalamic burst firing by enhancing a hyperpolarization-activated cation current. *Nature* 1989;340:715–718.

156. Bobker DH, Williams JT. Serotonin augments in the cationic current I_h in central neurons. *Neuron* 1989;2:1535–1540.

157. Nedergaard S, Flatman JA, Engberg I. Excitation of substantia nigra pars compacta neurones by 5-hydroxytryptamine *in vitro. Neuroreport* 1991;2:329–332.

158. Gasparini S, DeFrancesco D. Action of serotonin on the hyperpolarization-activated cation current (I_h) in rat CA1 hippocampal neurons. *Eur J Neurosci* 1999;11:3093–3100.

159. Cardenas CG, Mar LP, Vysokanov AV, et al. Serotoninergic modulation of hyperpolarization-activated current in acutely isolated rat dorsal root ganglion cells. *J Physiol (Lond)* 1999;518:507–523.

160. Liu YF, Ghahremani MH, Rasenick MM, et al. Stimulation of cAMP synthesis by G_i-coupled receptors upon ablation of distinct $G_{\alpha i}$ protein expression. G_i subtype specificity of the 5-HT_{1A} receptor. *J Biol Chem* 1999;274:16444–16450.

161. Albert PR, Sajedi N, Lemonde S, et al. Constitutive $G_{(i2)}$-dependent activation of adenylyl cyclase type II by the 5-HT_{1A} receptor. Inhibition by anxiolytic partial agonists. *J Biol Chem* 1999;274:35469–35474.

162. Andrade R. Enhancement of beta-adrenergic responses by G_i-linked receptors in rat hippocampus. *Neuron* 1993;10:83–88.

163. Baker LP, Nielsen MD, Impey S, et al. Stimulation of type 1 and type 8 Ca^{2+}/calmodulin-sensitive adenylyl cyclases by the G_s-coupled 5-hydroxytryptamine subtype 5-HT_{7A} receptor. *J Biol Chem* 1998;273:17469–17476.

164. Markstein R, Matsumoto M, Kohler C, et al. Pharmacological characterisation of 5-HT receptors positively coupled to adenylyl cyclase in the rat hippocampus. *Naunyn Schmiedebergs Arch Pharmacol* 1999;359:454–459.

165. Aiyar J, Grissmer S, Chandy KG. Full-length and truncated Kv1.3 K^+ channels are modulated by 5-HT_{1c} receptor activation and independently by PKC. *Am J Physiol* 1993;265:C1571–C1578.

166. Alberts GL, Pregenzer JF, Im WB, et al. Agonist-induced GTPγ^{35}S binding mediated by human 5-$HT_{(2C)}$ receptors expressed in human embryonic kidney 293 cells. *Eur J Pharmacol* 1999;383:311–319.

167. Chang M, Zhang L, Tam JP, et al. Dissecting G protein-coupled receptor signaling pathways with membrane-permeable blocking peptides. Endogenous 5-$HT_{(2c)}$ receptors in choroid plexus epithelial cells. *J Biol Chem* 2000;275:7021–7029.

168. Gohla A, Offermanns S, Wilkie TM, et al. Differential involvement of $G_{\alpha 12}$ and $G_{\alpha 13}$ in receptor-mediated stress fiber formation. *J Biol Chem* 1999;274:17901–17907.

169. Berg KA, Clarke WP, Sailstad C, et al. Signal transduction differences between 5-hydroxytryptamine type 2A and type 2C receptor systems. *Mol Pharmacol* 1994;46:477–484.

170. Banes A, Florian JA, Watts SW. Mechanisms of 5-hydroxytryptamine$_{(2A)}$ receptor activation of the mitogen-activated protein kinase pathway in vascular smooth muscle. *J Pharmacol Exp Ther* 1999;291:1179–1187.

171. Watts SW. Activation of the mitogen-activated protein kinase pathway via the 5-HT_{2A} receptor. *Ann N Y Acad Sci* 1998;861:162–168.

172. Mitchell R, McCulloch D, Lutz E, et al. Rhodopsin-family receptors associate with small G proteins to activate phospholipase D. *Nature* 1998;392:411–414.

173. Vaidya VA, Marek GJ, Aghajanian GK, et al. 5-HT_{2A} receptor-mediated regulation of brain-derived neurotrophic factor mRNA in the hippocampus and the neocortex. *J Neurosci* 1997;17:2785–2795.

174. Vaidya VA, Terwilliger RM, Duman RS. Role of 5-HT_{2A} receptors in the stress-induced down-regulation of brain-derived neurotrophic factor expression in rat hippocampus. *Neurosci Lett* 1999;262:1–4.

175. Grewal JS, Mukhin YV, Garnovskaya MN, et al. Serotonin 5-HT_{2A} receptor induces TGF-β_1 expression in mesangial cells via ERK: proliferative and fibrotic signals. *Am J Physiol* 1999;276:F922–F930.

176. Pei Q, Lewis L, Sprakes ME, et al. Serotoninergic regulation of mRNA expression of *Arc*, an immediate early gene selectively localized at neuronal dendrites [In Process Citation]. *Neuropharmacology* 2000;39:463–470.

177. Tournois C, Mutel V, Manivet P, et al. Cross-talk between 5-hydroxytryptamine receptors in a serotoninergic cell line. Involvement of arachidonic acid metabolism. *J Biol Chem* 1998;273:17498–17503.

178. Blier P, Lista A, de Montigny C. Differential properties of pre- and postsynaptic 5-hydroxytryptamine$_{1A}$ receptors in the dorsal raphe and the hippocampus: II. Effect of pertussis and cholera toxins. *J Pharmacol Exp Ther* 1993;265:16–23.

179. Innis RB, Nestler EJ, Aghajanian GK. Evidence for G protein mediation of serotonin- and GABA$_B$-induced hyperpolarization of rat dorsal raphe neurons. *Brain Res* 1988;459:27–36.

180. Andrade R, Malenka RC, Nicoll RA. A G protein couples serotonin and GABA$_B$ receptors to the same channels in hippocampus. *Science* 1986;234:1261–1265.

181. Kofuji P, Davidson N, Lester HA. Evidence that neuronal G-protein-gated inwardly rectifying K$^+$ channels are activated by G$_{\beta\gamma}$ subunits and function as heteromultimers. *Proc Natl Acad Sci U S A* 1995;92:6542–6546.

182. Reuveny E, Slesinger PA, Inglese J, et al. Activation of the cloned muscarinic potassium channel by G protein βγ subunits. *Nature* 1994;340:143–146.

183. Wickman KD, Iniguez-Lluhi JA, Davenport PA, et al. Recombinant G-protein βγ subunits activate the muscarinic-gated atrial potassium channel. *Nature* 1994;368:255–257.

184. Karschin C, Schreibmayer W, Dascal N, et al. Distribution and localization of a G protein-coupled inwardly rectifying K$^+$ channel in the rat. *FEBS Lett* 1994;348:139–144.

185. Aghajanian GK. Serotonin-induced current in rat facial motoneurons: evidence for mediation by G proteins but not protein kinase C. *Brain Res* 1990;524:171–174.

186. Smrcka AV, Hepler JR, Brown KO, et al. Regulation of polyphosphoinositide-specific phospholipase C activity by purified G$_q$. *Science* 1991;251:804–808.

187. Torres G, Chaput Y, Andrade R. Cyclic AMP and protein kinase A mediate 5-hydroxytryptamine type 4 receptor regulation of calcium-activated potassium current in adult hippocampal neurons. *Mol Pharmacol* 1995;47:191–197.

188. Fagni L, Dumuis A, Sebben M, et al. The 5-HT$_4$ receptor subtype inhibits K$^+$ current in colliculi neurones via activation of a cyclic AMP-dependent protein kinase. *Br J Pharmacol* 1992;105:973–979.

189. Torres G, Arfken C, Andrade R. 5-Hydroxytryptamine$_4$ receptors reduce afterhyperpolarization in hippocampus by inhibiting calcium ion-induced calcium release. *Mol Pharmacol* 1996;50:1316–1322.

190. Kenakin T. Agonist-receptor efficacy. II. Agonist trafficking of receptor signals. *Trends Pharmacol Sci* 1995;16:232–238.

191. Backstrom JR, Chang MS, Chu H, et al. Agonist-directed signaling of serotonin 5-HT$_{2C}$ receptors: differences between serotonin and lysergic acid diethylamide (LSD). *Neuropsychopharmacology* 1999;21:77S–81S.

192. Berg KA, Maayani S,. Goldfarb J, et al. Effector pathway-dependent relative efficacy at serotonin type 2A and 2C receptors: evidence for agonist-directed trafficking of receptor stimulus. *Mol Pharmacol* 1998;54:94–104.

193. Gettys TW, Fields TA, Raymond JR. Selective activation of inhibitory G-protein alpha subunits by partial agonists of the human 5-HT$_{1A}$ receptor [published erratum appears in *Biochemistry* 1994;33:11404]. *Biochemistry* 1994;33:4283–4290.

194. Aghajanian GK. *Use of brain slices in the study of serotoninergic pacemaker neurons of the brainstem raphe nuclei*. New York: John Wiley and Sons, 1990.

195. Jacobs BL, Fornal CA. Activity of serotoninergic neurons in behaving animals. *Neuropsychopharmacology* 1999;21:9S–15S.

196. Liu W, Alreja M. Atypical antipsychotics block the excitatory effects of serotonin in septohippocampal neurons in the rat. *Neuroscience* 1997;79:369–382.

197. Murphy DL, Wichems C, Li Q, et al. Molecular manipulations as tools for enhancing our understanding of 5-HT neurotransmission. *Trends Pharmacol Sci* 1999;20:246–252.

198. Saudou F, Amara DA, Dierich A, et al. Enhanced aggressive behavior in mice lacking 5-HT$_{1B}$ receptor. *Science* 1994;265:1875–1878.

199. Crabbe JC, Phillips TJ, Feller DJ, et al. Elevated alcohol consumption in null mutant mice lacking 5-HT$_{1B}$ serotonin receptors. *Nat Genet* 1996;14:98–101.

200. Rocha BA, Scearce-Levie K, Lucas JJ, et al. Increased vulnerability to cocaine in mice lacking the serotonin-1B receptor. *Nature* 1998;393:175–178.

201. Dulawa SC, Hen R, Scearce-Levie K, et al. 5-HT$_{1B}$ receptor modulation of prepulse inhibition: recent findings in wild-type and 5-HT$_{1B}$ knockout mice. *Ann N Y Acad Sci* 1998;861:79–84.

202. Boutrel B, Franc B, Hen R, et al. Key role of 5-HT$_{1B}$ receptors in the regulation of paradoxical sleep as evidenced in 5-HT$_{1B}$ knock-out mice. *J Neurosci* 1999;19:3204–3212.

203. Tecott LH, Sun LM, Akana SF, et al. Eating disorder and epilepsy in mice lacking 5-HT$_{2c}$ serotonin receptors. *Nature* 1995;374:542–546.

204. Brennan TJ, Seeley WW, Kilgard M, et al. Sound-induced seizures in serotonin 5-HT$_{2C}$ receptor mutant mice. *Nat Genet* 1997;16:387–390.

205. Nonogaki K, Strack AM, Dallman MF, et al. Leptin-independent hyperphagia and type 2 diabetes in mice with a mutated serotonin 5-HT$_{2C}$ receptor gene. *Nat Med* 1998;4:1152–1156.

206. Tecott LH, Logue SF, Wehner JM, et al. Perturbed dentate gyrus function in serotonin 5-HT$_{2C}$ receptor mutant mice. *Proc Natl Acad Sci U S A* 1998;95:15026–15031.

207. Heisler LK, Chu HM, Brennan TJ, et al. Elevated anxiety and antidepressant-like responses in serotonin 5-HT$_{1A}$ receptor mutant mice [see Comments]. *Proc Natl Acad Sci U S A* 1998;95:15049–15054.

208. Parks CL, Robinson PS, Sibille E, et al. Increased anxiety of mice lacking the serotonin$_{1A}$ receptor. *Proc Natl Acad Sci U S A* 1998;95:10734–10739.

209. Ramboz S, Oosting R, Amara DA, et al. Serotonin receptor 1A knockout: an animal model of anxiety-related disorder [see Comments]. *Proc Natl Acad Sci U S A* 1998;95:14476–14481.

210. Bengel D, Murphy DL, Andrews AM, et al. Altered brain serotonin homeostasis and locomotor insensitivity to 3, 4-methylenedioxymethamphetamine ("Ecstasy") in serotonin transporter-deficient mice. *Mol Pharmacol* 1998;53:649–655.

211. Sora I, Wichems C, Takahashi N, et al. Cocaine reward models: conditioned place preference can be established in dopamine- and in serotonin-transporter knockout mice. *Proc Natl Acad Sci U S A* 1998;95:7699–7704.

212. Grailhe R, Waeber C, Dulawa SC, et al. Increased exploratory activity and altered response to LSD in mice lacking the 5-HT$_{(5A)}$ receptor. *Neuron* 1999;22:581–591.

213. Cases O, Vitalis T, Seif I, et al. Lack of barrels in the somatosensory cortex of monoamine oxidase A-deficient mice: role of a serotonin excess during the critical period. *Neuron* 1996;16:297–307.

214. Cases O, Seif I, Grimsby J, et al. Aggressive behavior and altered amounts of brain serotonin and norepinephrine in mice lacking MAOA [see Comments]. *Science* 1995;268:1763–1766.

215. Ghavami A, Stark KL, Jareb M, et al. Differential addressing of 5-HT_{1A} and 5-HT_{1B} receptors in epithelial cells and neurons. *J Cell Sci* 1999;112:967–976.

216. Lesch KP, Bengel D, Heils A, et al. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region [see Comments]. *Science* 1996;274:1527–1531.

217. Heils A, Mossner R, Lesch KP. The human serotonin transporter gene polymorphism—basic research and clinical implications. *J Neural Transm* 1997;104:1005–1014.

218. Greenberg BD, McMahon FJ, Murphy DL. Serotonin transporter candidate gene studies in affective disorders and personality: promises and potential pitfalls [Editorial]. *Mol Psychiatry* 1998;3:186–189.

219. Lappalainen J, Zhang L, Dean M, et al. Identification, expression, and pharmacology of a Cys23-Ser23 substitution in the human 5-HT_{2c} receptor gene *(HTR2C). Genomics* 1995;27:274–279.

220. Marshall SE, Bird TG, Hart K, et al. Unified approach to the analysis of genetic variation in serotoninergic pathways. *Am J Med Genet* 1999;88:621–627.

3

Opioid Peptides and Their Receptors: Overview and Function in Pain Modulation

Gavan P. McNally

Huda Akil

Gavan P. McNally and Huda Akil: Mental Health Research Institute, University of Michigan, Ann Arbor, Michigan.

Few neurotransmitter systems have fascinated the general public as much as the endorphins, otherwise known as the *endogenous opioid peptides*. They have been termed the “heroin within” and endowed with the power to relieve pain and allow one to experience “runner’s high” or enjoy the taste of chocolate. Although these powers may or may not withstand close scientific scrutiny, there is little question that endogenous opioid systems play a critical role in modulating a large number of sensory, motivational, emotional, and cognitive functions. As inhibitory neuropeptide transmitters, they fine-tune neurotransmission across a wide range of neuronal circuits, setting thresholds or upper limits. In addition, they have served as prototypes for understanding many structural and functional features of peptidergic systems. Thus, the first neuronal receptor binding assays were conducted on opioid receptors. The first peptides to be discovered and identified after the hypothalamic neurohormones (oxytocin and vasopressin) were the endogenous opioids. The first mammalian cyclic DNA (cDNA) to be cloned was an opioid precursor (proopiomelanocortin), which also served as the prototype for genes that encode multiple active substances and process them in a tissue-specific and situation-specific manner.

Scientific studies of these systems during the last 30 years have uncovered a complex and subtle system that exhibits impressive diversity in terms of the number of endogenous ligands (more than a dozen) yet amazing convergence at the level of receptors (only three major types). Based on the results of these studies, the endogenous opioids have been implicated in circuits involved in the control of sensation, emotion, and affect, and a role has been ascribed to them in addiction—not only to opiate drugs, such as morphine and heroin, but also to other highly abused drugs, such as alcohol. This chapter cannot do justice to the rich body of information we possess on the endogenous opioid system. However, we attempt to give the reader key information about the biochemical nature of the system, along with an update on our understanding of the recently cloned receptors and their functions. Finally, we describe the regulation of pain responsiveness as one example of a function mediated by opioids to illustrate the complexity of their role.

OPIOID PEPTIDES AND THEIR RECEPTORS

Part of "3 - Opioid Peptides and Their Receptors: Overview and Function in Pain Modulation "

Genes and Proteins

The opioid peptide precursors are encoded by three genes: pre-proopiomelanocortin, pre-proenkephalin, and pre-prodynorphin. Each precursor is subject to complex post-translational modifications that result in the synthesis of multiple active peptides. These peptides share the common N-terminal sequence of Tyr-Gly-Gly-Phe-(Met or Leu), which has been termed the *opioid motif*; this is followed by various C-terminal extensions yielding peptides ranging from 5 to 31 residues in length. The major opioid peptide encoded by pre-proopiomelanocortin is ß-endorphin. In addition to ß-endorphin, the proopiomelanocortin precursor encodes the nonopioid peptides adrenocorticotropic hormone (ACTH), α-melanocyte-stimulating hormone (α-MSH), and ß-lipotropic pituitary hormone (ß-LPH). Pre-proenkephalin encodes multiple copies of Met-enkephalin, including two extended forms of Met-enkephalin (a heptapeptide and an octapeptide), and a single copy of Leu-enkephalin. Pre-prodynorphin encodes three opioid peptides of various lengths that all begin with the Leu-enkephalin sequence: dynorphin A, dynorphin B, and neoendorphin (Fig. 3.1).

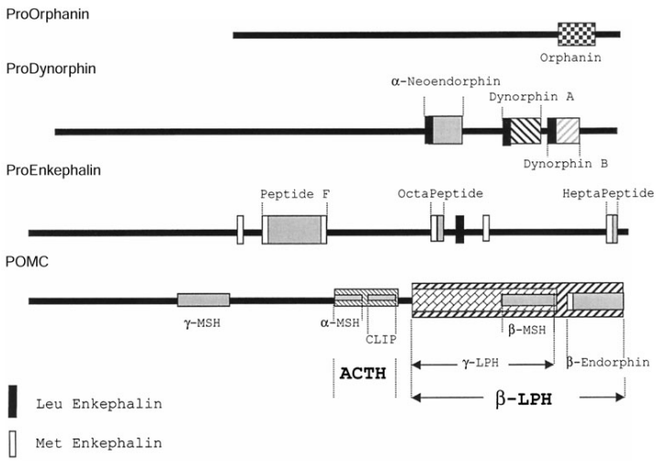


FIGURE 3.1. The opioid-peptide precursors. (From Akil H, Owens C, Gutstein H, et al. Endogenous opioids: overview and current issues. *Drug Alcohol Depend* 1998;51:127–140, with permission.)

The μ-opioid receptors (MORs), δ-opioid receptors (DORs), and κ-opioid receptors (KORs) have been isolated and cloned. The mouse DOR receptor was the first opioid receptor cloned (1 ,2), and this initial cloning facilitated the rapid cloning of MOR and KOR from various rodent species (3 ,4 ,5 ,6 ,7 ,8 and 9). The coding regions of human genes for these

receptors were subsequently isolated and chromosomally assigned (10 ,11 and 12). These studies confirmed earlier pharmacologic data indicating that all three receptors belong to the superfamily of seven transmembrane-spanning G protein-coupled receptors. A high degree of structural similarity exists between the three opioid receptors, which is highest in transmembrane domains 2, 3, and 7 and the first and second intracellular loops. The extracellular loops diverge considerably among the three receptor classes, and this divergence may explain differences in ligand selectivity among the opioid receptors (Fig. 3.2).

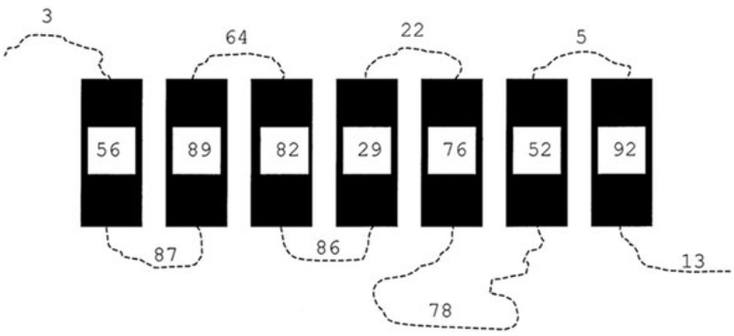


FIGURE 3.2. The opioid receptors display a high degree of structural similarity. Numbers refer to the percentages of amino acid identity between the cloned μ-, δ-, and κ-opioid receptors.

The relationship between the opioid peptides and their receptors is complex. This has been reviewed in detail elsewhere (13), and we will note only some salient features. It is clear from studies of the cloned receptors that high-affinity interactions between each of the precursor and receptor families are possible (14). For example, the proenkephalin peptide Tyr-Gly-Gly-Phe-Met-Arg-Phe binds with subnanomolar affinity to each of the cloned receptors. Similarly, although binding with greater affinity to the KOR, several of the shorter prodynorphin peptides bind with reasonable affinity to the MOR and DOR. By contrast, the binding of shorter proenkephalin peptides Leu-Enk (Tyr-Gly-Gly-Phe-Leu) and Met-Enk (Tyr-Gly-Gly-Phe-Met) readily discriminates between the three receptor families. Overall, the KOR displays the greatest selectivity across the endogenous ligands, with an approximately 1000-fold difference in affinity between the most preferred (Dyn A 1–7) and least preferred (Leu-Enk) ligand, whereas the MOR and DOR differ only across a 10-fold range (14). These differences in selectivity could indicate the existence of distinct mechanisms for ligand recognition, such that MOR and DOR recognize the common Tyr-Gly-Gly-Phe core, whereas the KOR discriminates among the larger variation in C-terminal regions. Indeed, elegant studies in which receptor chimeras were used have identified the critical domains in the three opioid receptors that help discriminate among the endogenous ligands (13).

Further attempts to detect novel opioid receptors resulted in the isolation of a clone with high structural homology to the opioid clones but little or no binding affinity for the opioid ligands (15 ,16). The structural similarity between this orphan (or opioid receptor-like) opioid receptor

(ORL-1) and the opioid receptors is highest in the transmembrane regions and cytoplasmic domain and lowest in the extracellular domains critical for ligand selectivity (see below). A ligand for the receptor was subsequently identified by two groups using chromatographic fractionation techniques coupled to ORL-1-mediated inhibition of adenylyl cyclase (17 ,18). This 17-amino acid peptide is identical in length and C-terminal sequence to dynorphin A. Curiously, the N-terminal is slightly modified (*Phe*-Gly-Gly-Phe) from the opioid core described above. It has been termed *orphanin-FQ* or *nociceptin* because of its putative ability to lower pain thresholds. The orphanin-FQ/nociceptin (OFQ/N) precursor has been cloned from mouse, rat, and human and has been localized on the short arm of human chromosome 8 (19 ,20). In addition to OFQ/N, evidence suggests that this precursor may encode other biologically active peptides. Immediately downstream to OFQ/N is a 17-amino acid peptide (OFQ-2) that also starts with phenylalanine and ends with glutamine but is otherwise distinct from OFQ/N, and a putative peptide upstream from OFQ/N may be liberated on post-translational processing (nocistatin). The OFQ/N system is a distinct neuropeptide system with a high degree of sequence identity to the opioids. This slight change in structure results in a profound alteration in function. Thus, OFQ/N is motivationally neutral, as indexed by conditioned place preference (21), and has pain modulators distinct from those of the opioid peptides (see below). However, changes in as few as four amino acids endow the ORL-1 receptor with the ability to recognize prodynorphin products while still retaining recognition of OFQ/N (22). These findings suggest that unique mechanisms may have evolved to ensure selectivity against the opioids versus selectivity for OFQ/N.

Issues and Complexities Revealed by Multiple Pharmacologic Forms, Splice Variants, and Receptor Dimers

The molecular cloning studies described above identified only a single gene encoding each of the opioid receptors. These findings contrast with results of pharmacologic studies indicating the existence of two subtypes of MOR and DOR and up to four subtypes of KOR, and suggesting additional receptor families (23 ,24 ,25 and 26). For example, the DOR_1 subtype is held to display greater affinity for the agonist DPDPE, whereas the DOR_2 subtype displays greater affinity for the agonist deltorphin 2 (26). These two subtypes may also make independent contributions to DOR antinociception (27). It is possible that further molecular cloning will identify unique genes encoding these receptor subtypes. However, several authors have suggested that if multiple opioid-receptor subtypes exist, they could be derived from a single gene, and multiple mechanisms might exist to achieve these distinct pharmacologic profiles. In this section, we consider two such pathways to opioid-receptor diversity: alternative splicing of receptor RNA and dimerization of receptor proteins.

Alternative splicing of receptor heteronuclear RNA (e.g., exon skipping and intron retention) has been accorded an important role in producing *in vivo* diversity within many members of the superfamily of seven transmembrane-spanning receptors (28). For example, alternative splicing of the coding region for the N-terminus of the corticotropin-releasing hormone CRH-2 receptor results in α, β, and γ variants, each with a unique tissue distribution (see ref. 28 for review). It follows that splice variants may exist within each of the three opioid-receptor families and that this alternative splicing of receptor transcripts may be critical for the diversity of opioid receptors. A technique used extensively to identify potential sites of alternative splicing is antisense oligodeoxynucleotide (ODN) mapping. The ability of antisense ODNs to target specific regions of cDNA allows systematic evaluation of the contribution of individual exons to the observed properties of a receptor. Antisense ODN targeting of exon 1 of the cloned rat and mouse MOR prevents morphine analgesia in these species (29 ,30 and 31). By contrast, administration of antisense ODNs targeting exon 2, which are inactive against morphine analgesia, prevents the analgesia produced by heroin, fentanyl, and the morphine metabolite morphine-6-β-glucuronide (M6G) (29 ,30 and 31). A similar disruption of M6G but not morphine analgesia is observed following administration of antisense ODNs targeting exon 3 (29). These results suggest that unique MOR mechanisms may mediate the analgesic effects of a variety of opiate alkaloids, and are consistent with the claim that these unique receptor mechanisms could be achieved via alternative splicing. The use of antisense ODNs has also resulted in the identification of potential sites for splice variation in the KOR and DOR (32). Central to the claim that these results reflect the existence of opioid-receptor splice variants is the *in vivo* isolation of such variants. A splice variant of the MOR has been identified that differs considerably within its C-terminus (33). As might be expected on the basis of the location of the alternative splicing, this variant exhibits a binding profile similar to that of the cloned MOR but does not readily undergo the desensitization frequently observed following agonist exposure. Thus, although it differs in composition, the existence of this splice variant cannot explain the results described above. However, just such a variant was detected in mice subjected to targeted disruption of exon 1 (34). Thus, transcripts of the MOR that contained exons 2 and 3 were identified in exon 1-deficient mice. Moreover, whereas morphine analgesia was abolished, heroin and M6G analgesia was retained in these mice (see below).

The physical interaction of receptors to form a unique structure (dimerization) has also been accorded an important role in regulating receptor function. For example, dimerization of $GABA_BR1$ (γ-aminobutyric acid subtype B receptor 1) and $GABA_BR2$ subunits is required for the formation

of a functional $GABA_B$ receptor (35). Both the cloned KORs and DORs have been found to exist *in vitro* as homodimers (36). However, the most important demonstrations of this kind are those showing dimerization between different functional opioid-receptor types. Jordan and Devi (37) coexpressed tagged KOR and DOR or tagged KOR and MOR and used coprecipitation techniques to show that KOR and DOR can exist as heterodimers *in vitro*. Dimerization of these receptors profoundly altered their properties. The affinity of the heterodimers for highly selective KOR and DOR agonists and antagonists was greatly reduced. Instead, the heterodimers showed greatest affinity for partially selective agonists such as bremazocine. This pharmacologic profile is similar to that claimed for the KOR_2-receptor subtype, which suggests that receptor dimerization may explain at least part of the discrepancy between the molecular and pharmacologic properties of opioid receptors. Heterodimerization thus offers a mechanism for the formation of novel receptor forms and a possible explanation for the *in vivo* diversity of opioid receptors. It will be of particular interest to identify those factors governing formation of opioid-receptor heterodimers, to determine if and how frequently opioid receptors dimerize *in vivo*, and to generate ligands that selectively recognize the dimerized form of the receptors.

Signal Transduction Mechanisms and Their Adaptation after Chronic Stimulation

The opioid receptors couple to G proteins inhibiting adenylyl cyclase, activating inwardly rectifying K^+ channels and decreasing the conductance of voltage-gated Ca^{2+} channels (38). These mechanisms of signal transduction have been verified by studies of the cloned receptors expressed in a variety of host cells (see ref. 39 for review). Studies with the cloned receptors have also indicated that the opioid receptors may couple to an array of other second messenger systems, which include activation of the mitogen-activated protein (MAP) kinases and the phospholipase C-mediated cascade leading to the formation of IP_3 (inositol-1,4,5-triphosphate) and diacyl glycerol (see ref. 40 for review). Prolonged exposure to opiates results in adaptations at multiple levels within these signaling cascades. The significance of these adaptations rests, at least in part, in the causal relationship that may exist between them and those seen at the organismic level following iterated exposure to opiates, such as tolerance, sensitization, and withdrawal. We consider two such adaptations: those related to cyclic adenosine monophosphate (cAMP) signaling and those related to receptor densensitization and internalization.

In an elegant series of experiments using rat locus ceruleus (LC) neurons, Nester and co-workers identified increased levels of protein $G_{i\alpha}$ and protein $G_{o\alpha}$, adenylyl cyclase, and protein kinase A following prolonged *in vivo* exposure to morphine (41 ,42 and 43). This up-regulation of the cAMP signaling pathway mediates the increased excitability of LC neurons observed after prolonged exposure to morphine and has been invoked as a causal mechanism for the increased or “rebound” activity of LC neurons frequently observed when the drug is withdrawn (44 ,45). It follows that these compensatory adaptations in cAMP signaling could be causal to the opiate withdrawal syndrome observed at the organismic level. Consistent with this possibility, infusions of a protein kinase A inhibitor into the LC reduced the severity of the antagonist-precipitated withdrawal syndrome in rats (46 ,47). Importantly, these adaptations are not unique to opioid receptors in the LC. Increased levels of adenylyl cyclase and protein kinase A in response to chronic morphine exposure have also been detected in the nucleus accumbens and amygdala (48). However, the changes in levels of G-protein subunits induced by this treatment are more complex, with decreased levels of $G_{i\alpha}$ detected in the nucleus accumbens and increased levels of $G_{i\alpha}$ and G_{io} in the amygdala. These widespread changes are also of significance at the organismic level. For example, infusions of a protein kinase A inhibitor into the periacqueductal gray (PAG) reduced the severity of the antagonist-precipitated withdrawal syndrome (47), and inactivation of G_i and G_o proteins in the nucleus accumbens reduced heroin self-administration in rats (49).

The desensitization, internalization, and sequestration of opioid receptors following their activation may also constitute mechanisms for adaptation in signaling relevant for understanding alterations in the physiologic impact of opiates. For example, phosphorylation of MOR and DOR via protein kinase C results in a transient desensitization that could subserve acute tolerance to opiates (50 ,51 ,52 and 53). Similarly, the internalization of opioid receptors via a classic endocytic pathway may have important implications for the physiologic impact of opiates. The internalization of opioid receptors occurs in a ligand-specific manner. For example, DAMGO and methadone promote internalization of the MOR, but morphine does not (53 ,54). This ligand-specific internalization is determined, at least in part, by differences in the conformational changes induced by the ligand and is independent of its ability to stimulate G-protein signaling (55). These findings may offer a novel explanation of differences in the efficacy and abuse potential of various opiates. However, at the time of this writing, few attempts have been made to study the relevance of these alterations in signaling to the adaptations seen in response to opiate exposure *in vivo*. Perhaps the most interesting are demonstrations that acute morphine analgesia is enhanced in mice in which the gene encoding ß-arrestin 2 was disrupted (56). Opioid-receptor internalization is mediated, at least in part, by the actions of the G protein-receptor kinases (GRKs). The GRKs selectively phosphorylate the agonist-bound receptor promoting interactions with ß-arrestins, which interfere with G-protein coupling and promote receptor internalization (56).

Demonstrations that acute morphine analgesia is enhanced in mice lacking ß-arrestin 2 are consistent with a role for the GRKs and arrestins in regulating alterations in responsivity to opiates *in vivo*. This finding is even more intriguing given the inability of morphine to support arrestin translocation and receptor internalization *in vitro* (57).

RECEPTOR AND LIGAND KNOCKOUTS: INSIGHTS, ISSUES, AND COMPLEXITIES

Part of "3 - Opioid Peptides and Their Receptors: Overview and Function in Pain Modulation "

Advances in understanding the molecular biology of the opioid family, coupled with developments in recombinant technology, have resulted in the generation of mice with targeted disruptions of various opioid genes. The study of these animals offers unique insights into opioid-receptor function. The initial study of these mice has allowed evaluation of the critical receptor subtypes mediating the effects of a variety of opiate alkaloids and the selective peptide agonists. In addition, they have identified potential interactions between receptor subtypes and suggested novel functions for opioids (e.g., reproductive function).

MOR Knockouts

The *MOR* gene has been disrupted via targeted deletion of exon 1 (34 ,58 ,59), exon 2 (34), or exons 2 and 3 (60). Disruption of exons 2 and 3 had no detectable effect on development, health, and fertility (60), whereas disruptions of exon 1 impaired sexual function in male mice, manifested by reduced mating activity, decreased sperm count and motility, and smaller litter size (59). Evidence was also found for alterations in hematopoiesis—specifically, increased proliferation of granulocyte-macrophages and erythroid and multipotential progenitor cells—in exon 1 knockout mice (59). Assessment of these mice has revealed that the MOR is absolutely necessary for the analgesic effects of morphine. Thus, systemic, intracerebal ventricular, and intrathecal administration of morphine failed to produce analgesia as assayed by tail flick, hot plate, and paw withdrawal tests across a wide dose range. For example, doses of morphine as high as 56 mg/kg failed to produce analgesia in exon 1 knockout mice (58), and the median effective dose (ED_{50}) for morphine analgesia in exon 2 knockout mice exceeded 100 mg/kg (a potency shift of two orders of magnitude) (34). The MOR is also required for the rewarding (indexed by levels of conditioned place preference) and immunosuppressive effects of injections of morphine, and for the physical dependence induced by such injections (indexed by somatic signs of morphine withdrawal) (60 ,61). By contrast, the analgesic efficacy of heroin and the major morphine metabolite M6G remains intact in exon 1-deficient mice (34). This result is consistent with the antisense mapping studies described above. Although successfully identifying the critical receptor substrate for the therapeutic and recreational uses of morphine, these experiments have failed to address the involvement of MOR in basal pain sensitivity convincingly. For example, considerable controversy has surrounded the ability of an injection of naloxone to produce hyperalgesia in otherwise intact animals, and this has not been resolved by studies of the MOR knockout mice. Sora et al. (58) reported that MOR knockout mice displayed increased sensitivity to noxious stimulation, but this hyperalgesia was not readily detected by others (60). This difference could be related to differences in the impact of specific exon deletion, as in measurements of reproductive function. Alternatively, a stress-induced analgesia, such as that provoked by exposure to novel handling procedures or contextual cues, may have decreased basal pain sensitivity among control animals.

In addition to providing insight into the mechanisms of actions of the opiate alkaloids, studies of MOR knockout mice have allowed systematic investigation into the potential interactions between the three opioid-receptor families *in vivo*. Studies of DOR function in MOR knockout mice have failed to detect compensatory changes in either the number or localization of DORs (62). Similarly, no significant alteration in DOR signal transduction, as indexed by G-protein and adenylyl cyclase activity, has been observed (63). By contrast, the analgesic efficacy of DOR agonists in these mice may be slightly reduced. Specifically, a reduction in DPDPE analgesia appears most robust, whereas the analgesic effects deltorphin 2 have been found intact or slightly attenuated (63 ,64). This evidence for MOR-mediated effects of DOR agonists is intriguing and consistent, at least in part, with the possibility of interactions between MOR and DOR *in vivo*. However, it is worth bearing in mind that these studies uniformly indicate the preservation of a large component of DOR function in MOR knockout mice. Studies of KOR function in MOR knockout mice have also failed to detect significant alterations in receptor number, distribution, and signal transduction (62 ,63). However, no evidence has been found of a reduction in the analgesic efficacy of KOR agonists, unlike that of DOR agonists, in MOR knockout mice (64).

DOR Knockouts

The *DOR* gene has been disrupted in mice via targeted disruption of exon 2 (65). This deletion had no detectable effects on the health or reproductive function of the mice. Deletion of exon 2 completely abolished [^{3}H]DPDPE and [^{3}H]deltorphin 2 binding in the brain, which indicates that the putative subtypes of the DOR are encoded by the same gene product. Studies of pain sensitivity in these mice indicate that basal pain sensitivity is unaffected by disruption of the *DOR* gene. Spinal DPDPE and deltorphin 2 analgesia is significantly reduced in the DOR knockout mice. By contrast, the analgesic efficacy of intracerebral ventricular infusions of DPDPE and deltorphin 2 remains intact. The

retention of supraspinal but not spinal DOR analgesia in DOR knockout mice is surprising. This could be evidence for a novel receptor mechanism because this residual supraspinal analgesia is reduced by naltrexone but not by selective MOR or KOR antagonists. Disruption of the *DOR* gene has no significant effect on the levels and distribution of either MOR or KOR, nor is any effect noted on the levels and distribution of proenkephalin, prodynorphin, and proopiomelanocortin. Similarly, no significant alterations occur in the analgesic effectiveness of morphine, M6G, and the κ agonist U50,488H.

KOR Knockouts

The *KOR* gene has been disrupted in mice via targeted deletion of the initiation codon and N-terminal coding region (66). This disruption had no detectable effects on the health of the mice but increased litter size. The deletion completely abolished [^{3}H]CI-977 binding in the brain. Studies of pain sensitivity revealed that KOR knockout mice are hyperalgesic when assayed by the acetic acid writhing test but not the formalin, tail pressure, tail flick, and hot plate tests. This finding is consistent with the important role accorded KOR in the regulation of visceral nociception. Systemic injection of the KOR agonist U50,488H failed to produce an analgesic response as assayed by the tail flick and hot plate tests. Similarly, the locomotor depressive effects and aversive motivational effects of the injection (indexed by conditioned place aversion learning) were abolished. These results indicate that the analgesic and motivational effects of the prototypical KOR agonist are mediated via actions at the receptor(s) encoded by the *KOR* gene and are consistent with results of antisense mapping studies indicating that ODNs directed against each of the three exons of the *KOR* gene disrupt the analgesic efficacy of U50,488H. The effects of disruption of the *KOR* gene on the activity of dynorphin B and α-neodynorphin, whose selectivity in antisense mapping studies differs considerably from that of U50,488H (67), remains unclear. Disruption of the *KOR* gene had no significant effect on the levels and distribution of either MOR or DOR (68), nor was any effect noted on the level and distribution of proenkephalin, prodynorphin, and proopiomelanocortin (67). Interestingly, the analgesic efficacy of morphine was retained, but the aversive motivational effects of the dependence induced by iterated exposures to morphine were reduced. This finding supports demonstrations of a role for dynorphin and the KOR in opiate withdrawal (69).

Pre-proenkephalin Knockouts

Mice with targeted deletions of exon 3 of the pre-proenkephalin gene have been created (70). Although this disruption had little effect on levels of prodynorphin- and proopiomelanocortin-derived peptides, a large up-regulation of MOR binding in the striatum was observed (71). Neither fertility nor gross abnormalities developed in the enkephalin knockout animals. These mice displayed increased anxiety and fear-related behaviors (indexed by freezing, hiding, and performance in an open field and elevated O maze). These results suggest that enkephalins are important in the negative feedback control of anxiety and aversive motivation. Enkephalin knockout mice appeared hyperalgesic when tested with the hot plate, but not the tail flick, test. However, because the procedure for this test involved repeated exposure to the hot plate apparatus, it is again unclear whether the experimental mice were hyperalgesic or whether the control mice were hypoalgesic as a consequence of repeated testing in the hot plate apparatus. The enkephalin knockout mice also showed altered sensitivity when assayed by the formalin test. Specifically, a decrease in recuperative behaviors (lifting and licking the injected paw) could be mimicked by injection of naloxone (10 mg/kg) in wild-type control mice, which suggests that the proenkephalin-derived peptides may regulate responding in the formalin test. This result is also difficult to interpret because naloxone does not modulate formalin pain in rats under resting conditions. In short, across three measures of pain sensitivity, three different influences of the deletion of the pre-proenkephalin gene were detected: no effect in the tail flick test, increased sensitivity (hyperalgesia) in the hot plate test, and decreased sensitivity (indexed by recuperative responding) in the formalin test. Although dissociations between these measures are not uncommon, the pattern of responding across the three measures is difficult to interpret and underscores the complexity of pain modulation by aversive motivational states such as anxiety and fear. Indeed, these mice displayed intact analgesic responses to stressors (swim stress) that produce naloxone-reversible analgesia. This result is consistent with the binding studies reviewed above, indicating the potential for high-affinity interactions between peptides derived from the proopiomelanocortin and prodynorphin precursors and each receptor class.

Orphanin-FQ-Receptor and Ligand Knockouts

The gene encoding the ORL-1 receptor has been disrupted via targeted deletion of exon 1 (72), and the OFQ/N precursor has been disrupted via targeted deletion of exon 2 (73). Studies of these mice have proved particularly interesting. First, they have confirmed that the role of OFQ/N in pain modulation is quite distinct from that of the other opioids. Thus, disruption of the ORL-1 receptor had no effect on basal pain sensitivity in the tail flick test but prevented the development of tolerance to morphine analgesia (72), whereas disruption of the OFQ/N precursor consistently decreased pain sensitivity on the same measure. This discrepancy between the effects of receptor and precursor disruption could be interpreted to mean that post-translational

processing of the OFQ/N precursor may result in the presence of multiple active peptides that interact with unique receptors to produce different physiologic effects (see above). However, these studies uniformly indicate that if OFQ/N and the ORL-1 receptor have any role in pain modulation, it is facilitative (or pronociceptive) rather than inhibitory (or antinociceptive). Second, these studies have confirmed that OFQ/N serves an important role in the regulation of emotional responsiveness. Specifically, OFQ/N knockout mice display increased anxiety (indexed by performance in the elevated plus maze, open field, and light–dark box) and enhanced basal and post-stress glucocorticoid levels. Interestingly, these findings contrast with the effects of administration in rats. Devine et al. (74) have shown that infusions of OFQ/N increase plasma ACTH and glucocorticoid levels in the unstressed animal and prolong the stress response in the stressed rat. The reasons for this discrepancy are unclear. Finally, these studies have suggested an important role for the OFQ/N system in learning and memory processes. Thus, OFQ/N-receptor knockout mice show enhanced hippocampal long-term potentiation and a moderately enhanced performance in tests of spatial learning (75). However, the OFQ/N-precursor knockout mice do not show enhanced performance in the same test of spatial learning (73). Regardless of the reason for the discrepancy between the OFQ/N and ORL-1 knockout mice, the interpretation of these effects on learning and memory is difficult. For example, the spatial task used in these experiments can be mediated by several learning strategies. Clearly, a more sophisticated characterization of the nature of the potential learning and memory deficits in these mice is required, and the results described above provide an important starting point.

Summary

Studies of mice with targeted disruptions of opioid-receptor and peptide genes have enabled important insights into the function of the opioid family. Chiefly, they have made possible the identification of the critical receptor substrates for a variety of opiate alkaloids and opioid peptides. These studies have also provided insights into the functional diversity of each receptor class. For example, it is clear that the two subtypes of DOR identified in pharmacologic studies are encoded by the single cloned *DOR* gene. Furthermore, the retention of MOR- and KOR-independent supraspinal DPDPE analgesia in these mice raises the possibility that further, unidentified opioid-receptor variants may exist. A similar possibility is raised by the retention of heroin and M6G analgesia in mice with targeted deletions of the MOR. These results suggest that complex post-transcriptional modifications play an important role in producing the *in vivo* diversity of opioid-receptor pharmacology. At the time of this writing, only OFQ/N-receptor and OFQ/N-precursor knockout mice have been studied in more complex behavioral tasks. However, the widespread distribution of opioid peptides and their receptors in the central nervous system, in addition to their critical role in controlling an animal's interaction with its environment, ensure that it is only a matter of time before mice are studied with more behaviorally sophisticated and ecologically relevant measures of attention, learning, memory, and motivation. Finally, it is worth noting that this first generation of genetic manipulations are neither tissue-specific nor conditional. Compensatory adaptation within the opioid-peptide and receptor family following the targeted disruption of one of its members appears to be minimal. Indeed, in the studies reviewed above, the only evidence for such compensation has been obtained for measures of receptor binding in pre-proenkephalin knockout animals. Nonetheless, the possibility of widespread adaptation in nonopioid systems cannot be discounted. Thus, the application of tissue-specific and inducible knockout techniques to the opioid receptors and their peptides remains an exciting area of research.

AN EXAMPLE OF OPIOID FUNCTION AND ITS CLINICAL IMPLICATIONS: OPIOIDS AND PAIN CONTROL

Part of "3 - Opioid Peptides and Their Receptors: Overview and Function in Pain Modulation "

Understanding the role of opioids in pain modulation is not only of clinical importance but also of historical interest. Demonstrations that microinjections of morphine into various brainstem regions are analgesic (76), and that injections of naloxone partially reverse the analgesia produced by focal electric stimulation in these regions (77 ,78), provided the first physiologic evidence for an endogenous opioid system. In this section, we briefly review the neural circuits subserving opioid analgesia and discuss recent findings relevant to these actions. Many excellent reviews of this topic are available (79 ,80).

Functional Anatomy of Opioids in Descending Pain Control Circuits

It is well established that the analgesic effects of opioids arise from their ability to inhibit directly the ascending transmission of nociceptive information from the spinal cord dorsal horn, and from their ability to activate pain control circuits that descend from the midbrain, via the rostral ventromedial medulla (RVM), to the spinal cord dorsal horn. Opioid peptides and their receptors are distributed throughout these descending pain control circuits (81 ,82). MOR messenger RNA (mRNA) or binding has been detected throughout the PAG, pontine reticular formation, median raphe, nucleus raphe magnus and adjacent gigantocellular reticular nucleus in the RVM, and spinal cord. Inspection of the discrepancies between levels of receptor binding and mRNA expression provide important insights into the mechanisms of MOR analgesia. For example, the

presence of significant MOR binding in the superficial dorsal horn but scarcity of mRNA expression suggests that the majority of these spinal MOR binding sites are located presynaptically on the terminals of primary afferent nociceptors. This conclusion is consistent with the high levels of MOR mRNA expression in dorsal root ganglia (DRG). A similar mismatch between MOR binding and mRNA expression can be found in the dorsolateral PAG (strong binding vs. sparse mRNA). DOR mRNA and binding have been detected in the ventral and ventrolateral quadrants of the PAG, pontine reticular formation, and gigantocellular reticular nucleus, but only at low levels in the median raphe and nucleus raphe magnus. Like MOR binding sites, DOR binding sites are present in significant numbers in the dorsal horn without detectable mRNA expression, which suggests an important role for presynaptic actions of DOR in spinal analgesia. Finally, KOR mRNA and binding are widely distributed throughout the PAG, pontine reticular formation, median raphe, and nucleus raphe magnus and adjacent gigantocellular reticular nucleus. Again, significant levels of KOR binding but sparse levels of mRNA have been found in the dorsal horn. Although all three receptor mRNAs are found in the DRG, they are localized on different classes of primary afferent nociceptors. Thus, MOR mRNA has been detected in medium- and large-diameter DRG cells, DOR mRNA in large-diameter cells, and KOR mRNA in small- and medium-diameter cells. This differential localization could be linked to functional differences in pain modulation.

The distribution of opioid receptors in descending pain control circuits indicates substantial overlap between MOR and KOR. The largest differentiation between these two receptors and DOR is in the PAG, median raphe, and nucleus raphe magnus (82). A similar differentiation of MOR and KOR from DOR is observed in the thalamus, which suggests that interactions between KOR and MOR may be important for modulating nociceptive transmission from the dorsal horn as well as in higher nociceptive centres. The actions of MOR agonists are invariably antinociceptive, whereas those of KOR agonists can be either antinociceptive or pronociceptive. Consistent with the anatomic overlap between the MOR and KOR, the pronociceptive actions of the KOR appear to be mediated by a functional antagonism of the actions of the MOR. The MOR produces antinociception within descending pain control circuits, at least in part, via the removal of GABAergic inhibition of RVM projecting neurons in the PAG and spinally projecting neurons in the RVM (79). Pan et al. (83) have presented evidence from both *in vitro* slice preparations and *in vivo* pain responding that the pain modulatory effects of the KOR in the brainstem oppose those of the MOR. Thus, activation of the KOR hyperpolarized the same RVM neurons hypopolarized by the MOR, and microinjections of a κ agonist into the RVM antagonized the analgesia produced by microinjections of DAMGO into this region. These data are among the strongest that opioids can have pronociceptive in addition to antinociceptive effects and could explain behavioral evidence for a reduction in hyperalgesia following injections of naloxone.

As described above, significant opioid-receptor binding, little detectable expression of receptor mRNA in the spinal cord dorsal horn, but large levels of this mRNA in DRG have been observed. The anatomy of spinal opioid receptors suggests that their actions relevant to analgesia at this level are predominantly presynaptic. At least one presynaptic mechanism viewed as having clinical significance is the inhibition of spinal tachykinin signaling. Indeed, it is well established that opioids decrease the noxious stimulant-evoked release of tachykinins from primary afferent nociceptors (84 ,85). Recently, the significance of this action has been questioned. Measuring the internalization of neurokinin receptors following noxious stimulation, Trafton et al. (86) demonstrated that at least 80% of tachykinin signaling remains intact after the intrathecal administration of large doses of opioids. These results indicate that although opioid administration may reduce tachykinin release from primary afferent nociceptors, the reduction has little functional impact on the actions of tachykinins on postsynaptic nociceptive neurons. The obvious implication of this finding is that either tachykinin signaling is not central to nociception and/or opioid antinociception at the spinal level, or that, contrary to the conclusions suggested by anatomic studies, the presynaptic actions of opioids are of little analgesic significance.

Just as important insights have been made into brainstem and spinal mechanisms for opioid analgesia, so too have insights been made into forebrain mechanisms for such analgesia. It is well established that the actions of opioids in bulbospinal pathways are critical to their analgesic efficacy. It has been less clear what role should be accorded forebrain actions and whether these actions are independent of those in bulbospinal pathways. There can be little doubt that opioid actions in the forebrain contribute to analgesia because decerebration prevents analgesia when rats are tested for pain sensitivity with the formalin test (87), and microinjections of opioids into the several forebrain regions are analgesic in this test (88). However, because these manipulations frequently leave intact the analgesic efficacy of opioids in measures of phasic nociception, such as the tail flick test, a distinction has been drawn between forebrain-dependent mechanisms for morphine analgesia in the presence of tissue injury and bulbospinal mechanisms for this analgesia in the absence of tissue injury. In an important series of experiments, Manning and Mayer (89 ,90) have shown that this distinction is not absolute and that opioid actions in the forebrain are also important to analgesia, both in measures of tissue damage and in acute, phasic nociception. Thus, systemic morphine analgesia in both the tail flick and formalin tests was disrupted by either lesioning or reversible inactivation of the central nucleus of the amygdala. The involvement

of the amygdala in morphine analgesia is particularly interesting because this structure has been implicated in the environmental activation of pain control circuits, and it projects extensively to those brainstem regions involved in descending pain control (80).

Role of OFQ/N and ORL-1 in Pain Modulation

OFQ/N mRNA and peptide are present throughout the descending pain control circuits described above. For example, OFQ/N-containing neurons are present in the PAG, the median raphe, throughout the RVM, and the superficial dorsal horn (91) This distribution overlaps with that of the opioid peptides, but the degree of colocalization remains unclear. ORL-1 binding and mRNA can be detected in the PAG, median raphe, and RVM (92). In the spinal cord, ORL-1 mRNA expression is stronger in the ventral than in the dorsal horn, but levels of binding are higher in the dorsal horn. High levels of ORL-1 mRNA are also found in the DRG. Despite this clear anatomic evidence for a role of the orphanin system in pain modulation, its function remains unclear. As reviewed above, targeted disruption of the ORL-1 receptor had little effect on basal pain sensitivity according to several measures, whereas targeted disruption of the OFQ/N precursor consistently elevated pain sensitivity in the tail flick test, findings that suggest an important role for OFQ/N in regulating basal pain sensitivity. Intrathecal injections of OFQ/N have been reported to be analgesic as assayed by the tail flick and formalin tests (93 ,94). Similarly, these injections attenuated the hyperalgesia produced by constriction injury of sciatic nerve (95). However, the profound motor effects of these injections render interpretation of changes in response latency difficult. The effects of supraspinally administered OFQ/N are also difficult to interpret; hyperalgesia has been detected across a variety of measures, but failures to detect hyperalgesia have also been reported.

Three interesting results may explain at least part of the variations noted in the effects of the orphanin opioid system in modulating pain. First, Rossi et al. (95 ,96) reported a biphasic effect of OFQ/N administration, characterized initially by hyperalgesia and later by analgesia. Second, Grisel et al. (97) reported that OFQ/N does not affect basal pain sensitivity but does reduce analgesia according to the site of administration. Finally, Heinricher et al. (98) reported that OFQ/N exerts an inhibitory effect on several classes of RVM neurons whose activity has been implicated in producing analgesia and hyperalgesia at the spinal level. These results suggest that the effects on pain modulation observed following administration of OFQ/N in the intact animal are influenced by route, time since administration, the presence of stressors that provoke analgesia (e.g., novel handling or test procedures), and the current balance of activity in pain modulatory neurons in the RVM. The development of specific ORL-1-receptor antagonists will undoubtedly enable a rapid clarification of the role of the orphanin opioid system in pain modulation.

CONCLUSIONS AND FUTURE DIRECTIONS

Part of "3 - Opioid Peptides and Their Receptors: Overview and Function in Pain Modulation "

The interplay between the orphanin system and the endogenous opioids represents a prime example of evolutionary changes that have led to subtle diversity in structure and significant alteration in function. Indeed, this entire peptidergic family exemplifies the way in which an increase in genetic diversity can lead from simple on/off signaling to a complex pattern of signaling wherein multiple, coordinately secreted peptides interact with multiple receptors to effect a complex regulation of functions as diverse as pain responsiveness, stress regulation, control of feeding, and modulation of development, learning, and memory. Many questions remain to be answered in the context of the opioid family. At the most basic level is the question of whether additional members of the family exist. The completion of sequencing of the human genome and the rat or mouse genome should help answer this question. We should be able to lay to rest the questions of whether additional opioid-receptor types or subtypes exist, and whether other endogenous ligands that are uniquely selective for a particular receptor type exist. In particular, endomorphin 1 (Tyr-Pro-Trp-Phe) and endomorphin 2 (Tyr-Pro-Phe-Phe) have been proposed by Zadina et al. (99) to be endogenous, highly selective μ ligands. However, their precursor remains uncloned, although the genome project should help clarify the matter. Further, as we obtain full sequences of the genomes of other species, we should be able to track the fascinating evolutionary history of this peptide family.

At functional levels, many questions remain, especially concerning the exact role of endogenous opioids in addictive and emotional behavior and psychiatric disorders. Because these disorders are typically of a complex genetic nature, involving the interaction of multiple genes with one another and with the environment, it is likely that the endogenous opioid genes are involved in vulnerability to certain brain-related illnesses. Here again, progress in genomics and complex genetics should open new avenues for investigating the likely role of the opioid molecules in a range of psychiatric disorders.

REFERENCES

1. Evans CJ, Keith DE, Morrison H, et al. Cloning of a delta opioid receptor by functional expression. *Science* 1992;258:1952–1955.

2. Kieffer B, Befort K, Gaveriaux-Ruff C, et al. The delta opioid receptor: isolation of a cDNA by expression cloning and pharmacological characterization. *Proc Natl Acad Sci U S A* 1992;89:12048–12052.

3. Chen Y, Mestek A, Liu J, et al. Molecular cloning and functional expression of a μ-opioid receptor from rat brain. *Mol Pharmacol* 1993;44:8–12.

4. Kong, H, Raynor R, Yano H, et al. Agonists and antagonists bind to different domains of the cloned kappa receptor. *Proc Natl Acad Sci U S A* 1994;91:8042–8046.

5. Meng F, Xie GX, Thompson RC, et al. Cloning and pharmacological characterization of a rat kappa opioid receptor. *Proc Natl Acad Sci U S A* 1993;90:9954–9958.

6. Minami M, Toya T, Katao Y, et al. Cloning and expression of a cDNA for the rat κ-opioid receptor. *FEBS Lett* 1993;329:291–295.

7. Thompson RC, Mansour A, Akil H, et al. Cloning and pharmacological characterization of a rat mu-opioid receptor. *Neuron* 1993;11:903–913.

8. Wang JB, Imai Y, Eppler CM, et al. μ Opiate receptor: cDNA cloning and expression. *Proc Natl Acad Sci U S A* 1993;90:10230–10234.

9. Yasuda K, Raynor K, Kong H, et al. Cloning and functional expression comparison of kappa and delta opioid receptors from mouse brain. *Proc Natl Acad Sci U S A* 1993;90:6736–6740.

10. Befort K, Mattei MG, Roeckel N, et al. Chromosomal localization of the δ opioid receptor gene to human 1p34.3-p36.1 and mouse 4D bands by *in situ* hybridization. *Genomics* 1994;20:143–145.

11. Yasuda K, Espinoza R, Takeda J, et al. Localization of the kappa opioid receptor gene to human chromosome band 8q11.2. *Genomics* 1994;19:596–597.

12. Wang JB, Johnson PS, Perisco AM, et al. Human μ-opiate receptor cDNA and genomic clones, pharmacologic characterization and chromosomal assignment. *FEBS Lett* 1994;338:217–222.

13. Akil H, Owens C, Gutstein H, et al. Endogenous opioids: overview and current issues. *Drug Alcohol Depend* 1998;51:127–140.

14. Mansour A, Hoversten M, Taylor LP, et al. The cloned μ, δ, or κ receptors and endogenous ligands: evidence for two opioid peptide recognition cores. *Brain Res* 1995;700:89–98.

15. Bunzow JR, Saez C, Mortrud M, et al. Molecular cloning and tissue distribution of a putative member of the rat opioid receptor gene family that is not a μ, δ, or κ receptor type. *FEBS Lett* 1994;347:279–282.

16. Mollereau C, Parmentier M, Mailleux P, et al. ORL1, a novel member of the opioid receptor family—cloning, functional expression and localization. *FEBS Lett* 1994;341:33–38.

17. Meunier JC, Mollereau C, Toll L, et al. Isolation and structure of the endogenous agonist of the opioid receptor like ORL_1 receptor. *Nature* 1995;377:532–535

18. Reinscheid RK, Nothacker HP, Bourson A, et al. Orphanin FQ: a neuropeptide that activates an opioid-like G-protein-coupled receptor. *Science* 1995;270:792–794.

19. Nothacker HP, Reinscheid RK, Mansour A, et al. Primary structure and tissue distribution of the orphanin FQ precursor. *Proc Natl Acad Sci U S A* 1996;93:8622–8677.

20. Pan YX, Xu J, Pasternak GW, Cloning and expression of a cDNA encoding a mouse brain orphanin FQ/nociceptin precursor. *Biochem J* 1996;315:11–13.

21. Devine DP, Reinscheid RK, Monsma FJ, et al. The novel neuropeptide orphanin FQ fails to produce conditioned place preference or aversion. *Brain Res* 1996;727:225–229.

22. Meng F, Taylor LP, Hoversten M, et al. Moving from the orphanin FQ receptor to an opioid receptor using four point mutations. *J Biol Chem* 1996;271:32016–32020.

23. Zukin RZ, Eghali M, Olive D, et al. Characterization and visualization of rat and guinea pig brain κ opioid receptors: evidence for κ1 and κ2 opioid receptors. *Proc Natl Acad Sci U S A* 1998;85:4061–4065.

24. Paul D, Bodnar RJ, Gistrak MA, et al. Different μ receptor subtypes mediate spinal and supraspinal analgesia in mice. *Eur J Pharmacol* 1989;168:307–314.

25. Sofuoglu M, Portoghese PS, Takemori AE. Differential antagonism of δ opioid agonists by naltrindole and its benzofuran analog (NTB) in mice: evidence for δ opioid receptor subtypes. *J Pharmacol Exp Ther* 1991;257:676–680.

26. Zaki PA, Bilsky EJ, Vanderah TW, et al. Opioid receptor types and subtypes: the δ receptor as a model. *Ann Rev Pharmacol Toxicol* 1996;36:379–401.

27. Stewart PE, Hammond DL. Evidence for δ opioid receptor subtypes in rat spinal cord: studies with intrathecal naltriben, cyclic (D-Pen^2, D-Pen^5)-enkephalin and (D-Ala^2, Glu^4)-deltorphin. *J Pharmacol Exp Ther* 1993;266:820–828.

28. Kilpatrick GJ, Dautzenburg FM, Martin GR, et al. 7TM receptors: the splicing on the cake. *Trends Pharmacol Sci* 1999;20:294–301.

29. Rossi GC, Leventhal L, Pan YX, et al. Antisense mapping of MOR-1 in rats: distinguishing between morphine and morphine-6β-glucuronide antinociception. *J Pharmacol Exp Ther* 1997;281:109–144.

30. Rossi GC, Brown GP, Leventhal L, et al. Novel receptor mechanisms for heroin and morphine-6β-glucuronide analgesia. *Neurosci Lett* 1996;216:1–4.

31. Rossi GC, Pan YX, Brown GP, et al. Antisense mapping of the MOR-1 opioid receptor: evidence for alternative splicing and a novel morphine-6β-glucuronide receptor. *FEBS Lett* 1995;369:192–196.

32. Pasternak GW, Standifer KM. Mapping opioid receptors using antisense oligodeoxynucleotides: correlating their molecular biology and pharmacology. *Trends Pharmacol Sci* 1995;16:344–350.

33. Zimprich A, Simon T, Hollt V. Cloning and expression of an isoform of the rat μ opioid receptor (rMOR1B) which differs in agonist induced desensitization from rMOR1. *FEBS Lett* 1995;359:142–146.

34. Schuller AGP, King MA, Zhang J, et al. Retention of heroin and morphine-6β-glucuronide analgesia in a new line of mice lacking exon 1 or MOR-1. *Nat Neurosci* 1999;2:151–156.

35. Jones KA, Borowsky B, Tamm J, et al. $GABA_B$ receptors function as a heteromeric assembly of the subunits $GABA_BR1$ and $GABA_BR1$. *Nature* 1998;396:674–679.

36. Cejic S, Devi L. Dimerization of the delta opioid receptor: implications for a function in receptor internalization. *J Biol Chem* 1997;272:26959–26964.

37. Jordan BA, Devi LA, G-protein coupled receptor heterodimerization modulates receptor function. *Nature* 1999;299:697–700.

38. Childers SR. Opioid receptor-coupled second messenger systems. *Life Sci* 1991;48:1991–2003.

39. Kieffer BL. Recent advances in molecular recognition and signal transduction of active peptides: receptors for opioid peptides. *Cell Mol Neurobiol* 1995;15:615–634.

40. Akil H, Meng F, Devine DP, Watson SJ, Molecular and neuroanatomical properties of the endogenous opioid system: implications for the treatment of opiate addiction. *Semin Neurosci* 1997;9:70–83.

41. Nestler EJ, Erdos JJ, Terwilliger RZ, et al. Regulation by chronic morphine in the rat locus coeruleus. *Brain Res* 1989;476:230–239.

42. Duman RS, Tallman JF, Nestler EJ. Acute and chronic opiate regulation of adenylate cyclase in brain: specific effects in locus coeruleus. *J Pharmacol Exp Ther* 1988;246:1033–1039.

43. Nestler EJ, Tallman JF. Chronic morphine treatment increases cyclic AMP-dependent protein kinase activity in rat locus coeruleus. *Mol Pharmacol* 1988;33:127–132.

44. Nestler EJ. Molecular mechanisms of drug addiction. *J Neurosci* 1992;12:2439–2450.

45. Nestler EJ. Drug addiction: a model for the molecular basis of neural plasticity. *Neuron* 1993;11:995–1006.

46. Maldonado R, Valverde O, Garbay C, et al. Protein kinases in the locus coeruleus and periaqueductal gray matter are involved in the expression of opiate withdrawal. *Naunyn Schmiedebergs Arch Pharmacol* 1995;352:565–575.

47. Punch LJ, Self DW, Nestler EJ, et al. Opposite modulation of opiate withdrawal behaviors on microinfusion of a protein kinase A inhibitor versus activator into the locus coeruleus or periaqueductal gray. *J Neurosci* 1997;17:8520–8527.

48. Terwilliger RZ, Beitner-Johnson D, Sevarino KA, et al. A general role for adaptations in G-proteins and the cyclic AMP system in mediating the chronic actions of morphine and cocaine on neuronal function. *Brain Res* 1991;548:100–110.

49. Self DW, Genova LM, Hope BT, Barnhart W, et al. Involvement of cAMP-dependent protein kinase in the nucleus accumbens in cocaine self-administration and relapse of cocaine-seeking behavior. *J Neurosci* 1998;18:1848–1859.

50. Mestek A, Hurley JH, Bye LS, et al. The human mu opioid receptor: modulation of functional desensitization by calcium calmodulin-dependent protein kinase and protein kinase C. *Can J Neurosci* 1995;15:2396–2406.

51. Narita M, Narita M, Mizoguchi H, et al. Inhibition of protein kinase C, but not of protein kinase A, blocks the development of acute antinociceptive tolerance to intrathecally administered mu-opioid receptor agonist in the mouse. *Eur J Pharmacol* 1995;280:R1–R3.

52. Ueda H, Miyame T, Hayashi C, et al. Protein kinase C involvement in homologous desensitization of delta-opioid coupled to G_{i1}-phospholipase C activation in *Xenopus* oocytes. *J Neurosci* 1995;15:78485–78499.

53. Keith DE, Murray SR, Zaki PA, et al. Morphine activates opioid receptors without causing their rapid internalization. *J Biol Chem* 1996;271:19021–19024.

54. Keith DE, Anton B, Murray SR, et al. Mu-opioid receptor internalization: opiate drugs have differential effects on a conserved endocytic mechanism *in vitro* and in the mammalian brain. *Mol Pharmacol* 1998;53:377–384.

55. Whistler J, Chuang H, Chu P, et al. Functional dissociation of μ opioid receptor signaling and endocytosis: implications for the biology of opiate tolerance and addiction. *Neuron* 1999;23:737–746.

56. Bohn LM, Lefkowitz RJ, Gainetdinov RR, et al.. Enhanced morphine analgesia in mice lacking beta-arrestin 2. *Science* 1999;286:2495–2498.

57. Whistler JL, von Zastrow M. Morphine activated opioid receptors elude desensitization by beta-arrestin. *Proc Natl Acad Sci U S A* 1998;95:9914–9919.

58. Sora I, Takahashi N, Funada M, et al. Opiate receptor knockout mice define μ receptor roles in endogenous nociceptive responses and morphine-induced analgesia. *Proc Natl Acad Sci U S A* 1997;94:1544–1549.

59. Tian M, Broxmeyer HE, Fan Y, et al. Altered hematopoiesis, behavior, and sexual function in μ opioid receptor-deficient mice. *J Exp Med* 1997;185:1517–1522.

60. Matthes HWD, Maldonado R, Simonin F, et al. Loss of morphine-induced analgesia, reward effect and withdrawal symptoms in mice lacking the μ-opioid receptor gene. *Nature* 1996;383:819–823

61. Gaveriaux-Ruff C, Matthes HWD, Peluso J, et al. Abolition of morphine-immunosuppression in mice lacking the μ-opioid receptor gene. *Proc Natl Acad Sci U S A* 1998;95:6236–6330.

62. Kitchen I, Slowe SJ, Matthes HW, et al. Quantitative audioradiographic mapping of mu-, delta-, and kappa-opioid receptors in mice lacking the mu-opioid receptor gene. *Brain Res* 1997;778:73–88.

63. Matthes HWD, Smadja C, Valverde O, et al. Activity of the δ-opioid receptor is partially reduced, whereas activity of the κ-opioid receptor is maintained in mice lacking the μ-receptor. *J Neurosci* 1998;18:7285–7295.

64. Fuchs PN, Roza C, Sora I, et al. Characterization of mechanical withdrawal responses and the effect of μ-, δ-, and κ-opioid agonists in normal and μ-opioid receptor knockout mice. *Brain Res* 1999;821:480–486.

65. Zhu Y, King MA, Alwin GP, et al. Retention of supraspinal delta-like analgesia and loss of morphine tolerance in δ opioid receptor knockout mice. *Neuron* 1999;24:243–252.

66. Simonin F, Valverde O, Smadja C, et al. Disruption of the κ-opioid receptor gene in mice enhances sensitivity to chemical visceral pain, impairs the pharmacological actions of the selective κ-agonist U-50,488H and attenuates morphine withdrawal. *EMBO J* 1998;17:886–897.

67. Pasternak KR, Rossi GC, Zuckerman A, et al. Antisense mapping KOR-1: evidence for multiple kappa analgesic mechanisms. *Brain Res* 1999;826:289–292.

68. Slowe SJ, Simonin F, Kieffer B, et al. Quantitative autoradiography of mu, delta, and kappa1 opioid receptors in kappa-opioid receptor deficient mice. *Brain Res* 1999;818:335–345.

69. McGinty JF, Wang JQ. Drugs of abuse and striatal gene expression. *Adv Pharmacol* 1998;42:1017–1019.

70. Konig M, Zimmer AM, Steiner H, et al. Pain responses, anxiety and aggression in mice deficient in pre-proenkephalin. *Science* 1996;383:535–538.

71. Brady LS, Herkenham M, Rothman B, et al. Region-specific up-regulation of opioid receptor binding in enkephalin knockout mice. *Mol Brain Res* 1999;68:193–197.

72. Nishi M, Houtani T, Noda Y, et al. Unrestrained nociceptive response and dysregulation of hearing ability in mice lacking the nociceptin/orphanin FQ receptor. *EMBO J* 1997;16:1858–1864.

73. Koster A, Montkowski A, Schulz A, et al. Targeted disruption of the orphanin FQ/nociceptin gene increases stress susceptibility and impairs stress adaptation in mice. *Proc Natl Acad Sci U S A* 1998;96:10444–10449.

74. Devine DP, Watson SJ, Akil H. Nociceptin/orphanin FQ regulates neuroendocrine function of the limbic-hypothalamic-pituitary-adrenal axis. *Neuroscience* 2001;102:541–553.

75. Manabe T, Noda Y, Mamimya T, et al. Facilitation of long-term potentiation and memory in mice lacking nociceptin receptors. *Nature* 1998;394:577–581.

76. Yaksh TL, Yeung JC, Rudy TA. Systematic examination in the rat of brain sites sensitive to the direct application of morphine: observation of differential effect within the periaqueductal gray. *Brain Res* 1976;114:83–103.

77. Akil H, Mayer DJ, Liebeskind JC, et al. Comparison chez le rat entre l'analgésie induite par stimulation de la substance grise peri-aqueductale et l'analgésie morphinique. *C. R. Acadéemie de Science Paris* 1974;274:3603–3605.

78. Akil H, Mayer DJ, Liebeskind JC. Antagonism of stimulation-produced analgesia by naloxone, a narcotic antagonist. *Science* 1976;191:961–962.

79. Fields HL, Heinricher MM, Mason P. Neurotransmitters in nociceptive modulatory circuitries. *Annu Rev Neurosci* 1991;14:219–245.

80. Harris JA. Descending antinociceptive mechanisms in the brainstem: their role in an animal's defensive system. *J Physiol Paris* 1996;20:15–25.

81. Mansour A, Fox CA, Akil H, et al. Opioid-receptor mRNA expression in the rat CNS: anatomical and functional implications. *Trends Neurosci* 1995;18:22–29.

82. Gutstein HB, Mansour A, Watson SJ, et al. Mu and kappa receptors in periaqueductal gray and rostral ventromedial medulla. *Neuroreport* 1998;9:1777–1781.

83. Pan ZZ, Tershner SA, Fields HL. Cellular mechanisms for anti-analgesic action of agonists of the κ-opioid receptor. *Nature* 1997;389:382–384.

84. Jessel T, Iversen L, Opiate analgesics inhibit substance P release from the rat trigeminal nucleus. *Nature* 1977;268:549–551.

85. Yaksh T, Jessell T, Gamse R, et al. Intrathecal morphine inhibits substance P release from mammalian spinal cord *in vivo. Nature* 1980;286:155–157.

86. Trafton JA, Abbadie XC, Marchand S, et al. Spinal opioid analgesia: how critical is the role substance P signaling? *J Neurosci* 1999;19:9642–9653.

87. Matthes BK, Franklin K. Formalin pain is expressed in decerebrate rats but is not attenuated by morphine. *Pain* 1992;51:199–206.

88. Manning BH, Morgan MJ, Franklin K. Morphine analgesia in the formalin test: evidence for forebrain and midbrain sites of action. *Neuroscience* 1994;63:289–294.

89. Manning BH, Mayer DJ. The central nucleus of the amygdala contributes to the production of morphine antinociception in the rat tailflick test. *J Neurosci* 1995;15:8199–8213.

90. Manning BH, Mayer DJ. The central nucleus of the amygdala contributes to the production of morphine analgesia in the formalin test. *Pain* 1995;63:141–152.

91. Neal CR, Mansour A, Reinscheid R, et al. Opioid receptor-like distribution in the rat central nervous system: Comparison of ORL1 receptor mRNA expression with ^{125}I-(^{14}Tyr)-orphanin FQ binding. *J Comp Neurol* 1999;412:563–605.

92. Neal CR, Mansour A, Reinscheid R, et al. Localization of orphanin FQ (nociceptin) peptide and messenger RNA in the central nervous system of the rat. *J Comp Neurol* 1999;406:503–547.

93. Yammamoto T, Nozaki-Taguchi N, Kimura S. Analgesic effect of intrathecally administered nociceptin, an opioid-receptor like agonist, in the rat formalin test. *Neuroscience* 1997;81:249–254.

94. Xu XJ, Hao JX, Wiesendelf-Hallin Z. Nociceptin or antinociceptin: potent antinociceptive effect of orphanin FQ/nociceptin in the rat. *Neuroreport* 1996;7:2092–2094.

95. Rossi GC, Leventhal L, Pasternak GW. Naloxone-sensitive orphanin FQ-induced analgesia in mice. *Eur J Pharmacol* 1996;311:R7–R8.

96. Rossi, GC, Leventhal L, Bolan E. Pharmacological characterization of orphanin FQ/nociceptin and its fragments. *J Pharmacol Exp Ther* 1997;282:858–865.

97. Grisel JE, Mogill JS, Belknap JK, et al. Orphanin FQ acts as at supraspinal, but not spinal, antiopioid peptide. *Neuroreport* 1996;7:2125–2129.

98. Heinricher MM, McGaraughty S, Grandy DK. Circuitry underlying antiopioid actions of orphanin FQ in the rostral ventromedial medulla. *J Neurophysiol* 1997;78:3351–3358.

99. Zadina JE, Hackler L, Ge LJ, et al. A potent and selective endogenous agonist for the mu-opiate receptor. *Nature* 1997;386:499–502.

4

Norepinephrine

Gary Aston-Jones

G. Aston-Jones: Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania.

This chapter reviews findings from basic research concerning brain norepinephrine (NE) systems. The focus is on work that is relevant to the mechanisms of psychiatric disorders, or the actions of drugs used to treat such disorders. The locus ceruleus (LC) system receives most of the attention here, but recent findings concerning the role of the A2/A1 medullary cell groups in drug abuse are also reviewed. Emphasis is placed on studies published since the last version of this volume. Space limitations prevent a thorough review of the involvement of any brain NE system in mental function and dysfunction, so that only a fraction of the relevant research can be covered. Apologies are offered to those whose work could not be included.

MOLECULAR–GENETIC STUDIES

Part of "4 - Norepinephrine "

Previous studies have revealed molecular properties of NE neurons and their effector systems that have extended our understanding of the function and pharmacology of this system. For example, Duman et al. (1) have shown that acute opiate administration decreases cyclic adenosine monophosphate (cAMP) and adenylate cyclase activity in LC neurons, whereas long-term use of opiates or opiate withdrawal results in elevated activity in this second messenger mechanism. Continuing studies in this vein have resulted in a more complete picture of molecular events and properties within LC neurons that help regulate their discharge activity. Thus, the adenylate cyclase/cAMP system is up-regulated with chronic stress but down-regulated with long-term antidepressant treatment (2). Additional studies indicate that impulse activity of LC neurons may be regulated in part by a nonspecific cation current that is activated by this second messenger system (2). These findings suggest a molecular mechanism whereby the overall excitability of LC neurons may be modulated in accordance with long-term environmental or pharmacologic conditions and may be involved in the mechanisms of action of antidepressant and other psychopharmacologic agents.

Recent genetic studies have also revealed important aspects of NE systems relevant to their role in psychopharmacology. Xu et al. (3) studied the brains of mice with a knockout of the NE transporter (3). These mice exhibited characteristics of animals treated with antidepressants (i.e., prolonged clearance of NE and elevated extracellular levels of this catecholamine). In a test for antidepressant drugs, the NE transporter knockouts behaved like antidepressant-treated wild-type mice, being hyperresponsive to locomotor stimulation by cocaine or amphetamine. Importantly, these animals also exhibited dopamine D2/D3-receptor supersensitivity. Thus, NE transporter function can alter midbrain dopaminergic systems, an effect that may be an important mechanism of action of antidepressants and psychostimulants.

NEUROANATOMY

Part of "4 - Norepinephrine "

Chemoanatomy of the LC

The neuroanatomy of the major brain NE systems has been recently reviewed in detail (4), and only the most salient features are described here. In the rat and primate (but not cat, guinea pig, and most other species), virtually all neurons located within the compact LC nucleus are noradrenergic. It is notable that LC neurons also often contain other possible neurotransmitters (e.g., neuropeptides), and subsets of rat NE neurons can be distinguished by neurotransmitter molecules that they co-localize (see ref. 4 for review). Additional work is needed to determine the functional significance of co-localization of other transmitter molecules within LC neurons.

Peri-LC Dendritic Shell

A prominent feature of LC neurons in all species is that their dendrites typically extend hundreds of micra from the parent cell body. Our recent studies have revealed that these dendrites in rat are organized into two prominent collections that project outside the nuclear core in the caudodorsal

and rostroventromedial directions (5). This work has also demonstrated that these dendrites receive numerous synaptic contacts, indicating that the extranuclear peri-LC processes serve as a substantial receptive surface for LC neurons.

Afferents to the LC

Prior studies indicated that prominent afferents to the LC include the nucleus paragigantocellularis (PGi) and the ventromedial aspect of the prepositus hypoglossi (PrH) in the rostroventrolateral and dorsomedial medulla, respectively (6 ,7). These nuclei provide strong excitatory and inhibitory influences on LC neurons, respectively, and are also sources of several neurotransmitter inputs to the LC nucleus (see below) (4 ,8). However, as previously stated, LC dendrites that extend outside the LC nucleus proper provide a prominent receptive surface for inputs to LC neurons (5). Studies of inputs to these peri-LC dendritic zones indicate several additional possible strong inputs to LC neurons, including the periaqueductal gray, medial preoptic nucleus, prefrontal cortex, and hypothalamus (4 ,8). Recent work has confirmed some of the proposed inputs, showing direct contacts onto peri-LC dendrites from amygdala (9) and nucleus tractus solitarius (NTS) (10). Additional work is needed to test some of the other possible inputs to LC distal dendrites. These dendritic inputs are important in revealing additional functional circuitry linked to the LC system (e.g., limbic, autonomic, and cognitive functions).

A host of immunohistochemically defined fibers have been found in LC afferents (see ref. 4 for review). The sources of some of these inputs have been determined. Strong glutamate (11) and epinephrine inputs (12) originate in the PGi, γ-aminobutyric acid (GABA) inputs arise from the PrH (13), and strong enkephalin projections to the LC originate in both the PGi and the PrH (14). Histamine fibers innervate the LC, presumably originating in the tuberomammillary nucleus (15). A particularly dense innervation by serotonin fibers also exists; the origin of this projection has not been determined. Ultrastructural analyses have shown that several of these inputs directly innervate LC neurons (16 ,17 ,18 ,19 and 20).

Most recently, the novel neuropeptide hypocretin (synonymous with orexin) has been shown to innervate the LC densely in rats and monkeys (21 ,22 ,23 and 24) (Fig. 4.1). This projection presumably originates in the hypothalamus (the sole location of hypocretin-producing cells) and is mirrored by dense projections to other nuclei associated with sleep and arousal functions (e.g., the raphe serotonin neurons, tuberomammillary histamine cells, and cholinergic neurons of the brainstem). Initial studies of this peptide suggested a role in feeding (24 ,25). However, more recent work has stimulated considerable interest in this neurotransmitter by closely linking its function to sleep regulation. Specifically, mutations of the gene that makes a hypocretin receptor (26), or of another gene that makes hypocretin itself (27), produced narcolepsy symptoms in animals. This finding supports the long-standing belief that the LC system is important in sleep–waking processes (28) and indicates that sleep disorders may involve anomalies in this hypocretin projection to the LC. These findings also offer a novel target for pharmacologic manipulation of the LC and other systems involved in sleep function.

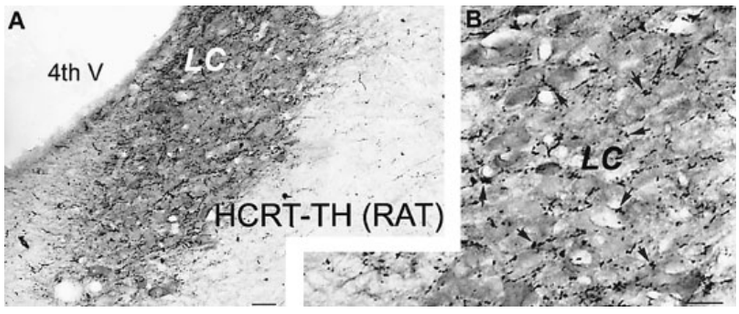


FIGURE 4.1. Photomicrograph showing dense innervation of the locus ceruleus (LC) by hypocretin/orexin Fibers. Low-power **(A)** and high-power **(B)** photographs of frontal sections through the rat LC after staining with antibodies for hypocretin and tyrosine hydroxylase (TH). Note the proximity of numerous black, punctate hypocretin fibers and brown TH-positive NE somata and dendrites. (From Horvath TL, Peyron C, Sabrina D, et al. Strong hypocretin (orexin) innervation of the locus coeruleus activates noradrenergic cells. *J Comp Neurol* 1999;415:145–159, with permission.) See color version of figure.

The functions of the different inputs to LC neurons or their dendrites are being revealed in behavioral and neurophysiologic studies. Stimulation of the PGi strongly excites LC neurons (11). The PGi has strong autonomic functions, an observation consistent with the marked parallel found between LC and sympathetic activities (29). These findings, together with the strong cortical projections of LC neurons, suggest that the LC acts as a cognitive component of a global sympathetic system (8). In contrast, strong inhibition is produced by PrH stimulation (13); the functional significance of this input is unclear. That inhibitory adrenergic input also arises from the PGi is revealed when the strong glutamate input is antagonized pharmacologically (30). Inputs to distal LC dendrites from the amygdala (9) or NTS (10) may convey limbic/emotional or autonomic information to the LC, respectively, although an influence of activity in these afferents on LC activity has not yet been found (8 ,31). Our unpublished studies in monkey indicate that the anterior cingulate cortex strongly innervates the LC (32). Some of our other recent results suggest that this input may modulate the mode of LC activity and thereby its influence on cognitive performance (described below) (33). Finally, our recent studies using transsynaptic retrograde tracing reveal that the suprachiasmatic nucleus is a prominent indirect afferent to the LC (34 ,35 and 36). This is the first demonstration of a circuit that links the circadian suprachiasmatic nucleus mechanism with the arousal/alerting LC system. Inasmuch as other studies have linked circadian disturbances with

depression (37), and the LC system is also associated with depression and other mood disorders (38), this pathway may also be important for affective function.

Topography of LC Efferents

It is well-known that LC axons are highly branched and have extensive efferents that ramify throughout the central nervous system, providing NE innervation at all levels of the neuraxis (see ref. 4 for review). Previous studies have found topography among these efferent projections (39), but the degree of specificity for projections of different LC neurons appears to be quite limited. Recent studies by Simpson et al. (40) have revealed topography of a novel type. They report that LC neurons selectively collateralize to different nuclei of the somatosensory system, so that individual neurons are more likely to send branches to thalamic and cortical areas within the somatosensory system than to, e.g., a somatosensory thalamic nucleus and a visual cortical area. This "functional topography" for projections of individual LC neurons provides a new dimension for the anatomic organization of this ubiquitous brain system and may indicate a means for coordination or synchronization of NE release along relays in serial functional pathways.

A2 NE Neurons of the Caudal Medulla

Norepinephrine neurons in the A2 group (caudal NTS) have recently been implicated in behavioral functions of psychiatric importance. Previously relegated solely to autonomic and visceral control (e.g., see ref. 41), the strong ascending projections of these NE cells to forebrain areas such as the hypothalamus (42), bed nucleus of the stria terminalis (BNST) (43), nucleus accumbens (44), and amygdala (45 ,46) have now been shown also to be important in affective and cognitive processes (43 ,47). As described below, these findings identify new circuits for understanding affective and mnemonic functions.

NEUROPHYSIOLOGY

Part of "4 - Norepinephrine "

Several recent findings regarding the neurophysiology of LC neurons have extended our understanding of this system. Notably, integration of studies at the cellular and behavioral levels indicates a potentially important role of coupling among LC neurons.

Electrotonic Coupling

Experiments by Christie and Williams and colleagues (48 ,49 and 50) showed that LC neurons may be regulated by electrotonic coupling, not only during development but also in adults. Additional studies by these workers indicate that such coupling may be modulated by inputs to LC neurons that alter cAMP (51). This is significant because electrotonic coupling allows rapid, powerful cell-to-cell communication (electrically and biochemically) via large transmembrane channels between neurons (called *gap junctions*). Once relegated to the domain of the esoteric but unimportant, electrotonic coupling is now being demonstrated in an increasing number of central neurons. Of great interest is the fact that such coupling is readily modulated by other inputs to coupled cells—for example, in the retina, coupling is strongly attenuated by dopamine inputs in a cAMP/protein kinase A manner. This line of work is very promising in neuropsychopharmacology because it suggests a novel set of targets (receptors that regulate electrotonic coupling) that could be used to develop new drugs to modulate the function of systems important in mental function and dysfunction (such as the LC). Our recent work (described below) shows how modulation of such coupling can have profound influences on behavior and cognitive performance (33). It is noteworthy that electrotonic coupling has been reported among striatal neurons in a dopamine-modulated manner (see Chapter 9 , *this volume*), as well as among interneurons in the cerebral cortex (52 ,53).

LC Activity, Electrotonic Coupling, and Cognitive Performance in Behaving Monkeys

A possible role for electrotonic coupling among LC neurons in cognitive performance was revealed by combining our recordings of LC neurons in monkeys performing a signal detection task with neural network modeling (33). In these recordings, LC neurons exhibited two modes of activity during task performance: a phasic mode, in which LC cells responded phasically to target stimuli, and a tonic mode, in which the tonic baseline activity of LC neurons was high but responses to target cues were absent. Moreover, the phasic mode corresponded closely to focused attention and good task performance, whereas the tonic mode was associated with scanning attentiveness and poor performance in this task, which requires focused attention. Task performance could be improved by systemic or local (intra-LC) injection of clonidine during poor performance, which indicates a causal influence of these patterns of LC activity on performance. A neural network model was constructed to investigate mechanisms involved in generating these modes of LC activity and the corresponding task performance. Space limitations prohibit a full discussion of the findings, which are reported and reviewed in recent publications (33 ,54). In brief, the model showed that modulated electrotonic coupling among LC neurons could produce the patterns of LC firing observed in the monkeys, and that known modulatory effects of NE could then translate these modes of LC activity into corresponding levels of task performance, also observed in the monkeys (Fig. 4.2 and Fig. 4.3). These findings have a number of implications for neuropsychopharmacology.

First, they support the view that the LC has an important role in attentional processes, and that pathology in LC function could contribute to mental disorders with attentional components [e.g., attention-deficit/hyperactivity disorder (ADHD), stress disorders, schizophrenia]. These results also indicate that alterations in coupling among widely projecting neurons can have profound mental and behavioral consequences, offering a new dimension for analyzing the function of highly divergent modulatory brain systems. Finally, these results, in view of other findings that electrotonic coupling can be rapidly modulated by neurotransmitter inputs (55), indicate that coupling may be a valuable new target for pharmaceutical development in neuropsychopharmacology.

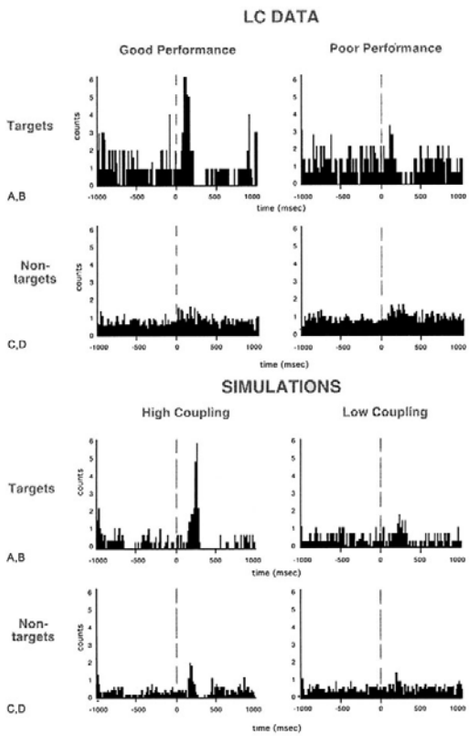


FIGURE 4.2. Simulation of locus ceruleus (LC) activity by modulated electrotonic coupling. **Upper:** Post-stimulus time histograms (PSTHs) for LC activity during the visual discrimination task. **A,B:** Response for targets. **C,D:** Response for distractors. **A,C:** Periods of good performance (phasic LC mode). **B,D:** Poor behavioral performance (false alarm rate typically > 7%; tonic LC mode). Stimuli occur at time zero. All histograms are normalized to a standard of 100 trials. Note that the phasic LC mode is found during periods of good performance, and that the tonic mode corresponds to poor performance on this task. Bin width, 10 ms. **Lower:** Simulation of LC responses. **A,B:** Response to targets. **C,D:** Response to distractors. **A,C:** Coupling among LC neurons. **B,D:** No coupling among LC neurons. These simulation PSTHs are normalized for 100 trials, as for the empiric data. Note that coupling reduces tonic (baseline) LC activity but increases phasic (transient) response to target stimuli, capturing the phasic mode of LC neurons in our recordings. See Fig. 4.3 for corresponding behavioral simulation results. (From Usher M, Cohen JD, Rajkowski J, et al. The role of locus coeruleus in the regulation of cognitive performance. *Science* 1999;283:549–554, with permission.)

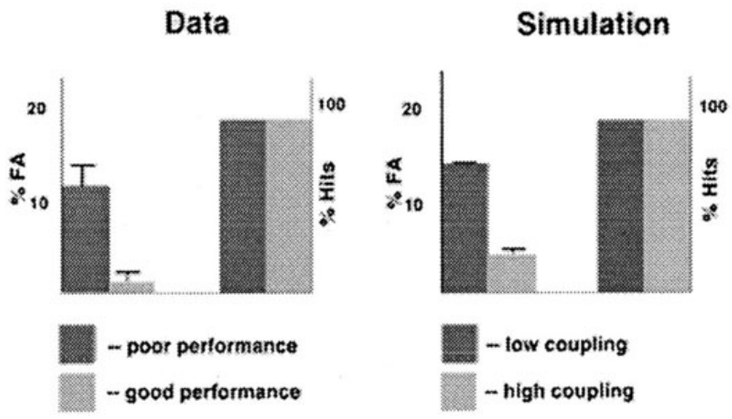


FIGURE 4.3. Simulation of behavioral performance by modulated coupling among locus ceruleus (LC) neurons. **Left:** Graphs showing higher rate of false alarm errors (% FA) during epochs of poor versus good performance by monkeys in the visual discrimination task (33). No differences were noted in the percentage of hit responses during the various levels of performance, as misses were rare. **Right:** Graphs showing higher % FA in the simulated data from our model (33) during epochs of low versus high coupling among LC neurons. Note similarity to empiric data at left. See Fig. 4.2 and ref. 33 for further details.

Opiate Withdrawal

A long series of studies has implicated the LC system in opiate withdrawal (see ref. 56 for review). Recent work has

shed light on molecular and cellular changes that occur in LC neurons during long-term opiate exposure that may underlie their strong activation during withdrawal (reviewed above). It is generally acknowledged that the bulk of this hyperactive LC response is mediated by glutamate inputs from the PGi (11 ,57 ,58). However, a possible intrinsic source of withdrawal-induced hyperactivity in LC neurons has been somewhat controversial. Although some studies find no evidence for withdrawal-induced activation of LC neurons in slices taken from morphine-dependent rats (59 ,60), others have presented evidence for such intrinsically mediated withdrawal responses in LC (61 ,62 ,63 and 64). Our study of local intra-LC microinfusion of opiate antagonists in morphine-dependent rats has confirmed the likelihood that intrinsic changes with dependence contribute to the hyperactivity of these neurons during withdrawal (65). Different studies have suggested different mechanisms for this locally mediated withdrawal effect. Lane-Ladd et al. (62) and Nestler and Aghajanian (66) have presented evidence from slice experiments consistent with the possibility that long-term morphine exposure causes a sustained increase in a tetrodotoxin-insensitive Na^+ current, linked to the increase in cAMP, adenylate cyclase activity, and cAMP response element-binding protein (CREB) that occurs in the LC during withdrawal. In their view, this inward current causes LC hyperactivity when the inhibitory influence of morphine is removed during withdrawal. Our recent *in vitro* studies suggest a different mechanism. These results indicate that long-term opiate administration produces a decrease in K^+ conductance in LC cells that leads to a state of increased excitability when the inhibitory influence of morphine is removed during withdrawal (63 ,64). The decreased K^+ conductance during long-term morphine administration may be a direct compensatory response to the increased K^+ conductance evoked by acute opiates (49). In either case, it seems clear that the local component of withdrawal-induced activation of LC neurons is small compared with the strong excitation evoked by the increased glutamate input from the PGi (see above).

Hypocretin/Orexin

As discussed above, the hypothalamic neuropeptide hypocretin, which is strongly implicated in sleep regulation, densely innervates the LC in rat and monkey (21). Recent studies have revealed that this peptide activates LC neurons both *in vitro* (21 ,67) and *in vivo* (68). The activation is associated with a mild depolarization but is independent of tetrodotoxin and Ca^{2+} (67). The results have led to the tentative conclusion that hypocretin activates LC neurons by decreasing a resting potassium conductance (67). Overall, the results are important because they indicate a possible pathway and transmitter mechanism by which the LC becomes activated during arousal from sleep, which may in turn help to drive a sleep-to-waking transition. This pathway could be involved also in the psychiatric disorders associated with sleep dysfunction (e.g., depression, stress disorders, ADHD).

Cortical Influences on LC Activity

Tract-tracing studies have revealed that the prefrontal cortex may directly innervate LC neurons. Our retrograde and anterograde studies in rat find a projection from the medial prefrontal cortex to the extranuclear peri-LC dendritic zone (69). Another of our studies confirms a projection from the cingulate cortex to the LC in the monkey (32). In line with these findings, additional experiments have revealed prominent effects of cortical stimulation on LC activity. As shown in Fig. 4.4 , we found that electric stimulation of the medial prefrontal cortex in rats activates LC neurons; similar results were obtained with chemical stimulation (70). We also found this activation to be mediated by glutamate release within the LC, as would be expected for a direct cortical (presumably glutamatergic) input (71). In contrast, Sara and Herve-Minvielle (72) reported that medial prefrontal stimulation in rats results in inhibition of LC activity. Procedural differences may underlie the different results. In particular, the study by Sara and Herve-Minvielle used ketamine anesthesia, a potent glutamate antagonist. Thus, the results may indicate an underlying inhibitory effect of prefrontal activation on LC activity when the more potent glutamate-mediated excitation is antagonized. In any case, the results reveal that the prefrontal cortex can strongly influence activity of LC neurons.

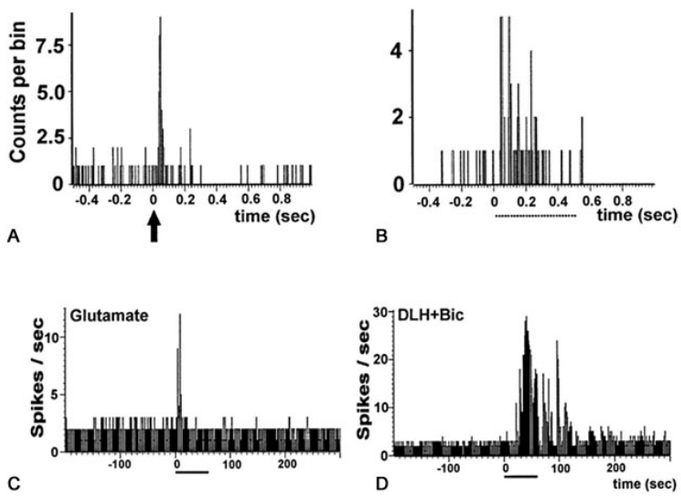


FIGURE 4.4. Activation of locus ceruleus (LC) neuron by stimulation of medial prefrontal cortex (PFC) in rat. **A:** Cumulative post-stimulus time histogram (PSTH) for single-pulse electric stimulation of the PFC. Stimulation presented at *arrow*. **B:** PSTH for train stimulation (20 Hz for 0.5 s) given during the epoch designated by *small dots*. Bin width in each PSTH, 5 ms. **C:** Response of an LC cell to stimulation of PFC with 100-mM glutamate (at *bar* below). **D:** Response of an LC neuron to stimulation of PFC with 10-mM D,L-homocysteic acid plus 50-µM bicuculline (DLH + bic; 60-nL injection). (From Jodo E, Chiang C, Aston-Jones G. Potent excitatory influence of prefrontal cortex activity on noradrenergic locus coeruleus neurons. *Neuroscience* 1998;83:63–80, with permission.)

Postsynaptic Actions of NE

The proposed role of the NE–LC system in arousal was confirmed by Berridge and Foote (73), who showed that local activation of LC neurons by microinjection of bethanechol produces EEG activation in the halothane-anesthetized rat. Similar studies demonstrated that LC inactivation by local microinfusion of clonidine decreases EEG arousal (74). Additional experiments revealed that the arousing effects of LC stimulation are mimicked by stimulation of ß adrenoceptors within the medial septum and are blocked by ß-receptor antagonists infused into this area (75). Continuing studies along these lines confirmed that local LC stimulation in waking animals increases EEG and behavioral indices of arousal (76). Additional studies found, however, that septal infusion of ß antagonists in unanesthetized animals does not decrease arousal (77). Thus, in the waking rat, actions at other NE or non-NE receptors may also be necessary for arousal. Together, these studies indicate that LC activity is an important regulator of EEG arousal, and that these effects are mediated, at least in part, by ß receptors in the medial septum area. Additional studies are needed to determine the precise location of these actions and what other systems and receptors may be important for maintaining the alert state.

Studies in intact animals have shown that ß-receptor activation from the LC can induce plasticity in hippocampal responses. Chaulk and Harley (78) found that *in vivo* or *in vitro* administration of ß- or α-receptor agonists significantly potentiates the population spike amplitude recorded in the dentate gyrus in response to perforant path stimulation. Because the LC is the sole source of NE in the hippocampus, these findings confirm previous results that LC stimulation also potentiates such dentate gyrus responses (79 ,80). These results indicate a role for NE from the LC in plasticity in hippocampal activity, and may provide evidence for a role of this system in memory consolidation (described below).

BEHAVIOR

Part of "4 - Norepinephrine "

Opiate Withdrawal and the LC

Several recent studies in which behavioral pharmacologic techniques were used have reexamined the role of the LC system in opiate withdrawal and abuse. The results of lesion studies by Chieng and Christie (81), Caille et al. (82), and Delfs et al. (43), in which different methods and approaches were used, all agree that the LC system is not necessary for physical signs of morphine withdrawal (Fig. 4.4). This finding contrasts with previous ideas and represents a significantly changed view of the role of the LC system in withdrawal. Although some studies involving microinjection of agents that alter LC activity (83) or molecular events within LC neurons (62 ,84) implicate the LC in withdrawal responses, their results must be viewed with caution because diffusion of injected substances from the small LC nucleus to adjacent areas that have been implicated in withdrawal, such as the periacqueductal gray (85), difficult to rule out. Further studies are needed to determine the behavioral consequences of LC hyperactivity during opiate withdrawal.

Critical Role of A2 NE Neurons Innervating the BNST in Aversion Induced by Opiate Withdrawal

Our recent work has demonstrated that NE innervation of the BNST from A2 noradrenergic neurons is critical for affective responses to opiate withdrawal (43 ,86). We demonstrated that antagonists of ß receptors injected into the BNST, or lesions of the ventral NE bundle that carries fibers from the A2 group to the BNST, eliminate aversive responses to withdrawal (Fig. 4.5). Interestingly, these same manipulations had almost no effect on the physical withdrawal response. These findings, and other results showing that aversive responses to withdrawal can occur in the absence of somatic responses (87 ,88), indicate that withdrawal aversion is not simply a consequence of physical symptoms, and that separate pathways are involved in physical and affective withdrawal responses (Fig. 4.5 and Fig. 4.6). This is important for neuropsychopharmacology because the affective response during withdrawal is the most potent motivator of further drug seeking (89). Thus, studies to develop pharmacotherapies for opiate abuse should focus on aversive withdrawal responses specifically, rather than examining only physical signs. Lesions of the LC system had no effect on aversive or physical signs of withdrawal, findings that corresponded to other recent results (discussed above).

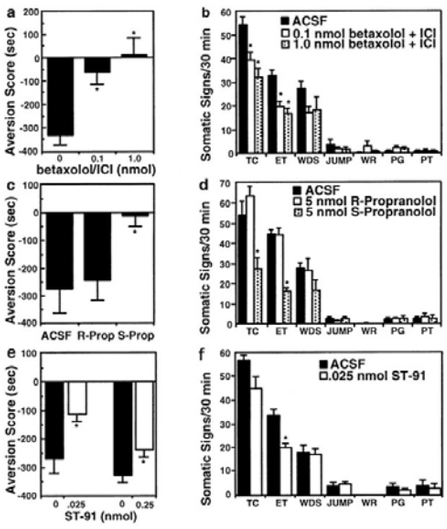


FIGURE 4.5. Effects of intra-BNST (bed nucleus of the stria terminalis) injection of noradrenergic drugs on conditioned place aversion and somatic signs of opiate withdrawal. **A–D:** Effects of the ß-antagonist cocktail betaxolol/ICI 118,551 (A,B) or propranolol isomers (C,D) on place aversion and somatic signs. **E,F:** Effects of ST-91 on place aversion and somatic signs. TC, teeth chatter; ET, eye twitch; WDS, wet dog shakes; JUMP, jumping; WR, writhing; PG, penile grooming; PT, paw tremor. All data are expressed as mean ± standard error of the mean (n = 6 to 8 animals per dose). For A–D, $p < .05$, analysis of variance followed by Fisher's PLSD test for multiple comparisons. For E,F, $p < .05$, Student's t-test. (From Delfs J, Zhu Y, Druhan J, et al. Noradrenaline in the ventral forebrain is critical for opiate withdrawal-induced aversion. *Nature* 2000;403:430–434, with permission.)

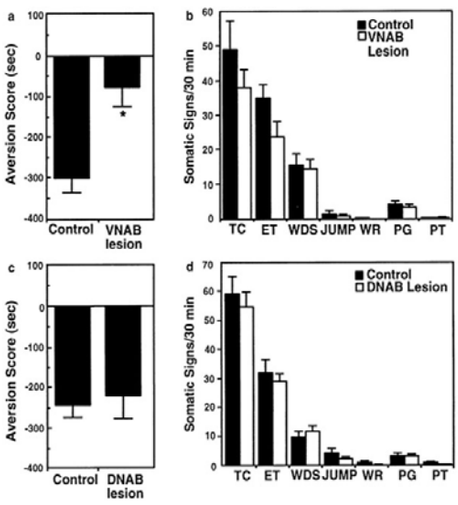


FIGURE 4.6. Effects of dorsal (DNAB) and ventral (VNAB) noradrenergic bundle lesions on aversive and somatic signs of opiate withdrawal. **A,C:** Aversion scores. Aversion score equals time in the naltrexone-paired side on the test day minus the preconditioning day. **B,D:** Number of somatic counts in 30 minutes. See Fig. 4.5 legend for details and abbreviations. Nondependent lesioned animals exhibited neither aversion nor somatic signs following naltrexone (data not shown). All data are mean ± standard error of the mean (n = 6 to 8 control, 10 to 11 lesioned animals per group). $p < .05$, analysis of variance followed by Fisher's PLSD test for multiple comparisons. (From Delfs J, Zhu Y, Druhan J, et al. Noradrenaline in the ventral forebrain is critical for opiate withdrawal-induced aversion. *Nature* 2000;403:430–434, with permission.)

Memory and the LC

Recent studies by Clayton and Williams (90) have indicated new evidence for involvement of the NE–LC system in memory. Inactivation of the PGi (a major input to the LC, described above) with either lidocaine or the GABA agonist muscimol immediately after acquisition in a one-trial inhibitory avoidance task produced marked deficits on a retention test given 48 hours later. Conversely, chemical stimulation

of the PGi with glutamate following training in either an inhibitory avoidance or spatial delayed matching to sample radial maze task enhanced retention performance when assessed 48 or 18 hours later, respectively (91). Given the excitatory connections between PGi and LC, these findings suggest that pharmacologic manipulation of PGi neuronal activity may affect memory formation via influences on LC and subsequent NE release in brain systems involved in the encoding of new information.

Recent studies by Przybyslawski et al. (92) have also indicated a role for the LC–NE system in memory. These experiments indicate that memories are normally reconsolidated each time they are reactivated by relevant cues. They found that blockade of ß adrenoceptors after memory reactivation, during the consolidation process, produced impairment on future tests of the same memory. These results indicate that reactivation of memory produces a ß receptor-dependent intracellular cascade that reenacts the consolidation process responsible for the initial memory acquisition. This NE-dependent lability of active memory traces indicates a novel mechanism to target in pharmacologic manipulation of memory-related disorders, such as posttraumatic stress disorder and Alzheimer's disease.

Studies by Mao et al. (93) have found a role for α_1 and α_2 NE receptors in the dorsolateral prefrontal cortex in memory. Infusion of α_1 agonists into the monkey prefrontal cortex produced deficits in working memory (93), whereas similar treatments with α_2 agonists improved memory performance (94).

Memory and the A2 NE System

Studies by McGaugh (95) during the last several years have established a role for NE stimulation of ß receptors in the amygdala in the strong memories that are established for emotionally salient events. Recently, this line of work has shown that the NTS is involved in this process, as lidocaine anesthesia of the NTS prevents the memory-enhancing effects of peripheral epinephrine (47). Because the A2 neurons of

the NTS strongly innervate the amygdala, this finding indicates that the A2 neurons may be importantly involved in memory modulation. These studies suggest that a “central nervous system–periphery–central nervous system long-loop” circuit may be involved, in which descending activity in response to emotional events produces a peripheral response (e.g., epinephrine release); this response in turn stimulates receptors on vagal afferents that then stimulate the NTS to release NE in its hypothalamic and forebrain targets. This possible route for enhancement of emotional memories and other cognitive processes has received little attention previously. Such a loop may also be involved in the activation of A2 neurons during opiate withdrawal that leads to the corresponding aversive response (described above) (43). This is potentially important clinically and psychopharmacologically because peripheral receptors on visceral afferent fibers that may be involved in mental disorders represent a novel mechanism and target for new pharmacotherapies.

PSYCHOPATHOLOGY

Part of "4 - Norepinephrine "

Depression

Recent work by Miller et al. (96) has increased our understanding of the role of NE systems in depression. In their studies, reduction of NE metabolites (presumably reflecting decreased NE turnover) after treatment with α-methyl-*p*-tyrosine (AMPT) caused no change in scores on the Hamilton Depression Rating Scale in normal human subjects. In contrast, AMPT administration and reductions in NE turnover in patients in remission from depression after treatment with desipramine or mazindol significantly increased the Hamilton Depression Rating Scale measures of depressive symptoms (97). This change was not seen in patients under treatment with serotonin antidepressants (fluoxetine or sertraline). The results indicate that monoamine deficiency alone may not produce depressive symptoms, but that different types of depression exist that respond to manipulations of different monoamine systems.

Advances in understanding the actions of antidepressant drugs have highlighted the possible role of NE systems in depression. New drugs such as venlafaxine, which inhibits reuptake of both serotonin and NE, have been found to be effective, particularly in refractory depression (98). In addition, the highly effective antidepressant paroxetine, which was previously thought to act selectively to block serotonin reuptake, has recently been found also to inhibit NE reuptake (99 ,100). These findings confirm long-held beliefs that NE is importantly involved in depression, and indicate that blockers of NE uptake, including drugs that selectively act at the NE transporter, such as reboxetine (101 ,102), may be effective in treating at least certain types of affective illness (103).

Anxiety

Brain NE has long been implicated in anxiety disorders (104). Our studies with cocaine- and morphine-dependent animals have provided new evidence for a role of central NE systems in anxiety. By means of a place-conditioning paradigm, we found that withdrawal from long-term administration of morphine or cocaine is associated with strong anxiety, measured by the conditioned burying paradigm (105). Importantly, the anxiogenic response to drug withdrawal is strongly attenuated by administration of the ß-receptor antagonist propranolol, and by similar doses of the lipophobic $ß_1$ antagonist atenolol, which is believed to act primarily peripherally. These findings indicate that at least some types of anxiety involve stimulation of peripheral ß adrenoceptors.

ADHD

The firing patterns of LC neurons in behaving monkeys indicate that this system plays an important role in attention and performance (reviewed above) (33 ,54 ,106). In particular, one mode of LC activity, characterized by elevated tonic discharge, corresponds to poor performance on a continuous performance task that requires focused attention, with a high rate of false alarm errors. These and other results have led us to propose that this tonic mode of LC activity promotes high behavioral flexibility and disables focused or selective attention (33 ,54). This view also implies that attentional disorders may be associated with LC dysregulation in which the proper mode of activity is not engaged adaptively for the context at hand. Specifically, several parallels have been noted between behaviors in monkeys during the tonic mode of LC activity and symptoms of ADHD, including hypervigilance, irritability, poor focused attentiveness, and a high false alarm rate in continuous performance tasks. These findings indicate that the LC may play an important role in ADHD, and that drugs that modulate LC mode, or switching between modes, may be helpful in treating this disorder. In fact, many of the stimulants that are effective in treating ADHD decrease tonic LC activity.

A role for the LC–NE system in attentional disorders is also indicated by behavioral pharmacology experiments by Arnsten and colleagues (107). These investigators have found that overstimulation of $α_1$ receptors in the prefrontal cortex produces deficits in behaviors that depend on prefrontal function (107). Because ADHD includes symptoms of prefrontal dysfunction, these findings raise the possibility that an overactive LC system may contribute to ADHD by overstimulation of $α_1$ receptors in prefrontal areas (108).

CONCLUSIONS

Part of "4 - Norepinephrine "

An impressive amount of research on NE systems has been performed since the previous edition of this volume was

published. This work is revealing an increasingly important role for brain NE in mental function and dysfunction. Mechanisms by which NE systems are involved in cognitive, addictive, stress-related, and other behavioral functions are being elucidated. This progress not only reinforces the importance of this system for neuropsychopharmacology, but also indicates that NE systems represent a promising area for discovering new and fruitful approaches to developing treatments for psychiatric disorders.

ACKNOWLEDGMENTS

Part of "4 - Norepinephrine "

This work was supported by PHS grants NS24698, DA06214, DA10088, MH55309, and MH59978. Comments on the manuscript by Drs. Glenda Harris and Jon Druhan are greatly appreciated.

REFERENCES

1. Duman RS, Tallman JF, Nestler EJ. Acute and chronic opiate-regulation of adenylate cyclase in brain: specific effects in locus coeruleus. *J Pharmacol Exp Ther* 1988;246:1033–1039.

2. Nestler EJ, Alreja M, Aghajanian GK. Molecular control of locus coeruleus neurotransmission. *Biol Psychiatry* 1999;46:1131–1139.

3. Xu F, Gainetdinov RR, Wetsel WC, et al. Mice lacking the norepinephrine transporter are supersensitive to psychostimulants. *Nat Neurosci* 2000;3:465–471.

4. Aston-Jones G, Shipley MT, Grzanna R. The locus coeruleus A5 and A7 noradrenergic cell groups. In: Paxinos G, ed. *The rat nervous system*, second ed. New York: Academic Press, 1995:183–214.

5. Shipley MT, Fu L, Ennis M, et al. Dendrites of locus coeruleus neurons extend preferentially into two pericoerulear zones. *J Comp Neurol* 1996;365:56–68.

6. Aston-Jones G, Ennis M, Pieribone VA, et al. The brain nucleus locus coeruleus: restricted afferent control of a broad efferent network. *Science* 1986;234:734–737.

7. Van Bockstaele EJ, Colago EEO, Aicher S. Light and electron microscopic evidence for topographic and monosynaptic projections from neurons in the ventral medulla to noradrenergic dendrites in the rat locus coeruleus. *Brain Res* 1998;784:123–138.

8. Aston-Jones G, Shipley MT, Chouvet G, et al. Afferent regulation of locus coeruleus neurons: anatomy, physiology and pharmacology. *Prog Brain Res* 1991;88:47–75.

9. Van Bockstaele EJ, Colago EE, Valentino RJ. Amygdaloid corticotropin-releasing factor targets locus coeruleus dendrites: substrate for the co-ordination of emotional and cognitive limbs of the stress response. *J Neuroendocrinol* 1998;10:743–757.

10. Van Bockstaele EJ, Peoples J, Telegan P. Efferent projections of the nucleus of the solitary tract to peri-locus coeruleus dendrites in rat brain: evidence for a monosynaptic pathway. *J Comp Neurol* 1999;412:410–428.

11. Ennis M, Aston-Jones G. Activation of locus coeruleus from nucleus paragigantocellularis: a new excitatory amino acid pathway in brain. *J Neurosci* 1988;8:3644–3657.

12. Pieribone VA, Aston-Jones G. Adrenergic innervation of the rat nucleus locus coeruleus arises from the C1 and C3 cell groups in the rostral medulla: an anatomic study combining retrograde transport and immunofluorescence. *Neuroscience* 1991;41:525–542.

13. Ennis M, Aston-Jones G. GABA-mediated inhibition of locus coeruleus from the dorsomedial rostral medulla. *J Neurosci* 1989;9:2973–2981.

14. Drolet G, Van Bockstaele EJ, Aston-Jones G. Robust enkephalin innervation of the locus coeruleus from the rostral medulla. *J Neurosci* 1992;12:3162–3174.

15. Panula P, Pirvola U, Auvinen S, et al. Histamine-immunoreactive nerve fibers in the rat brain. *Neuroscience* 1989;28:585–610.

16. Van Bockstaele EJ, Colago EE, Valentino RJ. Corticotropin-releasing factor-containing axon terminals synapse onto catecholamine dendrites and may presynaptically modulate other afferents in the rostral pole of the nucleus locus coeruleus in the rat brain. *J Comp Neurol* 1996;364:523–534.

17. van Bockstaele EJ, Colago EE, Pickel VM. Enkephalin terminals form inhibitory-type synapses on neurons in the rat nucleus locus coeruleus that project to the medial prefrontal cortex. *Neuroscience* 1996;71:429–442.

18. Van Bockstaele EJ. Morphological substrates underlying opioid, epinephrine and gamma-aminobutyric acid inhibitory actions in the rat locus coeruleus. *Brain Res Bull* 1998;47:1–15.

19. Van Bockstaele EJ, Chan J. Electron microscopic evidence for coexistence of leucine 5-enkephalin and gamma-aminobutyric acid in a subpopulation of axon terminals in the rat locus coeruleus region. *Brain Res* 1997;746:171–182.

20. Van Bockstaele EJ, Saunders A, Commons KG, et al. Evidence for coexistence of enkephalin and glutamate in axon terminals and cellular sites for functional interactions of their receptors in the rat locus coeruleus. *J Comp Neurol* 2000;417:103–114.

21. Horvath TL, Peyron C, Sabrina D, et al. Strong hypocretin (orexin) innervation of the locus coeruleus activates noradrenergic cells. *J Comp Neurol* 1999;415:145–159.

22. de Lecea L, Kilduff TS, Peyron C, et al. The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity. *Proc Natl Acad Sci U S A* 1998;95:322–327.

23. Peyron C, Tighe DK, van den Pol AN, et al. Neurons containing hypocretin (orexin) project to multiple neuronal systems. *J Neurosci* 1998;18:9996–10015.

24. Sakurai T, Amemiya A, Ishii M, et al. Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* 1998;92:573–585.

25. Horvath TL, Diano S, van den Pol AN. Synaptic interaction between hypocretin (orexin) and neuropeptide Y cells in the rodent and primate hypothalamus: a novel circuit implicated in metabolic and endocrine regulations. *J Neurosci* 1999;19:1072–1087.

26. Lin L, Faraco J, Li R, et al. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* 1999;98:365–376.

27. Chemelli RM, Willie JT, Sinton CM, et al. Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell* 1999;98:437–451.

28. Aston-Jones G, Chiang C, Alexinsky T. Discharge of noradrenergic locus coeruleus neurons in behaving rats and monkeys suggests a role in vigilance. *Prog Brain Res* 1991;88:501–520.

29. Svensson TH. Peripheral, autonomic regulation of locus coeruleus noradrenergic neurons in brain: putative implications for psychiatry and psychopharmacology. *Psychopharmacology (Berl)* 1987;92:1–7.

30. Aston-Jones G, Astier B, Ennis M. Inhibition of locus coeruleus noradrenergic neurons by C1 adrenergic cells in the rostral ventral medulla. *Neuroscience* 1992;48:371–382.

31. Aston-Jones G, Ennis M, Pieribone VA, et al. The brain nucleus locus coeruleus: restricted afferent control of a broad efferent network. *Science* 1986;234:734–737.

32. Rajkowski J, Lu W, Zhu Y, et al. Prominent projections from the anterior cingulate cortex to the locus coeruleus in rhesus monkey. *Soc Neurosci Abst* 2000;26:2230.

33. Usher M, Cohen JD, Rajkowski J, et al. The role of locus coeruleus in the regulation of cognitive performance. *Science* 1999;283:549–554.

34. Aston-Jones G, Chen S, Yang M, et al. The neural pathway from the suprachiasmatic nucleus to the locus coeruleus: a transsynaptic tracing study. *Soc Neurosci Abst* 1998;24:1596.

35. Aston-Jones G, Chen S, Zhu Y, et al. A neural circuit for circadian regulation of arousal. *Nat Neurosci* 2001;4:732–738.

36. Aston-Jones G, Card JP. Use of pseudorabies virus to delineate multisynaptic circuits in brain: opportunities and limitations. *J Neurosci Methods* 2000;103:51–61.

37. Gillin JC, Sitaram N, Duncan WC, et al. Sleep disturbance in depression: diagnostic potential and pathophysiology [Proceedings]. *Psychopharmacol Bull* 1980;16:40–42.

38. Siever LJ, Davis KL. Overview: toward a dysregulation hypothesis of depression. *Am J Psychiatry* 1985;142:1017–1031.

39. Loughlin SE, Foote SL, Grzanna R. Efferent projections of nucleus locus coeruleus: morphologic subpopulations have different efferent targets. *Neuroscience* 1986;18:307–319.

40. Simpson KL, Altman DW, Wang L, et al. Lateralization and functional organization of the locus coeruleus projection to the trigeminal somatosensory pathway in rat. *J Comp Neurol* 1997;385:135–147.

41. Blessing WW. *The lower brainstem and bodily homeostasis*. New York: Oxford University Press:1997.

42. Sawchenko PE, Swanson LW. The organization of noradrenergic pathways from the brainstem to the paraventricular and supraoptic nuclei in the rat. *Brain Res* 1982;257:275–325.

43. Delfs J, Zhu Y, Druhan J, et al. Noradrenaline in the ventral forebrain is critical for opiate withdrawal-induced aversion. *Nature* 2000;403:430–434.

44. Delfs JM, Zhu Y, Druhan JP, et al. Origin of noradrenergic afferents to the shell subregion of the nucleus accumbens: anterograde and retrograde tract-tracing studies in the rat. *Brain Res* 1998;806:127–140.

45. Zardetto-Smith AM, Gray TS. Catecholamine and NPY efferents from the ventrolateral medulla to the amygdala in the rat. *Brain Res Bull* 1995;38:253–260.

46. Emson PC, Bjorklund A, Lindvall O, et al. Contributions of different afferent pathways to the catecholamine and 5-hydroxytryptamine innervation of the amygdala: a neurochemical and histochemical study. *Neuroscience* 1979;4:1347–1357.

47. Williams CL, McGaugh JL. Reversible lesions of the nucleus of the solitary tract attenuate the memory-modulating effects of posttraining epinephrine. *Behav Neurosci* 1993;107:955–962.

48. Christie MJ, Williams JT, North RA. Electrical coupling synchronizes subthreshold activity in locus coeruleus neurons *in vitro* from neonatal rats. *J Neurosci* 1989;9:3584–3589.

49. Travalgi RA, Dunwiddie TV, Williams JT. Opioid inhibition in locus coeruleus. *J Neurophysiol* 1995;74:519–528.

50. Ishimatsu M, Williams JT. Synchronous activity in locus coeruleus results from dendritic interactions in pericoerulear regions. *J Neurosci* 1996;16:5196–5204.

51. Osborne PB, Williams JT. Forskolin enhancement of opioid currents in rat locus coeruleus neurons. *J Neurophysiol* 1996;76:1559–1565.

52. Gibson JR, Beierlein M, Connors BW. Two networks of electrically coupled inhibitory neurons in neocortex. *Nature* 1999;402:75–79.

53. Galarreta M, Hestrin S. A network of fast-spiking cells in the neocortex connected by electrical synapses. *Nature* 1999;402:72–75.

54. Aston-Jones G, Rajkowski J, Cohen J. Role of locus coeruleus in attention and behavioral flexibility. *Biol Psychiatry* 1999;46:1309–1320.

55. Baldridge W, Vaney DI, Weiler R. The modulation of intercellular coupling in the retina. *Semin Cell Dev Biol* 1998;9:311–318.

56. Aston-Jones G, Shiekhattar R, Rajkowski J, et al. Opiates influence noradrenergic locus coeruleus neurons by potent indirect as well as direct effects. In: Hammer R, ed. *The neurobiology of opiates*. New York: CRC Press, 1993:175–202.

57. Akaoka H, Aston-Jones G. Opiate withdrawal-induced hyperactivity of locus coeruleus neurons is substantially mediated by augmented excitatory amino acid input. *J Neurosci* 1991;11:3830–3839.

58. Rasmussen K, Aghajanian GK. Withdrawal-induced activation of locus coeruleus neurons in opiate-dependent rats: attenuation by lesions of the nucleus paragigantocellularis. *Brain Res* 1989;505:346–350.

59. Bell JA, Grant SJ. Locus coeruleus neurons from morphine-treated rats do not show opiate-withdrawal hyperactivity *in vitro. Brain Res* 1998;788:237–244.

60. Christie MJ, Williams JT, Osborne PB, et al. Where is the locus in opioid withdrawal? *Trends Pharmacol Sci* 1997;18:134–140.

61. Kogan JH, Nestler EJ, Aghajanian GK. Elevated basal firing rates and enhanced responses to 8-Br-cAMP in locus coeruleus neurons in brain slices from opiate-dependent animals. *Eur J Pharmacol* 1992;211:47–53.

62. Lane-Ladd SB, Pineda J, Boundy VA, et al. CREB (cAMP response element-binding protein) in the locus coeruleus: biochemical, physiological, and behavioral evidence for a role in opiate dependence. *J Neurosci* 1997;17:7890–7901.

63. Ivanov A, Aston-Jones G. Local opiate withdrawal in locus coeruleus neurons *in vitro. J Neurophysiol* 2001;85:2388–2397.

64. Ivanov AY, Aston-Jones G. Local opioid withdrawal in locus coeruleus (LC) neurons suppressed by protein kinase A (PKA) inhibitors. *Soc Neurosci Abst* 1998;24:1488.

65. Aston-Jones G, Hirata H, Akaoka H. Local opiate withdrawal in locus coeruleus *in vivo. Brain Res* 1997;765:331–336.

66. Nestler EJ, Aghajanian GK. Molecular and cellular basis of addiction. *Science* 1997;278:58–63.

67. Ivanov A, Aston-Jones G. Hypocretin/orexin depolarizes and decreases potassium conductance in locus coeruleus neurons. *Neuroreport* 2000;11:1755–1758.

68. Hagan JJ, Leslie RA, Patel S, et al. Orexin A activates locus coeruleus cell firing and increases arousal in the rat. *Proc Natl Acad Sci U S A* 1999;96:10911–10916.

69. Zhu Y, Aston-Jones G. The medial prefrontal cortex prominently innervates a peri-locus coeruleus dendritic zone in rat. *Soc Neurosci Abst* 1996;22:601.

70. Jodo E, Chiang C, Aston-Jones G. Potent excitatory influence of prefrontal cortex activity on noradrenergic locus coeruleus neurons. *Neuroscience* 1998;83:63–80.

71. Jodo E, Aston-Jones G. Activation of locus coeruleus by prefrontal cortex is mediated by excitatory amino acid inputs. *Brain Res* 1997;768:327–332.

72. Sara SJ, Herve-Minvielle A. Inhibitory influence of frontal cortex on locus coeruleus neurons. *Proc Natl Acad Sci U S A* 1995;92:6032–6036.

73. Berridge CW, Foote SL. Effects of locus coeruleus activation on electroencephalographic activity in the neocortex and hippocampus. *J Neurosci* 1991;11:3135–3145.

74. Berridge CW, Page ME, Valentino RJ, et al. Effects of locus coeruleus inactivation on electroencephalographic activity in neocortex and hippocampus. *Neuroscience* 1993;55:381–393.

75. Berridge CW, Bolen SJ, Manley MS, et al. Modulation of forebrain electroencephalographic activity in halothane-anesthetized rat via actions of noradrenergic B-receptors within the medial septal region. *J Neurosci* 1996;16:7010–7020.

76. Berridge CW, Foote SL. Enhancement of behavioral and electroencephalographic indices of waking following stimulation of noradrenergic B-receptors within the medial septal region of the basal forebrain. *J Neurosci* 1996;16:6999–7009.

77. Berridge CW, Wifler K. Contrasting effects of noradrenergic beta-receptor blockade within the medial septal area on forebrain electroencephalographic and behavioral activity state in anesthetized and unanesthetized rat. *Neuroscience* 2000;97:543–552.

78. Chaulk PC, Harley CW. Intracerebroventricular norepinephrine potentiation of the perforant path-evoked potential in dentate gyrus of anesthetized and awake rats: a role for both alpha- and beta-adrenoceptor activation. *Brain Res* 1998;787:59–70.

79. Klukowski G and Harley CW. Locus coeruleus activation induces perforant path-evoked population spike potentiation in the dentate gyrus of awake rat. *Exp Brain Res* 1994;102:165–170.

80. Harley C, Milway JS, Lacaille JC. Locus coeruleus potentiation of dentate gyrus responses: evidence for two systems. *Brain Res Bull* 1989;22:643–650.

81. Chieng B, Christie MJ. Lesions to terminals of noradrenergic locus coeruleus neurones do not inhibit opiate withdrawal behaviour in rats. *Neurosci Lett* 1995;186:37–40.

82. Caille S, Espejo EF, Reneric JP, et al. Total neurochemical lesion of noradrenergic neurons of the locus ceruleus does not alter either naloxone-precipitated or spontaneous opiate withdrawal nor does it influence ability of clonidine to reverse opiate withdrawal. *J Pharmacol Exp Ther* 1999;290:881–892.

83. Taylor JR, Punch LJ, Elsworth JD. A comparison of the effects of clonidine and CNQX infusion into the locus coeruleus and the amygdala on naloxone-precipitated opiate withdrawal in the rat. *Psychopharmacology (Berl)* 1998;138:133–142.

84. Punch LJ, Self DW, Nestler EJ, et al. Opposite modulation of opiate withdrawal behaviors on microinfusion of a protein kinase A inhibitor versus activator into the locus coeruleus or periaqueductal gray. *J Neurosci* 1997;17:8520–8527.

85. Chieng B, Christie MJ. Local opioid withdrawal in rat single periaqueductal gray neurons *in vitro*. *J Neurosci* 1996;16:7128–7136.

86. Aston-Jones G, Delfs J, Druhan J, et al. The bed nucleus of the stria terminalis: s target site for noradrenergic actions in opiate withdrawal. In: McGinty J, ed. *Advancing from the ventral striatum to the extended amygdala: implications for neuropsychiatry and drug abuse*. New York: New York Academy of Sciences, 1999:486–498.

87. Harris G, Aston-Jones G. Beta-adrenergic antagonists attenuate somatic and aversive signs of opiate withdrawal. *Neuropsychopharmacology* 1993;9:303–311.

88. Schulteis G, Markou A, Gold LH, et al. Relative sensitivity to naloxone of multiple indices of opiate withdrawal: a quantitative dose–response analysis. *J Pharmacol Exp Ther* 1994;271:1391–1398.

89. O'Brien CP, Childress AR, McLellan A, et al. Types of conditioning found in drug-dependent humans. *NIDA Res Monogr* 1988;84:44–61.

90. Clayton EC, Williams CL. Posttraining inactivation of excitatory afferent input to the locus coeruleus impairs retention in an inhibitory avoidance learning task. *Neurobiol Learn Mem* 2000;73:127–140.

91. Clayton EC, Williams CL. Glutamatergic influences on the nucleus paragigantocellularis: contribution to performance in avoidance and spatial memory tasks. *Behav Neurosci* 2000;114:707–712.

92. Przybyslawski J, Roullet P, Sara SJ. Attenuation of emotional and nonemotional memories after their reactivation: role of beta adrenergic receptors. *J Neurosci* 1999;19:6623–6628.

93. Mao ZM, Arnsten AF, Li BM. Local infusion of an alpha-1 adrenergic agonist into the prefrontal cortex impairs spatial working memory performance in monkeys. *Biol Psychiatry* 1999;46:1259–1265.

94. Franowicz JS, Arnsten AF. The alpha-2a noradrenergic agonist, guanfacine, improves delayed response performance in young adult rhesus monkeys. *Psychopharmacology (Berl)* 1998;136:8–14.

95. McGaugh JL. Memory—a century of consolidation. *Science* 2000;287:248–251.

96. Miller HL, Delgado PL, Salomon RM, et al. Effects of alpha-methyl-para-tyrosine (AMPT) in drug-free depressed patients. *Neuropsychopharmacology* 1996;14:151–157.

97. Miller HL, Delgado PL, Salomon RM, et al. Clinical and biochemical effects of catecholamine depletion on antidepressant-induced remission of depression. *Arch Gen Psychiatry* 1996;53:117–128.

98. Andrews JM, Ninan PT, Nemeroff CB. Venlafaxine: a novel antidepressant that has a dual mechanism of action. *Depression* 1996;4:48–56.

99. Owens MJ, Knight DL, Nemeroff CB. Paroxetine binding to the rat norepinephrine transporter *in vivo*. *Biol Psychiatry* 2000;47:842–845.

100. Owens MJ, Morgan WN, Plott SJ, et al. Neurotransmitter receptor and transporter binding profile of antidepressants and their metabolites. *J Pharmacol Exp Ther* 1997;283:1305–1322.

101. Wong EH, Sonders MS, Amara SG, et al. Reboxetine: a pharmacologically potent, selective, and specific norepinephrine reuptake inhibitor. *Biol Psychiatry* 2000;47:818–829.

102. Gorman JM, Sullivan G. Noradrenergic approaches to antidepressant therapy. *J Clin Psychiatry* 2000;61:13–16.

103. Ressler KJ, Nemeroff CB. Role of norepinephrine in the pathophysiology and treatment of mood disorders. *Biol Psychiatry* 1999;46:1219–1233.

104. Charney DS, Bremner JD, Redmond J. Noradrenergic neural substrates for anxiety and fear: clinical associations based on preclinical research. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:387–397.

105. Harris G, Aston-Jones G. Beta-adrenergic antagonists attenuate withdrawal anxiety in cocaine- and morphine-dependent rats. *Psychopharmacology* 1993;113:131–136.

106. Aston-Jones G, Rajkowski J, Kubiak P, et al. Locus coeruleus neurons in the monkey are selectively activated by attended stimuli in a vigilance task. *J Neurosci* 1994;14:4467–4480.

107. Arnsten AF, Mathew R, Ubriani R, et al. Alpha-1 noradrenergic receptor stimulation impairs prefrontal cortical cognitive function. *Biol Psychiatry* 1999;45:26–31.

108. Arnsten AF, Steere JC, Hunt RD. The contribution of alpha 2-noradrenergic mechanisms of prefrontal cortical cognitive function. Potential significance for attention-deficit hyperactivity disorder. *Arch Gen Psychiatry* 1996;53:448–455.

5

Regulation of G Protein-Coupled Receptors by Phosphorylation and Endocytosis

Mark Von Zastrow

Mark von Zastrow: Department of Psychiatry, Department of Cellular and Molecular Pharmacology, and Program in Cell Biology, University of California, San Francisco, California.

G protein-coupled receptors (GPCRs) comprise a large superfamily of heptahelical integral membrane proteins that mediate transmembrane signal transduction in response to a wide variety of hormones, neurotransmitters, and neuromodulators. GPCRs are extremely important targets for neuropsychopharmacology. Indeed, the vast majority of clinically relevant neuropsychiatric drugs either bind directly to specific GPCRs (e.g., many antipsychotic drugs) or function indirectly via GPCRs by influencing the amount of available native agonist (e.g., many antidepressant drugs).

A general feature of GPCRs is that they are extensively regulated in cells (1 ,2 ,3 and 4). Regulation of GPCRs is thought to play a fundamental role in maintaining physiologic homeostasis in the face of fluctuating internal and external stimuli. A number of pathologic states are associated with disturbances in the number or functional activity of certain GPCRs (5). In addition, many clinically important drugs influence the physiologic regulation of GPCRs (6). Together, these observations suggest that mechanisms of GPCR regulation may be of fundamental importance to neuropsychiatric disorders and to the actions of clinically relevant drugs.

The physiologic and biomedical importance of GPCR regulation has motivated an enormous amount of study into underlying molecular mechanisms of regulation. Progress in this area has been facilitated enormously by molecular and cell biological approaches applied to a variety of experimental model systems. Our understanding remains at an early stage and is limited, in most cases, to studies of a small number of GPCRs. Nevertheless, great progress has been made in elucidating certain mechanisms of GPCR regulation, to the extent that it is possible to begin to discern fundamental principles that control the number and functional activity of GPCRs in individual cells.

The present chapter discusses some of this progress, with an emphasis on developing a unified view of GPCR regulation. We have restricted our scope to a limited number of regulatory mechanisms that have been elucidated by detailed study of the some of the most extensively characterized GPCRs. First, we survey classic studies describing the general properties of the physiologic and pharmacologic regulation of receptor-mediated signaling; these have established a terminology and conceptual framework for our later focus on specific mechanisms of receptor regulation. Second, we discuss pioneering studies of “prototypic” GPCRs that have established paradigms for understanding the role of receptor phosphorylation in mediating rapid desensitization of GPCRs. Third, we focus on a specific mechanism mediating regulated endocytosis of certain GPCRs, and discuss how this endocytic mechanism can promote rapid desensitization and resensitization of receptor-mediated signal transduction. In this section, we also highlight the close interdependence between mechanisms of GPCR phosphorylation and membrane trafficking in mediating rapid regulation of receptor function. Finally, we discuss the functions of both phosphorylation and endocytic membrane trafficking in mediating longer-term regulation of the number of GPCRs present in cells, focusing on recent studies into mechanisms that control down-regulation of receptors via proteolytic degradation in lysosomes.

GENERAL PROCESSES OF GPCR REGULATION

Part of "5 - Regulation of G Protein-Coupled Receptors by Phosphorylation and Endocytosis "

Rapid Desensitization and Resensitization

It has been known for many years that multiple mechanisms can contribute to GPCR regulation (1 ,7). Early studies,

which preceded the elucidation of any of the biochemical machinery involved, distinguished general processes of receptor regulation according to differences in kinetics and reversibility. This is well illustrated by classic studies of the β_2-adrenergic receptor (B2AR), reviewed in detail elsewhere (2 ,3 and 4). Agonist-induced activation of the B2AR stimulates adenylyl cyclase via coupling to the G_s heterotrimeric G protein (Fig. 1A). Receptor-mediated signaling via this pathway occurs within seconds after agonist binding. However, after more prolonged activation, the ability of receptors to activate adenylyl cyclase via G_s diminishes greatly. This diminution of signal transduction is generally called *desensitization* and is mediated, at least in part, by regulation of the receptor itself. A process of *rapid desensitization* was so named because it occurs within seconds to minutes after agonist-induced activation. Rapid desensitization of the B2AR can be reversed within several minutes after removal of agonist in a process called *resensitization*. Rapid desensitization of the B2AR is not associated with a decrease in the total number of receptors present in cells or tissues, and resensitization does not require biosynthesis of new receptor protein. Therefore, rapid desensitization is thought to reflect a change in the *functional activity*, rather than absolute number, of receptors (Fig. 5.1B).

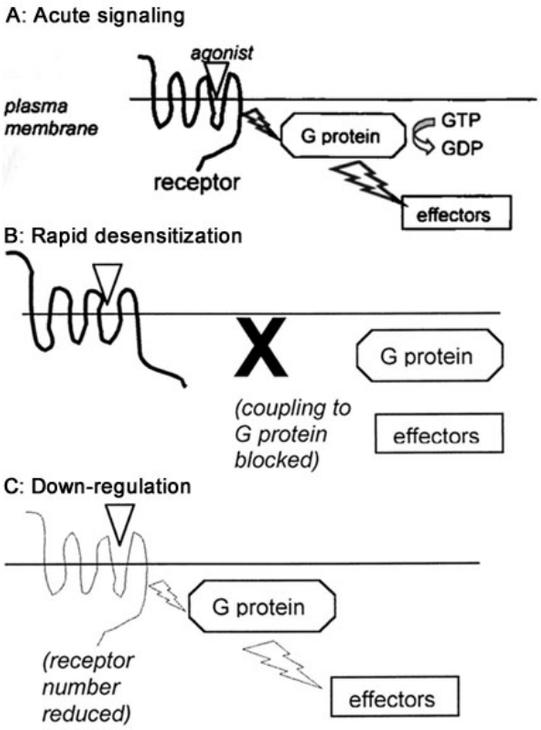


FIGURE 5.1. Desensitization and down-regulation of G protein-coupled receptors. **Panel A:** Within milliseconds to seconds after agonist binding, receptors present in the plasma membrane mediate signal transduction to effectors by functionally coupling (promoting guanine nucleotide exchange on) a heterotrimeric G protein. **Panel B:** Within several minutes after agonist binding, rapid desensitization occurs by functional uncoupling of the receptor from G protein. This represents a change in the functional activity of receptors, which inhibits signal transduction to the effector without changing the number of receptors in the cell. **Panel C:** After more prolonged agonist-induced activation of receptors (typically several hours to days), the number of receptors present in cells is greatly reduced, so signal transduction via G proteins to effectors is strongly attenuated. This process is called *receptor down-regulation* because it is thought to represent primarily a reduction in the number, rather than functional activity, of receptors.

Down-regulation and Up-regulation of Receptors

The kinetics of rapid desensitization and resensitization may be relevant to the physiologic action of catecholamines (endogenous agonist ligands for the B2AR), as these molecules can be released intermittently by vesicular exocytosis and are rapidly removed from the extracellular milieu by membrane transport, enzymatic degradation, or both. However, many clinically important drugs that activate GPCRs have a more prolonged duration of action. Studies of these drugs established the existence of a distinct process of receptor regulation that occurs much more slowly, typically within several hours to days after prolonged or repeated exposure of tissues to ligand. This process is called *down-regulation* because (in contrast to rapid desensitization) it is associated with a pronounced decrease in the total number of receptors present in cells or tissues, as typically detected by means of radioligand binding techniques. Further distinguishing the process of down-regulation from rapid desensitization, recovery of signaling activity after down-regulation is generally a slow process that requires biosynthesis of new receptor protein (8). Therefore, down-regulation is thought to reflect primarily a change in the *number*, rather than functional activity, of receptors present in cells or tissues (Fig. 5.1C). In most cases, down-regulation of GPCRs (like rapid desensitization) is induced by *agonists*, but not by antagonists. Moreover, certain antagonists can induce an opposite process of increased receptor number called *up-regulation* (9 ,10). These observations are consistent with a fundamental role of down-regulation and up-regulation as a negative feedback mechanism that seeks to maintain physiologic homeostasis of receptor signaling. However, in some cases, processes associated with GPCR down-regulation may be induced by antagonists (11 ,12). In other cases, up-regulation of receptors can be induced by drugs with partial agonist activity (10). These observations suggest that certain clinically relevant drugs may not simply mimic or block the effects of endogenous agonists. Indeed, it is proposed that such “paradoxical” regulatory effects may contribute to the pathologic or therapeutic actions of certain clinically relevant

neuropsychiatric drugs, including opiate analgesics and atypical antipsychotic agents (10 ,13 ,14 ,15 and 16).

Distinguishable Processes of Homologous and Heterologous Desensitization

Another important observation leading to our present view of GPCR regulatory mechanisms has come from studies investigating the pharmacologic specificity of receptor regulation (1 ,17). Many cell types express multiple types of GPCR (Fig. 5.2A). It has been observed that in some cases, prolonged activation of one type of GPCR causes attenuated signal transduction only by that receptor, without any detectable effect on signaling by other types of GPCR present in the same cell. In this case, the regulation of receptors is said to be *homologous* (Fig. 5.2B). The existence of homologous processes of regulation provided early evidence, before specific regulatory mechanisms were elucidated, that signal transduction can be modulated by modification of the receptor itself. In other cases, activation of one type of GPCR attenuates signaling not only by that receptor but also by other type(s) of GPCR present in the same cell. Regulation of this kind is said to be *heterologous* (Fig. 5.2C). Heterologous regulation of receptor signaling is consistent with modification of a "downstream" component in the signal transduction pathway that is involved in signaling by more than one type of GPCR. However, as discussed below, important examples of heterologous regulation mediated by modification of the receptor protein itself also exist. Homologous processes of receptor regulation are capable of modulating signal transduction in a highly specific manner, whereas heterologous processes of regulation may play an important role in facilitating functional "cross-talk" between pharmacologically distinct signaling pathways (18).

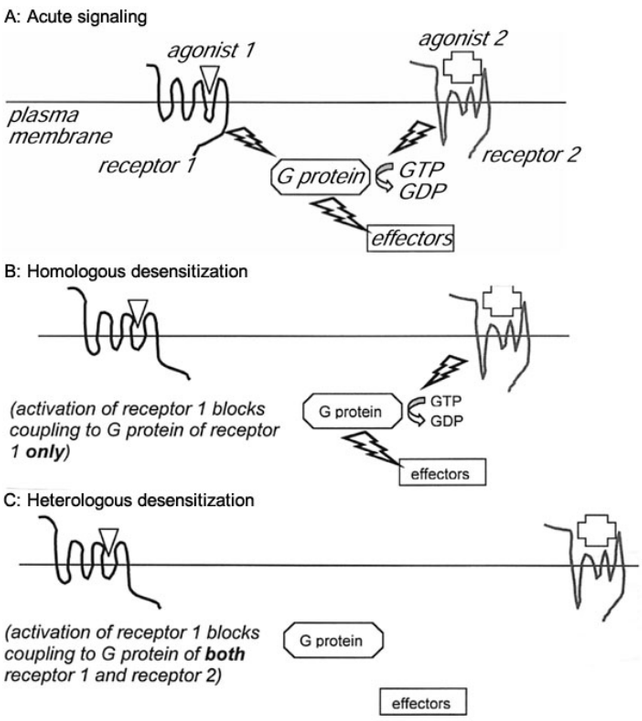


FIGURE 5.2. Homologous and heterologous desensitization. **Panel A:** Receptors 1 and 2 can signal via the same G protein-mediated pathway. **Panel B:** Agonist 1 causes uncoupling only of receptor 1; signaling induced by agonist 2 binding to receptor 2 is not affected. This is homologous desensitization of receptor 1. **Panel C:** Agonist 1 causes functional uncoupling of both receptor 1 and receptor 2, so that signaling induced by both agonists is blocked. This is heterologous desensitization.

SPECIFIC MECHANISMS OF GPCR REGULATION

Part of "5 - Regulation of G Protein-Coupled Receptors by Phosphorylation and Endocytosis "

Rapid Desensitization Mediated by Phosphorylation-Dependent Uncoupling of Receptor from Heterotrimeric G Protein

It is well established that many GPCRs are regulated by phosphorylation. Classic studies of rhodopsin (a light-activated GPCR) and the B2AR (a ligand-activated GPCR) provide examples of illustrating distinct molecular mechanisms that mediate homologous and heterologous desensitization of receptors. Because the principles behind these mechanisms have proven to be widely applicable to other GPCRs, rhodopsin, and the B2AR are often considered "prototypic" GPCRs that have established general paradigms for understanding phosphorylation-dependent regulation of GPCRs (19).

Phosphorylation of Rhodopsin: a Model for Functional Inactivation of GPCRs

Elegant studies of the vertebrate visual system identified a critical role of phosphorylation in inactivating rhodopsin following light-induced activation (20). Light-activated rhodopsin is a good substrate for phosphorylation by a cytoplasmic enzyme called *rhodopsin kinase*, whereas rhodopsin that has not been activated by light is a poor substrate (20). Phosphorylation of the carboxyl-terminal cytoplasmic domain of rhodopsin is sufficient to attenuate the ability of light-activated rhodopsin to couple functionally to its cognate heterotrimeric G protein (transducin). Studies of rhodopsin function in isolated membrane fractions indicated that rhodopsin kinase-mediated phosphorylation can strongly attenuate rhodopsin signaling, but it does so to a lesser extent than observed in the intact rod cell. A second cytoplasmic protein was identified that, when added to membrane preparations in combination with rhodopsin kinase, greatly accelerates the attenuation of rhodopsin signaling (21). This protein was proposed to act as a protein

cofactor that “arrests” signal transduction by phosphorylated rhodopsin and was therefore called *arrestin* (Fig. 5.3A).

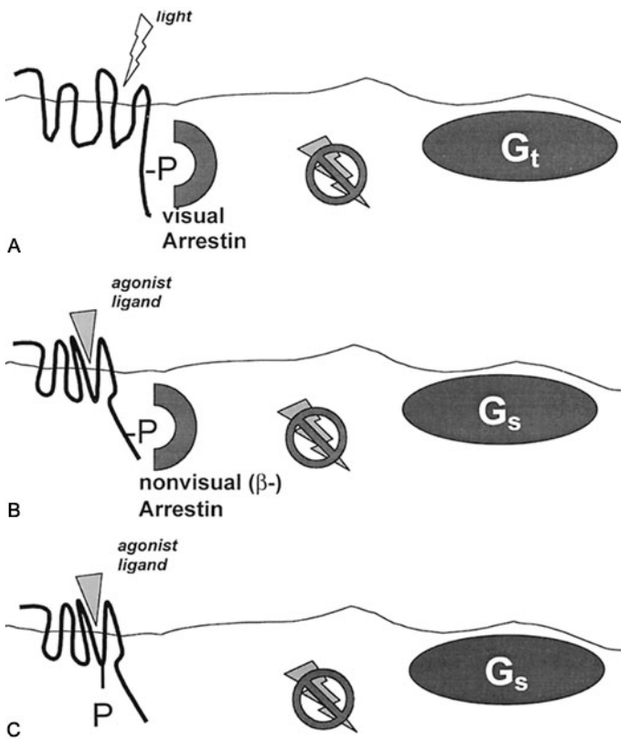


FIGURE 5.3. Paradigms for phosphorylation-dependent desensitization. **A:** GRK1-mediated inactivation of rhodopsin. **B:** GRK2-mediated desensitization of the B2AR. **C:** PKA-mediated desensitization of the B2AR.

GRK-Mediated Phosphorylation of the B2AR: a Model for Homologous Desensitization

Studies of functional reconstitution of B2AR-mediated activation of adenylyl cyclase provided compelling evidence for a role of phosphorylation in mediating rapid desensitization of a ligand-activated GPCR (22). Biochemical purification of the cytoplasmic activity responsible for this phosphorylation identified a kinase called *β-adrenergic receptor kinase* (BARK), and it was noted that this kinase preferentially phosphorylates agonist-occupied receptors (23). Protein sequencing and cyclic DNA (cDNA) cloning later indicated that this kinase is closely similar in structure to rhodopsin kinase (24). Furthermore, cDNA-cloning experiments in which reduced stringency hybridization was used identified additional kinases homologous to rhodopsin kinase and BARK (19). Collectively, these observations led to the discovery of the family of *G protein-coupled receptor kinases* (GRKs). According to this nomenclature, rhodopsin is denoted GRK-1, the originally identified BARK enzyme is denoted GRK-2, and other members of this family of protein kinases are numbered sequentially thereafter. Six members of the GRK family of receptor kinases have been identified to date.

Biochemical reconstitution studies indicated that increasingly purified fractions of BARK exhibit a reduced ability to attenuate B2AR-mediated signal transduction. Further analysis of this effect led to the identification of a distinct protein component that copurifies with BARK in initial stages of purification but is resolved from BARK in more highly purified fractions. This protein component reconstitutes strong attenuation of B2AR-mediated activation of adenylyl cyclase when added back to highly purified fractions of BARK (25 ,26). The protein cofactor involved in desensitization of the B2AR turned out to be a protein similar to visual arrestin and was therefore named *β-arrestin* or *barrestin*. cDNA cloning has identified additional proteins with similar structure, thus defining a family of protein cofactors for phosphorylation-dependent regulation of GPCR function (4). Two nomenclatures are currently in common use for these molecules. In one, the originally identified ß-arrestin is denoted barrestin 1, and additional homologues are named sequentially barrestin 2, and so on. In another nomenclature, all members of this protein family are referred to as *arrestins*, with visual arrestin denoted arrestin 1, ß-arrestin as arrestin 2, and subsequently identified family members numbered sequentially thereafter. Four members of the arrestin family of protein cofactors have been identified to date.

As noted above, an important feature of many GRKs is that their kinase activity is highly sensitive to the conformation of the receptor that they phosphorylate. This property of GRKs facilitates specific phosphorylation of only those receptors that are activated by ligand, whereas other receptors present in the same cells (but not activated by agonist) are not phosphorylated. Thus, GRK-mediated phosphorylation is generally considered to be a paradigm for *homologous* desensitization (Fig. 5.3B).

Protein Kinase A-Mediated Phosphorylation of the B2AR: a Model of Heterologous Desensitization

Other kinases, such as the so-called second messenger-regulated kinases, are also implicated in mediating desensitization of GPCRs. For example, the B2AR can be phosphorylated by cyclic adenosine monophosphate (cAMP)-dependent protein kinase (PKA). PKA-mediated phosphorylation of a single residue located in the third intracellular loop of the B2AR impairs the ability of the receptor to couple to G_s and subsequently activate adenylyl cyclase (27 ,28 and 29). Phosphorylation of this residue is thought to impair receptor–G protein coupling directly, without a requirement for any known protein cofactor such as an arrestin. An important feature of PKA is that this kinase can phosphorylate B2ARs whether or not they have been activated

by ligand, in contrast to the preferential phosphorylation of agonist-activated receptors by GRK-2. Because PKA is activated by cAMP, a signaling intermediate produced as a result of B2AR activation, PKA-mediated phosphorylation of the B2AR can be viewed as an example of feedback inhibition by a second messenger. In addition, because activation of any other receptor that stimulates adenylyl cyclase can also activate PKA, phosphorylation of the B2AR by PKA is generally considered to be a paradigm for *heterologous* desensitization (Fig. 5.3C).

Rapid Desensitization and Resensitization Mediated by Regulated Endocytosis of Receptors

Endocytosis of GPCRs by Clathrin-Coated Pits

Studies conducted in cultured cells indicate that many GPCRs can be regulated by ligand-induced endocytosis. The B2AR is perhaps the most extensively studied GPCR with respect to endocytic membrane trafficking. Early evidence for agonist-induced endocytosis of the B2AR was suggested by observations from subcellular fractionation and radioligand binding assays conducted with membrane-impermeant antagonists (30 ,31). These studies indicated that the number of B2AR binding sites detected in the plasma membrane can be reduced within several minutes after agonist-induced activation, a process called *sequestration*.

The development of receptor-specific antibodies allowed the application of immunocytochemical methods to visualize the subcellular localization of the B2AR and directly demonstrate agonist-induced internalization of the receptor protein. Internalization of the B2AR was observed to represent a steady state of a highly dynamic process involving continuous endocytosis and recycling of receptors through an endocytic pathway similar to that mediating constitutive (ligand-independent) endocytosis of nutrient receptors (32). This dynamic cycling of the B2AR was also suggested by elegant studies in which subcellular fractionation and radioligand binding techniques were used (33).

Regulation of B2AR endocytosis was shown to be mediated by a ligand-dependent lateral redistribution of receptors in the plasma membrane, from a relatively diffuse distribution throughout the plasma membrane to a pronounced concentration of agonist-activated receptors in structures resembling clathrin-coated pits when examined by immunoelectron microscopy. Furthermore, this process of ligand-regulated concentration of B2ARs in coated pits of the plasma membrane was shown to be mechanistically distinct from the subsequent endocytosis of receptors by membrane fission, which can occur even in the absence of continued ligand-induced activation of receptors (34). A protein that is required for this latter step of endocytic membrane fission is the cytoplasmic guanosine triphosphatase *dynamin* (35 ,36). Consistent with this, agonist-induced endocytosis of the B2AR is inhibited by overexpression of certain “dominant-negative” mutant forms of dynamin (37 ,38). Subsequent studies have demonstrated that regulated endocytosis of several other GPCRs is also mediated by a dynamin-dependent mechanism, which suggests a conserved role of clathrin-coated pits in the endocytosis of many GPCRs (Fig. 5.4).

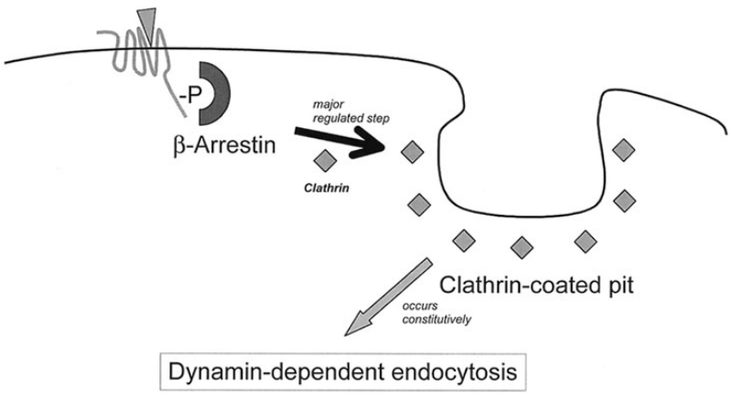


FIGURE 5.4. Regulated endocytosis of the β_2-adrenergic receptor (B2AR) by clathrin-coated pits. G protein-coupled receptor kinase-mediated phosphorylation of the B2AR promotes receptor interaction with nonvisual arrestins, which cause uncoupling of heterotrimeric G proteins and also promote interaction of arrestin–receptor complexes with clathrin coats. Once concentrated in clathrin-coated pits by this mechanism, receptors undergo endocytosis rapidly (even if agonist is removed from the receptor in the coated pit) via a constitutive (ligand-independent) mechanism of endocytic membrane fission that requires the cytoplasmic guanosine triphosphatase dynamin.

Role of GPCR Phosphorylation in Promoting Endocytosis by Clathrin-Coated Pits

Studies elucidating the mechanism mediating the key step of regulated concentration of GPCRs in clathrin-coated pits were initiated by the observation that GRK-mediated phosphorylation of the M2 muscarinic acetylcholine receptor can promote agonist-induced endocytosis of the receptor, whereas a kinase-defective mutant form of GRK inhibits this process (39). Similar observations were made for other GPCRs, including the B2AR (37 ,40). In an elegant series of

experiments, it was shown that certain arrestins can directly promote B2AR concentration in clathrin-coated pits by physically linking phosphorylated receptors with clathrin (40). This endocytic “adapter” function of arrestins can be distinguished from the function of arrestins as cofactors for functional uncoupling of receptor from G protein because the latter process can occur in the absence of endocytosis. Further distinguishing these functions, visual arrestin (arrestin 1) was shown to be unable to serve as an adapter for B2AR endocytosis even though it can serve as a cofactor for desensitization mediated by functional uncoupling of G protein (41). This distinction between visual and nonvisual arrestins led to the identification of a carboxyl-terminal clathrin-binding domain, present specifically in nonvisual arrestins, that is necessary for endocytosis of GPCRs but not for phosphorylation-dependent uncoupling of receptors from heterotrimeric G proteins (42 ,43).

Functional Role of Endocytosis in the Processes of Rapid Desensitization and Resensitization of GPCRs

Physiologic ligands are generally thought to bind to GPCRs in the plasma membrane. Biochemical and immunocytochemical studies suggest that certain GPCRs, such as the B2AR, interact with heterotrimeric G proteins primarily in the plasma membrane but not in intracellular membranes after endocytosis (44 ,45). Together, these observations suggest that endocytosis may, by itself, mediate *desensitization* of GPCR-mediated signal transduction by directly reducing the number of receptors present in the plasma membrane. Indeed, in some cases, endocytosis may be a principal mechanism of rapid desensitization (46). However, as discussed above, rapid desensitization by phosphorylation-dependent uncoupling of receptor from G protein does not require endocytosis of the receptor. This is consistent with the ability of certain GPCRs, such as the α_{2A}-adrenergic receptor, to desensitize in the absence of detectable endocytosis (47 ,48). Studies of μ-opioid receptors expressed in transfected cells suggest that the effectiveness of endocytosis as a means of attenuating signal transduction is inversely proportional to the number of “spare receptors” present in cells (46). Thus, the precise role of endocytosis in contributing to desensitization of GPCR-mediated signal transduction probably varies among systems and may be particularly important in cells expressing relatively low numbers of receptors.

Strong evidence is available to indicate that endocytosis of certain GPCRs serves a distinct function in promoting *resensitization*, rather than desensitization, of signal transduction. The most thoroughly studied example of this mechanism derives from elegant studies of the B2AR (2 ,49 ,50). As discussed above, agonist-induced phosphorylation of the B2AR by GRKs causes rapid desensitization by promoting receptor interaction with arrestins and functional uncoupling from heterotrimeric G proteins (Fig. 5.5 , step 1). This initial desensitization of receptors occurs in the plasma membrane and does not require endocytosis of the receptor protein. Within several minutes after this initial uncoupling of receptor from heterotrimeric G protein, arrestins promote the concentration of receptors in clathrin-coated pits and subsequent endocytosis (Fig. 5.5 , steps 2 and 3). Endocytic membranes containing internalized

B2ARs are associated with activity of a protein phosphatase (PP2A) that can catalyze dephosphorylation of receptors (51). Based on these observations, it is proposed that endocytosis of receptors promotes dephosphorylation of receptors, after which receptors can be recycled back to the plasma membrane in a dephosphorylated, fully active state (Fig. 5.5 , steps 4 and 5). Supporting this hypothesis, inhibitors of B2AR endocytosis do not block agonist-induced desensitization but strongly inhibit resensitization of receptor-mediated signal transduction following removal of agonist from the culture medium (52). Thus, agonist-induced regulation of the B2AR appears to involve two linked regulatory cycles: a biochemical cycle mediating phosphorylation and dephosphorylation of receptors, and a membrane trafficking cycle mediating endocytosis and recycling of receptors.

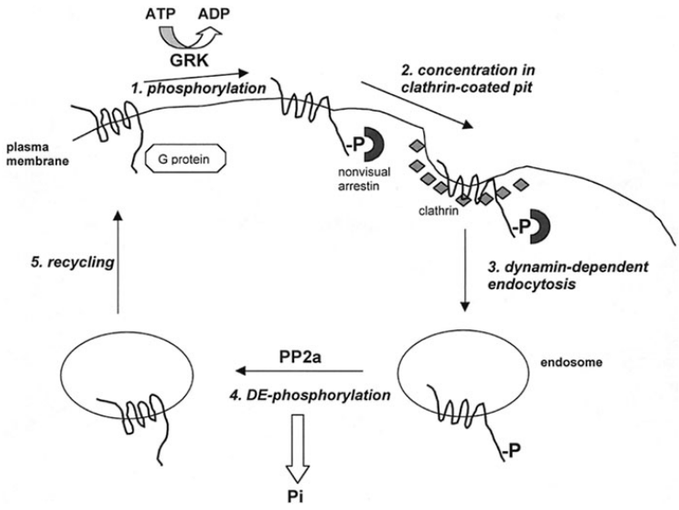


FIGURE 5.5. Linked cycles of G protein-coupled receptor phosphorylation and endocytosis mediating rapid desensitization and resensitization of the β_2-adrenergic receptor (B2AR). Agonist-induced activation of the B2AR causes G protein-coupled receptor kinase-mediated phosphorylation, which promotes receptor interaction with nonvisual arrestins and uncoupling of heterotrimeric G protein (step 1). Arrestin binding to the phosphorylated receptor also promotes receptor concentration in clathrin-coated pits (step 2), promoting rapid endocytosis of receptors by dynamin-dependent fission of coated pits from the plasma membrane and subsequent formation of endocytic vesicles (step 3). Together, these steps cause profound functional desenstization of signal transduction. Endocytic membranes containing internalized B2ARs are associated with a protein phosphatase (PP2A) that can catalyze dephosphorylation of receptors (step 4). Dephosphorylation of receptors followed by recycling to the plasma membrane (step 5) mediates the return of receptors to the plasma membrane in a fully functional state, promoting functional resensitization of the signal transduction system.

Down-regulation of GPCRs by Regulated Proteolysis

Evidence for Regulated Proteolysis of GPCRs

Down-regulation of GPCRs is defined from saturation binding analysis by a decrease in total specific binding sites (*B*max) without a change in apparent affinity (*K*d), which suggests that down-regulation reflects a decreased number of total receptors present cells or tissues (1 ,53). In principle, this process could be mediated by modulation of receptor biosynthesis or degradation. In practice, evidence suggests that both classes of mechanism contribute to the GPCR down-regulation observed physiologically. The role of transcriptional regulatory mechanisms in controlling biosynthesis have been characterized in some detail (see Chapter 17). As discussed in greater detail below, there is also considerable evidence supporting the importance of proteolysis of the receptor itself in mediating down regulation of a number of GPCRs (8).

Mechanisms of GPCR Proteolysis

Multiple mechanisms can mediate GPCR proteolysis. Proteolysis of endocytosed receptors in lysosomes is perhaps the best-established mechanism contributing to GPCR down-regulation in mammalian cells, including certain neurally derived cell types (54 ,55). However, it is apparent that other mechanisms of GPCR proteolysis also exist, some of which may not involve membrane trafficking of receptors at all. For example, the V2 vasopressin receptor can undergo ligand-induced proteolysis at the cell surface by a nonendocytic mechanism mediated by a plasma membrane-associated metalloprotease (56). Recent studies of B2AR down-regulation support the idea that regulated proteolysis of GPCRs can occur without endocytosis in some cell types (57).

In mammalian cells, ubiquitination is well established to promote degradation of various cytoplasmic proteins by a nonlysosomal mechanism mediated by proteasomes (58). Emerging evidence also suggests a role of ubiquitination in promoting endocytosis and proteolytic degradation of certain membrane proteins, including GPCRs in yeast (59 ,60). The role of such a mechanism in mediating down-regulation of mammalian signaling receptors comes from studies of receptor tyrosine kinases (61). Alternate mechanisms of GPCR proteolysis in mammalian cells have been reported to be mediated by a distinct, nonproteasomal mechanism (56) or have been shown to be insensitive to inhibitors of proteasome-mediated proteolysis (57). Thus, to our knowledge, it is not yet clear to what extent ubiquitination or proteasomes may contribute to down-regulation of GPCRs in mammalian cells.

Membrane Pathway Mediating Receptor Delivery to Lysosomes

The delivery of membrane proteins from the plasma membrane to lysosomes is a multiple-step process that is mediated by endocytosis of receptors from the plasma membrane followed by shuttling to lysosomes via a specific series of membrane transport reactions (62 ,63). It is well-known that many GPCRs undergo ligand-induced endocytosis. However, specific mechanisms and pathways mediating subsequent stages of receptor trafficking to lysosomes are poorly understood.

Early studies demonstrated that ligand-induced sequestration and down-regulation of the B2AR can be differentially affected by pharmacologic manipulations and receptor mutation, which suggests that these processes may be mediated by distinct mechanisms (64 ,65 ,66 and 67). Furthermore, naturally occurring subtypes of α_2-adrenergic receptor down-regulate with similar rates (68) despite significant differences in their ability to undergo rapid endocytosis (47 ,48). Analogous processes of rapid sequestration and more gradual down-regulation have also been observed in studies of opioid receptors, where pharmacologic differences between the effects of individual agonists are very pronounced (69 ,70 and 71) and appear to be relevant to the physiologic effects of opiate drugs in native neurons (14 ,72). Compelling evidence for the existence of distinguishable membrane trafficking mechanisms comes from recent studies of mutant thrombin and substance P receptors, in which divergent residues in the carboxyl-terminal cytoplasmic domain specify differences in receptor trafficking between lysosomal and recycling pathways (73). Analyses based on kinetic modeling techniques are consistent either with completely separate pathways mediating rapid endocytosis and proteolytic degradation of GPCRs or with the operation of partially overlapping pathways that differ in their rate-limiting step (53). These models differ in whether sorting of GPCRs is proposed to occur before or after endocytosis (Fig. 5.6).

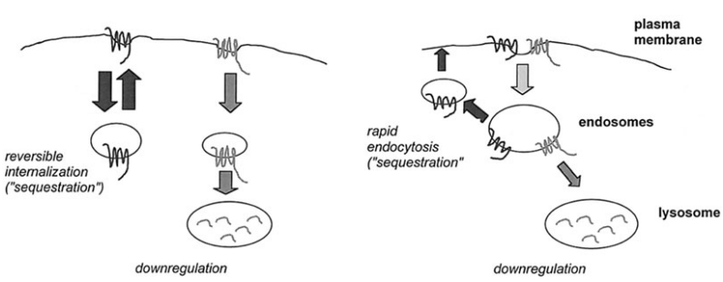


FIGURE 5.6. Models for sorting of G protein-coupled receptors (GPCRs) between pathways mediating reversible internalization ("sequestration") and proteolytic down-regulation of GPCRs. **Panel A:** This model proposes that distinct GPCRs are segregated in the plasma membrane and are subsequently endocytosed by different membrane vesicles that either mediate reversible internalization or deliver receptors to lysosomes. **Panel B:** This model proposes that receptors can be endocytosed by the same membrane mechanism and delivered to the same early endocytic vesicles. Receptors are sorted after endocytosis between distinct pathways that mediate either recycling or lysosomal delivery of receptors.

Molecular Sorting of GPCRs after Endocytosis

Extensive previous studies of receptor-mediated uptake of extracellular ligands indicate that sorting of receptors between recycling and lysosomal pathways can occur after endocytosis (74). Recent studies of adrenergic and opioid receptors suggest that this is also true for certain GPCRs. For example, it has been shown recently that both agonist-induced sequestration and down-regulation of the B2AR are specifically inhibited by a dominant-negative mutant form of dynamin, which suggests that endocytosis of receptors by clathrin-coated pits is an obligate first step common to membrane pathways leading to recycling of endosomes and lysosomes (75). Recent studies of membrane trafficking of opioid receptors, which are also endocytosed by a dynamin-dependent mechanism involving clathrin-coated pits (76), support this conclusion. These studies demonstrate that distinct GPCRs can be sorted in a type-specific manner between recycling and degradative pathways after endocytosis by the same mechanism (Fig. 5.6B). These studies also suggest that the machinery that sorts GPCRs can function very rapidly (within several minutes) after the initial endocytosis of receptors (77).

Insight into a Mechanism Controlling Endocytic Sorting of the B2AR

The membrane machinery controlling postendocytic sorting of internalized GPCRs to lysosomes remains poorly understood. A recent study of B2AR trafficking has provided initial insight into a specialized mechanism that controls sorting of a specific GPCR between recycling and lysosomal pathways. As noted above, the B2AR is capable of recycling rapidly to the plasma membrane following endocytosis in cultured cells. It was observed that a series of mutations in the distal part of the carboxyl-terminal cytoplasmic tail strongly inhibit recycling of receptors and cause receptors to be “mistargeted” to lysosomes (78). All the receptor mutations that cause this phenotype disrupt a specific interaction of the B2AR with NHERF/EBP50/E3KARP-family proteins (78 ,79 and 80), a previously described

family of proteins that interact with the B2AR via PDZ (PSD95/Discs large/ZO-1-homologous) domains and also associate with the cortical actin cytoskeleton (81). Overexpression of a mutant form of EBP50/NHERF that cannot interact with the cortical actin cytoskeleton, or chemical disruption of the actin cytoskeleton itself, also inhibits recycling of internalized B2ARs. These observations suggest that sorting of internalized B2ARs between a distinct recycling and degradative pathway can be mediated by a protein complex associated with the cortical actin cytoskeleton. Moreover, these studies suggest that phosphorylation of a specific serine residue (Ser411), a potential substrate for a subset of GRKs (82), can modulate the sorting of receptors by this mechanism (78). Thus, sorting of receptors after endocytosis, like the initial endocytosis of receptors from the plasma membrane, may be closely linked to the cycle of receptor phosphorylation and dephosphorylation involved in mediating desensitization and resensitization of signal transduction. A proposed model summarizing our current understanding of distinct stages of B2AR endocytosis and sorting is summarized in Fig. 5.7 .

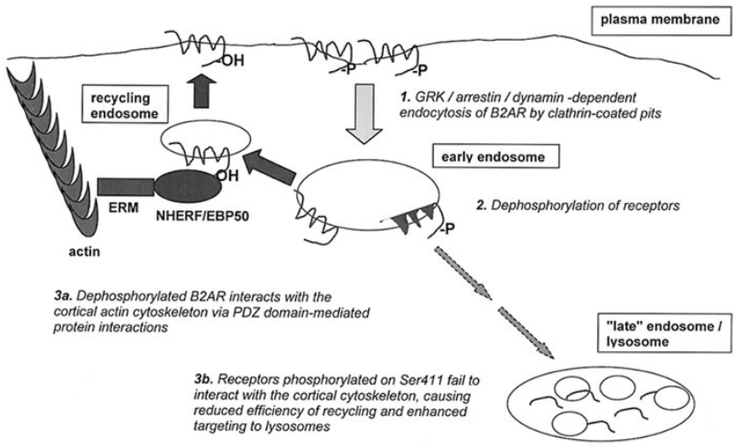


FIGURE 5.7. Model for endocytic sorting of the $ß_2$-adrenergic receptor (B2AR) by phosphorylation-regulated association of receptors with the cortical actin cytoskeleton. B2ARs undergo rapid desensitization and endocytosis (step 1). If receptors are dephosphorylated (step 2), they are able to interact with a cytoskeleton-associated protein complex via a PDZ domain-mediated protein interaction with the carboxyl-terminal cytoplasmic domain of the receptor, which promotes rapid recycling of receptors to the plasma membrane (step 3a). If serine 411 in the receptor tail remains phosphorylated, this protein interaction does not occur, and recycling of receptors is inhibited. Receptors that fail to recycle can later be delivered to lysosomes, with down-regulation of receptors accomplished by proteolysis (step 3b).

CONCLUSIONS AND FUTURE PERSPECTIVES

Part of "5 - Regulation of G Protein-Coupled Receptors by Phosphorylation and Endocytosis "

In this chapter, we have surveyed general processes of GPCR regulation and focused on selected mechanisms that are understood in molecular detail. A general principle emerging from this approach is that phosphorylation and membrane trafficking mechanisms are of fundamental importance to multiple processes of GPCR regulation. A second important principle is that distinct mechanisms of regulatory phosphorylation and membrane trafficking are closely interdependent, which leads to an appreciation of linked cycles coordinating the functional activation and regulation of receptors. An interesting corollary of this principle, suggested by certain recent studies of atypical antipsychotics and opiate analgesics (15), is that the linkage between specific mechanisms of receptor activation and regulation may be modified or disrupted by certain drugs. Therefore, it may be possible to target specific mechanisms of GPCR regulation, or the biochemical linkages between these mechanisms, to develop novel therapeutics that may influence the regulation of GPCRs in a manner quite different from that of classic agonists or antagonists.

We have restricted our focus to a subset of mechanisms that are, arguably, among the best established experimentally. It is important to note that these studies are in large part limited to model cell systems rather than native tissues. Nevertheless, the available data derived from transgenic and homologous recombination methodologies, and studies based on immunocytochemical localization of endogenously expressed GPCRs, suggest that phosphorylation (83 ,84) and endocytosis of receptors do occur in native tissues in response to physiologic stimulation (85 ,86) and certain drugs (14 ,72). Thus, it is reasonable to propose that mechanisms elucidated with model cell systems may indeed be relevant to receptor regulation *in vivo*. However, it is not yet possible to predict precisely how specific mechanism(s) of regulation modulate the central nervous system under physiologic or pathophysiologic conditions, or how specific mechanisms of regulation may contribute to the *in vivo* effects of clinically important drugs. Thus, an important future goal is to examine the regulation of GPCRs in native cell types and tissues, and to determine how specific mechanisms of regulation contribute to physiologic and pathophysiologic states.

Finally, it is important to note that by restricting our scope to a limited subset of GPCR regulatory mechanisms, we have underrepresented the diversity of GPCR regulation and the high degree of specificity with which individual receptors are regulated in various cell types. For example, although we have discussed only endocytosis of GPCRs by clathrin-coated pits, substantial evidence indicates that other mechanism(s) also can mediate regulated endocytosis of receptors (37 ,87 ,88 and 89). Moreover, it is increasingly apparent that mechanisms of GPCR phosphorylation and endocytosis discussed in this chapter serve additional important functions in signal transduction, such as controlling the interaction of “classic” GPCR signaling pathways with mitogenic kinase cascades (see also Chapter 16 and Chapter 22) (90 ,91 and 92). Indeed, we anticipate that our present understanding of the array of GPCR signaling and regulatory mechanisms is at a relatively early stage of rapid development, a view consistent with the results of recent studies identifying an unexpected diversity of cellular proteins that interact with specific GPCRs (93). Future studies of these unexplored mechanisms of GPCR regulation may lead to important new insights relevant to neuropsychiatric disease and may identify exciting new targets for the development of novel therapeutic drugs.

REFERENCES

1. Clark RB. Receptor desensitization. *Adv Cyclic Nucl Prot Phos Res* 1986;20:151–209.

2. Ferguson SS, Zhang J, Barak LS, et al. Molecular mechanisms of G protein-coupled receptor desensitization and resensitization. *Life Sci* 1998;62:1561–1565.

3. Lefkowitz RJ, Pitcher J, Krueger K, et al. Mechanisms of beta-adrenergic receptor desensitization and resensitization. *Adv Pharmacol* 1998;42:416–420.

4. Carman CV, Benovic JL. G-protein-coupled receptors: turn-ons and turn-offs. *Curr Opin Neurobiol* 1998;8:335–344.

5. Lefkowitz RJ. G protein-coupled receptors and receptor kinases: from molecular biology to potential therapeutic applications. *Nat Biotechnol* 1996;14:283–286.

6. Roth BL, Willins DL, Kroeze WK. G protein-coupled receptor (GPCR) trafficking in the central nervous system: relevance for drugs of abuse. *Drug Alcohol Depend* 1998;51:73–85.

7. Perkins JP, Hausdorff WP, Lefkowitz RJ. Mechanisms of ligand-induced desensitization of beta-adrenergic receptors. In: Perkins JP, ed. *The beta-adrenergic receptor*. Clifton, NJ: Humana Press, 1991:73–124.

8. Doss RC, Perkins JP, Harden TK. Recovery of beta-adrenergic receptors following long-term exposure of astrocytoma cells to catecholamine: role of protein synthesis. *J Biol Chem* 1981;256:12281–12286.

9. Ng GY, Varghese G, Chung HT, et al. Resistance of the dopamine D2L receptor to desensitization accompanies the up-regulation of receptors on to the surface of Sf9 cells. *Endocrinology* 1997;138:4199–4206.

10. Zaki PA, Keith DE, Brine GA, et al. Ligand-induced changes in surface mu-opioid receptor number: relationship to G protein activation? *J Pharmacol Exp Ther* 2000;292:1127–1134.

11. Pfeiffer R, Kirsch J, Fahrenholz F. Agonist and antagonist-dependent internalization of the human vasopressin V2 receptor. *Exp Cell Res* 1998;244:327–339.

12. Willins DL, Berry SA, Alsayegh L, et al. Clozapine and other 5-hydroxytryptamine-2A receptor antagonists alter the subcellular distribution of 5-hydroxytryptamine-2A receptors *in vitro* and *in vivo*. *Neuroscience* 1999;91:599–606.

13. von Zastrow M, Keith DE, Zaki P, et al. Morphine and opioid peptide cause opposing effects on the endocytic trafficking of opioid receptors. *Regul Pept* 1994;54:315.

14. Keith DE, Anton B, Murray SR, et al. Mu-opioid receptor internalization: opiate drugs have differential effects on a conserved endocytic mechanism *in vitro* and in the mammalian brain. *Mol Pharmacol* 1998;53:377–384.

15. Roth BL, Willins DL. What's all the RAVE about receptor internalization? *Neuron* 1999;23:629–631.

16. Whistler JL, Chuang HH, Chu P, et al. Functional dissociation of mu opioid receptor signaling and endocytosis: implications for the biology of opiate tolerance and addiction. *Neuron* 1999;23:737–746.

17. Chuang TT, Iacovelli L, Sallese M, et al. G protein-coupled receptors: heterologous regulation of homologous desensitization and its implications. *Trends Pharmacol Sci* 1996;17:416–421.

18. Bouvier M. Cross-talk between second messengers. *Ann N Y Acad Sci* 1990;594:120–129.

19. Krupnick JG, Benovic JL. The role of receptor kinases and arrestins in G protein-coupled receptor regulation. *Annu Rev Pharmacol Toxicol* 1998;38:289–319.

20. McDowell JH, Kuhn H. Light-induced phosphorylation of rhodopsin in cattle photoreceptor membranes: substrate activation and inactivation. *Biochemistry* 1977;16:4054–4060.

21. Bennett N, Sitaramayya A. Inactivation of photoexcited rhodopsin in retinal rods: the roles of rhodopsin kinase and 48-kDa protein (arrestin). *Biochemistry* 1988;27:1710–1715.

22. Sibley DR, Strasser RH, Caron MG, et al. Homologous desensitization of adenylate cyclase is associated with phosphorylation of the beta-adrenergic receptor. *J Biol Chem* 1985;260:3883–3886.

23. Benovic JL, Strasser RH, Caron MG, et al. Beta-adrenergic receptor kinase: identification of a novel protein kinase that phosphorylates the agonist-occupied form of the receptor. *Proc Natl Acad Sci U S A* 1986;83:2797–2801.

24. Benovic JL, Stone WC, Huebner K, et al. cDNA cloning and chromosomal localization of the human beta-adrenergic receptor kinase. *FEBS Lett* 1991;283:122–126.

25. Lohse MJ, Benovic JL, Codina J, et al. Beta-arrestin: a protein that regulates beta-adrenergic receptor function. *Science* 1990;248:1547–1550.

26. Benovic JL, Kuhn H, Weyand I, et al. Functional desensitization of the isolated beta-adrenergic receptor by the beta-adrenergic receptor kinase: potential role of an analog of the retinal protein arrestin (48-kDa protein). *Proc Natl Acad Sci U S A* 1987;84:8879–8882.

27. Hausdorff WP, Lohse MJ, Bouvier M, et al. Two kinases mediate agonist-dependent phosphorylation and desensitization of the beta 2-adrenergic receptor. *Symp Soc Exp Biol* 1990;44:225–240.

28. Bouvier M, Hausdorff WP, De Blasi A, et al. Removal of phosphorylation sites from the beta 2-adrenergic receptor delays onset of agonist-promoted desensitization. *Nature* 1988;333:370–373.

29. Benovic JL, Bouvier M, Caron MG, et al. Regulation of adenylyl cyclase-coupled beta-adrenergic receptors. *Annu Rev Cell Biol* 1988;4:405–428.

30. Staehelin M, Simons P. Rapid and reversible disappearance of beta-adrenergic cell surface receptors. *EMBO J* 1982;1:187–190.

31. Toews ML, Perkins JP. Agonist-induced changes in beta-adrenergic receptors on intact cells. *J Biol Chem* 1984;259:2227–2235.

32. von Zastrow M, Kobilka BK. Ligand-regulated internalization and recycling of human beta 2-adrenergic receptors between the plasma membrane and endosomes containing transferrin receptors. *J Biol Chem* 1992;267:3530–3538.

33. Kurz JB, Perkins JP. Isoproterenol-initiated beta-adrenergic receptor diacytosis in cultured cells. *Mol Pharmacol* 1992;41:375–381.

34. von Zastrow M, Kobilka BK. Antagonist-dependent and -independent steps in the mechanism of adrenergic receptor internalization. *J Biol Chem* 1994;269:18448–18452.

35. van der Bliek AM, Redelmeier TE, Damke H, et al. Mutations in human dynamin block an intermediate stage in coated vesicle formation. *J Cell Biol* 1993;122:553–563.

36. Herskovits JS, Burgess CC, Obar RA, et al. Effects of mutant rat dynamin on endocytosis. *J Cell Biol* 1993;122:565–578.

37. Zhang J, Ferguson S, Barak LS, et al. Dynamin and beta-arrestin reveal distinct mechanisms for G protein-coupled receptor internalization. *J Biol Chem* 1996;271:18302–18305.

38. Cao TC, Mays RW, von Zastrow M. Regulated endocytosis of G protein-coupled receptors by a biochemically and functionally distinct subpopulation of clathrin-coated pits. *J Biol Chem* 1998;273:24592–24602.

39. Tsuga H, Kameyama K, Haga T, et al. Sequestration of muscarinic acetylcholine receptor m2 subtypes. Facilitation by G protein-coupled receptor kinase (GRK2) and attenuation by a dominant-negative mutant of GRK2. *J Biol Chem* 1994;269:32522–32527.

40. Goodman OJ, Krupnick JG, Santini F, et al. Beta-arrestin acts as a clathrin adaptor in endocytosis of the beta2-adrenergic receptor. *Nature* 1996;383:447–450.

41. Goodman OB Jr, Krupnick JG, Santini F, et al. Role of arrestins in G-protein-coupled receptor endocytosis. *Adv Pharmacol* 1998;42:429–433.

42. Krupnick JG, Santini F, Gagnon AW, et al. Modulation of the arrestin-clathrin interaction in cells. Characterization of beta-arrestin dominant-negative mutants. *J Biol Chem* 1997;272:32507–32512.

43. Krupnick JG, Goodman OJ, Keen JH, et al. Arrestin/clathrin interaction. Localization of the clathrin-binding domain of nonvisual arrestins to the carboxy terminus. *J Biol Chem* 1997;272:15011–15016.

44. Wedegaertner PB, Bourne HR, von Zastrow M. Activationinduced subcellular redistribution of Gs alpha. *Mol Biol Cell* 1996;7:1225–1233.

45. Ransnas LA, Jasper JR, Leiber D, et al. Beta-adrenergic-receptor-mediated dissociation and membrane release of the Gs protein in S49 lymphoma-cell membranes. Dependence on Mg^{2+} and GTP. *Biochem J* 1992;281:519–524.

46. Pak Y, Kouvelas A, Scheideler MA, et al. Agonist-induced functional desensitization of the mu-opioid receptor is mediated by loss of membrane receptors rather than uncoupling from G protein. *Mol Pharmacol* 1996;50:1214–1222.

47. von Zastrow M, Link R, Daunt D, et al. Subtype-specific differences in the intracellular sorting of G protein-coupled receptors. *J Biol Chem* 1993;268:763–766.

48. Daunt DA, Hurt C, Hein L, et al. Subtype-specific intracellular trafficking of alpha2-adrenergic receptors. *Mol Pharmacol* 1997;51:711–720.

49. Pippig S, Andexinger S, Daniel K, et al. Overexpression of beta-arrestin and beta-adrenergic receptor kinase augment desensitization of beta 2-adrenergic receptors. *J Biol Chem* 1993;268:3201–3208.

50. Yu SS, Lefkowitz RJ, Hausdorff WP. Beta-adrenergic receptor sequestration. A potential mechanism of receptor resensitization. *J Biol Chem* 1993;268:337–341.

51. Pitcher JA, Payne ES, Csortos C, et al. The G-protein-coupled receptor phosphatase: a protein phosphatase type 2A with a distinct subcellular distribution and substrate specificity. *Proc Natl Acad Sci U S A* 1995;92:8343–8347.

52. Pippig S, Andexinger S, Lohse MJ. Sequestration and recycling of beta 2-adrenergic receptors permit receptor resensitization. *Mol Pharmacol* 1995;47:666–676.

53. Koenig JA, Edwardson JM. Endocytosis and recycling of G protein-coupled receptors. *Trends Pharmacol Sci* 1997;18:276–287.

54. Law P-Y, Hom DS, Loh HH. Down-regulation of opiate receptor in neuroblastoma x glioma NG108-15 hybrid cells: chloroquine promotes accumulation of tritiated enkephalin in the lysosomes. *J Biol Chem* 1984;259:4096–4104.

55. Ko JL, Arvidsson U, Williams FG, et al. Visualization of time-dependent redistribution of delta-opioid receptors in neuronal cells during prolonged agonist exposure. *Brain Res Mol Brain Res* 1999;69:171–185.

56. Kojro E, Fahrenholz F. Ligand-induced cleavage of the V2 vasopressin receptor by a plasma membrane metalloproteinase. *J Biol Chem* 1995;270:6476–6481.

57. Jockers R, Angers S, Da Silva A, et al. Beta(2)-adrenergic receptor down-regulation. Evidence for a pathway that does not require endocytosis. *J Biol Chem* 1999;274:28900-28908.

58. Bochtler M, Ditzel L, Groll M, et al. The proteasome. *Annu Rev Biophys Biomol Struct* 1999;28:295–317.

59. Strous GJ, van Kerkhof P, Govers R, et al. The ubiquitin conjugation system is required for ligand-induced endocytosis and degradation of the growth hormone receptor. *EMBO J* 1996;15:3806–3812.

60. Hicke L. Gettin' down with ubiquitin: turning off cell-surface receptors, transporters and channels. *Trends Cell Biol* 1999;9:107–112.

61. Levkowitz G, Waterman H, Ettenberg SA, et al. Ubiquitin ligase activity and tyrosine phosphorylation underlie suppression of growth factor signaling by c-Cbl/Sli-1. *Mol Cell* 1999;4:1029–1040.

62. Gruenberg J, Maxfield FR. Membrane transport in the endocytic pathway. *Curr Opin Cell Biol* 1995;7:552–563.

63. Law P-Y, Hom DS, Loh HH. Loss of opiate receptor activity in neuroblastoma x glioma NG108-15 cells after chronic etorphine treatment: a multiple step process. *Mol Pharmacol* 1982;72:1–4.

64. Campbell PT, Hnatowitch M, O'Dowd BF, et al. Mutational analysis of adrenergic receptor sequestration. *Mol Pharmacol* 1991;39:192–198.

65. Bouvier M, Collins S, O'Dowd BF, et al. Two distinct pathways for cAMP-mediated down-regulation of the beta 2-adrenergic receptor. Phosphorylation of the receptor and regulation of its mRNA level. *J Biol Chem* 1989;264:16786–16792.

66. Hausdorff WP, Bouvier M, O'Dowd BF, et al. Phosphorylation sites on two domains of the beta 2-adrenergic receptor are involved in distinct pathways of receptor desensitization. *J Biol Chem* 1989;264:12657–12665.

67. Hausdorff WP, Campbell PT, Ostrowski J, et al. A small region of the beta-adrenergic receptor is selectively involved in its rapid regulation. *Proc Natl Acad Sci U S A* 1991;88:2979–2983.

68. Heck DA, Bylund DB. Differential down-regulation of alpha-2 adrenergic receptor subtypes. *Life Sci* 1998;62:1467–1472.

69. von Zastrow M, Keith DJ, Evans CJ. Agonist-induced state of the delta-opioid receptor that discriminates between opioid peptides and opiate alkaloids. *Mol Pharmacol* 1993;44:166–172.

70. Arden JR, Segredo V, Wang Z, et al. Phosphorylation and agonist-specific intracellular trafficking of an epitope-tagged mu-opioid receptor expressed in HEK 293 cells. *J Neurochem* 1995;65:1636–1645.

71. Keith DE, Murray SR, Zaki PA, et al. Morphine activates opioid receptors without causing their rapid internalization. *J Biol Chem* 1996;271:19021–19024.

72. Sternini C, Spann M, Anton B, et al. Agonist-selective endocytosis of mu opioid receptor by neurons *in vivo*. *Proc Natl Acad Sci U S A* 1996;93:9241–9246.

73. Trejo J, Coughlin SR. The cytoplasmic tails of protease-activated receptor-1 and substance P receptor specify sorting to lysosomes versus recycling. *J Biol Chem* 1999;274:2216–2224.

74. Mellman I. Endocytosis and molecular sorting. *Annu Rev Cell Dev Biol* 1996;12:575–625.

75. Gagnon AW, Kallal L, Benovic JL. Role of clathrin-mediated endocytosis in agonist-induced down-regulation of the beta2-adrenergic receptor. *J Biol Chem* 1998;273:6976-6981.

76. Chu P, Murray S, Lissin D, et al. Delta and kappa opioid receptors are differentially regulated by dynamin-dependent endocytosis when activated by the same alkaloid agonist. *J Biol Chem* 1997;272:27124–27130.

77. Tsao PI, von Zastrow M. Type-specific sorting of G protein-coupled receptors after endocytosis. *J Biol Chem* 2000;275:11130–11140.

78. Cao TT, Deacon HW, Reczek D, et al. A kinase-regulated PDZ-domain interaction controls endocytic sorting of the beta2-adrenergic receptor. *Nature* 1999;401:286–290.

79. Hall RA, Premont RT, Chow CW, et al. The beta2-adrenergic receptor interacts with the Na^+/H^+-exchanger regulatory factor to control Na^+/H^+ exchange. [Nature] 1998;392:626-630.

80. Hall RA, Ostedgaard LS, Premont RT, et al. A C-terminal motif found in the beta2-adrenergic receptor, P2Y1 receptor and cystic fibrosis transmembrane conductance regulator determines binding to the Na^+/H^+ exchanger regulatory factor family of PDZ proteins. *Proc Natl Acad Sci U S A* 1998;95:8496–8501.

81. Reczek D, Berryman M, Bretscher A. Identification of EBP50: a PDZ-containing phosphoprotein that associates with members of the ezrin–radixin–moesin family. *J Cell Biol* 1997;139:169–179.

82. Fredericks ZL, Pitcher JA, Lefkowitz RJ. Identification of the G protein-coupled receptor kinase phosphorylation sites in the human beta2-adrenergic receptor. *J Biol Chem* 1996;271:13796–13803.

83. Gainetdinov RR, Bohn LM, Walker JK, et al. Muscarinic supersensitivity and impaired receptor desensitization in G protein-coupled receptor kinase 5-deficient mice. *Neuron* 1999;24:1029–1036.

84. Eckhart AD, Duncan SJ, Penn RB, et al. Hybrid transgenic mice reveal *in vivo* specificity of G protein-coupled receptor kinases in the heart. *Circ Res* 2000;86:43–50.

85. Mantyh PW, DeMaster E, Malhotra A, et al. Receptor endocytosis and dendrite reshaping in spinal neurons after somatosensory stimulation. *Science* 1995;268:1629–1632.

86. Mantyh PW, Allen CJ, Ghilardi JR, et al. Rapid endocytosis of a G protein-coupled receptor: substance P evoked internalization of its receptor in the rat striatum *in vivo. Proc Natl Acad Sci U S A* 1995;92:2622–2626.

87. Pals-Rylaarsdam R, Gurevich VV, Lee KB, et al. Internalization of the m2 muscarinic acetylcholine receptor. Arrestin-independent and -dependent pathways. *J Biol Chem* 1997;272:23682–23689.

88. Lee KB, Pals RR, Benovic JL, et al. Arrestin-independent internalization of the m1, m3, and m4 subtypes of muscarinic cholinergic receptors. *J Biol Chem* 1998;273:12967–12972.

89. Vickery RG, von Zastrow M. Distinct dynamin-dependent and -independent mechanisms target structurally homologous dopamine receptors to different endocytic membranes. *J Cell Biol* 1999;144:31–43.

90. Luttrell LM, Ferguson SS, Daaka Y, et al. Beta-arrestin-dependent formation of beta2 adrenergic receptor-Src protein kinase complexes. *Science* 1999;283:655–661.

91. DeFea KA, Zalevsky J, Thoma MS, et al. Beta-arrestin-dependent endocytosis of proteinase-activated receptor 2 is required for intracellular targeting of activated ERK1/2. *J Cell Biol* 2000;148:1267–1281.

92. Miller WE, Maudsley S, Ahn S, et al. Beta-arrestin1 interacts with the catalytic domain of the tyrosine kinase c-SRC. Role of beta-arrestin 1-dependent targeting of c-SRC in receptor endocytosis. *J Biol Chem* 2000;275:11312–11319.

93. Hall RA, Premont RT, Lefkowitz RJ. Heptahelical receptor signaling: beyond the G protein paradigm. *J Cell Biol* 1999;145:927–932.

6

The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction

Joseph T. Coyle

Michael L. Leski

John H. Morrison

Joseph T. Coyle: Department of Psychiatry, Harvard Medical School, Boston, Massachusetts.

Michael Leski: Department of Psychiatry, McLean Hospital, Belmont, Massachusetts.

John H. Morrison: Mount Sinai School of Medicine, New York, New York.

L-Glutamic acid (Glu) is accepted as the major excitatory neurotransmitter in the nervous system, although other acidic amino acids such as L-aspartic acid and L-homocysteic acid may also participate (1). Nevertheless, ongoing research reveals that the functions of Glu are much more diverse and complex than simply generating excitatory postsynaptic currents (EPSCs). It plays a major role in brain development, affecting neuronal migration, neuronal differentiation, axon genesis, and neuronal survival (2, 3 and 4). In the mature nervous system, Glu is central to neuroplasticity, in which there are use-dependent alterations in synaptic efficacy as well as changes in synaptic structure. These latter actions are intimately implicated in memory and related cognitive functions. Finally, persistent or overwhelming activation of glutamate-gated ion channels can cause neuronal degeneration (5) depending on the circumstances, this occurs by means of necrosis or apoptosis (6). Known as “excitotoxicity,” this phenomenon has been linked to the final common pathway of neuronal death in a range of disorders including Huntington’s disease, Alzheimer’s disease, amyotrophic lateral sclerosis (ALS), and stroke (7, 8).

This chapter provides an overview of the physiology and pharmacology of brain glutamatergic systems. There is a special emphasis on glutamate receptors because their rich diversity confers physiologic and pharmacologic specificity for this single neurotransmitter, which is used by up to 40% of all brain synapses. Finally, the potential role of glutamatergic system dysfunction in the pathophysiology of neuropsychiatric disorders is addressed.

AMPA-KAINATE RECEPTORS

Part of "6 - The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction "

Glutamate receptors mediating fast EPSCs have been distinguished from the voltage-dependent NMDA receptors through the effects of conformationally restricted agonists. The former glutamate-gated ion channels (iGluRs) have been segregated into two types: the α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptors and KA receptors; however, cloning of the genes that encode the proteins comprising these iGluRs and the analysis of their pharmacology and biophysics in various expression systems indicate that each family of receptors represents complex heteromeric proteins consisting of multiple subunits with differential representation resulting in diverse functional attributes (9).

The AMPA receptor family consists of four genes that encode proteins of approximately 900 amino acids with 70% homology among themselves (GluR1-4) (1); however, cell-specific exon splice variants and posttranscriptional editing result in a complex range of physiologic responses. Two exon splice variants of GluR1-4 known as “flip” and “flop” affect desensitization, with the flip form associated with larger, more sustained currents (10). Splice variants truncated at the carboxy terminus have been described for GluR2 and GluR4 as well as the kainate subunits, GluR5–7 (9). Furthermore, nuclear editing of the mRNA encoding GluR2 transforms this receptor channel from one permeable to Ca^{2+} to one impermeable to the cation. Double-stranded RNA adenosine deaminase converts the adenosine in a CAG codon for glutamine (Q) to an inosine, thereby creating a CIG codon for arginine (R) (11, 12). The regulatory site for the editing process is located downstream in an intron that aligns with the exon to form the secondary structure recognized by the enzyme. Although the other AMPA receptor subunits are generally Ca^{2+} permeable, the presence of edited GluR2 dominates a heteromeric receptor

complex in that the multisubunit AMPA receptor behaves like GluR2 (i.e., has low calcium permeability). In the mature brain, the vast majority of GluR2 is edited and the majority of AMPA receptors have low calcium permeability, suggesting that GluR2 is reasonably ubiquitous (see the following). Mice in which the GluR2 editing process has been inactivated by a null mutation exhibit increased Ca^{2+} permeability with AMPA receptor activation, epilepsy, and early death (13). Stroke is associated with the suppression of GluR2 expression in the penumbra, resulting in increased Ca^{2+} permeability through the remaining AMPA receptors, thereby causing delayed neuronal degeneration (14).

Five separate genes encode the components of the kainate receptor: GluR5–GluR7 and KA1–KA2 (1). Homomeric complexes of GluR5, GluR6, and GluR7 form ion channels in the Xenopus expression system that are activated by KA but not by AMPA. However, homomeric complexes of KA1 or KA2 do not generate functional ion channels, although they exhibit high-affinity binding for kainic acid. It appears that KA1 or KA2 in conjunction with GluR5, GluR6, or GluR7 form functional KA receptors. GluR5 and GluR6 possess the Q/R editing site in their second transmembrane (M2) domain, which modulates Ca^{2+} permeability as for GluR2. Additional editing sites on GluR5–7 have complex effects on the receptor-channel function (15, 16). GluR5–7 are also widely represented in the brain, particularly in neocortex and hippocampus.

A number of modulators of AMPA/KA receptors have been identified that act by attenuating their rapid and profound desensitization. Cyclothiazide selectively inhibits desensitization of AMPA receptors, whereas the lectin, concanavalin A, blocks desensitization at KA receptors (17). “Ampakines” also enhance AMPA receptor activity by attenuating desensitization and have been shown to facilitate glutamatergic neurotransmission *in vivo*, thereby improving performance in several cognitive tasks (18). Polyamines, first noted for their effects on NMDA receptors, mediate rectification of Ca^{2+}-permeable AMPA and KA receptors by inducing a voltage-dependent channel blockade (19).

A number of conformationally restricted analogues have assisted in characterizing the activity of subsets of AMPA/KA receptors. Prior to development of AMPA, quisqualic acid was used as a non–NMDA receptor agonist; however, its specificity for AMPA receptors is poor as it also activates mGluR 1 and 5 metabotropic receptors and inhibits glutamate carboxypeptidase II (GCP II) (20, 21). The conformational rigidity of AMPA provides good specificity for AMPA receptors, whereas the more flexible kainic acid interacts with KA receptors as well as other types of iGluRs. The orientation, length, and saturation of the side chain of kainic acid and its analogues play a critical role in their binding to the receptor site with domoic acid exhibiting higher affinity and dihydrokainic acid having much lower affinity than kainate (22). The first potent selective antagonists at AMPA/KA receptors with negligible effects at NMDA receptors to be developed were the dihydroxyquinoxalines: CNQX, DNQX, and NBQX. With the exception of NBQX, which has a somewhat higher selectivity for AMPA receptors, they do not distinguish between AMPA and KA receptors. More recently 2,3-benzodiazepines have been demonstrated to act as selective noncompetitive AMPA receptor antagonists. The most potent in this family, GYKI53655, blocks AMPA receptors with an IC_{50} of 1 μM and has a 200-fold lower affinity for KA receptors (23). The recently developed LY294486 has tenfold specificity for inhibiting GluR5, indicating that iGluR subtype specific ligands can be developed.

Most of our initial knowledge concerning the regional brain expression of KA/AMPA receptors was based on autoradiographic studies. Specific binding of [^{3}H]-KA is relatively enriched in the hippocampal CA 3, striatum, deep layers of the neocortex, the reticular nucleus of the thalamus, and the cerebellar granular cell layer. When considered in aggregate, immunohistochemical staining for the five KA receptor subunits shares this ligand binding distribution (26).

Audioradiographic studies of the distribution of [^{3}H]-AMPA binding show high density in the CA 1 stratum radiatum, the dentate gyrus, superficial layers of the neocortex, and the molecular layer of the cerebellum (25). Consistent with its role in mediating EPSCs, moderate levels of binding are observed throughout the rest of the central nervous system (CNS). The distribution of [^{3}H]-AMPA binding corresponds with the regional expression of GluR1 and GluR2, because levels of GluR3 and GluR4 are much lower in the adult rat brain and have a more restricted distribution (27).

Although the autoradiographic approaches have been very informative with respect to general distribution of the KA/AMPA receptors, they have been less informative with respect to cellular and synaptic distribution and differential subunit distribution. It has recently become possible to overcome these limitations on both spatial and biochemical resolution, and begin the arduous process of delineating the GluR subunit profile of identified neurons and circuits in specific brain structures such as the neocortex and hippocampus. This capacity is a direct result of molecular neurobiological analyses that have fostered a detailed understanding of the subunit protein constituents of the three classes of inotropic GluRs, and thus allowed for *in situ* hybridization to be used to localize specific mRNAs (28) and with the use of class and subunit-specific antibodies for immunocytochemistry, it is now feasible to analyze GluR distribution at the highest levels of cellular and synaptic resolution (29, 30).

AMPA receptor subunit distribution in the hippocampus and neocortex offer an instructive example of such an approach. Although early studies demonstrated a wide distribution of AMPA subunits GluRs 2/3 in the brain and spinal cord (31, 32, 33 and 34), it became clear early on that the relationship

between GluRs and specific circuits needed to be analyzed at a high level of resolution; colocalization studies directed at subsets of neurons (29, 35, 36, 37 and 38) and ultrastructural dissection of the synapse (39, 40, 41, 42, 43 and 44). A key theme that emerges from these studies is that regional distribution and cellular colocalization patterns should not be extended to a synaptic interpretation: Such interpretations must be founded on ultrastructural data as seen in the following examples.

The cellular distribution of GluR2 has been linked to heterogeneity in calcium permeability of AMPA receptors (45, 46). For example, electrophysiological analyses have demonstrated that pyramidal cells have AMPA receptors with low calcium permeability, and interneurons have AMPA receptors with relatively high calcium influx (47, 48), and these properties are linked to the relative abundance of GluR1 versus GluR2 mRNAs (45, 46). In addition, early reports using polyclonal antisera that did not differentiate among GluR2, 3, and 4c, obtained results implying that GABAergic interneurons might not contain GluR2, 3, and 4C (49, 50 and 51).

Definitive conclusions regarding selective distribution of GluR2-specific protein were not possible until 1996, when an exclusively GluR2-specific monoclonal antibody (52), followed by a rabbit polyclonal (53) were developed. The GluR2 antibodies showed that virtually all pyramidal cells and the majority of GABAergic interneurons in the neocortex (e.g., S1) contain GluR2. A similar pattern was found in hippocampus, suggesting that the majority of the GABAergic interneurons in hippocampus are GluR2-positive, although a subset of GABAergic neurons lacks any detectable GluR2 (52, 53), as in neocortex. These results are in excellent accord with the GluR2 mRNA results obtained by single cell RT-PCR studies (45, 46), and suggest that a minority of the GABAergic interneurons lack GluR2 mRNA/protein. Thus, the differences in calcium permeability between GABAergic interneurons and pyramidal cells could not be the result of a widespread lack of GluR2 in GABAergic interneurons.

A double label GABA/GluR2 analysis that was extended to the ultrastructural level further clarified the issue of GluR2 representation in GABAergic interneurons (54). It was hypothesized that if the difference in calcium permeability between pyramidal and GABAergic neurons was related to differences in GluR2 expression, then it would likely be more apparent on the synaptic than the cellular level (54). Ultrastructural analysis revealed that there is a consistently lower number of immunogold particles at the labeled asymmetric synapses on GABAergic dendrites than those on pyramidal cell dendrites or spines, suggesting that a cell class-specific difference in synaptic abundance of GluR2 is the substrate for the observed differences in calcium permeability across these two cell classes revealed electrophysiologically (45, 46).

As demonstrated in the GluR2 studies discussed in the preceding, cellular colocalization may not adequately reflect the localization patterns at the synaptic level. This was reinforced in studies of GluR2/NR1 colocalization in hippocampus and neocortex, designed to delineate the degree of synaptic colocalization of NMDA and AMPA receptors in asymmetrical synapses (39, 40, 41, 42 and 43). Although NR1 and GluR2 are broadly colocalized on a cellular level, extensive synaptic heterogeneity exists in their representation. NR1 and GluR2 are often colocalized at the same synapse; however, there are also a large group of NR1-containing synapses that lack GluR2 labeling (33% in [39]), many of which were on spines. These may be candidates for the “silent synapses” that have been described electrophysiologically (55, 56) that might be activated by insertion of AMPA subunits (40, 42, 55, 56 and 57).

It has been generally recognized that AMPA receptors play a dominant role in mediating EPSCs. A physiologic role for KA receptors has been elucidated only recently with the development of more selective agonists and antagonists. In the hippocampal slice in which the AMPA, NMDA, and GABA receptors have been blocked pharmacologically, stimulation of the mossy fibers generates a slow excitatory synaptic current system with the biophysical properties of the KA receptor (58). This current is absent in mice homozygous for null mutation of the GluR6 subunit and less vulnerable to the epileptogenic effects of systemic KA (59). The presynaptic inhibitory effect on GABA release in the CA 1 region of the hippocampus is mediated by the GluR5 subunit (60).

Although KA subunits have not been localized as extensively at the ultrastructural level as have AMPA or NMDA receptors, immunocytochemical studies have demonstrated their broad distribution in the hippocampus and neocortex and broad colocalization with AMPA and NMDA receptor subunits (38, 61) in both pyramidal and GABAergic interneurons (35, 36).

NMDA RECEPTORS

Part of "6 - The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction "

The NMDA receptor, as its name indicates, was identified by the selective excitatory effects of the synthetic analogue of glutamate, N-methyl-D-aspartic acid (1). A number of properties distinguishes the NMDA receptor from the non-NMDA iGluRs. First, its activity is voltage dependent. At resting membrane potential, the channel is blocked by Mg^{2+}, which is relieved by membrane depolarization. Second, the receptor requires occupancy of another ligand binding site, the so-called glycine modulatory site, in order for glutamate to gate channel opening. Recent evidence indicates that not only glycine but also D-serine, which is synthesized in astrocytes by serine racemase, is a potent endogenous agonist at the glycine site (62). Third, the NMDA receptor possesses a number of modulatory sites of physiologic significance. Zn^{2+} is a potent inhibitor of NMDA receptor conductance, especially those containing the NR2A subunit (24). Zn^{2+} is concentrated in some glutamatergic

terminals (e.g., the mossy fibers) and released with glutamate (63). A binding site for polyamines, when occupied, enhances conductance in part through increasing the affinity of the glycine modulatory site on the NMDA receptor (64). Receptor function is also modulated by redox status (65). Within the channel, there is a binding site for the dissociative anesthetics such as phencyclidine (PCP), MK-801, and ketamine, which serve as noncompetitive inhibitors (66). The effects of the dissociative anesthetics occur only with open channels, thereby causing a use-dependent inhibition. Finally, the NMDA channel provides ready passage of Ca^{2+}, a cation involved in a number of intracellular signaling processes.

Molecular cloning has disclosed at least six genes that comprise a family of polypeptides that form the various subtypes of the NMDA receptor (1). NR1 was the first component cloned and, when expressed in Xenopus oocytes, was shown to possess the primary electrophysiologic and pharmacologic features of the NMDA receptor–channel complex. Seven splice variants of NR1 have been described, which reflect the exclusion or inclusion of three exons, two in the C terminal and one in the N terminal portion (1). These splice variants significantly impact the biophysical characteristics of the receptor. The NR2 subunits, NR2A-D and the recently identified NMDARL or NR3A, do not form channels (67); however, when coexpressed with NR1, the heteromeric channels exhibit a markedly increased current as compared to the homomeric NR1 channels. Each of the NR2 subunits, when complexed with the NR1 subunit, exhibits different biophysical and pharmacologic properties. The NR2A and NR2B subunits, more highly expressed in adult cortex in contrast to NR2C/D receptors, appear to be less sensitive to NMDA receptor antagonists, not as vulnerable to Mg^{2+} blockade, and have lower Ca^{2+} conductance (1).

NR1 is expressed in most neurons, whereas the NR2 mRNAs exhibit different regional and developmental patterns of expression (68, 69). The NR2A subunit is highly expressed in the neocortex, hippocampus, cerebellum, and several thalamic nuclei. The NR2B subunit is found in the neocortex, hippocampus, striatum, septum, and thalamic nuclei of the adult rat brain. The expression of the NR2C subunit is much more restricted in the adult brain, being enriched in the olfactory bulb, thalamic nuclei, and cerebellum. Finally, the NR2D subunit is enriched in brainstem nuclei, midline thalamic nuclei, and bipolar cells of the retina. The NR2B and NR2D subunits appear early in brain development, followed by a decline in NR2D expression in the third week after birth of the rat, whereas the acquisition of NR2A and NR2C subunits appears primarily postnatally in the rat (1).

With respect to NMDA receptor localization, as in the case of AMPA/kainate receptors, the early *in situ* hybridization studies discussed in the preceding offered important information as to the regional distribution of NMDA receptors in the brain. These studies have been followed with very extensive immunocytochemical analyses, particularly of the obligatory subunit NR1. With respect to hippocampus and neocortex, NR1 is very broadly distributed and present in virtually all pyramidal neurons and nonpyramidal GABAergic interneurons (29, 38), and in fact, appears to be present in over 90% of asymmetric synapses (30, 41). NR1 distribution has also been shown to be modifiable on both the cellular and synaptic level with respect to plasticity. For example, deafferentation causes rearrangements of NR1 distribution at the cellular and synaptic level in a matter of days (70, 71). In addition, although NMDA receptors have a very broad distribution on a cellular level, they can display a high degree of specificity on a synaptic basis. The most dramatic example of this is in CA3 of the hippocampus, where NMDA receptors are present postsynaptically in the distal dendrites (i.e., stratum moleculare), yet are absent in the stratum lucidum terminal zone, which is noted for LPT being NMDA receptor-independent (37, 56, 72). This suggests that there are intracellular trafficking mechanisms or local synthesis that can position NMDA receptors in a subset of synapses receiving a particular input to a given cell while not mediating other inputs.

The data on the distribution of other subunits are less well developed, and this is partly owing to the fact that it has been very difficult to develop antibodies that differentiate NR2A from NR2B, the two dominant subunits in the NR2 group in the hippocampus and neocortex. In general, it appears that NR2A and NR2B overlap in their distribution with NR1 to a large degree (73, 74 and 75); however, there are regions such as CA3 where they differ in their distribution with NR1 (75, 76), but this has yet to be worked out at the synaptic level. The detailed delineation of the synaptic distribution of NR2A and NR2B is an important task for the future given that the presence of these subunits confer different functional attributes on the receptor, and the analysis of genetically manipulated mice have suggested that up- or down-regulation of one of these subunits can profoundly impact their function. For example, NR2B overexpression in mouse enhances learning and memory (77). Thus, as is the case for AMPA and kainate receptors, delineating the subunit representation and potential stoichiometry at specific circuits and synapses for the NMDA receptor is of paramount importance if this receptor is to be definitively linked to circuits that mediate specific behaviors and suffer under certain pathologic conditions.

Several endogenous amino acids, aside from glutamate, are selective agonists at the NMDA receptor, including L-homocysteic acid, L-aspartic acid, L-cysteine sulfate, L-serine-O-sulfate, L-cysteic acid, and quinolinic acid in order of decreasing potency (78). Given the multiple modulatory sites and agonist binding sites for the NMDA receptor, it is not surprising that the antagonist pharmacology for this receptor is complex. Several phosphonate analogues of glutamate, including D-aminophosphonovaleric acid (APV),

D-aminophosphonoheptanoic (APH) acid, D-aminoadipic acid, and the cyclic analogue of AHP, (2-carboxypiperazin-4-yl)-propyl-1-phosphonate (CPP), are competitive inhibitors to glutamate. As described, the NMDA receptor channel has a binding site for dissociative anesthetics, which shares pharmacologic features with the σ receptor. The glycine modulatory site, which must be occupied for glutamate gating of the ion channel, is subject to inhibition by the endogenous metabolite of tryptophan, kynurenic acid as well as synthetic analogues such as 7-chlorokynurenate (79). For greater selectivity, attention is now directed at developing subtype specific antagonists such as ifenprodil, which inhibits NMDA receptors bearing the NR2B subunit (80).

METABOTROPIC GLUTAMATE RECEPTORS

Part of "6 - The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction "

It was generally believed that the neurophysiologic effects of glutamate were mediated exclusively by iGluR, until Sladeczek and colleagues (81) reported that glutamate catalyzed phosphoinositide hydrolysis through a receptor coupled to a G-protein. Since then, research has disclosed the existence of a family of glutamate receptors whose effects are largely mediated by G-proteins, the so-called metabotropic glutamate receptors (mGluRs). The mGluRs appear to play an important role modulating both presynaptically and postsynaptically the effects of glutamate at glutamatergic synapses where iGluRs are also engaged (82).

Cloning of mGluR 1 indicated that the predicted amino acid sequence shared negligible homology with any of the other G-protein coupled receptor (GPCR) families except the parathyroid Ca^{2+}-sensing receptor (83) and the GABA-B receptor (84). On the basis of pharmacology, physiologic effects, and sequence homology, the eight mGluRs have been subdivided into three groups (85). Group I includes mGluR 1 and mGluR 5, which act via phospholipase C; group II includes mGluR 2 and mGluR 3, which are negatively coupled to adenylyl cyclase; and group III, which includes mGluR 4, 6, 7, and 8, and are also negatively coupled to adenylyl cyclase. Adding to this complexity, mGluR 1 has four splice variants, whereas mGluR 4 and 5 each have two. The mGluRs within a group exhibit greater than 70% sequence homology, whereas the homology falls to approximately 45% between groups (82).

Like other GPCRs, the mGluRs have seven putative transmembrane domains separated by short intracellular and extracellular loops and an unusually large extracellular domain that contains nearly a score of cysteine residues (86). In further contradistinction to other GPCRs, the agonist binding sites for the mGluRs are located in the extracellular domain. As a family, mGluRs are broadly expressed in the nervous system, with group I having primarily a postsynaptic localization, whereas groups II and III primarily have a presynaptic localization where they serve as autoreceptors and heteroreceptors that negatively regulate neurotransmission (82). mGluR 1 is prominently expressed in the hippocampal granule and pyramidal cells, the Purkinje cells of the cerebellum, the thalamus, and the lateral septum (87). mGluR 5 exhibits somewhat of a complementary distribution with high levels expressed in the neocortex, the pyramidal cells in CA1 sector of the hippocampus, lateral septum, striatum, and nucleus accumbens (88). Although studies *in vivo* suggest a low level of expression of mGluR 5 in glia, cultured astrocytes exhibit high expression (89). The expression of group II receptors is more restricted than group I, with mGluR 2 found in the cerebellum, pyramidal cells in the entorhinal cortex and the dentate gyrus, and presynaptically on corticostriatal afferents (90). The majority of the cerebellar Golgi cells have mGluR 2 but a minor subset express mGluR 5 in a complementary fashion (91). mGluR 3 is found in astrocytes throughout the brain and in neurons in the neocortex, caudate putamen, thalamic reticular nucleus, and granule cells of the dentate gyrus (92). With regard to group III, mGluR 4 is found in the thalamus, lateral septum, dentate gyrus, and cerebellar granule cells (92). mGluR 7 has a widespread distribution with prominent representation in sensory afferent systems including the dorsal root ganglia and the trigeminal nucleus in addition to the cerebral cortex, hippocampus, striatum, thalamus, and Purkinje cells (93). mGluR 6 is restricted to on-bipolar cells in the inner nuclear layer of the retina (94), with a highly specific synaptic distribution (95). Finally, mGluR8 is found in the mitral cells of the olfactory bulb, piriform cortex, and lateral thalamic reticular nucleus (96).

Ultrastructural studies have demonstrated that mGluRs differ in their predominant synaptic distribution from AMPA and NMDA receptors. Although AMPA and NMDA receptors are localized predominantly in the postsynaptic density, mGluRs are principally localized presynaptically and perisynaptically (97, 98 and 99). In addition, the mGluRs exhibit a high degree of specificity in this regard with mGluR7 (100, 101) and mGluR4 (102) (which are primarily present presynaptically) and mGluRs 1 and 5 (which are primarily present perisynaptically) (103, 104 and 105). These patterns suggest that synaptic localization is partially segregated by mGluR group (99), with group I primarily perisynaptic, surrounding the postsynaptic density, group II primarily presynaptic or extrasynaptic, and group III optimally situated as an autoreceptor on the presynaptic terminal. These synaptic distribution patterns position mGluRs to play a critical role in modulating excitatory neurotransmission.

Group I mGluRs stimulate phospholipase C and the hydrolysis of phosphoinositide. These two receptors have been reported to stimulate cAMP formation in different model systems (106). They also increase the excitability of neurons by reducing the K^+ currents, through a mechanism that appears to be independent of G-protein action (107). Although the AMPA receptor agonist, quisqualic acid, is the

most potent agonist at the group I mGluRs, 3,5-dihydroxyphenylglycine (3,5-DHPG) is the most specific agonist. Group II mGluRs inhibit adenylyl cyclase by coupling with the G_i-protein. (See ref. 43 for a detailed review of mGluR pharmacology.) Group II receptors also inhibit the N-type Ca^{2+} channels through GI-protein coupling. The group III mGluRs inhibit adenylyl cyclase via GI-protein, although they may utilize other transduction mechanisms.

The most specific agonist at group I mGluRs appears to be 2R, 4R-4-aminio pyrrolidine-2, 4 dicarboxylate (APDC). LY354740 is an exceptionally potent and selective agonist at group II mGluRs and is effective with systemic administration (108). Notably, N-acetylaspartyl glutamate (NAAG), an endogenous neuropeptide, is a relatively potent agonist at mGluR 3, although it also serves as an antagonist at the NMDA receptor (109). L-aminophosphonobutyric acid (L-AP4) is the most potent and selective agonist at the group III mGluRs with the exception of mGluR 7, where L-serine-O-phosphate has greater potency. Presently, 2-carboxy-3-phenylcyclopropyl glycine (CPCG) is the most potent antagonist against the group II mGluRs, and α-methyl-4-phosphonophenyl-glycine appears to be an effective antagonist against the group III mGluRs, although the pharmacology of these receptors remains less well developed than the group I mGluRs.

The heterogeneity of the mGluRs and their role in modulating glutamatergic neurotransmission make them attractive potential therapeutic targets for drug development. Thus, activation of group II and group III mGluRs has been associated with protection against excitotoxicity, whereas activation of group I mGluRs may actually enhance NMDA receptor-mediated neuronal degeneration (110, 111). Recently it has been shown that inhibition of GCP II, the enzyme that degrades the selective mGluR 3 agonist NAAG, provides potent protection against neuronal degeneration caused by transient occlusion of the middle cerebral artery (112). A similar reciprocal relationship has been observed with regard to the effects on epilepsy with group I agonists exacerbating and group II and III agonists attenuating seizures (113). Finally, metabotropic receptors, both in the dorsal root and thalamus, have been implicated in modulating neuropathic pain (114).

GLUTAMATE TRANSPORTERS

Part of "6 - The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction "

The demonstration of sodium-dependent high-affinity transport of glutamate in synaptosomal preparations was the first evidence supporting the hypothesis that glutamate serves as a neurotransmitter (115). The presence of these transporters on excitatory nerve terminals was exploited to label putative glutamatergic pathways such as the climbing fibers of the cerebellum through the autoradiographic visualization of retrogradely transported radiolabel in axons and cell bodies. The transporters are capable of maintaining extraordinary gradients with the extracellular concentration of glutamate in brain low μM (i.e., below the threshold for iGluRs) in the face of a tissue concentration in the mM range (116). Energy deprivation not only collapses the sodium gradient across the membrane that drives the transporter, thereby inhibiting glutamate uptake, but also results in reverse transport with massive efflux of glutamate stores (117). Pharmacologic inhibition of glutamate transport in tissue culture models has been shown to promote excitotoxic neurodegeneration (118).

Early pharmacologic studies pointed to the existence of subtypes of sodium-dependent glutamate transporters in brain with cerebellum exhibiting a form that is much more sensitive to inhibition by L-α-aminoadipate and forebrain sensitive to dihydrokainate (119). Both the pharmacologic heterogeneity as well as the diverse cellular distribution of the sodium-dependent glutamate transporters have been illuminated recently by their cloning and molecular characterization. (See refs. 120 and 121 for review.) Although acronyms have varied with the sequence of discovery, five excitatory acidic amino acid transporters (EAAT) that are sodium dependent and chloride independent have been cloned. EAAT 1 or GLAST has its highest expression in brain but is also found in peripheral tissue and placenta. The cerebellum appears to have the highest level within brain, depending on the species. EAAT 2 (GLT 1) is primarily expressed in brain, although low levels have been reported in pancreas and placenta; the highest expression occurs in the forebrain, the lowest in the cerebellum. Its expression is predominantly if not exclusively astroglial in localization. The predominant neuronal transporter is EAAT 3, which is also expressed in kidney and to a lesser extent in other peripheral tissues. Consistent with the broad distribution of glutamatergic neuronal systems in brain, the levels of EAAT 3 are fairly uniform. EAAT 3 is not consistently expressed in all glutamatergic systems, and some nonglutamatergic systems express it (122). Thus, EAAT 3 does not appear to be a specific marker for glutamatergic neurons. Two additional minor forms have been identified: EAAT 4, which is expressed in cerebellar Purkinje cells; and EAAT 5, which is limited to the retina.

Like the mGluRs, the neuronal transporter EAAT3 (i.e., EAAC1) has been shown to have a synaptic distribution different from the AMPA or NMDA receptors in that it is primarily perisynaptic and presynaptic (44, 123). A double-label postembedding immunogold study demonstrated the value of such ultrastructural data for revealing the importance of differential distribution of these proteins with respect to the synapse (44). The double-label analysis targeted the AMPA subunit GluR2 and the glutamate transporter, EAAT3 (i.e., EAAC1), the neuron-specific glutamate transporter. This study revealed differential spatial distribution of these two proteins very clearly. At the light microscopic level, as expected, GluR2 was broadly colocalized with EAAC1 in hippocampal projection neurons, and both proteins

had substantial cytoplasmic pools. However, the postembedding immunogold localization offered additional insights into the spatial relationships between EAAT3 and GluR2 localization in and near the synapse, revealing morphologic and molecular constraints on excitatory synaptic transmission. Specifically, synaptic GluR2 was present primarily in the postsynaptic specialization, appropriately positioned to mediate the synaptic effects of glutamate. EAAT3 was not intermingled with GluR2 postsynaptically, but was generally present perisynaptically, often immediately outside the synaptic specialization, with a small but significant presynaptic pool as well (44). This arrangement positions EAAT3 to both confine glutamate to the synaptic site that contains the inotropic receptor molecules, as well as to regulate its levels in immediately adjacent presynaptic and postsynaptic domains. This distribution is also interesting with respect to the distribution of mGluRs, which are located perisynaptically and extrasynaptically (see the preceding). This suggests that EAAT3 is also optimally positioned to regulate the exposure of perisynaptic and presynaptic mGluRs to glutamate (44).

The EAATs exhibit affinities for glutamate in the low μM range, two orders of magnitude more avid than the vesicular transporter for glutamate, which is not sodium dependent. A generally accepted model for transport involves the binding of three Na^+, one H^+, and glutamic acid, which is linked to the counter transport of one K^+ (124). In addition, there is increasing evidence that individual EAAT subtypes may also subserve signal transduction activity; thus, EAATs have been implicated in inhibition of adenylyl cyclase, altering Ca^{2+} levels and Cl^- flux by mechanisms independent of glutamate transport (121).

Pharmacologic inhibition of EAAT activity potentiates excitotoxic effects in tissue culture (118). Mice homozygous for the nul mutation of EAAT 2 develop fatal epilepsy and exhibit increased vulnerability to excitotoxic insults (125), and mice homozygous for the nul mutation for EAAT 1 exhibit cerebellar dysfunction (126). Thus, in contrast to the original hypothesis of the essential role of neuronal glutamate transport in terminating excitatory neurotransmission, it is now apparent that the glial transporters play the dominant role in neuroprotection from glutamate. In this regard, astroglial processes are tightly interdigitated with glutamatergic and GABAergic synapses, where the transporters are expressed in high density, which accounts for the previous identification of the “synaptosomal” localization of glutamate transport (127).

The activity of the glutamate transporters is regulate by transcriptional as well as posttranslational mechanisms. (See ref. 128 for review.) GLT-1 expression, which is normally low in primary astrocyte cultures, is significantly increased in astrocytes by coculture with neurons (129). Furthermore, subsequent destruction of the neurons results in down-regulation of GLT-1 but an up-regulation of GLAST protein (130). Fornix transection results in down-regulation of both GLT-1 and GLAST in the innervation field (131). The neuronal signal mediating this interaction has proved to be somewhat elusive but may in fact be glutamate itself acting at AMPA/KA receptors on astrocytes. PKC has also been implicated in the regulation of glutamate transporter activity both directly and by altering trafficking (136).

GLUTAMATE AND GLIA

Part of "6 - The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction "

The neuroprotective action of glutamate transporters expressed by astroglia represents only one facet of the critical role astroglia play in modulating glutamatergic neurotransmission. Astrocytes express a high-affinity Na^+-dependent transporter for glycine, GlyT-1, which maintains concentrations of glycine that are subsaturating for its modulatory site (K_d = 20 nM) on the NMDA receptor in spite of μM levels in cerebrospinal fluid (CSF) (133). In addition, forebrain astroglia express serine racemase, which generates D-serine, a potent agonist at the glycine modulatory site (31). The metabolism of tryptophan by a series of enzymes expressed in glia generates both positive and negative modulators of the NMDA receptor (134). Of these, quinolinic acid is an agonist at NMDA receptors. Although its affinity for the receptor is relatively low, lack of efficient clearance mechanisms renders it a pathophysiologically significant NMDA receptor agonist. For example, the levels can achieve toxic concentrations in HIV encephalopathy with the activation of microglia and infiltration of macrophages highly expressing quinoline (135). Kynurenic acid, another metabolite of tryptophan, is a noncompetitive antagonist at the NMDA receptor acting at the glycine modulatory site. The synthesis of kynurenic acid in brain takes place almost exclusively in astrocytes and is regulated by cellular energy status, ionic environment, local 2-oxo acid concentration, and dopaminergic neurotransmission (136). Altered levels of kynurenic acid in disease states such as Huntington’s disease (HD) and schizophrenia appear to correlate with NMDA receptor dysregulation (137).

N-acetylaspartyl glutamate is an abundant neuropeptide found in many but not all glutamatergic systems and several well-characterized nonglutamatergic systems such as the locus ceruleus and motor neurons. Located in storage vessels and subject to evoke release by a Ca^{2+}-dependent process, NAAG is a noncompetitive antagonist at NMDA receptors (138) and a selective agonist at the mGluR 3 (60). It is metabolized to N-acetylaspartate and glutamate by GCP II, which is selectively expressed by astrocytes (139). Postmortem studies indicate reduced activity of GCP II in hippocampus, temporal cortex, and frontal cortex in schizophrenia, a disorder potentially involving hypofunction of NMDA receptors (140), and increased activity in motor cortex and dorsal horn, which would increase extracellular glutamate generated from NAAG in the neurodegenerative disorder ALS (141).

GLUTAMATE AND NEURODEGENERATION

Part of "6 - The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction "

The term “excitotoxicity” was coined by Olney 30 years ago to designate a selective form of neurodegeneration caused by systemic treatment of newborn rodents with monosodium glutamate (5). Degeneration of neurons with their perikarya in the arcuate nucleus of the hypothalamus was apparent 90 minutes after injection; notably, axons passing through the lesion and glia were spared. As the neurotoxic effects of glutamate analogues correlated with their excitatory potency, Olney hypothesized that the neuronal degeneration resulted from excessive activation of glutamate receptors expressed on neurons. Through the use of potent excitatory analogues of glutamate, the relevance of this phenomenon to neurodegenerative diseases became apparent. Thus, Coyle and Schwarcz reported that intrastriatal injection of kainic acid in the rat replicated the neuropathologic and synaptic neurochemical alterations that occur in HD (142). Systemic treatment with kainic acid produced a behavioral syndrome and histopathology associated with temporal lobe epilepsy. GluR-mediated neurodegeneration now has been implicated in a broad range of neurodegenerative conditions including stroke, ALS, cerebellar degeneration, and trauma (5, 143, 144).

Our knowledge of the role of iGluRs in neurodegenerative diseases has greatly increased since the realization that excitotoxicity is linked to programmed cell death (PCD) in many neurodegenerative disorders. A seminal paper by Kerr and colleagues (145) illustrated the morphological differences between apoptosis and necrosis. Subsequently, several genes that regulate cell death in *C. elegans* were identified, leading to the discovery of the caspases and Bcl-2 and homologues (146). As a consequence, the idea that cells could control their own death through the synthesis of new proteins was formulated. Subsequently, inhibitors of macromolecular synthesis were proved to prevent the naturally occurring cell death in both sensory and motor neurons. Recently, numerous molecular pathways and their components that activate or prevent neuronal cell death in response to iGluR activation have been identified.

A major effort has been undertaken to identify the specific iGluRs and the downstream events following receptor stimulation that mediate the death processes. Nevertheless, effective therapies to prevent or limit neuronal damage in neurodegenerative diseases remain elusive, reflecting an incomplete understanding of the mechanisms of neuronal death *in vivo*. It has become apparent that the boundary between apoptosis and necrosis is not well defined, leading to the realization that there exists a gradual shift from an apoptotic to a necrotic cell death in many cases, referred to as the “apoptotic-necrotic” continuum (6). Weaker insults typically promote apoptosis, whereas stronger ones favor necrosis. In other cases the apoptotic mechanism is activated along with the necrotic one, hampering attempts to distinguish the two (147).

Apoptotic cell death is typically associated with caspase activation, chromatin condensation, DNA laddering, and cell membrane blebbing that lead to cell shrinkage (6). Necrosis, on the other hand, is usually associated with the failure of ion pumps that causes cells to swell and burst, and is identified in tissues by the presence of invading macrophages and disruption of plasma membrane integrity, whereas in cell culture the absence of apoptotic markers and the rapid time course of death are the best indicators. Proteolysis by calpain and the caspases is often an early event following iGluR activation (148). Calcium overload mediated by iGluRs has a significant role in neurodegeneration (149). In addition, reactive oxygen species (ROS) play an important role in neuronal death mediated by iGluRs (150).

Mitochondria are almost invariably involved in the pathways triggered by iGluR activation. There are several general mechanisms at play, including release of proteins that activate caspases, disruption of electron transport, and alteration of cellular redox potential. The discovery that Bcl-2 is localized to the mitochondria directed attention toward the role of this relationship in cell death (151). Subsequently, it was found that Bcl-2 could prevent cytochrome c release, an inducer of apoptosis, from mitochondria. Bcl-2 also blocks the onset of the mitochondrial permeability transition (MPT) (152). The MPT represents an increase in permeability of the mitochondrial inner membrane to solutes of 1,500 daltons or less that results in membrane depolarization, uncoupling of oxidative phosphorylation, ion release, and mitochondrial swelling (153). Many of the insults that lead to the opening of the MPT function in a positive feedback manner (154). Glutamate can induce either early necrosis or delayed apoptosis in cultures of cerebellar granule cells, with mitochondrial function a critical factor that determines the mode of neuronal death (6).

Excitotoxicity is more commonly associated with necrosis because activation of iGluRs results in cation flux into the cell; however, apoptosis can follow stimulation of both AMPA/KA and NMDA receptors. The iGluRs may activate apoptosis using existing cellular components; for example, Simonian and colleagues demonstrated that kainate neurotoxicity in cerebellar granule cells was apoptotic but independent of protein synthesis (155). Alternatively, cells may require a prior insult or the addition or withdrawal of a trophic factor to become sensitive to iGluR activation that ordinarily would not be toxic. Cultured cerebellar granule cells remain resistant to AMPA–receptor-mediated toxicity when maintained in medium containing serum or insulin-like growth factor I (IGF-I), but become sensitive 4 to 5 days following the removal of trophic factors (156). In other cases, iGluR activation triggers PCD that utilizes a signaling pathway. Jiang and associates reported that NMDA receptor-mediated influx of extracellular Ca^{2+} rapidly and transiently activated ERK1/2, leading to apoptosis in cultured rat cortical neurons (157). Activation of the NMDA receptor

up-regulated p53 expression in cultured cerebellar granule cells, whereas blockade of p53 induction by an antisense oligonucleotide resulted in a complete inhibition of apoptosis (158). Similarly, systemic administration of kainate increased p53 mRNA levels in neurons exhibiting morphological features of damage within kainate-vulnerable brain regions (159).

Apoptosis may be a favored route in PCD partly because the reactive microglia that usually accompany necrosis often stimulate secondary cell death. Caspase-mediated degradation of AMPA receptor subunits occurs early during periods of cell stress in cultured rat hippocampal neurons (160). This may favor apoptosis, because levels of the AMPA receptor subunits GluR1 and GluR4 are rapidly decreased in neurons undergoing apoptosis in response to withdrawal of trophic support, whereas levels of NMDA receptor subunits NR1, NR2A, and NR2B are unchanged. Activation of calpain I by NMDA in cultured hippocampal neurons prevented the entry of cells into a caspase-dependent cell death program after the mitochondrial release of cytochrome c, possibly by inhibiting the processing of procaspase-3 and -9 into their active subunits (161). Thus, moderate NMDA receptor activation can prevent apoptosis without stimulating caspase-independent cell death, whereas a more severe stimulus favors apoptosis.

Necrotic cell death following iGluR activation is often attributed to alterations in receptor desensitization, subunit expression or other regulatory mechanisms. Human NT2-N neurons, which express calcium-permeable AMPA receptors, become vulnerable to excitotoxicity when desensitization is blocked with cyclothiazide (162). Necrosis is induced by insulin treatment within 48 hours in cultured mouse cortical neurons (163). Insulin exposure increased the level of the NR2A subunit of the NMDA receptor without altering NR1 or NR2B levels. Macromolecular synthesis inhibitors and NMDA antagonists blocked cell death, suggesting that an activity-dependent emergence of excitotoxicity contributed to insulin neurotoxicity. Cultured rat hippocampal neurons pretreated with BDNF exhibited increased levels of NR1 and NR2A, greater calcium responses to NMDA, and enhanced vulnerability to excitotoxic necrosis and reduced vulnerability to apoptosis (164). Cultured cerebellar granule cells, which show primarily an apoptotic death following KA treatment, undergo necrosis when L-type voltage-dependent calcium channels are blocked (147).

GLUTAMATE AND BRAIN DISORDERS

Part of "6 - The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction "

Neurodegenerative Diseases

HD is an autosomal dominant, progressive neurodegenerative disease that typically has its symptomatic onset in midlife. Its manifestations include chorea, dementia, and death 15 to 20 years after onset. Afflicted individuals have an expanded CAG repeat in the gene encoding huntingtin on chromosome 4, resulting in an elongated series of glutamines. The number of CAG repeats is 10 to 34 in normal individuals and 37 to 100 in HD patients (165). The identification of the HD gene has enabled the production of mouse models transgenic for huntingtin such as line R2/6, which has exon 1 with 92 repeats as well as transgenic cell lines that reiterate cellular characteristics of the disease (see the following).

HD was the first neurodegenerative disease for which iGluR-mediated neurodegeneration was implicated. Intrastriatal injection of kainate in the rat caused a striatal neuronal degeneration resembling HD (142); however, subsequent studies revealed that NMDA receptor agonists replicated the selective neuronal vulnerability in HD much more faithfully (166). Chronic treatment of rats with the mitochondrial toxin 3-nitropropionic acid elevated striatal lactate and selective striatal neuronal degeneration mediated by NMDA receptors (167). In this regard, lactate is elevated in the cerebral cortex and basal ganglia of HD patients. There is also reduced phosphocreatine/inorganic phosphate in resting muscle of HD patients, and mitochondrial electron transport enzymes are reduced in HD postmortem tissue. (See ref. 168 for review.) Consistent with these observations, mitochondria from HD lymphoblasts and fibroblasts display an increased tendency to depolarize (169). It may be that as a consequence of lowered energy levels, striatal neurons in HD can not maintain the resting membrane potential (thereby relieving the Mg^{2+} block), leading to increased $[Ca^{2+}]_i$ via NMDA receptors and ultimately cell death. Noteworthy is a report by Ferrante and associates that dietary creatine supplementation improved survival, slowed the development of striatal atrophy, and delayed the formation of huntingtin-positive aggregates in mice transgenic for huntingtin exon 1 with the expanded CAG repeat (170). Creatine may exert neuroprotective effects by increasing phosphocreatine levels or stabilizing the MPT, either of which could mitigate excitotoxicity mediated by GluRs.

Altered expression or composition of iGluR subunits may also contribute to neuronal death in HD. The editing of GluR2 mRNA is compromised in a region-specific manner in HD as well as in schizophrenia and AD, although there is still a large excess of edited GluR2 in each of these disorders (171). Chen and co-workers found that coexpression of huntingtin containing 138 repeats with NMDA receptors resulted in an increased number of functional NR1/NR2B-type receptors at the cell surface as compared to cells with normal huntingtin (172). Striatal spiny neurons are selectively vulnerable in HD and ischemia, whereas large aspiny (LA) cholinergic interneurons of the striatum are spared in these pathologic conditions. Because NR1/NR2B is the predominant NMDA receptor expressed in medium spiny neostriatal neurons, this may contribute to the selective vulnerability of these neurons in HD (172). Calabresi and associates found that membrane depolarization and inward currents produced by AMPA, KA, and NMDA were

much larger in spiny neurons than LA interneurons (173); moreover, concentrations of agonists producing reversible membrane changes in LA interneurons caused irreversible depolarization in spiny cells. The striatal and cortical neurons of R6/2 mice and mice with 94 CAG repeats displayed more rapid and increased swelling following NMDA treatment than controls, whereas AMPA and KA treatments had no differential effects. These findings suggest that NMDA antagonists or compounds that alter sensitivity of NMDA receptors may be useful in the treatment of HD (174).

The mGluRs may adversely affect iGluR function within the striatum in HD. The selective group I mGluR agonist 3,5-DHPG strongly enhanced membrane depolarization and intracellular calcium accumulation induced by NMDA application in striatal spiny neurons but not in LA interneurons, indicating a positive interaction between NMDA receptors and group I mGluRs, which are differently expressed between these two neuronal subtypes (175). Cha and colleagues found that 12-week-old R6/2 mice displayed decreased expression of AMPA- and KA- but not NMDA-type iGluR receptors compared to age-matched littermate controls. These mice also had decreased expression of mGluR1-3 that preceded the appearance of motor symptoms; therefore, altered mGluR function may contribute to subsequent pathology (176).

Approximately half of the variation in onset age for HD can be explained by the size of the repeat expansion. MacDonald and associates examined a TAA repeat polymorphism, which is closely linked to the GluR6 gene, in 258 unrelated HD-affected persons and found that younger onset age of HD was associated with linkage disequilibrium for this polymorphism (177). Rubinsztein and co-workers found that 13% of the variance in the age of onset of HD that was not accounted for by the CAG repeat size could be attributed to GluR6 genotype variation (178). These data implicate GluR6-mediated excitotoxicity in the pathogenesis of HD in addition to NMDAR-mediated neurodegeneration.

Alzheimer's disease (AD) is a progressive dementia characterized by a cortical neurodegeneration, particularly in the entorhinal cortex, hippocampal CA1 region, and subiculum. The etiology of AD is complex, with age, trauma, health, and environmental and genetic factors all playing a role (179). iGluR-mediated excitotoxicity is postulated to play a role in the neurodegeneration of AD (5). Mutations in the presenilin-1 (PS1) gene are causally linked to many cases of early-onset inherited autosomal dominant AD. Mice transgenic for the PS1M146V gene are hypersensitive to seizure-induced synaptic degeneration and necrotic neuronal death in the hippocampus (180). Cultured hippocampal neurons from PS1M146V knock in mice display increased vulnerability to glutamate, which is correlated with perturbed calcium homeostasis, increased oxidative stress, and mitochondrial dysfunction. Glutamate toxicity is potentiated by ROS mediated inhibition of EAATs; two studies have shown that ROS generated by Aß peptide inhibits astrocyte glutamate uptake (181, 182).

Immunocytochemical studies indicate that virtually all projection neurons in the hippocampus express iGluR subunits from each receptor class; however, regional differences in immunoreactivity were apparent in AD versus normal brain. In the vulnerable regions (i.e., CA1), GluR1, GluR2(4), GluR5/6/7, and NR1 were reduced, presumably owing to cell loss (183). In contrast, GluR2(4) immunolabeling appeared to be increased in the inner portion of the molecular layer of the dentate gyrus. Quantitative receptor autoradiography was also used to measure the laminar distribution of NMDA and AMPA receptors in three areas of visual cortex in control and AD postmortem human brains. The hierarchical pattern of the laminar loss of NMDA receptor binding correlated with the increasing complexity of associational visual cortices and increasing numbers of neurofibrillary tangles; however, AMPA receptor losses did not directly correlate with the pathology (184). Hyman and colleagues found no difference for the pattern of immunostaining between control and AD in either hippocampi or adjacent temporal cortices for GluR1, GluR2/3, and GluR4 (185); however, age-related loss of GluR2/3 immunoreactivity prior to degeneration has been reported in nucleus basalia of Meynert (186) and entorhinal cortex (187), suggesting that an increase in calcium permeability of AMPA receptors may leave these neurons vulnerable to degeneration in AD. Western blot analysis revealed average reductions of approximately 40% for GluR1 and GluR2/3 in the entorhinal cortex of patients with AD pathology versus age-matched controls, but neither GluR1 nor GluR2/3 protein concentration correlated significantly with tangle density (188). Thus, the relationship between excitotoxicity and neuronal loss in AD is complex and requires additional investigation.

Amyotrophic lateral sclerosis is a disorder characterized by a selective and progressive degeneration of motor neurons in the spinal cord and pyramidal neurons in the motor cortex, with onset in midlife (189). Death results from complications of the progressive paralysis. The two forms of ALS, sporadic and familial (FALS), have similar clinical symptoms and neuropathology, although the latter only accounts for 10% of the cases. Rosen and associates first reported a tight genetic linkage between FALS and the gene encoding Cu/Zn superoxide dismutase (SOD1), and identified 11 different SOD1 missense mutations in 13 different FALS families (190). Expression of high levels of a mutant form of human SOD1 for which the glycine at position 93 was replaced with an alanine (G93ASOD1; a change that has little effect on enzyme activity) caused a progressive motor neuron disease resulting in death by 6 months in transgenic mice (191). Because the mouse gene for SOD1 is unaffected in the transgenic mice, the results indicate that these mutations in SOD1 cause a gain-of-function that results in motor neuron death.

The reason for the selective vulnerability of motor neurons in ALS is unknown. Various molecular and neurochemical features of human motor neurons may render this cell group differentially vulnerable to such insults. Motor neurons are large cells with long axonal processes that require a high level of mitochondrial activity and have greater neurofilament content than other neuronal groups. Motor neurons have a very high expression of the cytosolic free radical scavenging enzyme Cu/ZnSOD1, which may render this cell group more vulnerable to genetic or posttranslational alterations interfering with the function of this protein. The low expression of calcium binding proteins and GluR2 AMPA receptor subunit by vulnerable motor neuron groups may render them unduly susceptible to calcium-mediated toxic events following GluR activation (192).

High levels of mRNA for GluR1, GluR3, and GluR4 are expressed in normal human spinal motor neurons; however, GluR2 subunit mRNA was not detectable in this cell group, predicting that normal human spinal motor neurons express calcium-permeable AMPA receptors unlike most neuronal groups in the human CNS (193); however, this has not been borne out in studies of mouse spinal cord in the context of the mouse models of ALS, where GluR2 is well represented in spinal cord motor neurons (194). AMPA or kainate exposure triggers substantial mitochondrial calcium loading in motor neurons, but causes little mitochondrial accumulation in forebrain GABAergic interneurons, neurons that express large numbers of calcium-permeable AMPA/kainate channels but do not degenerate in ALS. Brief exposure to either AMPA or kainate caused mitochondrial depolarization in motor neurons, whereas these effects were only observed in the GABAergic neurons after exposure to the nondesensitizing AMPA receptor agonist kainate. Finally, blocking mitochondrial calcium uptake attenuated AMPA/kainate receptor-mediated motor neuron injury. Thus, mitochondrial calcium uptake and consequent ROS generation may be central to the injury process (195). Quantification of mRNA expression in spinal cord showed a significant widespread loss of NR2A from both dorsal and ventral horns with losses of 55% and 78%, respectively, in ALS as compared to control. These results were substantiated by analysis of spinal cord homogenates, which showed a significant total decrease of 50% in NR2A message for ALS as compared to control (196).

Riluzole, which attenuates the glutamate neurotransmitter system, has been shown to prolong survival in patients with ALS (197). Riluzole affects neurons using three mechanisms: by inhibiting excitatory amino acid release, inhibiting events following stimulation of iGluRs, and stabilizing the inactivated state of voltage-dependent sodium channels. Mennini and associates studied inotropic glutamate receptor subtypes and the effect of chronic treatment with NBQX in the spinal cord of motor neuron disease (mnd) mice. NBQX significantly improved the behavioral scores in mnd mice. These findings suggest that selective antagonism of inotropic non–NMDA receptors may be of value in the treatment of motor neuron disease (198). Further research may allow the development of therapies that target specific glutamate receptor subunits and modulate “downstream” events within motor neurons, aimed at protecting vulnerable molecular targets in specific populations of ALS patients.

SCHIZOPHRENIA

Part of "6 - The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction "

Kim and associates reported diminished concentrations of glutamate in CSF of patients with schizophrenia and first proposed that hypofunction of glutamatergic systems might cause the disorder (199). This finding has been replicated by some studies but not by several others (200, 201, 202 and 203). In a postmortem study, Tsai and associates (140) studied eight brain regions and found decreased concentrations of glutamate and aspartate in the frontal cortex and decreased concentration of glutamate in the hippocampus of patients with schizophrenia as compared to controls. Furthermore, the concentration of NAAG was increased in the hippocampus and the activity of GCPII was selectively reduced in the frontal cortex, temporal cortex, and hippocampus of people with schizophrenia. Subsequent studies with magnetic resonance spectroscopy have revealed significant reductions in the level of N-acetylaspartate (NAA), the product of NAAG by GCPII, in the very same regions—frontal cortex, temporal cortex, and hippocampus (204).

Initial ligand binding studies in postmortem schizophrenic brain have revealed increases in the non–NMDA iGLURs in the prefrontal cortex (205, 206) and decreases in the hippocampus (207, 208). Strychnine-insensitive binding, which labels the glycine modulatory site on the NMDA receptor, is increased throughout the primary sensory cortex and related associational fields in schizophrenia (209). Molecular approaches have shown a reduction in mRNA encoding luR2 in the hippocampus and parahippocampus of people with schizophrenia, and reduced editing of GluR2 in the prefrontal cortex (210, 211). Although the density of NMDA receptors in the prefrontal cortex of people with schizophrenia was normal, the relative subunit composition differed significantly from controls with a large increase observed for NR2D (212).

A convincing link between glutamatergic dysfunction and schizophrenia came from anecdotal and subsequent controlled studies of the neuropsychologic effects of dissociative anesthetics, which are noncompetitive antagonists of the NMDA receptor (213). When chronically abused, PCP produces a syndrome in normal individuals that closely resembles schizophrenia and exacerbates symptoms in patients with chronic schizophrenia. Subanesthetic doses of ketamine administered to normal subjects produces positive symptoms, including delusions and thought disorder, negative symptoms, and frontal lobe cognitive impairments characteristic of schizophrenia (214). When administered

to schizophrenic subjects, subanesthetic doses of ketamine exacerbate delusion, hallucinations, and thought disorders that are consistent with the patient's typical pattern of psychotic relapse (215, 216). These effects are attenuated by the atypical antipsychotic clozapine but not haloperidol.

Although acute administration of ketamine to normal subjects causes increased (217) prefrontal cortical perfusion, chronic exposure to PCP is associated with the classical "hypofrontality" of schizophrenia (218, 219). Chronic PCP treatment produced more perseveration and fewer nonspecific cognitive deficits in monkeys that persisted after discontinuation. Notably, these memory deficits were prevented by clozapine treatment (220).

Administration of NMDA receptor antagonists markedly increased the release of dopamine and glutamate in prefrontal cortex and subcortical structures in rats (221, 222), which was associated with impaired performance on a memory task sensitive to prefrontal cortical function (217); these alterations could be ameliorated by treatment with an AMPA/KA receptor antagonist. Furthermore, administration of a group I mGluR agonist blocked PCP-induced glutamate release without affecting dopamine release (223). These effects of NMDA receptor antagonism observed in the rodent have been shown to be comparable to humans in a positron emission tomographic study in which [^{11}C]-raclopride binding in striatum was used to measure dopamine release; subanesthetic doses of ketamine cause increased dopamine release in human subjects (224).

If the symptoms of schizophrenia result from hypofunction of NMDA receptors, then agents that enhance NMDA receptor function would be predicted to reduce symptoms. Because full agonists could be excitotoxic, studies have focused primarily on agents that act via the glycine modulatory site (225). Electrophysiologic studies in the hippocampal slice indicate that the glycine modulatory site is not fully occupied because of efficient transport of glycine by the GLYT-1 transporter on astroglia so that the modulatory site is subject to pharmacologic manipulation (133). In most of the studies, the drugs were added to typical antipsychotics in stable patients with prominent negative symptoms. Javitt and colleagues have performed a series of placebo-controlled crossover trials in which high doses of glycine (30 to 60 g per day) were added to antipsychotic drugs. They demonstrated improvement in negative symptoms and cognitive function without effects on psychotic symptoms or extrapyramidal side effects (226, 227 and 228). Tsai and associates added D-serine at a dose of 30 mg per kg to typical antipsychotic drugs for 8 weeks and found significant improvements in negative symptoms, cognitive function (as measured by the Wisconsin Card Sorting test), and psychosis (229). The more robust effect of D-serine may reflect the fact that it has better penetrance of the blood–brain barrier, is a full agonist and has a higher affinity than glycine.

Another drug that has been extensively studied is the antitubercular agent, D-cycloserine, which is a partial agonist at the glycine modulatory site with 60% efficacy and readily crosses the blood–brain barrier (230). A blinded dose finding study in patients receiving typical antipsychotics and exhibiting prominent negative symptoms revealed a U-shaped dose–response curve with significant reductions in negative symptoms and improvement in cognitive function at 50 mg per day (231). van Berckel and associates observed improvement in negative symptoms at a D-cycloserine dose of 100 mg per day in a small open trial with medication-free schizophrenics (232). D-cycloserine at 50 mg per day significantly improved negative symptoms when added to conventional antipsychotics in an 8-week fixed dose placebo-controlled parallel group trial with patients meeting criteria for deficit syndrome of schizophrenia (233); however, performance on a cognitive battery did not change. Notably, full response for negative symptoms was not achieved until after 4 to 6 weeks of treatment.

It was of interest to determine whether the addition of D-cycloserine would have further ameliorative effects in clozapine responders because clozapine has substantial effects on negative symptoms in many patients who respond poorly to typical antipsychotics. To the contrary, two separate trials of D-cycloserine at 50 mg per day added to clozapine resulted in worsening of negative symptoms (234, 235). In contrast, trials in which the full agonists, glycine or D-serine, were added to clozapine yielded no additional change in negative symptoms or cognitive function (236, 237). A plausible explanation for these findings is that clozapine may exert its effects on negative symptoms and cognitive functions in part by increasing occupancy of the glycine modulatory site on the NMDA receptor, thereby transforming the partial agonist D-cycloserine into an antagonist. Support for this inference comes from electrophysiologic studies in the hippocampal slice where clozapine enhances NMDA receptor currents (238).

As hippocampal interneurons appear more sensitive to NMDA receptor antagonists owing to the presence of NR2C (239), hypofunction of these NMDA receptors because of an excess of endogenous antagonists such as NAAG or kynurenic acid could account for many features of schizophrenia. The disinhibition of cortico-hippocampal efferents appears to increase subcortical dopamine release associated with positive symptoms (240). This would also interfere with the precision of cortical/hippocampal activations consistent with schizophrenic subjects' inability to increase hippocampal neuronal activity in a memory task because of a ceiling effect (241). The effects of glycine modulatory site activation, particularly on negative symptoms and cognitive impairment, are consistent with this model. Finally, the fact that ketamine reproduces the eye tracking impairments found in schizophrenics and some of their first-degree relatives suggest that NMDA receptor hypofunction could be part of the endophenotype (242).

Age-Associated Memory Impairment

Glutamate receptors have also been implicated in the functional decline seen in normal aging in the absence of neurodegeneration. Spatial memory is particularly vulnerable to aging (243), and is also disrupted by pharmacologic blockade of NMDA receptor function (244) or hippocampal knockout of NR1 (245). Electrophysiologic investigations of aging in rat hippocampus have revealed that certain aspects of excitatory synaptic transmission are unaffected or even compensatory, whereas others are compromised (246). One component of synaptic transmission that is compromised is maintenance and induction of long-term potentiation (LTP) that could be related to impaired NMDA receptor-mediated processes and the decreased stability of spatial information coding by “place cell” in aged rats (247). At the regional level, receptor binding studies have reported decreases in NMDA binding in hippocampus (248) with confirmatory declines in mRNA expression (249); however, other studies suggested that there is no change in NMDA receptors with aging (250).

In studying age-related changes in receptors, it is particularly important to be able to take the analysis from the regional level to that of cell classes, circuits, individual neuronal compartments, and synapses, because the changes are very likely to be cell-, circuit-, and synapse-specific and therefore difficult to resolve at the regional level; for example, age-related shifts in NR1, have been reported in the molecular layer of the dentate gyrus (251). The projection from the entorhinal cortex (ERC) to the DG is strictly confined to the outer molecular layer (OML), that is, the distal dendrites of granule cells; whereas other excitatory inputs terminate in a nonoverlapping fashion in the inner molecular layer (IML), the proximal dendrites. Aged monkeys, compared to young adult monkeys, exhibit a decrease in the fluorescence intensity for NR1 in the OML of the DG as compared to the IML. Given the tight laminar organization of these circuits, this pattern means that decreased NR1 levels primarily affect the input from the ERC, pointing to the ERC input to the hippocampus as a key element in age-related changes, and suggests that the intradendritic distribution of a neurotransmitter receptor is modified in an age-related and circuit-specific manner (251).

Although these results suggest that age-related circuit-specific shifts in NMDA receptors might underlie memory defects, they were not done in behaviorally characterized animals, and need to be followed up in the context of behavior. In addition, the data on NR1 changes were limited to the dendrite and did not directly address the GluRs at the synapse. Thus, these data need to be extended, particularly in the nonhuman primate model. Species issues may be particularly relevant because in rat hippocampus decreases in presynaptic markers such as synaptophysin correlate with age-associated memory impairment more directly than do any age-related shifts in GluRs (252, 253).

CONCLUSION

Part of "6 - The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction "

In closing, glutamate sits at the epicenter of signal transduction in brain, not only mediating excitatory neurotransmission, but also modulating neuroplasticity at the genetic, synaptic, and structural levels. Furthermore, insufficient glutamatergic signaling causes the degeneration of immature neurons through apoptosis (254), whereas excessive activation of iGluRs kills neurons through necrosis and/or apoptosis. Dysregulation of glutamatergic neurotransmission has been implicated in an expanding number of neuropsychiatric disorders. Although the clinical pharmacology of glutamate is currently embryonic, the remarkable advances in the molecular characterization of this system hold promise for the development of a rich array of specific drugs in the future.

ACKNOWLEDGMENT

Part of "6 - The Diverse Roles of L-Glutamic Acid in Brain Signal Transduction "

Dr. Coyle serves as a consultant for Jansson, Abbott, and Prestwick and owns stock in Merck, Celera, Genzyme, and PESystems.

REFERENCES

1. Hollmann M, Heinemann S. Cloned glutamate receptors. *Annu Rev Neurosci* 1994;17:31–108.

2. Balázs R, Jorgensen OS, Hack N. NMDA promotes the survival of cerebellar granule cells in culture. *Neuroscience* 1988;27:437–451.

3. Wilson MT, Keith CH. Glutamate modulation of dendrite outgrowth: alterations in the distribution of dendritic microtubules. *J Neurosci Res* 1998;52:599–611.

4. Komuro H, Rakic P. Modulation of neuronal migration by NMDA receptors. *Science* 1993;260:95–97.

5. Olney JW, Wozniak DF, Farber NB. Excitotoxic neurodegeneration in Alzheimer disease. New hypothesis and new therapeutic strategies. *Arch Neurol* 1997;54:1234–1240.

6. Ankarcrona M, Dypbukt JM, Bonfoco E, et al. Glutamate-induced neuronal death: a succession of necrosis or apoptosis depending on mitochondrial function. *Neuron* 1995;15:961–973.

7. Choi DW. Glutamate neurotoxicity and diseases of the nervous system. *Neuron* 1988;1:623–634.

8. Robinson MB, Coyle JT. Glutamate and related acidic excitatory neurotransmitters: from basic science to clinical application. *Faseb J* 1987;1:446–455.

9. Dingledine R, Borges K, Bowie D, et al. The glutamate receptor ion channels. *Pharmacol Rev* 1999;51:7–61.

10. Monyer H, Seeburg PH, Wisden W. Glutamate-operated channels: developmentally early and mature forms arise by alternative splicing. *Neuron* 1991;6:799–810.

11. Rueter SM, Burns CM, Coode SA, et al. Glutamate receptor RNA editing in vitro by enzymatic conversion of adenosine to inosine. *Science* 1995;267:1491–1494.

12. Seeburg PH. The role of RNA editing in controlling glutamate receptor channel properties. *J Neurochem* 1996;66:1–5.

13. Brusa R, Zimmermann F, Koh DS, et al. Early-onset epilepsy and postnatal lethality associated with an editing-deficient GluR-B allele in mice. *Science* 1995;270:1677–1680.

14. Pellegrini-Giampietro DE, Gorter JA, Bennett MV, et al. The GluR2 (GluR-B) hypothesis: Ca(2+)-permeable AMPA receptors in neurological disorders. *Trends Neurosci* 1997;20:464–470.

15. Kohler M, Burnashev N, Sakmann B, et al. Determinants of Ca2+ permeability in both TM1 and TM2 of high affinity kainate receptor channels: diversity by RNA editing. *Neuron* 1993;10:491–500.

16. Burnashev N, Villarroel A, Sakmann B. Dimensions and ion selectivity of recombinant AMPA and kainate receptor channels and their dependence on Q/R site residues. *J PhysioI (Lond)* 1996;496:165–173.

17. Wong LA, Mayer ML. Differential modulation by cyclothiazide and concanavalin A of desensitization at native alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid- and kainate-preferring glutamate receptors. *MoI PharmacoI* 1993;44:504–510.

18. Staubli U, Rogers G, Lynch G. Facilitation of glutamate receptors enhances memory. *Proc NatI Acad Sci USA* 1994;91:777–781.

19. Bowie D, Mayer ML. Inward rectification of both AMPA and kainate subtype glutamate receptors generated by polyamine-mediated ion channel block. *Neuron* 1995;15:453–462.

20. Aramori I, Nakanishi S. Signal transduction and pharmacological characteristics of a metabotropic glutamate receptor, mGluR1, in transfected CHO cells. *Neuron* 1992;8:757–765.

21. Luthi-Carter R, Berger UV, Barczak AK, et al. Isolation and expression of a rat brain cDNA encoding glutamate carboxypeptidase II. *Proc NatI Acad Sci USA* 1998;95:3215–3220.

22. Coyle JT. Neurotoxic action of kainic acid. *J Neurochem* 1983;41:1–11.

23. Wilding TJ, Huettner JE. Differential antagonism of alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid-preferring and kainate-preferring receptors by 2,3-benzodiazepines. *MoI PharmacoI* 1995;47:582–587.

24. Fayyazuddin A, Villarroel A, Le Goff A, et al. Four residues of the extracellular N-terminal domain of the NR2A subunit control high-affinity Zn2+ binding to NMDA receptors. *Neuron* 2000;25:683–694.

25. Monaghan DT, Bridges RJ, Cotman CW. The excitatory amino acid receptors: their classes, pharmacology, and distinct properties in the function of the central nervous system. *Annu Rev PharmacoI ToxicoI* 1989;29:365–402.

26. Wisden W, Seeburg PH. A complex mosaic of high-affinity kainate receptors in rat brain. *J Neurosci* 1993;13:3582–3598.

27. Keinanen K, Wisden W, Sommer B, et al. A family of AMPA-selective glutamate receptors. *Science* 1990;249:556–560.

28. Watanabe M, Inoue Y, Sakimura K, Mishina M. Distinct distributions of five N-methyl-D-aspartate receptor channel subunit mRNAs in the forebrain. *J Comp NeuroI* 1993;338:377–90.

29. Huntley GW, Vickers JC and Morrison JH. Cellular and synaptic localization of NMDA and non-NMDA receptor subunits in neocortex: organizational features related to cortical circuitry, function and disease. *Trends Neurosci* 1994;17:536–543.

30. Conti F, Weinberg RJ. Shaping excitation at glutamatergic synapses. *Trends Neurosci* 1999;22(10):451–8

31. Rogers SW, Hughes TE, Hollmann M, et al. The characterization and localization of the glutamate receptor subunit GluR1 in the rat brain. *J Neurosci* 1991;11:2713–2724.

32. Petralia RS, Wenthold RJ. Light and electron Immunocytochemical localization of AMPA-selective glutamate receptors in the rat brain. *J Comp NeuroI* 1992;318:329–354.

33. Martin LJ, Blackstone CD, Levey AI, et al. AMPA glutamate receptor subunits are differentially distributed in rat brain. *Neurosci* 1993; 53:327–358.

34. Tachibana M, Wenthold RJ, Morioka H, et al. Light and electron microscopi immunocytochemical localizaton of AMPA-selective glutamate receptors in the rat spinal cord. *J Comp NeuroI* 1994;344:431–54.

35. Vickers JC, Huntley GW, Edwards AM, et al.. Quantitative localization of AMPA/kainate and kainate glutamate receptor subunit immunoreactivity in neurochemically identified subpopulations of neurons in the prefrontal cortex of the macaque monkey. *J Neurosci* 1993;13:2982–2992.

36. Huntley GW, Rogers SW, Moran T, et al. Selective distribution of kainate receptor subunit immunoreactivity in monkey neocortex revealed by a monoclonal antibody that recognizes glutamate receptor subunits GluR5/6/7. *J Neurosci* 1993;13:2965–2981.

37. Siegel SJ, Brose N, Janssen W, et al. Regional, cellular, and ultrastructural distribution of N-methyl-D-aspartate receptor subunit 1 in monkey hippocampus. *Proc NatI Acad Sci USA* 1994;91:564–568.

38. Siegel SJ, Janssen WG, Tullai JW, et al. Distribution of the excitatory amino acid receptor subunits GluR2(4) in monkey hippocampus and colocalization with subunits GluR5-7 and NMDAR1. *J Neurosci* 1995;15:2707–2719.

39. He Y, Janssen WGM, Morrison JH. Synaptic coexistence of AMPA and NMDA receptors in the rat hippocampus: A post-embedding immunogold study. *J Neurosci Res* 1998;54:444–449.

40. Nusser Z, Lujan R, Laube G, et al. Cell type and pathway dependence of synaptic AMPA receptor number and vairability in the hippocampus. *Neuron* 1998;21:545–559.

41. Takumi Y, Ramirez-Leon V, Laake P, et al. Different modes of expression of AMPA and NMDA receptors in hippocampal synapses. *Nature Neurosci* 1999;2:618–624.

42. Petralia RS, Esteban JA, Wang YX, et al. Selective acquition of AMPA receptors over postnatal development suggests a molecular basis of silent synapses. *Nature Neurosci* 1999;2:31–36.

43. Kharazia VN, Weinberg RJ. Immunogold localization of AMPA and NMDA receptors in somatic sensory cortex of albino rat. *J Comp NeuroI* 1999;412:292–302.

44. He Y, Janssen WGM, Rothstein JD, et al. Differential synaptic localization of the glutamate transporter EAAC1 and glutamate receptor subunit GluR2 in the rat hippocampus. *J Comp NeuroI* 2000;418:255–269.

45. Jonas P, Racca C, Sakmann B, et al. Differences in calcium permeability of AMPA type glutamate receptor channels in neocortical neurons caused by differential GluR B subunit expression. *Neuron* 1994;12:1281–1289.

46. Geiger JRP, Melcher T, Koh DS, et al. Relative abundance of subunit mRNAs determines gating and CA^{2+} permeability of AMPA receptors in principal neurons and interneurons and interneurons in rat CNS. *Neuron* 1995;15:193–204.

47. Koh DS, Geiger JRP, Jonas P, et al. Ca2+ permeable AMPa and NMDA receptor channels in basket cells of rat hippocampal dentate gyrus. *J PhysioI* 1995;485:383–402.

48. Spruston N, Jonas P, Sakmann B. Dendritic glutamate receptor channels in rat hippocampal CA3 and CA1 pyramidal neurons. *J PhysioI* 1995;482:325–352.

49. Leranth C, Szeidemann Z, Hsu M, et al. AMPA receptors in the rat and primate hippocampus; A posible absence of GluR2/3 subunits in most interneurons. *Neuroscience* 1996;70:631–652.

50. Martin LJ, Blackstone CD, Levey AI, et al. AMPA glutamate receptor subunits are differentially distributed in the rat brain. *Neuroscience* 1993;53:327–358.

51. Vickers JC, Huntley GW, Edwards AM, et al. Quantitative localization of AMPa/kainae and kainate glutamate receptor subunit immunoreactivity in neurochemically identified subpopulations of neurons in the prefrontal cortex of the macaque monkey. *J Neurosci* 1993;13:2982–2992.

52. Vissavajjhala P, Janssen WGM, Hu Y, et al. (1996) Synaptic distribution of AMPA-GluR2 subunit and its colocalization with calcium-binding proteins in rat cerebral cortex: an immunohistochemical study using a GluR2-specific monoclonal antibody. *Exp Neurol* 1996;142:296–312.

53. Petralia RS, Wang YX, Mayat E, et al. Glutamate receptor subunit 2-selective antibody shows a differential distribution of calxium-impermeable AMPA receptors among populations of neurons. *J Comp Neurol* 1997;385:456–476.

54. He Y, Janssen WGM, Vissavajjhala P, et al. Synaptic distribution of GluR2 in hippocampal GABAergic interneurons and pyramidal cells: a double label immunogold analysis. *Exp Neurol* 1998;150:1–13.

55. Isaac JTR, Nicoll RA, Malenka RC. Evidence for silent synapses: implications for the expression of LTP. *Neuron* 1995;15:427–434.

56. Liao D, Hessler NA, Malinow R. Activation of postsynaptically silent synapses during pairing-induced LTP in CA1 region of hippocampal slice. *Nature* 1995;375:400–404.

57. Shi SH, Hayashi Y, Petralia RS, et al. Rapid spine delivery and redistribution of AMPA receptors after synaptic NMDA receptor activation. *Science* 1999;284:1811–1816.

58. Castillo PE, Malenka RC, Nicoll RA. Kainate receptors mediate a slow postsynaptic current in hippocampal CA3 neurons. *Nature* 1997;388:182–186.

59. Mulle C, Sailer A, Perez-Otano I, et al. Altered synaptic physiology and reduced susceptibility to kainate-induced seizures in GluR6-deficient mice. *Nature* 1998;392:601–605.

60. Clarke VRJ, Ballyk BA, Hoo KH, et al. A hippocampal GluR5 kainate receptor regulating inhibitory synaptic transmission. *Nature* 1997;389:599–602.

61. Petralia RS, Wang YX, Wenthold RJ. Histological and ultrastructural localization of the kainate receptor subunits, KA2 and GluR6/7, in the rat neurvous system using selective antipeptide antibodies. *J Comp Neurol* 1994;349:85–110.

62. Wolosker H, Blackshaw S, Snyder SH. Serine racemase: a glial enzyme synthesizing D-serine to regulate glutamate-N-methyl-D-aspartate neurotransmission. *Proc Natl Acad Sci USA* 1999;96:13409–414.

63. Choi DW, Koh JY. Zinc and brain injury. *Annu Rev Neurosci* 1998;21:347–375.

64. Rock DM, Macdonald RL. Polyamine regulation of N-methyl-D-aspartate receptor channels. *Annu Rev Pharmacol Toxicol* 1995;35:463–482.

65. Sullivan JM, Traynelis SF, Chen HS, et al. Identification of two cysteine residues that are required for redox modulation of the NMDA subtype of glutamate receptor. *Neuron* 1994;13:929–936.

66. Jentsch JD, Roth RH. The neuropsychopharmacology of phencyclidine: from NMDA receptor hypofunction to the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology* 1999;20:201–225.

67. Ciabarra AM, Sullivan JM, Gahn LG, et al. Cloning and characterization of chi-1: a developmentally regulated member of a novel class of the inotropic glutamate receptor family. *J Neurosci* 1995;15:6498–6508.

68. Laurie DJ, Seeburg PH. Regional and developmental heterogeneity in splicing of the rat brain NMDAR1 mRNA. *J Neurosci* 1994;14:3180–3194.

69. Monyer H, Burnashev N, Laurie DJ, et al. Developmental and regional expression in the rat brain and functional properties of four NMDA receptors. *Neuron* 1994;12:529–540.

70. Gazzaley AH, Benson DL, Huntley GW, et al. Differential subcellular regulation of NMDAR1 protein and mRNA in dendrites of dentate gyrus granule cells after perforant path transection. *J Neurosci* 1997;17:2006–2017.

71. Janssen WGM, Adams MM, Andrews G, et al. An ultrastructural analysis of NMDA receptor plasticity in the molecular layer of the dentate gyrus following entorhinal cortex lesions: A quantitative immunogold study. *Soc Neurosci Abstr* 2000;26:357.

72. Watanabe M, Fukaya M, Sakimura K, et al. Selective scarcity of NMDA receptor channel subunits in the stratum lucidum (mossy fiber-recipient layer) of the mouse hippocampal CA3 subfield. *Eur J Neurosci* 1998;10:478–487.

73. Petralia RS, Wang YX, Wenthold RJ. The NMDA receptor subunits NR2A and NR2B show histological and ultrastructural localization patterns similar to those of NR1. *J Neurosci* 1994;14:6102–20.

74. Conti F, Barbaresi P, Melone M, et al. Neuronal and glial localization of NR1 and NR2A/B subunits of the NMDA receptor in the human cerebral cortex. *Cereb Cortex* 1999;9:110–120.

75. Fritschy JM, Weinmann O, Wenzel A, et al. Synapse-specific localization of NMDA and GABA(A) receptor subunits revealed by antigen-retrieval immunohistochemistry. *J Comp Neurol* 1998;390:194–210.

76. Watanabe M, Fukaya M, Sakimura K, et al. Selective scarcity of NMDA receptor channel subunits in the stratum lucidum (mossy fibre-recipient layer) of the mouse hippocampal CA3 subfield. *Eur J Neurosci* 1998;10:478–487.

77. Tang YP, Shimizu E, Dube GR, et al. Genetic enhancement of learning and memory in mice. *Nature* 1999;401:63–69.

78. Mayer ML, Benveniste M, Patneau DK, et al. Pharmacologic properties of NMDA receptors. *Ann NY Acad Sci* 1992;648:194–204.

79. Lodge D, Johnson KM. Noncompetitive excitatory amino acid receptor antagonists. *Trends Pharmacol Sci* 1990;11:81–86.

80. Williams K. Ifenprodil discriminates subtypes of the N-methyl-D-aspartate receptor: selectivity and mechanisms at recombinant heteromeric receptors. *Mol Pharmacol* 1993;44:851–859.

81. Sladeczek F, Pin JP, Recasens M, et al. Glutamate stimulates inositol phosphate formation in striatal neurones. *Nature* 1985;317:717–719.

82. Conn PJ, Pin JP. Pharmacology and functions of metabotropic glutamate receptors. *Annu Rev Pharmacol Toxicol* 1997;37:205–237.

83. Brown EM, Gamba G, Riccardi D, et al. Cloning and characterization of an extracellular Ca(2+)-sensing receptor from bovine parathyroid. *Nature* 1993;366:575–580.

84. Kaupmann K, Huggel K, Heid J, et al. Expression cloning of GABA(B) receptors uncovers similarity to metabotropic glutamate receptors. *Nature* 1997;386:239–246.

85. Nakanishi S. Molecular diversity of glutamate receptors and implications for brain function. *Science* 1992;258:597–603.

86. O'Hara PJ, Sheppard PO, Thogersen H, et al. The ligand-binding domain in metabotropic glutamate receptors is related to bacterial periplasmic binding proteins. *Neuron* 1993;11:41–52.

87. Martin LJ, Blackstone CD, Huganir RL, et al. Cellular localization of a metabotropic glutamate receptor in rat brain. *Neuron* 1992;9:259–270.

88. Romano C, Sesma MA, McDonald CT, et al. Distribution of metabotropic glutamate receptor mGluR5 immunoreactivity in rat brain. *J Comp Neurol* 1995;355:455–469.

89. Balazs R, Miller S, Romano C, et al. Metabotropic glutamate receptor mGluR5 in astrocytes: pharmacological properties and agonist regulation. *J Neurochem* 1997;69:151–163.

90. Ohishi H, Shigemoto R, Nakanishi S, et al. Distribution of the messenger RNA for a metabotropic glutamate receptor, mGluR2, in the central nervous system of the rat. *Neuroscience* 1993;53:1009–1018.

91. Neki A, Ohishi H, Kaneko T, et al. Metabotropic glutamate receptors mGluR2 and mGluR5 are expressed in two non-overlapping populations of Golgi cells in the rat cerebellum. *Neuroscience* 1996;75:815–826.

92. Tanabe Y, Nomura A, Masu M, et al. Signal transduction, pharmacological properties, and expression patterns of two rat metabotropic glutamate receptors, mGluR3 and mGluR4. *J Neurosci* 1993;13:1372–1378.

93. Bradley SR, Rees HD, Yi H, et al. Distribution and developmental regulation of metabotropic glutamate receptor 7a in rat brain. *J Neurochem* 1998;71:636–645.

94. Saugstad JA, Kinzie JM, Shinohara MM, et al. Cloning and expression of rat metabotropic glutamate receptor 8 reveals a distinct pharmacological profile. *Mol Pharmacol* 1997;51:119–125.

95. Vardi N, Duvoisin R, Wu G, et al. Localization of mGluR6 to dendrites of ON bipolar cells in primate retina. *J Comp Neurol* 2000;423:402–412.

96. Nakajima Y, Iwakabe H, Akazawa C, et al. Molecular characterization of a novel retinal metabotropic glutamate receptor mGluR6 with a high agonist selectivity for L-2-amino-4-phosphonobutyrate. *J Biol Chem* 1993;268:11868–11873.

97. Petralia RS, Wang YX, Niedzielski AS, et al. The metabotropic glutamate receptors, mGluR2 and mGluR3, show unique postsynaptic, presynaptic and glial localizations. *Neuroscience* 1996;71:949–976.

98. Shigemoto R, Kinoshita A, Wada E, et al. Differential presynaptic localization of metabotropic glutamate receptor subtypes in the rat hippocampus. *J Neurosci* 1997;17:7503–7522.

99. Takumi Y, Matsubara A, Rinvik E, Ottersen OP. The arrangement of glutamate receptors in excitatory synapses. *Ann N Y Acad Sci* 1999;868:474–482.

100. Kinoshita A, Shigemoto R, Ohishi H, et al. Immunohistochemical localization of metabotropic glutamate receptors, mGluR7a and mGluR7b, in the central nervous system of the adult rat and mouse: a light and electron microscopic study. *J Comp Neurol* 1998;393:332–352.

101. Kosinski CM, Risso Bradley S, Conn PJ, et al. Localization of metabotropic glutamate receptor 7 mRNA and mGluR7a protein in the rat basal ganglia. *J Comp Neurol* 1999;415:266–284.

102. Bradley SR, Standaert DG, Rhodes KJ, et al. Immunohistochemical localization of subtype 4a metabotropic glutamate receptors in the rat and mouse basal ganglia. *J Comp Neurol* 1999;407:33–46.

103. Lujan R, Nusser Z, Roberts JD, et al. Perisynaptic location of metabotropic glutamate receptors mGluR1 and mGluR5 on dendrites and dendritic spines in the rat hippocampus. *Eur J Neurosci* 1996;8:1488–500.

104. Petralia RS, Wang YX, Singh S, et al. A monoclonal antibody shows discrete cellular and subcellular localizations of mGluR1 alpha metabotropic glutamate receptors. *J Chem Neuroanat* 1997;13:77–93.

105. Mateos JM, Benitez R, Elezgarai I, et al. Immunolocalization of the mGluR1b splice variant of the metabotropic glutamate receptor 1 at parallel fiber-Purkinje cell synapses in the rat cerebellar cortex. *J Neurochem* 2000;74:1301–1309.

106. Joly C, Gomeza J, Brabet I, et al. Molecular, functional, and pharmacological characterization of the metabotropic glutamate receptor type 5 splice variants: comparison with mGluR1. *J Neurosci* 1995;15:3970–3981.

107. Gereau RWT, Conn PJ. Roles of specific metabotropic glutamate receptor subtypes in regulation of hippocampal CA1 pyramidal cell excitability. *J Neurophysiol* 1995;74:122–129.

108. Bond A, Monn JA, Lodge D. A novel orally active group 2 metabotropic glutamate receptor agonist: LY354740. *Neuroreport* 1997;8:1463–1466.

109. Wroblewska B, Wroblewski JT, Pshenichkin S, et al. N-acetylaspartylglutamate selectively activates mGluR3 receptors in transfected cells. *J Neurochem* 1997;69:174–181.

110. Buisson A, Yu SP, Choi DW. DCG-IV selectively attenuates rapidly triggered NMDA-induced neurotoxicity in cortical neurons. *Eur J Neurosci* 1996;8:138–143.

111. Bruno V, Copani A, Knopfel T, et al. Activation of metabotropic glutamate receptors coupled to inositol phospholipid hydrolysis amplifies NMDA-induced neuronal degeneration in cultured cortical cells. *Neuropharmacology* 1995;34:1089–1098.

112. Slusher BS, Vornov JJ, Thomas AG, et al. Selective inhibition of NAALADase, which converts NAAG to glutamate, reduces ischemic brain injury. *Nat Med* 1999;5:1396–1402.

113. Tizzano JP, Griffey KI, Schoepp DD. Induction or protection of limbic seizures in mice by mGluR subtype selective agonists. *Neuropharmacology* 1995;34:1063–1067.

114. Young MR, Fleetwood-Walker SM, Mitchell R, et al. The involvement of metabotropic glutamate receptors and their intracellular signalling pathways in sustained nociceptive transmission in rat dorsal horn neurons. *Neuropharmacology* 1995;34:1033–1041.

115. Divac I, Fonnum F, Storm-Mathisen J. High affinity uptake of glutamate in terminals of corticostriatal axons. *Nature* 1977;266:377–378.

116. Robinson MB, Dowd LA. Heterogeneity and functional properties of subtypes of sodium-dependent glutamate transporters in the mammalian central nervous system. *Adv Pharmacol* 1997;37:69–115.

117. Szatkowski M, Barbour B, Attwell D. Non-vesicular release of glutamate from glial cells by reversed electrogenic glutamate uptake. *Nature* 1990;348:443–446.

118. Rothstein JD, Jin L, Dykes-Hoberg M, et al. Chronic inhibition of glutamate uptake produces a model of slow neurotoxicity. *Proc Natl Acad Sci USA* 1993;90:6591–6595.

119. Ferkany J, Coyle JT. Heterogeneity of sodium-dependent excitatory amino acid uptake mechanisms in rat brain. *J Neurosci Res* 1986;16:491–503.

120. Sims KD, Robinson MB. Expression patterns and regulation of glutamate transporters in the developing and adult nervous system. *Crit Rev Neurobiol* 1999;13:169–197.

121. Gegelashvili G, Schousboe A. Cellular distribution and kinetic properties of high-affinity glutamate transporters. *Brain Res Bull* 1998;45:233–238.

122. Rothstein JD, Martin L, Levey AI, et al. Localization of neuronal and glial glutamate transporters. *Neuron* 1994;13:713–725.

123. Dehnes Y, Chaudhry FA, Ullensvang K, et al. The glutamate transporter EAAT4 in rat cerebellar Purkinje cells: a glutamate-gated chloride channel concentrated near the synapse in parts of the dendritic membrane facing astroglia. *J Neurosci* 1998;15:3606–3619.

124. Levy LM, Warr O, Attwell D. Stoichiometry of the glial glutamate transporter GLT-1 expressed inducibly in a Chinese hamster ovary cell line selected for low endogenous Na+-dependent glutamate uptake. *J Neurosci* 1998;18:9620–9628.

125. Tanaka K, Watase K, Manabe T, et al. Epilepsy and exacerbation of brain injury in mice lacking the glutamate transporter GLT-1. *Science* 1997;276:1699–1702.

126. Watase K, Hashimoto K, Kano M, et al. Motor discoordination and increased susceptibility to cerebellar injury in GLAST mutant mice. *Eur J Neurosci* 1998;10:976–988.

127. Harris KM, Rosenberg PA. Localization of synapses in rat cortical cultures. *Neuroscience* 1993;53:495–508.

128. Gegelashvili G, Schousboe A. High affinity glutamate transporters: regulation of expression and activity. *Mol Pharmacol* 1997;52:6–15.

129. Swanson RA, Liu J, Miller JW, et al. Neuronal regulation of glutamate transporter subtype expression in astrocytes. *J Neurosci* 1997;17:932–940.

130. Schlag BD, Vondrasek JR, Munir M, et al. Regulation of the glial Na+-dependent glutamate transporters by cAMP analogs and neurons. *Mol Pharmacol* 1998;53:355–369.

131. Ginsberg SD, Rothstein JD, Price DL, et al. Fimbria-fornix transections selectively down-regulate subtypes of glutamate transporter and glutamate receptor proteins in septum and hippocampus. *J Neurochem* 1996;67:1208–1216.

132. Davis KE, Straff DJ, Weinstein EA, et al. Multiple signaling pathways regulate cell surface expression and activity of the excitatory amino acid carrier 1 subtype of Glu transporter in C6 glioma. *J Neurosci* 1998;18:2475–2485.

133. Bergeron R, Meyer TM, Coyle JT, et al. Modulation of N-methyl-D-aspartate receptor function by glycine transport. *Proc Natl Acad Sci USA* 1998;95:15730–15734.

134. Schwarcz R, Guidotti P, Roberts RC. Quinolinic acid and kynurenic acid: glia derived modulations of excitotoxic brain injury. In: *The role of glia in neurotoxicity*. Boca Raton, FL: CRC Press, 1996:245–262.

135. Heyes MP, Saito K, Lackner A, et al. Sources of the neurotoxin quinolinic acid in the brain of HIV-1-infected patients and retrovirus-infected macaques. *Faseb J* 1998;12:881–896.

136. Caresoli-Borroni G, Guidetti R, Schwarcz R. Acute and chronic changes in kyurenate formation following an intrastriatal quinolinate injections in rats. *J Neural Trans* 1999;106:229–242.

137. Coyle JT, Schwarcz R. Mind glue: implications of glial cell biology for psychiatry. *Arch Gen Psychiatry* 2000;57:90–93.

138. Coyle JT. The nagging question of the function of N-acetylaspartylglutamate. *Neurobiol Dis* 1997;4:231–238.

139. Berger UV, Luthi-Carter R, Passani LA, et al. Glutamate carboxypeptidase II is expressed by astrocytes in the adult rat nervous system. *J Comp Neurol* 1999;415:52–64.

140. Tsai G, Passani LA, Slusher BS, et al. Abnormal excitatory neurotransmitter metabolism in schizophrenic brains. *Arch Gen Psychiatry* 1995;52:829–836.

141. Tsai GC, Stauch-Slusher B, Sim L, et al. Reductions in acidic amino acids and N-acetylaspartylglutamate in amyotrophic lateral sclerosis CNS. *Brain Res* 1991;556:151–156.

142. Coyle JT, Schwarcz R. Lesion of striatal neurones with kainic acid provides a model for Huntington's chorea. *Nature* 1976;263:244–246.

143. Shaw PJ, Ince PG. Glutamate, excitotoxicity and amyotrophic lateral sclerosis. *J Neurol* 1997;244:S3–14.

144. Myseros JS, Bullock R. The rationale for glutamate antagonists in the treatment of traumatic brain injury. *Ann NY Acad Sci* 1995;765:262–271.

145. Kerr JF, Wyllie AH, Currie AR. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br J Cancer* 1972;26:239–257.

146. Horvitz HR. Genetic control of programmed cell death in the nematode Caenorhabditis elegans. *Cancer Res* 1999;59:1701s–1706s.

147. Leski ML, Valentine SL, Coyle JT. L-Type voltage gated calcium channels modulate kainic acid neurotoxicity in cerebellar granule cells. *Brain Res* 1999;828:27–40.

148. Schulz JB, Weller M, Moskowitz MA. Caspases as treatment targets in stroke and neurodegenerative diseases. *Ann Neurol* 1999;45:421–429.

149. Choi DW. Calcium and excitotoxic neuronal injury. *Ann NY Acad Sci* 1994;747:162–171.

150. Coyle JT, Puttfarcken P. Oxidative stress, glutamate, and neurodegenerative disorders. *Science* 1993;262:689–695.

151. Hockenbery D, Nunez G, Milliman C, et al. Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. *Nature* 1990;348:334–336.

152. Kluck RM, Bossy-Wetzel E, Green DR, et al. The release of cytochrome c from mitochondria: a primary site for Bcl-2 regulation of apoptosis. *Science* 1997;275:1132–1136.

153. Zamzami N, Susin SA, Marchetti P, et al. Mitochondrial control of nuclear apoptosis. *J Exp Med* 1996;183:1533–1544.

154. Kroemer G. The proto-oncogene Bcl-2 and its role in regulating apoptosis. *Nat Med* 1997;3:614–620.

155. Simonian NA, Getz RL, Leveque JC, et al. Kainate induces apoptosis in neurons. *Neuroscience* 1996;74:675–683.

156. Leski ML, Valentine SL, Baer JD, et al. Insulin-like growth factor I prevents the development of sensitivity to kainate neurotoxicity in cerebellar granule cells. *J Neurochem* 2000;75:1548–1556.

157. Jiang Q, Gu Z, Zhang G, et al. Di-phosphorylation and involvement of extracellular signal-regulated kinases (ERK1/2) in glutamate-induced apoptotic-like death in cultured rat cortical neurons. *Brain Res* 2000;857:71–77.

158. Uberti D, Belloni M, Grilli M, et al. Induction of tumor-suppressor phosphoprotein p53 in the apoptosis of cultured rat cerebellar neurones triggered by excitatory amino acids. *Eur J Neurosci* 1998;10:246–254.

159. Sakhi S, Bruce A, Sun N, et al. p53 induction is associated with neuronal damage in the central nervous system. *Proc Natl Acad Sci USA* 1994;91:7525–7529.

160. Glazner GW, Chan SL, Lu C, et al. Caspase-mediated degradation of AMPA receptor subunits: a mechanism for preventing excitotoxic necrosis and ensuring apoptosis. *J Neurosci* 2000;20:3641–3649.

161. Lankiewicz S, Marc Luetjens C, Truc Bui N, et al. Activation of calpain I converts excitotoxic neuron death into a caspase-independent cell death. *J Biol Chem* 2000;275:17064–17071.

162. Itoh T, Itoh A, Horiuchi K, et al. AMPA receptor-mediated excitotoxicity in human NT2-N neurons results from loss of intracellular Ca2+ homeostasis following marked elevation of intracellular Na+. *Neuropharmacology* 1998;37:1419–1429.

163. Noh KM, Lee JC, Ahn YH, et al. Insulin-induced oxidative neuronal injury in cortical culture: mediation by induced N-methyl-D-aspartate receptors. *IUBMB Life* 1999;48:263–269.

164. Glazner GW, Mattson MP. Differential effects of BDNF, ADNF9, and TNFalpha on levels of NMDA receptor subunits, calcium homeostasis, and neuronal vulnerability to excitotoxicity. *Exp Neurol* 2000;161:442–452.

165. Group THsDCR. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 1993;72:971–983.

166. Bruyn RP, Stoof JC. The quinolinic acid hypothesis in Huntington's chorea. *J Neurol Sci* 1990;95:29–38.

167. Brouillet E, Hantraye P, Ferrante RJ, et al. Chronic mitochondrial energy impairment produces selective striatal degeneration and abnormal choreiform movements in primates. *Proc Natl Acad Sci USA* 1995;92:7105–7109.

168. Beal MF. Mitochondrial dysfunction in neurodegenerative diseases. *Biochim Biophys Acta* 1998;1366:211–223.

169. Sawa A, Wiegand GW, Cooper J, et al. Increased apoptosis of Huntington disease lymphoblasts associated with repeat length-dependent mitochondrial depolarization. *Nat Med* 1999;5:1194–1198.

170. Ferrante RJ, Andreassen OA, Jenkins BG, et al. Neuroprotective effects of creatine in a transgenic mouse model of Huntington's disease. *J Neurosci* 2000;20:4389–4397.

171. Akbarian S, Smith MA, Jones EG. Editing for an AMPA receptor subunit RNA in prefrontal cortex and striatum in Alzheimer's disease, Huntington's disease and schizophrenia. *Brain Res* 1995;699:297–304.

172. Chen N, Luo T, Wellington C, et al. Subtype-specific enhancement of NMDA receptor currents by mutant huntingtin. *J Neurochem* 1999;72:1890–1898.

173. Calabresi P, Centonze D, Pisani A, et al. Striatal spiny neurons and cholinergic interneurons express differential ionotropic glutamatergic responses and vulnerability: implications for ischemia and Huntington's disease. *Ann Neurol* 1998;43:586–597.

174. Levine MS, Klapstein GJ, Koppel A, et al. Enhanced sensitivity to N-methyl-D-aspartate receptor activation in transgenic and knock in mouse models of Huntington's disease. *J Neurosci Res* 1999;58:515–532.

175. Calabresi P, Centonze D, Pisani A, et al. Metabotropic glutamate receptors and cell-type-specific vulnerability in the striatum: implication for ischemia and Huntington's disease. *Exp Neurol* 1999;158:97–108.

176. Cha JH, Kosinski CM, Kerner JA, et al. Altered brain neurotransmitter receptors in transgenic mice expressing a portion of an abnormal human Huntington disease gene. *Proc Natl Acad Sci USA* 1998;95:6480–6485.

177. MacDonald ME, Vonsattel JP, Shrinidhi J, et al. Evidence for the GluR6 gene associated with younger onset age of Huntington's disease. *Neurology* 1999;53:1330–1332.

178. Rubinsztein DC, Leggo J, Chiano M, et al. Genotypes at the GluR6 kainate receptor locus are associated with variation in the age of onset of Huntington disease. *Proc Natl Acad Sci USA* 1997;94:3872–3876.

179. Selkoe DJ. Alzheimer's disease: genotypes, phenotypes, and treatments. *Science* 1997;275:630-631.

180. Guo Q, Fu W, Sopher BL, et al. Increased vulnerability of hippocampal neurons to excitotoxic necrosis in presenilin-1 mutant knockin mice. *Nat Med* 1999;40:713–720.

181. Harris ME, Carney JM, Cole PS, et al. beta-Amyloid peptide-derived, oxygen-dependent free radicals inhibit glutamate uptake in cultured astrocytes: implications for Alzheimer's disease. *Neuroreport* 1995;6:1875–1879.

182. Lauderback CM, Harris-White ME, Wang Y, et al. Amyloid beta-peptide inhibits Na+-dependent glutamate uptake. *Life Sci* 1999;65:1977–1981.

183. Aronica E, Dickson DW, Kress Y, et al. Non-plaque dystrophic dendrites in Alzheimer hippocampus: a new pathological structure revealed by glutamate receptor immunocytochemistry. *Neuroscience* 1998;82:979–991.

184. Carlson MD, Penney JB Jr, Young AB. NMDA, AMPA, and benzodiazepine binding site changes in Alzheimer's disease visual cortex. *Neurobiol Aging* 1993;14:343–352.

185. Hyman BT, Penney JB Jr, Blackstone CD, et al. Localization of non-N-methyl-D-aspartate glutamate receptors in normal and Alzheimer hippocampal formation. *Annals Neurol* 1994;35:31–37.

186. Ikonomovic MD, Nocera R, Mizukami K, et al. Age-related loss of the AMPA receptor subunits GluR2/3 in the human nucleus basalis of meynert. *Exp Neurol* 2000;166:363–375.

187. Ikonomovic MD, Mizukami K, Davies P, et al. The loss of GluR2(3) immunoreactivity precedes neurofibrillary tangle formation in the entorhinal cortex and hippocampus of Alzheimer brains. *J Neuropathol Exp Neurol* 1997;56:1018–1027.

188. Yasuda RP, Ikonomovic MD, Sheffield R, et al. Reduction of AMPA-selective glutamate receptor subunits in the entorhinal cortex of patients with Alzheimer's disease pathology: a biochemical study. *Brain Res* 1995;678:161-167.

189. Brown RH Jr. Amyotrophic lateral sclerosis. Insights from genetics. *Arch Neurol* 1997;54:1246–1250.

190. Rosen DR, Siddique T, Patterson D, et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 1993;362:59–62.

191. Gurney ME, Pu H, Chiu AY, et al. Motor neuron degeneration in mice that express a human Cu, Zn superoxide dismutase mutation. *Science* 1994;264:1772--1775.

192. Shaw PJ, Williams TL, Slade JY, et al. Low expression of GluR2 AMPA receptor subunit protein by human motor neurons. *Neuroreport* 1999;10:261–265.

193. Williams TL, Day NC, Ince PG, et al. Calcium-permeable alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors: a molecular determinant of selective vulnerability in amyotrophic lateral sclerosis. *Ann Neurol* 1997;42:200–207.

194. Morrison BM, Janssen WGM, Gordon JW, et al. Light and electron microscopic distribution of the AMPA receptor subunit, GluR2, in the spinal cord of control and G86R mutant superoxide dismutase transgenic mice. *J Comp Neurol* 1998;395:523–534.

195. Carriedo SG, Sensi SL, Yin HZ, et al. AMPA exposures induce mitochondrial Ca(2+) overload and ROS generation in spinal motor neurons in vitro. *J Neurosci* 2000;20:240–250.

196. Samarasinghe S, Virgo L, de Belleroche J. Distribution of the N-methyl-D-aspartate glutamate receptor subunit NR2A in control and amyotrophic lateral sclerosis spinal cord. *Brain Res* 1996;727:233–237.

197. Miller RG, Mitchell JD, Moore DH. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database Syst Rev* 2000;2:CD001447.

198. Mennini T, Cagnotto A, Carvelli L, et al. Biochemical and pharmacological evidence of a functional role of AMPA receptors in motor neuron dysfunction in mnd mice. *Eur J Neurosci* 1999;11:1705–1710.

199. Kim JS, Kornhuber HH, Schmid-Burgk W, et al. Low cerebrospinal fluid glutamate in schizophrenic patients and a new hypothesis on schizophrenia. *Neurosci Lett* 1980;20:379–382.

200. Bjerkenstedt L, Edman G, Hagenfeldt L, et al. Plasma amino acids in relation to cerebrospinal fluid monoamine metabolites in schizophrenic patients and healthy controls. *Br J Psychiatry* 1985;147:276–282.

201. Gattaz WF, Gattaz D, Beckmann H. Glutamate in schizophrenics and healthy controls. *Arch Psychiatr Nervenkr* 1982;231:221–225.

202. Perry TL. Normal cerebrospinal fluid and brain glutamate levels in schizophrenia do not support the hypothesis of glutamatergic neuronal dysfunction. *Neurosci Lett* 1982;28:81–85.

203. Tsai G, van Kammen DP, Chen S, et al. Glutamatergic neurotransmission involves structural and clinical deficits of schizophrenia. *Biol Psychiatry* 1998;44:667–674.

204. Kegeles LS, Humaran TJ, Mann JJ. In vivo neurochemistry of the brain in schizophrenia as revealed by magnetic resonance spectroscopy. *Biol Psychiatry* 1998;44:382–398.

205. Deakin JF, Slater P, Simpson MD, et al. Frontal cortical and left temporal glutamatergic dysfunction in schizophrenia. *J Neurochem* 1989;52:1781–1786.

206. Nishikawa T, Takashima M, Toru M. Increased [^{3}H]kainic acid binding in the prefrontal cortex in schizophrenia. *Neurosci Lett* 1983;40:245–250.

207. Kerwin RW, Patel S, Meldrum BS, et al. Asymmetrical loss of glutamate receptor subtype in left hippocampus in schizophrenia. *Lancet* 1988;1:583–584.

208. Kerwin R, Patel S, Meldrum B. Quantitative autoradiographic analysis of glutamate binding sites in the hippocampal formation in normal and schizophrenic brain post mortem. *Neuroscience* 1990;39:25–32.

209. Ishimaru M, Kurumaji A, Toru M. Increases in strychnine-insensitive glycine binding sites in cerebral cortex of chronic schizophrenics: evidence for glutamate hypothesis. *Biol Psychiatry* 1994;35:84–95.

210. Harrison PJ, McLaughlin D, Kerwin RW. Decreased hippocampal expression of a glutamate receptor gene in schizophrenia. *Lancet* 1991;337:450–452.

211. Eastwood SL, McDonald B, Burnet PW, et al. Decreased expression of mRNAs encoding non–NMDA glutamate receptors GluR1 and GluR2 in medial temporal lobe neurons in schizophrenia. *Brain Res Mol Brain Res* 1995;29:211–223.

212. Akbarian S, Sucher NJ, Bradley D, et al. Selective alterations in gene expression for NMDA receptor subunits in prefrontal cortex of schizophrenics. *J Neurosci* 1996;16:19–30.

213. Javitt DC, Zukin SR. Recent advances in the phencyclidine model of schizophrenia. *Am J Psychiatry* 1991;148:1301–1308.

214. Krystal JH, Karper LP, Seibyl JP, et al. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry* 1994;51:199–214.

215. Lahti AC, Koffel B, LaPorte D, et al. Subanesthetic doses of ketamine stimulate psychosis in schizophrenia. *Neuropsychopharmacology* 1995;13:9–19.

216. Malhotra AK, Pinals DA, Adler CM, et al. Ketamine-induced exacerbation of psychotic symptoms and cognitive impairment in neuroleptic-free schizophrenics. *Neuropsychopharmacology* 1997;17:141–150.

217. Moghaddam B, Adams B, Verma A, et al. Activation of glutamatergic neurotransmission by ketamine: a novel step in the pathway from NMDA receptor blockade to dopaminergic and cognitive disruptions associated with the prefrontal cortex. *J Neurosci* 1997;17:2921–2927.

218. Hertzmann M, Reba RC, Kotlyarov EV. Single photon emission computed tomography in phencyclidine and related drug abuse. *Am J Psychiatry* 1990;147:255–256.

219. Wu JC, Buchsbaum MS, Bunney WE. Positron emission tomography study of phencyclidine users as a possible drug model of schizophrenia. *Yakubutsu Seishin Kodo* 1991;11:47–48.

220. Jentsch JD, Redmond DE Jr, Elsworth JD, et al. Enduring cognitive deficits and cortical dopamine dysfunction in monkeys after long-term administration of phencyclidine. *Science* 1997;277:953–955.

221. Deutch AY, Tam SY, Freeman AS, et al. Mesolimbic and mesocortical dopamine activation induced by phencyclidine: contrasting pattern to striatal response. *Eur J Pharmacol* 1987;134:257–264.

222. Bowers MB Jr, Bannon MJ, Hoffman FJ Jr. Activation of forebrain dopamine systems by phencyclidine and footshock stress: evidence for distinct mechanisms. *Psychopharmacology (Berl)* 1987;93:133–135.

223. Moghaddam B, Adams BW. Reversal of phencyclidine effects by a group II metabotropic glutamate receptor agonist in rats. *Science* 1998;281:1349–1352.

224. Smith T, Groom A, Zhu B, et al. Autoimmune encephalomyelitis ameliorated by AMPA antagonists. *Nat Med* 2000;6:62–66.

225. Lawlor BA, Davis KL. Does modulation of glutamatergic function represent a viable therapeutic strategy in Alzheimer's disease? *Biol Psychiatry* 1992;31:337–350.

226. Javitt DC, Zylberman I, Zukin SR, et al. Amelioration of negative symptoms in schizophrenia by glycine. *Am J Psychiatry* 1994;151:1234–1236.

227. Heresco-Levy U, Javitt DC, Ermilov M, et al. Double-blind, placebo-controlled, crossover trial of glycine adjuvant therapy for treatment-resistant schizophrenia. *Br J Psychiatry* 1996;169:610–617.

228. Heresco-Levy U, Javitt DC, Ermilov M, et al. Efficacy of high-dose glycine in the treatment of enduring negative symptoms of schizophrenia. *Arch Gen Psychiatry* 1999;56:29–36.

229. Tsai G, Yang P, Chung LC, et al. D-serine added to antipsychotics for the treatment of schizophrenia. *Biol Psychiatry* 1998;44:1081–1089.

230. Watson GB, Bolanowski MA, Baganoff MP, et al. D-cycloserine acts as a partial agonist at the glycine modulatory site of the NMDA receptor expressed in Xenopus oocytes. *Brain Res* 1990;510:158–160.

231. Goff DC, Tsai G, Manoach DS, et al. Dose-finding trial of D-cycloserine added to neuroleptics for negative symptoms in schizophrenia. *Am J Psychiatry* 1995;152:1213-1215.

232. van Berckel BN, Hijman R, van der Linden JA, et al. Efficacy and tolerance of D-cycloserine in drug-free schizophrenic patients. *Biol Psychiatry* 1996;40:1298–1300.

233. Kirkpatrick B, Buchanan RW, McKenney PD, et al. The Schedule for the Deficit syndrome: an instrument for research in schizophrenia. *Psychiatry Res* 1989;30:119–123.

234. Goff DC, Tsai G, Manoach DS, et al. D-cycloserine added to clozapine for patients with schizophrenia. *Am J Psychiatry* 1996;153:1628–1630.

235. Goff DC, Henderson DC, Evins AE, et al. A placebo-controlled crossover trial of D-cycloserine added to clozapine in patients with schizophrenia. *Biol Psychiatry* 1999;45:512–514.

236. Potkin SG, Jin Y, Bunney BG, et al. Effect of clozapine and adjunctive high-dose glycine in treatment-resistant schizophrenia. *Am J Psychiatry* 1999;156:145–147.

237. Tsai GE, Yang P, Chung LC, et al. D-serine added to clozapine for the treatment of schizophrenia. *Am J Psychiatry* 1999;156:1822–1825.

238. Arvanov VL, Liang X, Schwartz J, et al. Clozapine and haloperidol modulate N-methyl-D-aspartate- and non-N-methyl-Daspartate receptor-mediated neurotransmission in rat prefrontal cortical neurons in vitro. *J Pharmacol Exp Ther* 1997;283:226–234.

239. Grunze HC, Rainnie DG, Hasselmo ME, et al. NMDA-dependent modulation of CA1 local circuit inhibition. *J Neurosci* 1996;16:2034–2043.

240. Smith GS, Schloesser R, Brodie JD, et al. Glutamate modulation of dopamine measured in vivo with positron emission tomography (PET) and 11C-raclopride in normal human subjects. *Neuropsychopharmacology* 1998;18:18–25.

241. Heckers S, Rauch SL, Goff D, et al. Impaired recruitment of the hippocampus during conscious recollection in schizophrenia. *Nat Neurosci* 1998;1:318–323.

242. Radant AD, Bowdle TA, Cowley DS, et al. Does ketamine-mediated N-methyl-D-aspartate receptor antagonism cause schizophrenia-like oculomotor abnormalities? *Neuropsychopharmacology* 1998;19:434–444.

243. Gallagher M, Rapp PR. The use of animal models to study the effects of aging on cognition. *Annu Rev Psychol* 1997;48:339–370

244. Kentros C, Hargreaves EL, Hawkins RD, et al. Abolition of long-term stability of new hippocampal place cell maps by NMDA receptor blockade. *Science* 1998;280:2121–2126.

245. Tsien JZ, Herta PT, Tonegawa S. The essential role of hippocampal CA1 NMDA receptor-dependent synaptic plasticity in spatial memory. *Cell* 1996;87:1327–1338.

246. Barnes CA. Normal aging: regionally specific changed in hippocampal synaptic transmission. *Trends Neurosci* 1994;17:13–18.

247. Barnes CA, Suster MS, Shen J, et al. Multistability of cognitive maps in the hippocampus of old rats. *Nature* 1997;388:272–275.

248. Magnusson KR, Cotman CW. Age-related changes in excitatory amino acid receptors in two mouse strains. *Neurobiol Aging* 1993;14:197–206.

249. Magnusson KR. Declines in mRNA expression of different subunits may account for differential effects of aging on agonist and antagonist binding to the NMDA receptor. *J Neurosci* 2000;20:1666–1674.

250. Nicolle MM, Bizon JL, Gallagher M. In vitro autoradiography of ionotropic glutamate receptors in hippocampus and striatum of aged long-evans rats: relationship to spatial learning. *Neuroscience* 1996;74:741–756.

251. Gazzaley AH, Siegel SJ, Kordower JH, et al. Circuit-specific alterations of N-methyl-D-aspartate subunit 1 in the dentate gyrus of aged monkeys. *Proc Natl Acad Sci USA* 1996;3121–3125.

252. Smith TD, Adams MM, Gallagher M, et al. Circuit-specific alterations in hippocampal synaptophysin immunoreactivity predict spatial learning impairment in aged rats. *J Neurosci* 2000;20:6587–6593.

253. Adams MM, Smith TD, Moga D, et al. Hippocampal dependent learning ability correlates with N-Methyl-D-Aspartate (NMDA) receptor level sin CA3 neurons of young and aged rats. *J Comp Neurol, (in press)*.

254. Ikonomidou C, Bosch F, Miksa M, et al. Blockade of NMDA receptors and apoptotic neurodegeneration in the developing brain. *Science* 1999;283:70–74.

7

Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders

Errol B. De Souza

Dimitri E. Grigoriadis

Errol B. De Souza: Aventis Pharmaceuticals, Inc., Bridgewater, New Jersey.

Dimitri E. Grigoriadis: Neurocrine Biosciences, Inc., San Diego, California.

HISTORICAL PERSPECTIVES

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

In 1948 Sir Geoffrey Harris first proposed the concept that the hypothalamus plays a primary role in the regulation of the pituitary-adrenocortical axis. Subsequently, during the 1950s, Guillemin and Rosenberg, and Saffran and Schally independently observed the presence of a factor in hypothalamic extracts (termed corticotropin-releasing factor, CRF) that could stimulate the release of adrenocorticotropic hormone (ACTH, corticotropin) from anterior pituitary cells *in vitro*. Although CRF was the first hypothalamic hypophysiotropic factor to be recognized, its chemical identity remained elusive until 1981, when Wylie Vale and colleagues at the Salk Institute reported the isolation, characterization, synthesis, and *in vitro* and *in vivo* biological activities of a 41-amino acid hypothalamic ovine CRF (1). Just over a decade later, Vale and colleagues were the first to report the cloning of the human CRF_1 receptor from a single human Cushing's corticotropic adenoma using an expression cloning techniques (2). This initial discovery led to the identification of a second receptor subtype (termed CRF_2), which has now been localized and characterized in a variety of species (see the following).

This chapter provides an overview of the CRF system and its related receptor targets. More detailed and comprehensive information on CRF is available in recent reviews (3 ,4) and books (5 ,6) on the topic.

CHARACTERISTICS OF THE CRF PEPTIDE AND GENE SEQUENCES

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

Amino Acid Sequence and Structure of CRF

The sequence of CRF has been determined in a variety of species including sheep, humans, rats, pigs, goats, and cows. In all species, CRF is a 41-amino acid residue single chain polypeptide (Fig. 7.1). Rat and human CRFs are identical to one another and differ from ovine CRF by seven amino acid residues. All three CRFs have close amino acid homology and share some biological properties with sauvagine, a 40-amino acid peptide that exists in frog skin, and urotensin I, a 41-amino acid peptide derived from fish urophysis. Caprine and ovine CRF are identical and differ from bovine CRF by one amino acid. Porcine CRF more closely resembles rat and human CRF. CRF and related peptides are amidated at their carboxy terminal; CRF COOH-terminal-free acid has less than 0.1% of the potency of native CRF, suggesting the importance of amidation to biological activity of the peptide. Studies to determine the solution structure of CRF using proton nuclear magnetic resonance suggest that human CRF comprises an extended N-terminal tetrapeptide connected to a well-defined α-helix between residues 6 and 36 (7). An α-helical oCRF(9-41) has been demonstrated to be an antagonist of CRF (8), which underscores the necessity of the α-helical conformation for receptor binding and biological activity.

FIGURE 7.1. Amino acid sequences of ovine and rat/human corticotropin-releasing factor (CRF) and related peptides including human and rat urocortin. Rat CRF and human CRF are identical and differ from ovine CRF by seven amino acid residues, which are denoted by the shading. All peptides are amidated at the carboxy terminus; the amidation is essential for biological activity. See color version of figure.

Organization of the CRF Gene and Protein Precursor

The nucleotide sequences encoding ovine and rat CRF cDNA precursors as well as human, rat, and ovine CRF

genes have been determined (9 ,10). The locus of the CRF gene is on chromosome 8q13 in the human. The CRF genes are quite similar to one another, containing two exons separated by intervening intron 686 to 800 base pairs long. The first exon encodes most of the 5′-untranslated region of the mRNA and the second encodes the entire prepro-CRF precursor polypeptide, which is 187 to 196 amino acids long; the carboxy end of the precursor contains the 41-amino acid peptide sequence. The high incidence of homology among species suggests that the gene has been highly conserved through evolution.

As previously demonstrated for other systems, the 5′-flanking DNA sequences are most likely to contain the DNA sequence elements responsible for glucocorticoid, cAMP and phorbol ester regulation, tissue-specific expression, and enhancer activity. Although a consensus cAMP response element has been identified, located 200 base pairs upstream from the major transcription initiation site, no obvious glucocorticoid response elements or activation protein (AP) 1-binding elements are present. A potential AP-2 binding site, which may mediate the responses to protein kinase A and C, is present 150 base pairs upstream from the major start site.

ANATOMY OF CRF

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

Distribution of CRF in the CNS

The distribution and localization of CRF mRNA in the central nervous system (CNS) have been evaluated using Northern blot analysis and *in situ* hybridization histochemistry, respectively. Radioimmunoassay and immunohistochemical studies have been critical in the determination of the neuroanatomic organization of CRF immunoreactive cells and fibers in the CNS. Overall, there is good concordance between studies demonstrating a widespread distribution of CRF cell bodies and fibers in the CNS. Detailed descriptions of the organization of CRF immunoreactive cells and fibers in rat brain have been published (11 ,12 and 13).

Morphologic data clearly indicate that the paraventricular nucleus of the hypothalamus (PVN) is the major site of CRF-containing cell bodies that influence anterior pituitary hormone secretion. These neurons originate in the parvocellular portion of the PVN and send axon terminals to the capillaries of the median eminence. CRF is also present in a small group of PVN neurons that project to the lower brainstem and spinal cord; this group of neurons may be involved in regulating autonomic nervous system function. Other hypothalamic nuclei that contain CRF cell bodies include the medial preoptic area, dorsomedial nucleus, arcuate nucleus, posterior hypothalamus, and mammillary nuclei.

The neocortex contains primarily CRF interneurons with bipolar, vertically oriented cell bodies predominantly localized to the second and third layers of the cortex and projections to layers I and IV. In addition, scattered cells are present in the deeper layers that appear to be pyramidal cells. Although CRF-containing neurons are found throughout the neocortex, they are found in higher densities in the prefrontal, insular, and cingulate areas. CRF neurons in the cerebral cortex appear to be important in several behavioral actions of the peptide, including effects on cognitive processing; furthermore, dysfunction of these neurons may contribute to many CNS disorders (see the following).

Large and discrete populations of CRF perikarya are present in the central nucleus of the amygdala, bed nucleus of the stria terminalis, and substantia innominata. CRF neurons in the central nucleus of the amygdala project to the parvocellular regions of the PVN, the parabrachial nucleus of the brainstem, and thus may influence both neuroendocrine and autonomic function in addition to behavioral activity. CRF neurons originating in the bed nucleus of the stria terminalis send terminals to brainstem areas such as the parabrachial nuclei and dorsal vagal complex, which coordinate autonomic activity. CRF fibers also interconnect the amygdala with the bed nucleus of the stria terminalis and hypothalamus. Scattered CRF cells with a few fibers are also present in telencephalic areas such as regions of the amygdala in addition to the central nucleus, septum, diagonal band of Broca, olfactory bulb, and all aspects of hippocampal formation, including the pyramidal cells, dentate gyrus, and subiculum.

Several groups of CRF cell bodies are present throughout the brainstem. In the midbrain, CRF perikarya are present in the periaqueductal gray, Edinger-Westphal nucleus, dorsal raphe nucleus, and ventral tegmental nucleus. Projections from the dorsal-lateral tegmental nucleus to a variety of anterior brain areas such as the medial frontal cortex, septum, and thalamus have also been described. In the pons, CRF cell bodies are localized in the locus ceruleus, parabrachial nucleus, medial vestibular nucleus, paragigantocellular

nucleus, and periaqueductal gray. CRF neurons originating in the parabrachial nucleus project to the medial preoptic nucleus of the hypothalamus. In the medulla, the large groups of cell bodies are present in the nucleus of the solitary tract and dorsal vagal complex with ascending projections to the parabrachial nucleus. Scattered groups of cell bodies are also present in the medullary reticular formation, spinal trigeminal nucleus, external cuneate nucleus, and inferior olive. The inferior olive gives rise to a well-defined olivocerebellar CRF pathway with projections to the Purkinje cells of the cerebellum. No CRF cell bodies are present in the cerebellar formation.

Within the spinal cord, CRF cell bodies are present in laminae V to VII and X and in the intermediolateral column of the thoracic and lumbar cord. CRF fibers originating in the spinal cord form an ascending system terminating in the reticular formation, vestibular complex, central gray, and thalamus. This ascending CRF system may play an important role in modulating sensory input. In addition, spinal cord CRF neurons may represent preganglionic neurons that modulate sympathetic outflow.

Distribution of CRF in Peripheral Tissues

In addition to its CNS distribution, CRF has been localized in a variety of peripheral tissues (14). CRF-like immunoreactive fibers are present in the intermediate lobe of the pituitary; these fibers originate in the hypothalamus. A physiologic role has been proposed for CRF in regulating pro-opiomelanocortin (POMC)-derived peptide secretion from the intermediate pituitary. CRF has also been localized in the adrenal medulla of a variety of species and is increased following stimulation of the splanchnic nerve stimulation and hemorrhagic stress. CRF-like immunoreactivity and CRF mRNA have been detected in lymphocytes, where they may play a role in regulating immune function. Other tissues in which CRF has been localized include the testis (Leydig cells and advanced germ cells), pancreas, stomach, and small intestine. Although CRF is not detected in the circulation under normal circumstances, very high levels have been measured in the plasma of pregnant women; the source of CRF in pregnancy appears to be the placenta. (See CRF-Binding Protein .)

UROCORTIN: A NOVEL MAMMALIAN CRF-RELATED PEPTIDE

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

Urocortin is the newest member of the CRF peptide family and has been demonstrated to possess many of the intrinsic properties of CRF itself as well as some unique properties of its own. Originally, the nonmammalian CRF-related analogues urotensin I (teleost fish) (15) and sauvagine (frog) (16) were thought to subserve the functions of CRF in the respective species; however, the discovery of peptides even closer to the structure of CRF in those species (17 ,18) led to the suggestion that other forms of CRF may exist in mammals. Furthermore, with the cloning of the CRF_2 receptor subtype, it became apparent that because sauvagine and urotensin had even higher affinity and activity at this subtype than CRF itself (19), a mammalian form of these peptides may exist that would serve as the endogenous ligand for the CRF_2 subtype. Indeed, a mammalian form of urotensin was recently discovered termed “urocortin” and the cDNA cloned from the rat (20) and human (21), respectively.

Amino Acid Sequence and Structure of Urocortin

The sequence of urocortin has been determined in both rats and humans. In rat, urocortin was identified using a library derived from rat midbrain and a carp urotensin cDNA probe. A full-length cDNA was described and encoded a putative 40-amino acid peptide that was related to CRF (20). The human form was subsequently identified using a cDNA probe encoding the peptide region of rat urocortin and screening a human genomic library (21). The resulting putative peptide demonstrated 88% identity to rat at the nucleotide level and 95% identity at the amino acid level (Fig. 7.1). In both species, urocortin is a 40-amino acid residue single-chain polypeptide; the two forms differ by only two amino acids at positions 2 (Asn to Asp) and 4 (Ser to Pro) (21). In addition to the homology between the species, the deduced amino acid sequence of rat and human urocortin exhibits sequence identity with urotensin I (63%) and human CRF (45%) (20). Consistent with the other CRF-related peptides, urocortin is also amidated at its carboxy terminal, again suggesting the importance of amidation to this family of peptides (Fig. 7.1).

Anatomic Distribution of Urocortin

The distribution of urocortin in rat tissues was elucidated first by examining the cellular localization of urotensin-like immunoreactivity and correlating that distribution with the *in situ* hybridization of urocortin mRNA. The highest areas of correlation and overlapping localization were the Edinger-Westphal nucleus and the lateral superior olive (20). In addition, cellular staining was observed in the external plexiform layer of the rat olfactory bulb and lateral hypothalamus. Interestingly, terminal projection fields in the lateral septum also demonstrated distinct localization where CRF_2 receptors have been uniquely localized (see the following). Although the localization of urocortin appears to be in very discrete brain regions, these regions demonstrate no CRF mRNA, suggesting that urocortin subserves some unique functionality within the CRF system. As described in the following, its affinity and functional activity at the

CRF_2 receptor subtype suggest that this may be one endogenous ligand for this subtype.

In Vitro and In Vivo Pharmacologic Effects

Urocortin binds with high affinity to all the known effectors of CRF function, including CRF_1, $CRF_{2\alpha}$, CRF_{2B} receptors, and CRF-binding protein (described in the sections that follow). This profile makes urocortin unique in the CRF system because endogenous r/hCRF has been shown to have relatively low affinity for the CRF_2 receptor subtypes and oCRF, which also has lower affinity for the CRF_2 subtype, also has very low affinity at the CRF-binding protein. Urocortin binds to cells stably transfected with the CRF_1, $CRF_{2\alpha}$, or CRF_{2B} receptors with affinities in the 100 to 500 pM range and has 100 pM affinity for the CRF-BP (20). In *in vitro* studies, urocortin stimulates cAMP accumulation in cells transfected with either CRF receptor subtype, and is extremely potent in stimulating ACTH release from cultured anterior pituitary cells (20). The effects on the CRF_1 receptor subtype are comparable to the effects of CRF itself; however, the activities observed at both CRF_2 receptor isoforms are approximately tenfold more potent than CRF itself (20 ,21). Furthermore, as has been shown for CRF, the presence of CRF-BP can decrease the ability of urocortin to stimulate ACTH release *in vitro*. Moreover, specific CRF-BP inhibitors such as r/hCRF(9-33) can restore the ability of urocortin to stimulate ACTH, further confirming the functional activity of urocortin at the CRF-BP (21).

In unanesthetized freely moving rats, urocortin administered IV was fivefold more potent than CRF in increasing plasma ACTH concentrations and demonstrated a longer duration of action. Similarly, urocortin reduced mean arterial pressure more potently and for a longer period of time than CRF or urotensin I (20). Thus, although capable of interacting with the CRF_1 receptor with equivalent potency and activity, the anatomic distribution, localization, and potency at the CRF_2 subtypes support the notion that urocortin is likely one endogenous ligand for this receptor subtype. Clearly, further study is required to determine the specific role that this novel endogenous peptide plays in the regulation of the CRF system.

CRF RECEPTORS

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

Cloning of CRF Receptor Subtypes

Molecular cloning studies have enabled the elucidation of receptor subtypes for the CRF system. Structurally, the CRF receptor subtypes all contain seven putative transmembrane domains and share considerable sequence homology with one another. These receptors are members of the family of “brain-gut” neuropeptide receptors, which includes receptors for calcitonin, vasoactive intestinal peptide, parathyroid hormone, pituitary adenylate cyclase-activating peptide, growth hormone-releasing factor, glucagon, and secretin. In addition, the known members of this neuropeptide receptor family also belong to the superfamily of G-protein coupled receptors; thus far, all have been shown to stimulate adenylate cyclase in response to their respective agonist activation.

The CRF_1 receptor was cloned first from a human Cushing’s corticotropic adenoma using an expression cloning technique and characterized as a 415-amino acid protein with potential N-linked glycosylation sites, protein kinase C phosphorylation sites in the first and second intracellular loops and in the C-terminal tail, as well as casein kinase II and protein kinase A phosphorylation sites in the third intracellular loop (2). Independently, this receptor subtype was also identified in mouse (22) and rat (23 ,24). In all three species, CRF_1 receptor mRNAs encode proteins of 415 amino acids, which are 98% identical to one another. The potential N-linked glycosylation sites on the N-terminal extracellular domain are characteristic of most G-protein coupled receptors and confirm the glycosylation profiles determined by chemical affinity cross-linking studies (25). In those studies, although the molecular weights of the proteins labeled from brain or pituitary appeared different when labeled with [^{125}I]oCRF, the deglycosylation and peptide mapping studies suggested that the protein itself was identical and that the differences were owing to posttranslational modification (25). Indeed, the molecular weight predicted from the deglycosylated forms of the CRF_1 receptor was virtually identical to that obtained from the cloned amino acid sequence. These data taken together established the fact that the CRF_1 receptor subtype is the dominant form in both the brain and pituitary.

CRF_2 Receptors

Following the cloning of the CRF_1 subtype, two forms of a second family member were discovered in the rat and termed $CRF_{2\alpha}$ and CRF_{2B}. The rat $CRF_{2\alpha}$ receptor (19) is a 411-amino acid protein with approximately 71% identity to the CRF_1 receptor. The CRF_{2B} receptor has been cloned from both rat (19) and mouse (24 ,26), and is a 431-amino acid protein that differs from the $CRF_{2\alpha}$ subtype in that the first 34 amino acids in the N-terminal extracellular domain are replaced by 54 different amino acids. The genomic structure and corresponding cDNA of the human $CRF_{2\alpha}$ receptor subtype was cloned and characterized. The cDNA sequence in the protein-coding region had 94% identity with the previously reported rat $CRF_{2\alpha}$ receptor (27). In addition, the human $CRF_{2\alpha}$ receptor protein was found to be a 411-amino acid protein that had an overall 70% identity with the human CRF_1 receptor sequence (less in the N-terminal extracellular domain; 47%). In stably transfected cells, the human $CRF_{2\alpha}$ receptor had the same pharmacologic characteristics as those demonstrated for the

rat and increased intracellular cAMP levels in response to either sauvagine or CRF (see the following for details). Very recently, the human form of the CRF_{2B} receptor was cloned from human amygdala and demonstrated 94% identity to human $CRF_{2\alpha}$ receptors at the protein level. Preliminary characterization of this novel human isoform indicated that this form also had higher affinity for sauvagine and urotensin than for r/hCRF (28). The $CRF_{2\gamma}$ receptor was the most recently identified isoform and has thus far only been found in human brain. This splice variant uses yet a different 5′ alternative exon for its amino terminus and replaces the first 34-amino acid sequence of the $CRF_{2\alpha}$ receptor with a unique 20 amino acid sequence. Thus, although the CRF_2 receptor exists as three isoforms, $CRF_{2\alpha}$, CRF_{2B}, and $CRF_{2\gamma}$, there are at present no known functional splice variants for the CRF_1 receptor. Figure 7.2 illustrates the differences among human $CRF_{2\alpha}$, CRF_{2B}, and $CRF_{2\gamma}$ in the N-terminal extracellular domain. Between the CRF_1 and CRF_2 receptors, there exist very large regions of amino acid identity, particularly between transmembrane domains five and six. This similarity strongly argues for conservation of biochemical function because this region is thought to be the primary site of G-protein coupling and signal transduction. All three CRF_2 receptor subtypes contain five potential N-glycosylation sites, which are analogous to those found on the CRF_1 receptor subtype. The genomic structure of the human CRF_2 receptor gene is similar to that of the mouse CRF_1 receptor described in the preceding and has 12 introns, the last ten of which interrupt the coding region in identical positions. These gene sequences, however, diverge significantly at the 5′ end, and the chromosomal mapping of the human CRF_2 gene has been localized to chromosome 7 p21-p15.

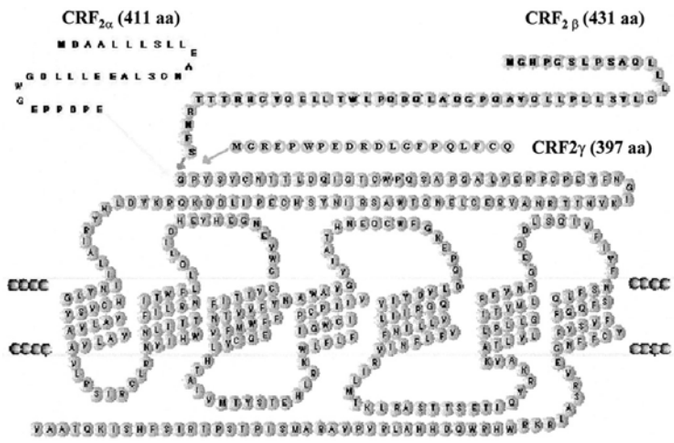


FIGURE 7.2. Amino acid sequences of the human $CRF_{2\alpha}$, CRF_{2B}, and $CRF_{2\gamma}$ receptor subtypes. The CRF_2 receptor isoforms are demonstrated with the seven putative transmembrane domains. The arrows indicate where the different CRF_2 receptor isoforms diverge at the N-terminus. See color version of figure.

Pharmacologic Characteristics

The literature is replete with information on the pharmacologic and biochemical characterization of CRF receptors in a variety of tissues and animal species. (See refs. 29 and 30 ,31 ,32 and 33 for reviews.) The radioligand binding characteristics of CRF receptors that have been performed thus far in brain, endocrine, and immune tissues have used the available radioligands at the time, which were [^{125}I]-Tyr^0 oCRF, [^{125}I]-Tyr^0 r/hCRF, and [^{125}I]-Nle^{21}-Tyr^{32} r/hCRF. These ligands have all demonstrated high affinity for the CRF_1 receptor subtype and lower affinity for the CRF_2 subtypes (as described in the following). Thus, the discovery of the CRF_2 receptor subtype and its isoforms has not confused the earlier literature owing to the apparent “selectivity” of the r/hCRF and oCRF analogue radioligands for the CRF_1 receptor. Recently, [^{125}I]-Tyr^0 sauvagine, a novel radioligand for the CRF_2 receptor, has been described that binds to both receptor subtypes with equal affinity and has become a useful tool in the study of CRF_2 receptors (34).

CRF receptors fulfill all of the criteria for bona fide receptors. The kinetic and pharmacologic characteristics of CRF_1 receptors are comparable in brain, pituitary, and spleen. The binding of [^{125}I]oCRF in a variety of tissue homogenates as well as in CRF_1 receptor-expressing cell lines is dependent on time, temperature, and tissue concentration, and is saturable, reversible, and of high affinity with K_d values of 200 to 400 pM. The pharmacologic rank order profile of these receptors from various tissues has been compared using closely related analogues of CRF. Bioactive analogues of CRF have high affinity for [^{125}I]oCRF binding sites, whereas biologically inactive fragments of the peptide and unrelated peptides are all without inhibitory binding activity in brain, endocrine, and immune tissues.

CRF_1 receptors exhibit the typical properties of neurotransmitter receptor systems linked to the adenylate cyclase system through a guanine nucleotide binding protein. In *in vitro* radioligand binding studies, divalent cations (e.g., magnesium ions) have been shown to enhance agonist binding to receptors coupled to guanine nucleotide binding proteins by stabilizing the high-affinity form of the receptor–effector complex. In contrast, guanine nucleotides have the ability to selectively decrease the affinity of agonists for their receptors by promoting the dissociation of the agonist high-affinity form of the receptor. Consistent with CRF receptors being coupled to a guanine nucleotide regulatory protein, the binding of [^{125}I]oCRF to pituitary, brain, and spleen homogenates is reciprocally increased by divalent cations such as Mg^{2+} and decreased by guanine nucleotides. Furthermore, in expressed cell lines using a ß-galactosidase reporter system, CRF and related analogues could stimulate the production of ß-galactosidase in whole cells with the same pharmacologic rank order of potencies as those in a variety of tissues from different species (35).

The cloning of the CRF_2 receptor subtype gave the first indication that other family members of this receptor system exist and have unique properties that could subserve functions that were previously undefined. As mentioned, a fundamental element in the characterization of any receptor system is the availability of high-affinity and selective ligands that can be used to label the proteins in a reversible manner. The initial observations clearly demonstrated that the CRF_2 receptor subtype recognized the nonmammalian analogues of CRF with high affinity (similar in profile to the CRF_1 subtype) but unlike the CRF_1 receptor, had low affinity for the endogenous CRF ligands (r/hCRF and its analogues) (19). Thus, the available radioligands used in the initial studies of CRF receptors were not useful in providing information about this subtype.

The development of a high-affinity radioligand suitable for the characterization of the CRF_2 receptor subtype was recently described (34). Using one of the high-affinity nonmammalian analogues of CRF (sauvagine), a radiolabel was developed, and its binding specificity and selectivity determined. [^{125}I]Tyr^0-sauvagine was found to bind reversibly, saturably, and with high affinity to both the human CRF_1

and $CRF_{s\alpha}$ receptor subtypes expressed in mammalian cell lines. The specific signal for the labeling of the human $CRF_{a\alpha}$ receptors was greater than 85% over the entire concentration range of the radioligand, which suggested very low nonspecific binding in the expressed cell line. The radioligand bound in a reversible, time- and protein-dependent manner, reaching equilibrium within 60 minutes with the binding being stable for at least 4 hours at 22°C. Scatchard analyses demonstrated an affinity of about 200 pM for the CRF_2 receptor subtype and a maximum receptor density in the expressing cells of about 180 fmol/mg protein (34).

The pharmacologic rank order of potencies for the CRF_2 receptor labeled with [^{125}I]sauvagine was essentially identical to the *in vitro* effects of the same unlabeled peptides in the production of cAMP in cells expressing the receptor as described. That is, the nonmammalian analogues sauvagine and urotensin I that were more potent in stimulation of cAMP production were also more potent at inhibiting the binding of [^{125}I]sauvagine than oCRF or r/hCRF. Interestingly, the putative antagonists for CRF receptors, D-PheCRF(12-41) and α-helical CRF(9-41) exhibited approximately equal affinity for the two receptor subtypes either in inhibiting [^{125}I]sauvagine binding or inhibiting sauvagine-stimulated cAMP production (34). These data clearly indicated that although distinct pharmacologic differences exist between the two receptor subtypes of the same family (in terms of their rank order profile), they still must share some structural similarities. Further study is required to determine the precise common structural features of these two family members.

In addition to the pharmacologic rank order profile, [^{125}I]sauvagine binding to the CRF_2 receptor was guanine nucleotide-sensitive, confirming the agonist activity of this peptide for the receptor. Although there is as yet no direct evidence, this modulation of the binding of [^{125}I]sauvagine to the human $CRF_{2\alpha}$ receptor by guanine nucleotides suggests that this receptor exists in two affinity states for agonists coupled through a guanine nucleotide binding protein to its second messenger system. Unfortunately to date, the only ligands available for the biochemical study of these receptors have been agonists, making it very difficult to examine the proportions and affinities of high- and low-affinity states of these receptors. Further study is required, possibly using labeled antagonists as tools in order to characterize the affinity states of these receptors.

The high affinity of the nonmammalian CRF analogues for this subtype has raised the possibility that other endogenous mammalian ligands exist that have high affinity and selectivity for this receptor subtype. As described, the recent discovery of urocortin (36), although not selective for the CRF_2 subtype, has provided the first evidence for one such endogenous molecule that has high affinity for the CRF_2 receptor. With the increase in the complexity of the CRF system recently elucidated, it is highly likely that more members, from both the receptor and ligand families, remain to be discovered that will lead to a much more comprehensive understanding of this system and its role in both normal and pathologic physiology.

Autoradiographic Localization of CRF Receptor Subtypes

Many studies to date have described the distribution of CRF receptors in various tissues, including the pituitary, brain, and spleen (29 ,31 ,32 and 33). The autoradiographic localization of CRF_1 receptors in the anterior pituitary demonstrates a clustering of binding sites that corresponds to the distribution of corticotrophs. The intermediate lobe shows a more uniform distribution of binding sites characteristic of the homogeneous population of POMC-producing cells in this lobe. Overall, the distribution pattern of CRF_1 receptors within the pituitary supports the functional role of CRF as the primary physiological regulator of POMC-derived peptide secretion from the anterior and intermediate lobes of the pituitary.

Receptor autoradiography and binding studies in discrete areas of rat and primate CNS demonstrate that, in general, the highest concentration of CRF binding sites are distributed in brain regions involved in cognitive function (cerebral cortex), limbic areas involved in emotion and stress responses (amygdala, nucleus accumbens, and hippocampus), brainstem regions regulating autonomic function (locus ceruleus and nucleus of the solitary tract), and olfactory bulb. In addition, there is a high density of CRF_1 receptor sites in the molecular layer of the cerebellar cortex and the spinal cord where the highest concentrations are present in the dorsal horn.

CRF receptors in spleen are primarily localized to the red pulp and marginal zones. The localization of [^{125}I]oCRF binding sites in mouse spleen to regions known to have a high concentration of macrophages suggests that CRF receptors are present on resident splenic macrophages. The absence of specific [^{125}I]oCRF-binding sites in the periarteriole and peripheral follicular white pulp regions of the spleen suggests that neither T nor B lymphocytes have specific high-affinity CRF receptors comparable to those localized in the marginal zone and red pulp areas of the spleen or in the pituitary and brain.

The availability of nucleotide sequences for CRF_1 and CRF_2 receptors has allowed a detailed examination of the regional and cellular distribution of CRF receptor subtype mRNA expression utilizing both RNAse protection assays and *in situ* hybridization histochemistry. A comparison of the distribution of CRF_1 and CRF_2 mRNA and receptor protein defined by ligand autoradiography is demonstrated in adjacent horizontal sections of rat brain (Fig. 7.3).

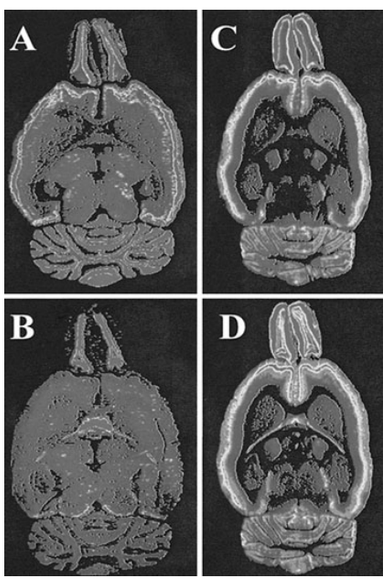


FIGURE 7.3. Digitized, color-coded images of CRF_1 *(Panel A)* and CRF_2 *(Panel B)* receptor mRNA expression and receptor autoradiography in adjacent horizontal sections of rat brain. The highest levels of mRNA expression are coded in red, whereas the lowest concentrations are coded in blue. Similarly, the highest densities of receptors labeled with either [^{125}I]oCRF (CRF_1 only; *Panel C*) or [^{125}I]sauvagine (CRF_1 and CRF_2; *Panel D*) are coded in red. There was a good correspondence between the message for a particular receptor subtype and its protein localization; the pharmacologic selectivity was retained for the two radioligands. (Compare *Panels C* and *D*.) See color version of figure.

The distribution of CRF_2 message clearly differs from that of the CRF_1 and exhibits a distinct subcortical pattern. In the rat brain, the CRF_1 mRNA was most abundant in neocortical, cerebellar, and sensory relay structures and generally

corresponded to the previously reported distribution of [^{125}I]oCRF binding sites (Fig. 7.3). On the other hand, the CRF_2 receptor mRNA was localized primarily in subcortical regions such as the lateral septal nuclei, hypothalamic nuclei, bed nucleus of the stria terminalis, and amygdaloid nuclei. Using the radioligand [^{125}I]sauvagine described, CRF_2 receptors could be localized to areas of high CRF_2 message. In addition, because [^{125}I]sauvagine has equal affinity for both receptor subtypes (34), the autoradiography revealed the localization of both the CRF_1 and CRF_2 receptor subtypes, demonstrating the utility of this novel radioligand. (See ref. 37 for a complete and detailed account of the mRNA distribution patterns of CRF_1 and CRF_2 receptors.)

The heterogeneous distribution of CRF_1 and CRF_2 receptor mRNA and protein suggests distinctive functional roles for each receptor within the CRF system. For example, the lateral septum, by virtue of widespread reciprocal connections throughout the brain, is implicated in a variety of physiologic processes. These range from higher cognitive functions such as learning and memory to autonomic regulation, including food and water intake (38). In addition, the septum plays a central role in classical limbic circuitry and thus is important in a variety of emotional conditions, including fear and aggression. Thus, the lack of CRF_1 receptor expression in these nuclei suggests that CRF_2 receptors may solely mediate the postsynaptic actions of CRF inputs to this region and strongly suggests a role for CRF_2 receptors in modulating limbic circuitry at the level of septal activity. In addition, the selective expression of CRF_2 receptor mRNA within hypothalamic nuclei indicates that the anxiogenic and anorexic actions of CRF in these nuclei may likely be CRF_2 receptor-mediated. In contrast, within the pituitary, there is a predominance of CRF_1 receptor expression with little or no CRF_2 expression in either the intermediate and anterior lobes, indicating that it is the CRF_1 receptor that is primarily responsible for CRF regulation of the HPA axis.

In addition to the differences in distribution between the CRF_1 and CRF_2 receptor subtypes, there exists a distinct pattern of distribution between the CRF_2 isoforms ($CRF_{2\alpha}$ and CRF_{2B}) as well. The $CRF_{2\alpha}$ isoform is primarily expressed within the CNS, whereas the CRF_{2B} form is found both centrally and peripherally. Within the brain, the $CRF_{2\alpha}$ form is the predominant one, whereas the CRF_{2B} form is localized primarily to non-neuronal structures, the choroid plexus of the ventricular system, and cerebral arterioles (37 ,39). The identification of the CRF_{2B} form in cerebral arterioles suggests a mechanism through which CRF may directly modulate cerebral blood flow. Peripherally, the highest detectable levels of mRNA were found in heart and skeletal muscle with lower levels detected in lung and intestine (24 ,39). Taken together, the results of these studies demonstrating a distinct heterogeneous distribution pattern of CRF receptor subtypes in brain and peripheral tissues, strongly suggest that these receptor subtypes subserve very specific physiological roles in CRF related function both centrally and peripherally.

Second Messengers Coupled to CRF Receptors

Radioligand binding studies have demonstrated that CRF receptors in the brain-endocrine-immune axis are coupled to a guanine nucleotide regulatory protein. In all of these tissues, the primary second messenger system involved in transducing the actions of CRF is stimulation of cAMP

production (29 ,31 ,32 and 33 ,40). CRF initiates a cascade of enzymatic reactions in the pituitary gland beginning with the receptor-mediated stimulation of adenylate cyclase, which ultimately regulates POMC-peptide secretion and possibly synthesis. POMC-derived peptide secretion mediated by the activation of adenylate cyclase in the anterior and neurointermediate lobes of the pituitary is dose-related and exhibits appropriate pharmacology. Similarly in the brain and spleen, the pharmacologic rank order profile of CRF-related peptides for stimulation of adenylate cyclase is analogous to the profile seen in pituitary and in keeping with the affinities of these compounds for receptor binding. In addition, the putative CRF receptor antagonist α-hel ovine CRF(9-41) inhibits CRF-stimulated adenylate cyclase in brain and spleen homogenates.

In addition to the adenylate cyclase system, other signal transduction mechanisms may be involved in the actions of CRF. For example, CRF has been shown to increase protein carboxyl methylation, and phospholipid methylation in AtT-20 cells (41). Preliminary evidence suggests that CRF may regulate cellular responses through products of arachidonic acid metabolism (42). Furthermore, although the evidence in anterior pituitary cells suggests that CRF does not directly regulate phosphatidylinositol turnover or protein kinase C activity (42), stimulation of protein kinase C either directly or by specific ligands (vasopressin or angiotensin II), enhances CRF-stimulated adenylate cyclase activity, ACTH release, and inhibits phosphodiesterase activity (42). Thus, the effects of CRF on anterior pituitary cells and possibly in neurons and other cell types expressing CRF receptors are likely to involve complex interactions among several intracellular second messenger systems.

CRF-BINDING PROTEIN

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

CRF and Its Binding Protein in Human Plasma

Under normal conditions, the plasma levels of CRF remain low; however, CRF levels are markedly elevated in plasma during the late gestational stages of pregnancy (43 ,44 and 45). The source of the pregnancy-associated CRF is most likely the placenta because previous studies have demonstrated that the human placenta synthesizes CRF (46). The CRF in the maternal plasma is bioactive in releasing ACTH from cultured pituitary cells (44). In spite of the high levels of CRF in the maternal plasma, there is no evidence of markedly increased ACTH secretion or hypercortisolism in pregnant women (43). A plausible explanation for this paradoxic situation could be the presence of a binding protein in the plasma of pregnant women that could specifically inhibit the biological actions of CRF (44 ,45). This hypothesis was validated by the isolation of a CRF-binding protein (CRF-BP) from human plasma and its subsequent cloning and expression (see the following).

cDNA and Amino Acid Sequences

The CRF-BP was first isolated and purified to near homogeneity for sequencing and generation of oligonucleotide probes (47). Screening a human liver cDNA library using probes generated from the original amino acid sequence revealed a full-length cDNA containing a 1.8-kb insert that coded for a novel protein of 322 amino acids (48). A single putative N-linked glycosylation site was found at amino acid 203, which agrees with the previous observation of the presence of asparagine-linked sugar moieties on the native protein (49). Subsequent screening of a rat cerebral cortical cDNA library, revealed the presence of a single clone containing a 1.85-kb insert predicting a protein of 322 amino acids, which was 85% identical to the human CRF-BP. The putative glycosylation site on the rat protein seems to be conserved between the rat and human sequences (48). The pharmacology of these proteins appears to be similar with both the rat and human binding proteins having high affinity for the rat/human CRF (K_d ~ 0.2 nM) and very low affinity for the ovine form of CRF (K_d ~ 250 nM). Although there may be some similarities in the binding domains of the binding protein and the CRF receptor (as evidenced by the equal affinity of r/hCRF), these are distinct proteins, each with unique characteristics and distributions.

Distribution in Brain and Pituitary

Although the human and rat forms of the CRF-BP are homologous (as indicated), there is a somewhat different anatomic distribution pattern in the two species. The human form of the binding protein has been found abundantly in areas including liver, placenta, and brain, whereas in the rat levels of mRNA for the binding protein have only been localized in the brain and pituitary (48). Peripheral expression of the binding protein may have its greatest utility in the modulation and control of the elevated levels of CRF in circulating plasma induced by various normal physiologic conditions (see the preceding). In addition, expression of this binding protein in the brain and pituitary offers additional mechanisms by which CRF-related neuronal or neuroendocrine actions may be modulated.

CRF-binding protein has been localized to a variety of brain regions including neocortex, hippocampus (primarily in the dentate gyrus), and olfactory bulb. In the basal forebrain, mRNA is localized to the amygdaloid complex with a distinct lack of immunostained cells in the medial nucleus. CRF-binding protein immunoreactivity is also present in the brainstem particularly in the auditory, vestibular, and trigeminal systems, raphe nuclei of the midbrain and pons, and reticular formation (50). In addition, high expression levels of binding protein mRNA are seen in the anterior pituitary, predominantly restricted to the corticotrope cells. Expression of this protein in the corticotropes strongly suggests that the CRF-BP is involved in the regulation of neuroendocrine

functions of CRF by limiting and/or affecting the interactions of CRF with its receptor, which is also known to reside on corticotropes; however, the detailed role of the binding protein in regulating pituitary–adrenal function remains to be elucidated.

CRF REGULATION OF NEUROENDOCRINE FUNCTION

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

Regulation of Pituitary Hormone Secretion

CRF is the major physiologic regulator of the basal and stress-induced release of ACTH, B-endorphin, and other POMC-derived peptides from the anterior pituitary. (See refs. 3 , 4 , and 51 for reviews.) CRF stimulates the release of POMC-derived peptides in anterior pituitary cells in culture and *in vivo;* these actions of CRF can be antagonized by the CRF receptor antagonist α-helical ovine CRF(9-41) or by immunoneutralization with an anti-CRF antibody. Several other lines of evidence support a critical role for endogenous CRF in regulating ACTH secretion. For example, increases in CRF in the hypophysial portal blood are observed following application of stress. Administration of CRF antisera or the CRF receptor antagonist results in attenuation of stress- or adrenalectomy-induced ACTH secretion further substantiating a role for CRF in regulating ACTH secretion from the anterior pituitary. In addition to effects in the anterior pituitary, CRF also has been reported to stimulate POMC-derived peptide secretion from the intermediate lobe of the pituitary gland.

Central administration of CRF inhibits the secretion of luteinizing hormone (LH) and growth hormone without any major effects on follicle-stimulating hormone, thyroid stimulating hormone, or prolactin secretion (3 ,4). The effects of CRF to inhibit LH secretion appear to be mediated at the hypothalamic level through effects of CRF to inhibit gonadotropin releasing hormone secretion. CRF-induced inhibition of LH secretion may also involve endogenous opioids since the effects are attenuated by administration of naloxone or antiserum to B-endorphin (3 ,4).

Regulation of Hypothalamic CRF Release

Plotsky and associates (52) and Owens and Nemeroff (4) provide a comprehensive review of the neurotransmitter regulation of hypothalamic CRF release. Most studies demonstrate stimulatory effects of cholinergic and serotonergic neurons on CRF release. The muscarinic and/or nicotinic cholinergic receptor subtypes involved in the stimulatory effects of acetylcholine on CRF secretion remain to be precisely elucidated. The effects of serotonin to stimulate CRF release appear to be mediated by a variety of receptor subtypes, including 5-HT_2, 5-HT_{1A}, and 5-HT_{1C} receptors. The effects of catecholamines and opioids on hypothalamic CRF release are less well defined. Norepinephrine has been reported to have both stimulatory and inhibitory effects on CRF release that may be a consequence of the dose administered as well as the receptor subtype involved. For example, in studies sampling hypophysial portal concentrations of CRF, Plotsky and colleagues (52) noted that low doses of norepinephrine stimulated CRF release *in vivo* via α-adrenergic receptors and inhibited CRF release at high doses via B-adrenergic receptors. Similarly, opioids have been reported to either inhibit or stimulate CRF release depending on the nature of the opioid tested, dose utilized, and receptor specificity (μ versus κ) involved. Drugs acting at GABA–benzodiazepine–chloride ionophore complex are potent inhibitors of CRF secretion.

Stress is a potent general activator of CRF release from the hypothalamus. The extent and time course of changes in CRF in the paraventricular nucleus and median eminence of the hypothalamus following application of stress are highly dependent on the nature of the stressor as well as the state of the animal. The effects of stress to increase the release and synthesis of CRF are mediated by many of neurotransmitter systems described in the preceding.

Glucocorticoids, which are involved in the negative feedback regulation the hypothalamic–pituitary–adrenocortical axis, are potent inhibitors of CRF release. Conversely, the absence of glucocorticoids following adrenalectomy results in marked elevations in the synthesis and release of CRF. The actions of glucocorticoids to inhibit CRF release are mediated directly at the level of the paraventricular nucleus of the hypothalamus as well as indirectly through actions on receptors in the hippocampus.

Modulation of Pituitary CRF Receptors

Stress (29 ,32 ,33 ,53) or adrenalectomy (29 ,32 ,33) result in hypersecretion of CRF and a consequent down-regulation of receptors in the anterior pituitary. The adrenalectomy-induced decreases in anterior pituitary receptors can be prevented by glucocorticoid replacement with corticosterone or dexamethasone (29 ,32 ,33). In addition, chronic administration of corticosterone has been reported to cause dose-dependent decreases in anterior pituitary CRF receptor number (29 ,32 ,33). An age-related decline in anterior pituitary CRF receptors has also been reported (54). In contrast, lesions of the paraventricular nucleus that result in dramatic reductions in hypothalamic CRF secretion have been reported to increase the density of pituitary CRF receptors (4). Thus, CRF receptors in the anterior pituitary appear to be reciprocally regulated by hypothalamic CRF release.

CRF REGULATION OF CNS ACTIVITY

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

Electrophysiologic Effects of CRF

CRF stimulates the electrical activity of neurons in various brain regions that contain CRF and CRF receptors, including

locus ceruleus (55), hippocampus (56), cerebral cortex, and hypothalamus as well as in lumbar spinal cord motor neurons (3 ,4). In contrast, CRF has inhibitory actions in the lateral septum, thalamus, and hypothalamic PVN (3 ,4). The electrophysiologic effects of CRF on spontaneous and sensory-evoked activity of locus ceruleus neurons are well documented (55). Activation of the locus ceruleus, a brainstem nucleus comprising of noradrenergic cells, results in arousal and increased vigilance. Furthermore, dysfunction of this nucleus has been implicated in the pathophysiology of depression and anxiety. Centrally administered CRF increases the spontaneous discharge rate of the locus ceruleus in both anesthetized and unanesthetized rats, while decreasing evoked activity in the nucleus (55). Thus, the overall effect of CRF in the locus ceruleus is to decrease the signal to noise ratio between evoked and spontaneous discharge rates.

The effects of CRF on EEG activity have been reviewed in detail (3 ,4 ,57). CRF causes a generalized increase in EEG activity associated with increased vigilance and decreased sleep time. At CRF doses below those affecting locomotor activity or pituitary–adrenal function, rats remain awake, vigilant and display decreases in slow-wave sleep compared to saline-injected controls (57). Higher doses of the peptide, on the other hand, cause seizure activity that is indistinguishable from seizures produced by electrical kindling of the amygdala, further confirming the role of CRF in brain activation.

Autonomic Effects of CRF

A great deal of anatomic, pharmacologic, and physiologic data support the concept that CRF acts within the CNS to modulate the autonomic nervous system (3 ,4 ,58 ,59). For example, central administration of CRF results in activation of the sympathetic nervous system resulting in stimulation of epinephrine secretion from the adrenal medulla and noradrenergic outflow to the heart, kidney, and vascular beds. Other consequences of central administration of CRF include increases in the mean arterial pressure and heart rate. These cardiovascular effects of CRF can be blocked by the ganglionic blocker chlorisondamine underscoring the sympathetic actions of the peptide. In contrast, CRF acts in brain to inhibit cardiac parasympathetic nervous activity. (See ref. 58 for review.) Peripheral administration of CRF causes vasodilation and hypotension in a variety of species including humans (3 ,4 ,58 ,59). The physiologic role of CRF in regulating the autonomic nervous system is supported by data demonstrating central effects of the CRF receptor antagonist, α-helical ovine CRF(9-41) to attenuate adrenal epinephrine secretion resulting from stressors such as insulin-induced hypoglycemia, hemorrhage, and exposure to ether vapor (59). Overall, these data substantiate a major role for CRF in coordinating the autonomic responses to stress.

Gastrointestinal Effects of CRF

Studies examining the gastrointestinal effects of CRF have determined that CRF modulates gastrointestinal activity by acting at central and possibly peripheral sites, and that these effects are qualitatively similar to those observed following exposure to various stressors. (See refs. 3 , 4 , and 60 for reviews.) CRF inhibits gastric acid secretion, gastric emptying, and intestinal transit while stimulating colonic transit and fecal excretion in a dose-dependent manner when administered centrally or systemically to dogs or rats. CRF is equipotent in inhibiting gastric emptying in both species following both central and peripheral routes of administration. The central effects of CRF on gastric acid secretion do not appear to result from leakage of the peptide into peripheral blood because measurable quantities of CRF are not present in the circulation following injection of CRF into the third ventricle of the dog. Furthermore, an intravenous injection of anti-CRF serum completely abolishes the peripheral but not the central effect of CRF on gastric acid secretion. These data strongly implicate CRF in the mechanisms through which various stressors alter gastrointestinal function and are consistent with its proposed role in integrating the autonomic nervous system’s response to stress.

Behavioral Effects of CRF

The behavioral effects of CRF in the CNS have been reviewed extensively (3 ,4 ,61). The effects of CRF on behavior are dependent on both the dose of peptide administered and the specific conditions under which the tests are performed. In a familiar or “home” environment, central administration of CRF produces a profound increase in locomotor activity. Although very low doses of CRF produce locomotor activation when tested in an open field test, higher doses produce a dramatic decrease in locomotor activity. CRF administered intracerebrally also produces additional behavioral effects including increases in sniffing, grooming, and rearing in a familiar environment, increased “emotionality” and assumption of a freeze posture in a foreign environment, decreased feeding and sexual behavior, and increased conflict behavior. The behavioral effects of CRF are not an indirect consequence of actions of the peptide to activate pituitary–adrenocortical hormone secretion because they are not seen following peripheral administration of CRF or following pretreatment with doses of dexamethasone that adequately block pituitary–adrenal activation. Of critical importance is the observation that these effects of CRF can all be blocked by administration of the peptide antagonist α-helical ovine CRF(9-41), strongly supporting a specific CRF receptor-mediated event in these behaviors. Furthermore, the CRF receptor antagonist by itself attenuates many of the behavioral consequences of stress underscoring the role of endogenous peptide in mediating many of the stress-related behaviors.

ROLE FOR CRF IN NEUROPSYCHIATRIC DISORDERS AND NEURODEGENERATIVE DISEASES

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

Major Depression and Anxiety Disorders

Many patients with major depression are hypercortisolemic and exhibit an abnormal dexamethasone suppression test. Given the primary role of CRF in stimulating pituitary–adrenocortical secretion, the hypothesis has been put forth that hypersecretion or hyperactivity of CRF in brain might underlie the hypercortisolemia and symptomatology seen in major depression. (See refs. 62 and 63 for review.) The concentration of CRF is significantly increased in the cerebrospinal fluid (CSF) of drug-free individuals (4 ,64 ,65), and a significant positive correlation is observed between CRF concentrations in the CSF and the degree of postdexamethasone suppression of plasma cortisol (66). Furthermore, the observation of a decrease in CRF binding sites in the frontal cerebral cortex of suicide victims compared to controls is consistent with the hypothesis that CRF is hypersecreted in major depression (67). The elevated CSF concentrations of CRF seen in depressed individuals are decreased following treatment with electroconvulsive therapy (68). In addition, a blunted ACTH response to intravenously administered ovine or human CRF is observed in depressed patients when compared to normal controls (69). The blunted ACTH response to exogenous CRF seen in depressed patients may be caused by the intact negative feedback of cortisol on the corticotrophs, a compensatory decrease in CRF receptors subsequent to chronic hypersecretion of the peptide, and/or desensitization of the pituitary corticotrophs to respond to CRF.

A number of studies suggest that anxiety-related disorders (e.g., panic disorder and generalized anxiety disorder) and depression are independent syndromes that share both clinical and biological characteristics. The role that has been proposed for CRF in major depressive disorders along with preclinical data in rats demonstrating effects of CRF administration to produce several behavioral effects characteristic of anxiogenic compounds (61) have led to the suggestion that CRF may also be involved in anxiety-related disorders. A role for CRF in panic disorder has been suggested by observations of blunted ACTH responses to intravenously administered CRF in panic disorder patients when compared to controls (70). The blunted ACTH response to CRF in panic disorder patients most likely reflects a process occurring at or above the hypothalamus, resulting in excess secretion of endogenous CRF.

Anorexia Nervosa

Anorexia nervosa is an eating disorder characterized by tremendous weight loss in the pursuit of thinness. There is similar pathophysiology in anorexia nervosa and depression, including the manifestation of hypercortisolism, hypothalamic hypogonadism, and anorexia. Furthermore, the incidence of depression in anorexia nervosa patients is high. Like depressed patients, anorexics show a markedly attenuated ACTH response to intravenously administered CRF (4 ,64 ,65). When the underweight anorexic subjects are studied after their body weight had been restored to normal, their basal hypercortisolism, increased levels of CRF in the CSF, and diminished ACTH response to exogenous CRF all return to normal at varying periods during the recovery phase (4 ,64 ,65). CRF can potently inhibit food consumption in rats, which further suggests that the hypersecretion of CRF may be responsible for the weight loss observed in anorexics. In addition, the observation that central administration of CRF diminishes a variety of reproductive functions (4 ,65) lends relevance to the clinical observations of hypogonadism in anorexics.

Alzheimer's Disease

Several studies have provided evidence in support of alterations in CRF in Alzheimer's disease (AD) (4 ,65 ,71 ,72 ,73 and 74). There are decreases in CRF content and reciprocal increases in CRF receptors in cerebral cortical areas affected in AD such as the temporal, parietal, and occipital cortex. The reductions in CRF and increases in CRF receptors are all greater than 50% of the corresponding control values. The up-regulation in cerebral cortical CRF receptors in AD under conditions in which the endogenous peptide is reduced suggests that CRF-receptive cells may be preserved in the cortex in AD. Chemical cross-linking studies have demonstrated a normal pattern of labeling of cerebral cortical CRF receptors in AD when compared to age-matched controls (75). Although these decreases in CRF content have a modulatory action on the receptors (up-regulation), there appears to be no effect on the concentration or levels of CRF-binding protein in cerebral cortical areas affected in AD (76). The reduction in cortical CRF content may be owing to selective degeneration of CRF neurons intrinsic to the cerebral cortex or dysfunction of CRF neurons innervating the cortex from other brain areas. Additional evidence for a role for CRF in AD is provided by observations of decreases in CRF in other brain areas including the caudate (71) and decreased concentrations of CRF in the CSF (77 ,78). Furthermore, a significant correlation is evident between CSF CRF and the global neuropsychological impairment ratings, suggesting that greater cognitive impairment is associated with lower CSF concentrations of CRF (79).

Immunocytochemical observations demonstrating morphologic alterations in CRF neurons in AD complement the studies described in the preceding. In AD, swollen, tortuous CRF-immunostained axons, termed fiber abnormalities, are clearly distinguishable from the surrounding normal neurons and are also seen in conjunction with amyloid deposits associated with senile plaques (80). Furthermore, the total

number of CRF-immunostained axons is reduced in the amygdala of Alzheimer's patients (80). Interestingly, the expression of CRF antigen in neurons is not globally reduced in Alzheimer's patients. CRF immunostaining of perikarya and axons located in the hypothalamic paraventricular nucleus is much more intense in AD patients than controls (80). Increased immunostaining of the paraventricular neurons in AD, if truly representative of increased content of CRF, could be related to increased amounts of CRF mRNA in these cells or increased translation of available mRNA. The increased expression and/or release of CRF from the paraventricular nucleus of the hypothalamus would provide a reasonable explanation for the hypercortisolemia often seen in Alzheimer's patients.

At present, the cerebral cortical cholinergic deficiency seems to be the most severe and consistent deficit associated with AD. Reductions in cerebral cortical CRF correlate with decreases in choline acetyl transferase (ChAT) activity (72). In Alzheimer's, there are significant positive correlations between ChAT activity and reduced CRF in the frontal, temporal, and occipital lobes. Similarly, significant negative correlations exist between decreased ChAT activity and increased number of CRF receptors in the three cortices. These data suggest that the reported reciprocal changes in presynaptic and postsynaptic markers in CRF in cerebral cortex of patients with AD may be, in part, a consequence of deficits in the cholinergic projections to the cerebral cortex. Additional studies are necessary to determine the functional significance of the interaction between CRF and cholinergic systems.

Other Neurological Disorders

Alternations in brain concentrations of CRF have been reported in other neurological diseases. For example, in cases of Parkinson's disease (PD) with dementia that also show pathologic features of AD, CRF content is decreased and shows a pattern similar to those cases exhibiting the pathology of AD alone (74 ,81). Specimens from patients with PD who did not have the histopathology characteristic of AD also demonstrate reductions of CRF content, although the reductions are less marked than in cases of combined AD and PD. Normal levels of CRF have been reported in the hypothalamus in PD (82), suggesting that the loss of CRF in the cerebral cortex is not generalized. CRF is deceased to approximately 50% of the control values in the frontal, temporal, and occipital lobes of patients with progressive supranuclear palsy (74 ,81), a rare neurodegenerative disorder that shares certain clinical and pathologic features with AD.

The similarity of the changes in CRF found in the context of the three neurological diseases associated with Alzheimer-type pathology raised the possibility that cerebral cortical reduction is nothing more than a nonspecific sequela of the disease process. In Huntington's disease (HD), a neurological disorder in which minimal cerebral cortical pathology is present, the CRF content in the frontal, temporal, parietal, occipital, and cingulate cortices and in the globus pallidus is not significantly different from that seen in neurologically normal controls (73). However, the CRF content in the caudate nucleus and putamen of the basal ganglia (a brain area that is severely affected in the disease) is less than 40% of the CRF concentrations seen in controls (73). The localization of the CRF changes to only affected brain regions in the four neurodegenerative disorders described suggest that CRF has an important role in the pathology of these dementias.

POTENTIAL THERAPEUTIC STRATEGIES DESIGNED FOR THE CRF SYSTEM

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

Nonpeptide CRF-Binding Protein Ligand Inhibitors

As mentioned, the CRF-binding protein has the capacity to bind and functionally inactivate CRF. Peptide CRF-BP ligand inhibitors have been shown *in vitro*, to release CRF from the binding protein making it available for binding to its receptor and have been theorized to have efficacy in diseases that are associated with low levels of CRF such as Alzheimer's disease (76). Interestingly, unlike the direct i.c.v. administration of CRF, inhibition of the CRF-BP by ligand inhibitors that release functional CRF does not cause anxiogenic-like activity in animal models, validating the approach for diseases that require an increase in CRF function (76). Thus, compounds that act to dissociate CRF from its binding protein complex will act to selectively increase synaptic concentrations of CRF in discrete brain regions and may provide a novel treatment opportunity for disorders associated with low levels of CRF. To date, however, no nonpeptide CRF-BP ligand inhibitors have been discovered to test this hypothesis. On the other hand, small molecule nonpeptide CRF receptor antagonists have shown encouraging preliminary results in this complex system.

Nonpeptide CRF Receptor Antagonists

As indicated in this and previous reviews of the CRF system, there are a number of neuropsychiatric indications where clinical and preclinical data have shown that the CRF system is hyperactivated as evidenced by abnormally high levels of CRF within the CNS. For example, in clinical studies of major depression, increased CRF concentrations in the cerebrospinal fluid, increased HPA activity, blunted ACTH responsiveness to exogenously administered CRF, and pituitary and adrenal hypertrophy have all indicated hypersecretion of CRF associated with this disorder. (See refs. 4 , 5 , 32 , 65 , and 83 for review.) It follows that for these indications the therapeutic strategy would involve functional blockade of the actions of CRF. This can be achieved

through inhibition of CRF synthesis and secretion, inactivation of CRF (either by antibody neutralization or increased metabolism), or direct antagonism by specific receptor blockade. Although anti-CRF antibodies have demonstrated antiinflammatory effects (84), these types of therapies are limited to peripheral use and could not be readily formulated for oral administration for central activity. Nevertheless, these therapeutics could be useful for the treatment of disorders with peripheral elevations of CRF such as rheumatoid arthritis.

The best strategy for the blockade of elevated CRF levels is to design specific and selective nonpeptide receptor antagonists. Particularly for use in the brain, these molecules can be designed to have receptor subtype specificity, good oral bioavailability, and rapid penetration across the blood–brain barrier; characteristics that are difficult to optimize with peptide therapies. The recent surge in combinatorial chemistry techniques, coupled with recent technological advances in robotic high-throughput screening and data management of large libraries of molecules have enabled the field of small molecule drug discovery. These advancements have led to the identification of several patented structural series of molecules known to antagonize the effects of CRF at the CRF_1 receptor subtype. (See ref. 85 for a complete review.) Several of these small molecules have recently appeared in the literature and reported to have good CRF_1 receptor antagonistic activity. Compounds such as CP 154,526 (86), NBI 27914 (87), Antalarmin (88), and most recently DMP 696 (89), all demonstrate a good *in vitro* profile, showing selectivity for the CRF_1 receptor subtype. Systemic administration of these compounds have been found to attenuate stress-induced elevations in plasma ACTH levels in rats, demonstrating that CRF_1 receptors can be blocked in the periphery. Furthermore, nonpeptide CRF_1 antagonists administered peripherally have also been demonstrated to inhibit CRF-induced seizure activity (90). Until recently, however, these compounds have suffered from poor solubility and pharmacokinetics, thus limiting their utility in *in vivo* characterization of the individual compounds as well as the overall proof of concept for the mechanism of CRF receptor efficacy. One compound has very recently been described as a water-soluble nonpeptide CRF_1 receptor antagonist (NBI 30775, also referred to as R121919) that demonstrates high affinity, and has a superior *in vitro* and *in vivo* profile compared to other nonpeptide CRF receptor antagonists (91).

NBI 30775, is a pyrazolopyrimidine with high affinity for the CRF_1 receptor and over 1,000-fold weaker activity at the CRF_2 receptor subtype. This compound does not interact at all with the CRF binding protein and was shown to be as potent as the peptide antagonist D-Phe CRF(12-41) at inhibiting the CRF-stimulated cAMP accumulation from cells that express the human CRF_1 receptor and CRF-stimulated ACTH release from cultured rat anterior pituitary cells *in vitro. In vivo*, this molecule potently attenuated the plasma elevations of ACTH observed following a stressful stimulus in the rat, and demonstrated both dose- and time-dependent CRF_1 receptor occupancy concomitant with the levels of drug measured in whole brain (91). Owing to the promising preclinical profile of this compound this particular compound was assessed in full Phase I clinical trials and an initial open label Phase IIa study where the compound was assessed in patients with major depressive disorder.

First Clinical Experience with CRF_1 Receptor Antagonists

Preclinically, studies have demonstrated that attenuation of the CRF system, either by decreasing synthesis and release or by selective blockade of the CRF receptor, results in decreased anxiety and behavioral activation in stressed animals; however, clinically it will probably not be beneficial to the overall outcome of the patient if the stress axis is maximally compromised. The preclinical studies described, prompted the development of NBI 30775 in Phase I safety studies in humans and in an open-label clinical trial in patients with major depressive disorder (92). In this latter study severely depressed patients were given NBI 30775 orally once daily in a dose-escalating manner and their hypothalamic-pituitary adrenal (HPA) function assessed. In addition, the patients' level of depression or anxiety was measured using the Hamilton depression (HAM-D) and anxiety (HAM-A) scales. The results demonstrated that this compound was safe and well tolerated under the conditions of this study. Moreover, the data suggested that blockade of the CRF_1 receptor in these patients did not result in an impairment of the HPA axis either at baseline or following an exogenous CRF challenge (92). This demonstration was critical in setting this potential therapy apart from existing therapies that blunt basal functioning of multiple neurotransmitter systems. Furthermore, although under the limited conditions of an open-label trial, there was a statistically significant dose-dependent reduction in the depression and anxiety scores using both clinician and patient ratings, suggesting that this mechanism may provide an exciting novel therapy in patients suffering with major depressive disorder. Although it is of great importance at this stage to develop these compounds as tools in the ultimate understanding of the CRF system and the role it plays in neuropsychiatric disorders, evidence is now beginning to emerge that compounds of this class, and more importantly of this mechanism, will prove beneficial in neuropsychiatric disorders such as depression or anxiety. Should compounds such as this continue to demonstrate efficacy in these disorders without severely compromising the stress axis as a whole, they would validate the CRF hypothesis for depression and anxiety and provide an entirely novel treatment for these devastating diseases.

CONCLUSION

Corticotropin-releasing factor is the key regulator of the organism's overall response to stress. CRF has hormone-like effects at the pituitary level to regulate ACTH secretion that, in turn coordinates the synthesis and secretion of glucocorticoids from the adrenal cortex. CRF also functions as a bona fide neurotransmitter in the CNS. CRF neurons and receptors are widely distributed in the CNS and play a critical role in coordinating the autonomic, electrophysiologic, and behavioral responses to stress.

Clinical data have implicated CRF in the etiology and pathophysiology of various endocrine, psychiatric, and neurological disorders. Hypersecretion of CRF in brain may contribute to the symptomatology seen in neuropsychiatric disorders such as depression, anxiety-related disorders, and anorexia nervosa. In contrast, deficits in brain CRF are apparent in neurodegenerative disorders such as AD, PD, and HD as they relate to dysfunction of CRF neurons in brain areas affected in the particular disorder. The recent discovery of novel receptor family members as well as novel alternative ligands for these subtypes serve not only to increase our understanding of the system but provide a basis for selective and rational drug design for the treatment of disorders that are associated with aberrant levels of CRF. Strategies directed at developing specific and selective CRF agents have yielded many nonpeptide small molecule CRF_1 receptor antagonists and a preliminary proof-of-concept has encouraged the further development of such agents. Compounds such as those described in this chapter may hold promise for novel therapies for the treatment of these various neuropsychiatric disorders without severely compromising this highly complex hormonal system. Clearly with the recent advances made within a very short period of time, it now seems possible to begin a full understanding of this increasingly complex neurohormone system.

ACKNOWLEDGMENT

Part of "7 - Corticotropin-Releasing Factor: Physiology, Pharmacology, and Role in Central Nervous System Disorders "

Dr. De Souza is a stockholder in Neurocrine Biosciences, Inc.

REFERENCES

1. Vale W, Spiess J, Rivier C, et al. Characterization of a 41-residue ovine hypothalamic peptide that stimulates secretion of corticotropin and B-endorphin. *Science* 1981;213:1394–1397.

2. Chen R, Lewis KA, Perrin MH, et al. Expression cloning of a human corticotropin-releasing-factor receptor. *Proc Natl Acad Sci USA* 1993;90:8967–8971.

3. Dunn AJ, Berridge CW. Physiological and behavioral responses to corticotropin-releasing factor administration: is CRF a mediator of anxiety of stress responses? *Brain Res Rev* 1990;15:71–100.

4. Owens MJ, Nemeroff CB. Physiology and pharmacology of corticotropin-releasing factor. *Pharmacol Rev* 1991;43:425–473.

5. De Souza EB, Nemeroff CB, eds. *Corticotropin-releasing factor: basic and clinical studies of a neuropeptide*. Boca Raton, FL: CRC Press, 1990.

6. Chadwick DJ, Marsh J, Ackrill K, eds. *Corticotropin-releasing factor*. Chichester, England: John Wiley and Sons, 1993.

7. Romier C, Bernassau J-M, Cambillau C, et al. Solution structure of human corticotropin releasing factor by ^{1}H NMR and distance geometry with restrained molecular dynamics. *Protein Eng* 1993;6:149–156.

8. Rivier J, Rivier C, Vale W. Synthetic competitive antagonist of corticotropin-releasing factor: effect on ACTH secretion in the rat. *Science* 1984;224:889–891.

9. Thompson RC, Seasholtz AF, Douglass JO, et al. Cloning and distribution of expression of the rat corticotropin-releasing factor (CRF) gene. In: *Corticotropin releasing factor: basic and clinical studies of a neuropeptide*. Boca Raton, FL: CRC Press, 1990:1–12.

10. Majzoub JA, Emanuel R, Adler G, et al. Second messenger regulation of mRNA for corticotropin-releasing factor. In: *Corticotropin releasing factor*. Chichester, England: John Wiley and Sons, 1993:30–43.

11. Swanson LW, Sawchenko PE, Rivier J, et al. The organization of ovine corticotropin-releasing factor immunoreactive cells and fibers in the rat brain: an immunohistochemical study. *Neuroendocrinology* 1983;36:165–186.

12. Sawchenko PE, Swanson LW Organization of CRF immunoreactive cells and fibers in the rat brain: immunohistochemical studies. In: *Corticotropin-releasing factor: basic and clinical studies of a neuropeptide*. Boca Raton, FL: CRC Press, 1990:29–52.

13. Petrusz P, Merchenthaler I. The corticotropin-releasing factor system. In: *Neuroendocrinology*. Boca Raton, FL: CRC Press, 1992:129–184.

14. Owens MJ, Nemeroff CB. Neurotransmitter regulation of the CRF secretion in vitro. In: *Corticotropin-releasing factor: basic and clinical studies of a neuropeptide*. Boca Raton, FL: CRC Press, 1990:107–114.

15. Letter A, McMaster B, Moore G, et al. Complete amino acid sequence of urotensin I, a hypotensive and corticotropin releasing neuropeptide from Catoftomus. *Science* 1982;218:162–164.

16. Montecucchi PC, Henschen A. Amino acid composition and sequence analysis of sauvagine, a new active peptide from the skin of *Phyllomedusa sauvagei. J Protein Res* 1981;18:113–120.

17. Okawara Y, Morley SD, Burzio LO, et al. Cloning and sequence analysis of cDNA for corticotropin-releasing factor precursor from the teleost fish Catostomus commersoni. *Proc Natl Acad Sci USA* 1988;85:8439–8443.

18. Stenzel-Poore MP, Heldwein KA, Stenzel P, et al. Characterization of the genomic corticotropin-releasing factor (CRF) gene from *Xenopus laevis:* two members of the CRF family exist in amphibians. *Mol Endocrinol* 1992;6:1716–1724.

19. Lovenberg TW, Liaw CW, Grigoriadis DE, et al. Cloning and characterization of a functionally distinct corticotropin-releasing factor receptor subtype from rat brain. *Proc Natl Acad Sci USA* 1995;92:836–840.

20. Vaughan J, Donaldson C, Bittencourt J, et al. Urocortin, a mammalian neuropeptide related to fish urotensin I and to corticotropin-releasing factor. *Nature* 1995;378:287–292.

21. Donaldson CJ, Sutton SW, Perrin MH, et al. Cloning and characterization of human urocortin. *Endocrinology* 1996;137:2167–2170.

22. Vita N, Laurent P, Lefort S, et al. Primary structure and functional expression of mouse pituitary and human brain corticotrophin releasing factor receptors. *Febs Letts* 1993;335:1–5.

23. Chang CP, Pearse RI, O'Connell S, et al. Identification of a seven transmembrane helix receptor for corticotropin-releasing factor and sauvagine in mammalian brain. *Neuron* 1993;11:1187–1195.

24. Perrin M, Donaldson C, Chen R, et al. Identification of a second corticotropin-releasing factor receptor gene and characterization of a cDNA expressed in heart. *Proc Natl Acad Sci USA* 1995;92:2969–2973.

25. Grigoriadis DE, De Souza EB. Heterogeneity between brain and pituitary corticotropin-releasing factor receptors is due to differential glycosylation. *Endocrinology* 1989;125:1877–1888.

26. Kishimoto T, Pearse II RV, Lin CR, et al. A sauvagine/corticotropin-releasing factor receptor expressed in heart and skeletal muscle. *Proc Natl Acad Sci USA* 1995;92:1108–1112.

27. Liaw CW, Lovenberg TW, Barry G, et al. Cloning and characterization of the human CRF2 receptor gene and cDNA. *Endocrinology* 1996;137:72–77.

28. Kostich W, Chen A, Sperle K, et al. Molecular cloning and expression analysis of human CRH receptor type 2 α and B isoforms. *Soc Neurosci Abs* 1996;22:1545.

29. Aguilera G, Millan MA, Hauger RL, et al. Corticotropin-releasing factor receptors: distribution in brain, pituitary and peripheral tissues. In: *The hypothalamic-pituitary-adrenal axis revisited*. New York: The New York Academy of Sciences, 1987:48–66.

30. De Souza EB. Corticotropin-releasing factor receptors in the rat central nervous system: characterization and regional distribution. *J Neurosci* 1987;7:88–100.

31. Webster EL, Grigoriadis DE, De Souza EB. Corticotropin-releasing factor receptors in the brain-pituitary-immune axis. In: *Stress, neuropeptides, and systemic disease*. San Diego: Academic Press, 1991:233–260.

32. De Souza EB. Corticotropin-releasing hormone receptors. In: *Handbook of chemical neuroanatomy: neuropeptide receptors in the CNS, Part III*. Amsterdam: Elsevier, 1992:145–185.

33. Grigoriadis DE, Heroux JA, De Souza EB. Characterization and regulation of corticotropin-releasing factor receptors in the central nervous, endocrine and immune systems. In: *Corticotropin-releasing factor*. Chichester, England: John Wiley and Sons, 1993:85–101.

34. Grigoriadis DE, Liu XJ, Vaughn J, et al. ^{125}I-Tyr^{0}-Sauvagine: a novel high affinity radioligand for the pharmacologic and biochemical study of human corticotropin-releasing factor$_{2\alpha}$ receptors. *Mol Pharmacol* 1996;50:679–686.

35. Liaw C, Grigoriadis DE, De Souza EB, et al. Colorimetric assay for rapid screening of corticotropin-releasing factor receptor ligands. *J Mol Neurosci* 1994;5:83–92.

36. Vaughan J, Donaldson C, Bittencourt J, et al. Characterization of a novel neuropeptide in rat brain related to CRF. In: *Proceedings of the 25th Annual Meeting of the Neuroscience Society*, San Diego, 1995.

37. Chalmers DT, Lovenberg TW, De Souza EB. Localization of novel corticotropin-releasing factor receptor (CRF2) mRNA to specific sub-cortical nuclei in rat brain: comparison with CRF1 receptor mRNA expression. *J Neurosci* 1995;15:6340–6350.

38. De France JF. *The septal nuclei*. New York: Plenum, 1976.

39. Lovenberg TW, Chalmers DT, Liu C, et al. CRF$_{2\alpha}$ and CRF$_{2B}$ receptor mRNAs are differentially distributed between the rat central nervous system and peripheral tissues. *Endocrinology* 1995;136:4139–4142.

40. Battaglia G, Webster EL, De Souza EB. Characterization of corticotropin-releasing factor receptor-mediated adenylate cyclase activity in the rat central nervous system. *Synapse* 1987;1:572–581.

41. Heisler S, Hook VYH, Axelrod J. Corticotropin-releasing factor stimulation of protein carboxylmethylation in mouse pituitary tumor cells. *Biol Pharmacol* 1983;32:1295–1299.

42. Abou-Samra A-B, Harwood JP, Catt KJ, et al. Mechanisms of action of CRF and other regulators of ACTH release in pituitary corticotrophs. In: *The hypothalamic-pituitary-adrenal axis revisited*. New York: New York Academy of Sciences, 1987:67–84.

43. Laatikainen T, Virtanen T, Raisanen I, et al. Immunoreactive corticotropin-releasing factor and corticotropin in plasma during pregnancy, labour and puerperium. *Neuropeptides* 1987;10:343–353.

44. Linton EA, Wolfe CDA, Behan DP, et al. A specific carrier substance for human corticotropin-releasing factor in late gestational maternal plasma which could mask the ACTH-releasing activity. *Clin Endocrinol* 1988;28:315–324.

45. Suda T, Iwashita M, Tozawa F, et al. Characterization of CRH binding protein in human plasma by chemical cross-linking and its binding during pregnancy. *J Clin Endocrinol Metab* 1988;67:1278–1283.

46. Shibasaki T, Odagiri E, Shizume K, et al. Corticotropin-releasing factor like activity in human placental extract. *J Clin Endocrinol Metab* 1982;55:384–386.

47. Behan DP, Linton EA, Lowry PJ. Isolation of the human plasma corticotrophin-releasing factor-binding protein. *J Endocrinol* 1989;122:23–31.

48. Potter E, Behan DP, Fischer WH, et al. Cloning and characterization of the cDNAs for human and rat corticotropin-releasing factor-binding proteins. *Nature* 1991;349:423–426.

49. Suda T, Sumitomo T, Tozawa F, et al. Corticotropin-releasing factor-binding protein is a glycoprotein. *Biochem Biophys Res Commun* 1989;165:703–707.

50. Potter E, Behan DP, Linton EA, et al. The central distribution of a corticotropin-releasing factor (CRF)-binding protein predicts multiple sites and modes of interaction with CRF. *Proc Natl Acad Sci USA* 1992;89:4192–4196.

51. Vale W, Rivier C, Brown MR, et al. Chemical and biological characterization of corticotropin-releasing factor. *Rec Prog Hormone Res* 1983;39:245–270.

52. Plotsky PM, Cunningham ETJ, Widmaier EP. Catecholaminergic modulation of corticotropin-releasing factor and adrenocorticotropin secretion. *Endocrinol Rev* 1989;10:437–458.

53. Anderson SM, Kant GJ, De Souza EB. Effects of chronic stress on anterior pituitary and brain corticotropin-releasing factor receptors. *Pharmacol Biochem Behav* 1993;44:755–761.

54. Heroux JA, Grigoriadis DE, De Souza EB. Age-related decreases in corticotropin-releasing factor (CRF) receptors in rat brain and anterior pituitary gland. *Brain Res* 1991;542:155–158.

55. Valentino RJ. Effects of CRF on spontaneous and sensory-evoked activity of locus ceruleus neurons. In: *Corticotropin-releasing factor: basic and clinical studies of a neuropeptide*. Boca Raton, FL: CRC Press, 1990:217–232.

56. Siggins GR. Electrophysiology of corticotropin-releasing factor in nervous tissue. In: *Corticotropin-releasing factor: basic and clinical studies of a neuropeptide*. Boca Raton, FL: CRC Press, 1990:205–216.

57. Ehlers CL. CRF effects on EEG activity: implications for the modulation of normal and abnormal brain states. In: *Corticotropin-releasing factor: basic and clinical studies of a neuropeptide*. Boca Raton, FL: CRC Press, 1990:233–252.

58. Fisher LA. Corticotropin-releasing factor: endocrine and autonomic integration of responses to stress. *Trends Pharmacol Sci* 1989;10:189–193.

59. Brown MR, Fisher LA. Regulation of the autonomic nervous system by corticotropin-releasing factor. In: *Corticotropin-releasing factor: basic and clinical studies of a neuropeptide*. Boca Raton, FL: CRC Press, 1990:291–298.

60. Tache Y, Gunion MM, Stephens R CRF. Central nervous system action to influence gastrointestinal function and role in the gastrointestinal response to stress. In: *Corticotropin-releasing factor: basic and clinical studies of a neuropeptide*. Boca Raton, FL: CRC Press, 1990:299–308.

61. Koob GF, Britton KT. Behavioral effects of corticotropin-releasing factor. In: *Corticotropin-releasing factor: basic and clinical studies of a neuropeptide*. Boca Raton, FL: CRC Press, 1990:253–265.

62. Nemeroff CB. Psychopharmacology of affective disorders in the 21st century. *Biol Psychiatry* 1998;44:517–525.

63. Holsboer F. The rationale for corticotropin-releasing hormone receptor (CRH-R) antagonists to treat depression and anxiety. *J Psychiatr Res* 1999;33:181–214.

64. Nemeroff CB, Widerlov E, Bissett G, et al. Elevated concentration of CSF corticotropin-releasing factor-like immunoreactivity in depressed patients. *Science* 1984;226:1342–1344.

65. De Souza EB. Role of corticotropin-releasing factor in neuropsychiatric disorders and neurodegenerative diseases. In: *Annual reports in medicinal chemistry*. San Diego: Academic Press, 1990:215–224.

66. Roy A, Pickar D, Paul S, et al. CSF corticotropin-releasing hormone in depressed patients and normal control subjects. *Am J Psychiatry* 1987;143:896–899.

67. Nemeroff CB, Owens MJ, Bissett G, et al. Reduced corticotropin-releasing factor receptor binding sites in the frontal cortex of suicide victims. *Arch Gen Psychiatry* 1988;45:577–579.

68. Nemeroff CB, Bissette G, Akil H, et al. Neuropeptide concentrations in the cerebrospinal fluid of depressed patients treated with electroconvulsive therapy: corticotropin-releasing factor, B-endorphin and somatostatin. *Br J Psychiatry* 1991;158:59–63.

69. Gold PW, Loriaux DL, Roy A, et al. Responses to corticotropin-releasing hormone in the hypercortisolism of depression and Cushing's disease. *New Engl J Med* 1986;314:1329–1334.

70. Roy-Byrne PP, Uhde T, Post R, et al. The corticotropin-releasing hormone stimulation test in patients with panic disorder. *Am J Psychiatry* 1986;143:896–899.

71. Bissette G, Reynolds GP, Kilts CD, et al. Corticotropin-releasing factor-like immunoreactivity in senile dementia of the Alzheimer type. *JAMA* 1985;254:3067–3069.

72. De Souza EB, Whitehouse PJ, Kuhar MJ, et al. Reciprocal changes in corticotropin-releasing factor (CRF)-like immunoreactivity and CRF receptors in cerebral cortex of Alzheimer's disease. *Nature* 1986;319:593–595.

73. De Souza EB, Whitehouse PJ, Price DL, et al. Abnormalities in corticotropin-releasing hormone (CRH) in Alzheimer's disease and other human disorders. *Ann NY Acad Sci* 1987;512:237–247.

74. De Souza EB. CRH defects in Alzheimer's and other neurological diseases. *Hosp Pract* 1988;23:59–71.

75. Grigoriadis DE, Struble RG, Price DL, et al. Normal pattern of labeling of cerebral cortical corticotropin-releasing factor (CRF) receptors in Alzheimer's disease: evidence from chemical cross-linking studies. *Neuropharmacology* 1989;28:761–764.

76. Behan DP, Heinrichs SC, Troncoso JC, et al. Displacement of corticotropin releasing factor from its binding protein as a treatment for Alzheimer's disease. *Nature* 1995;378:284–287.

77. Mouradian MM, Farah JM Jr, Mohr E, et al. Spinal fluid CRF reduction in Alzheimer's disease. *Neuropeptides* 1986;8:393–400.

78. May C, Rapoport SI, Tomai TP, et al. Cerebral spinal fluid concentrations of corticotropin-releasing hormone (CRH) and corticotropin (ACTH) are reduced in Alzheimer's disease. *Neurology* 1987;37:535–538.

79. Pomara N, Singh RR, Deptula D, et al. CSF corticotropin-releasing factor (CRF) in Alzheimer's disease: its relationship to severity of dementia and monoamine metabolites. *Biol Psychiatry* 1989;26:500–504.

80. Powers RE, Walker LC, De Souza EB, et al. Immunohistochemical study of neurons containing corticotropin-releasing factor in Alzheimer's disease. *Synapse* 1987;1:405–410.

81. Whitehouse PJ, Vale WW, Zweig RM, et al. Reduction in corticotropin-releasing factor-like immunoreactivity in cerebral cortex in Alzheimer's disease, Parkinson's disease, and progressive supranuclear palsy. *Neurology* 1987;37:905–909.

82. Conte-Devolx B, Grino M, Nieoullon A, et al. Corticoliberin, somatocrinin and amine contents in normal and parkinsonian human hypothalamus. *Neurosci Letts* 1985;56:217–222.

83. Owens MJ, Nemeroff CB. Preclinical and clinical studies with corticotropin-releasing factor: implications for affective disorders. *Psychopharmacol Bull* 1988;24:335–339.

84. Karalis K, Sano H, Redwine J, et al. Autocrine or paracrine inflammatory actions of corticotropin-releasing hormone in vivo. *Science* 1991;254:421–423.

85. McCarthy JR, Heinrichs SC, Grigoriadis DE. Recent advances with the CRF1 receptor: design of small molecule inhibitors, receptor subtypes and clinical indications. *Curr Pharm Design* 1999;5:289–315.

86. Chen YL, Mansbach RS, Winter SM, et al. Synthesis and oral efficacy of a 4-(butylethylamino)pyrrolo[2,3-d]pyrimidine: a centrally active corticotropin-releasing factor1 receptor antagonist. *J Med Chem* 1997;40:1749–1752.

87. Chen C, Dagnino R Jr, De Souza EB, et al. Design and synthesis of a series of nonpeptide high-affinity human corticotropin-releasing factor1 receptor antagonists. *J Med Chem* 1996;39:4358–4360.

88. Webster EL, Lewis DB, Torpy DJ, et al. In vivo and in vitro characterization of Antalarmin, a nonpeptide corticotropin-releasing hormone (CRH) receptor antagonist: suppression of pituitary ACTH release and peripheral inflammation. *Endocrinology* 1996;137:5747–5750.

89. He L, Gilligan PJ, Zaczek R, et al. 4-(1,3-Dimethoxyprop-2-ylamino)-2,7-dimethyl-8-(2,4-dichlorophenyl)pyrazolo[1,5-a]1,3,5-triazine: a potent, orally bioavailable CRF(1) receptor antagonist. *J Med Chem* 2000;43:449–456.

90. Baram TZ, Chalmers DT, Chen C, et al. The CRF1 receptor mediates the excitatory actions of corticotropin releasing factor (CRF) in the developing rat brain: in vivo evidence using a novel, selective, nonpeptide CRF receptor antagonist. *Brain Res* 1997;770:89–95.

91. Grigoriadis DE, Chen C, Wilcoxen K, et al. NBI 30775/R121919: a novel nonpeptide corticotropin-releasing factor1 (CRF1) receptor antagonist for the potential treatment of depression and anxiety-related disorders. Unpublished manuscript.

92. Zobel A, Nickel T, Kunzel H, et al. Effects of the high-affinity corticotropin-releasing hormone receptor 1 antagonist R121919 in major depression: the first 20 patients treated. *J Psychiatric Res* 2000;34:171–181.

8

Neurogenesis in Adult Brain

Fred H. Gage

Henriette Van Praag

Fred H. Gage and Henriette van Praag: The Laboratory of Genetics, The Salk Institute for Biological Studies, La Jolla, California.

HISTORIC PERSPECTIVE

Part of "8 - Neurogenesis in Adult Brain "

The idea that the adult brain retains the capacity to generate new neurons has been proposed several times over the last 40 years, and in each case both conceptual and technical constraints have led to resistance. Joseph Altman first reported that some dividing cells in the adult brain survived and differentiated into cells with morphology similar to neurons using tritiated thymidine autoradiography (1). Over the subsequent years he and his colleagues confirmed these initial observations and focused on the few areas where neurogenesis was apparent in the adult, at the light microscopic level, while systematically documenting the birth dates of neurons throughout the brain during development (2). When contrasting adult neurogenesis with the extensive neurogenesis in development, adult neurogenesis seemed almost like an epiphenomenon. The continued skepticism surrounding adult neurogenesis and the absence of definitive phenotypic markers limited the development of the field. In the mid-1970s and early 1980s Michael Kaplan reexamined the initial observations using the electron microscope and added substantial confidence that not only could neurogenesis occur in the adult brain, but also that the cells appear ultrastructurally, similar to sister cells in the dentate gyrus of the hippocampus, one of the structures shown to be neurogenic (3). In the mid-1980s, Fernando Nottebohm and his student Steve Goldman further stimulated this fledgling field by showing that songbirds experience a seasonal cell death and neurogenesis in a region of the brain important for song production (4). They have continued to reveal more about the environment and molecular regulators of this process in the adult avian brain (5). Despite these observations of neurogenesis in the adult brain, confusion over the mechanism origin of cell genesis in the adult brain persisted. In the early 1990s a series of papers in the adult mouse and rat revealed that cells with stem cell properties could be isolated and expanded in culture. Under a variety of culture conditions with different factors, these isolated cells can be induced to differentiate into glia and neurons (6 ,7 ,8 and 9). This later observation provided a mechanism for the neurogenesis in the adult. Mature committed neurons were not dividing, but rather a population of immature stem-like cells exists in the brain and it is likely that it is the proliferation and differentiation of this population that is resulting in neurogenesis. With this conceptual framework the original statement by Cajal that, “Once development was ended, the fonts of growth and regeneration of the axons and dendrites dried up irrevocably. In adult center, the nerve paths are something fixed and immutable: everything may die, nothing may be regenerated,” still holds true for most areas of the adult brain. The new lesson that we have learned is that development never ended in some areas of the brain.

CHARACTERIZATION OF CELL GENESIS *IN VIVO*

Part of "8 - Neurogenesis in Adult Brain "

Areas of Neurogenesis

Neurogenesis is a process that includes cell division, migration, and differentiation. There appear, at present, to be only two areas of the brain where stems cells initially reside and proliferate prior to migration and differentiation. Those areas are the lateral subventricular and subgranular zones of the dentate gyrus. The exact cell that corresponds to the initiating stem cell in the lateral ventricle is a point of contention. One view is that this cell is the ependymal cell facing the ventricle (10), whereas an alternative view is that a glial population one cell layer in from the ependyma are the stem cells (11). In any case, after cell division one of the stem cells begins to differentiate and migrate in what Alvarez-Buylla has described as “chain migration” along the rostral migratory stream toward the olfactory bulb where they differentiate into interneurons in the bulb (12). This process continues throughout life; the functional importance and consequences of this process are not understood. The stem cell in the adult mammalian subgranular zone of the dentate gyrus is likely an ectopically displaced cell

originating from the developing ventricular zone. It remains formally possible that a more primitive cell exists elsewhere in the adult brain in a quiescent state, and migrates to the dentate gyrus where the cells begin to divide. However, from the subgranular zone, one of the progeny migrates into the granule cell layer once the cells divide; there the majority become neurons with axons extending to the CA3 pyramidal neurons and receiving synaptic connections. The function of these newly born cells is being investigated.

Although these are the two principal areas where neurogenesis occurs in the mammalian brain, cell genesis occurs throughout the adult brain including cortex, optic nerve, spinal cord, and many brainstem and forebrain structures. To date the function of this cell genesis in the normal intact brain and spinal cord is not known, but some of these new cells can become glial cells (13). A clear challenge for the future is to document all the areas of the adult brain where cell genesis continues, and to understand the normal function as well as the factors that regulate this process.

Mammalian Species in Which Adult Neurogenesis Is Documented

The first studies demonstrating adult neurogenesis were in the rat. Subsequently, rabbit and cat were shown to exhibit similar characteristics, although little additional work has been done in those animals since the original publications (14 ,15). It was not until 1997 that Kempermann and associates showed the mice retain neurogenesis and that significant genetic variability exists among mouse strains (16). These findings, together with the important role for environmental stimulation as a regulator of neurogenesis, have placed adult neurogenesis as paradigm for examining the interactions of nature and nurture (17 ,18). Although there was debate in the mid-1980s as to whether nonhuman primates retained adult neurogenesis, a series of papers by Fuchs and associates beginning with tree shrew followed by marmosets and finally with Rhesus monkeys demonstrated and confirmed that neurogenesis in occurs in adult nonhuman primates (19 ,20). Recently, Gould presented data suggesting that neurogenesis occurs in the adult primate frontal cortex and concluded that the cells are derived from the subventricular zone where they migrate to specific cortical regions of the adult primate brain. This observation awaits confirmation (21). One of the markers for determining cell division is bromo-deoxyuridine (BrdU); a traceable analogy of uridine, which in incorporated into the genome of, cells undergoing cell division. Administering BrdU and then examining cell proliferation in tumor biopsies is occasionally used to monitor tumor progression in patients with cancer. Because BrdU is a small soluble molecule, it is distributed throughout the body including the brain, thus can be a marker for cell and neurogenesis in humans. In 1998 Eriksson and colleagues (22) reported that in five of the cancer patients they examined who received BrdU at between 15 days and over 2 years early, all of them showed neurogenesis as revealed by colabeling of BrdU with markers of mature neurons in the dentate gyrus. Together these studies clearly demonstrate that neurogenesis, at least in the dentate gyrus, is a process that persists throughout life and in all mammalian species (Fig. 8.1).

The extent to which or whether neurogenesis can occur in other brain areas remains an area of intense investigation.

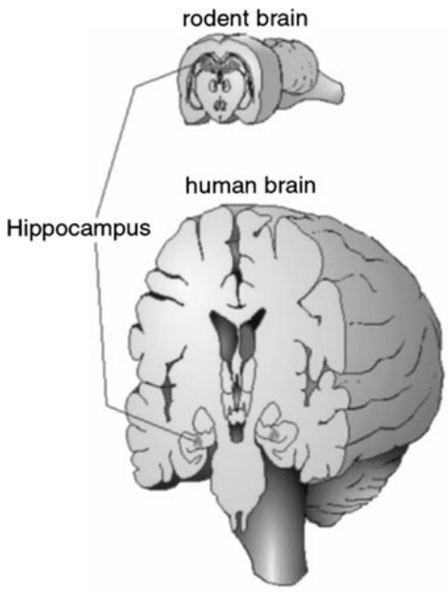


FIGURE 8.1. Birth of new neurons in the adult hippocampus has been documented in a variety of species, including rodents and humans. See color version of figure.

PROPERTIES OF STEM CELLS *IN VITRO*

Part of "8 - Neurogenesis in Adult Brain "

CNS Areas from Which Stem Cells Can Be Isolated

Neural stem cells can be derived from the adult brain and propagated *in vitro* (6 ,7 and 8 ,10 ,11 ,23 ,24 ,25 and 26). Ongoing adult neurogenesis, however, has only been documented convincingly in two brain areas, the subventricular zone/olfactory system and the dentate gyrus of the hippocampus. Interestingly, stem cells are found not only in these regions, but also have been isolated from areas that are non-neurogenic such as the septum, striatum (27), spinal cord (28), cerebral cortex, corpus callosum, and optic nerve (29) and eye (30). In culture, these cells are multipotent and can give rise to neurons and glia. However, cells isolated from areas outside of the hippocampus and subventricular zone require high levels of FGF-2 in order to give rise to neurons, rather than only glial cells (29). These findings suggest that either different

populations of stem cells exist in the nervous system or that they require unique culture conditions to become multipotent. Alternatively, the cells isolated from different CNS regions may already be committed toward a specific lineage. Indeed, there are no antigenic markers that allow unambiguous identification of stem cells in the nervous system. In the subventricular zone, stem cells are suggested to divide slowly, whereas and their offspring, progenitor cells, may divide more frequently (31). Stem cells in this area have been suggested to ependymal cells (10) or a subclass of glial cells in the subependymal zone (11). The location and identity of the hippocampal stem cell remains to be determined.

Factors That Affect Proliferation and Differentiation of Stem Cells In Vitro

A variety of cytokines, neurotrophins, and conditioned media are used to culture neural progenitor cells (32 ,33 and 34). The two major factors are EGF and FGF. Progenitor cells responsive to EGF have been isolated and cultured from adult mouse subventricular zone (6 ,7 ,31). FGF-2 has been found to be mitogenic for adult neural progenitors from brain and spinal cord (9 ,27 ,28). FGF-2, however, is member of a family of 10 related, but genetically and functionally distinct polypeptides. Among those, only FGF-2 and FGF-4 are mitogens for neural progenitor cells. Moreover, a comparison of amino acid sequences between the FGFs revealed a striking similarity between a 10-amino acid sequence of FGF-2 and FGF-4. This 10-amino acid sequence was been shown to elicit the mitogenic effects of FGF-2 and FGF-4 on neural progenitor cells, whereas similar regions in FGF-1 and -5 were found to be inactive (35).

Several factors have been found to be important for neuronal differentiation in cultured progenitor cells. In particular, retinoic acid and cAMP increase neuronal differentiation (27 ,36 ,37). In addition, neurotrophins such as NGF, BDNF, and NT-3 have been found to influence neuronal differentiation and transmitter phenotype (34 ,37 ,38), whereas CNTF can regulate glial differentiation of precursor cells (39).

Transplanted Cells and Responses to Local Cues

Progenitor cells may play an important role in brain or spinal cord repair. In particular, grafting of progenitor cells into degenerated or injured areas may be used to replace cells that are no longer functional. The phenotype and function that these cells acquire appears to be very much dependent on the specific environment into which they are transplanted. Thus, cultured hippocampal progenitors become granule cell neurons when grafted into the dentate gyrus, tyrosine hydroxylase, and calbindin positive neurons in the olfactory bulb after grafting into the rostral migratory stream (40). In addition, after implantation into the developing retina these cells showed properties of several types of retinal neurons (37). Moreover, progenitors isolated from a non-neurogenic area such as the spinal cord, acquired the morphologic characteristics of granule cell neurons when grafted into the dentate gyrus, and had a glial phenotype when grafted back in to the spinal cord (41). These studies suggest that neural stem cells derived from the adult mammalian brain retain multipotentiality. Recent research suggests that neuronal stem cells are multipotent outside the CNS as well. It was reported that neural progenitor cells repopulate experimentally depleted bone marrow and reconstitute the hematopoietic system (42). It remains to be determined what the local cues are, that are driving the neuronal precursor cells to acquire such specific fates when transplanted *in vivo*.

REGULATION OF PROLIFERATION AND DIFFERENTIATION *IN VIVO*

Part of "8 - Neurogenesis in Adult Brain "

The mechanisms that generate new granule cells in the dentate gyrus are poorly understood. A variety of environmental, behavioral, genetic, neuroendocrine, and neurochemical factors can regulate adult neurogenesis. Two critical processes that lead to neurogenesis, cell proliferation and the subsequent differentiation and survival of newborn neurons, can undergo differential regulation by these factors (Table 8.1).

Factor	Proliferation	Glia	Neurons	References
FGF	→	→	→	44,46
EGF	→	↑	↓	44
IGF	↑	→	↑	47
Estrogen	↑	→	→	67
Serotonin	↑	n.d.	n.d.	73,74,75
Glutamate	↓	n.d.	n.d.	65,70
MK801	↑	n.d.	↑	
Enriched environment	→	→	↑	17,18,80
Wheel running	↑	→	↑	55,56
Learning	→	n.d.	↑	83
	→	→	→	56
Stress	↓	n.d.	n.d.	60,61
Glucocorticoids	↓	n.d.	n.d.	57,58,59,64
Adrenalectomy	↑	n.d.	↑	
Stroke	↑	n.d.	↑	90,91
Epilepsy/seizures	↑	n.d.	↑	85,86,87,88,89,100
Vitamin E deficiency	n.d.	n.d.	↑	99
Aging	↓	n.d.	↓	63,64,81

TABLE 8.1. REGULATION OF CELL PROLIFERATION AND NEUROGENESIS IN THE DENTATE GYRUS *IN VIVO*

Genetics

In 1997 Kempermann and colleagues found that strains of mice differ with respect to rate of cell division and amount of cell survival and neurogenesis. Comparisons were made among C57BL/6, BalB/c, CD1, and 129/SVJ strains. Proliferation was found to be highest in C57BL/6 mice; however, net neurogenesis was highest in the CD1 strain. 129/SVJ produced relatively more astrocytes and fewer neurons than other strains (16). The degree to which environmental, behavioral, and biochemical factors can affect cell proliferation and neurogenesis may also differ depending on the species or strain of animal involved. Indeed, exposure to an enriched environment (18) had different effects on two of these strains of mice, C57BL/6 and 129/SVJ, respectively. In C57BL/6 mice enrichment promoted the survival of progenitor cells but did not affect proliferation, whereas the net increase in neurogenesis in 129/SVJ mice was accompanied by a twofold increase in proliferation (18). Thus, strain differences not only influence the baseline rate of adult hippocampal neurogenesis, but also influence how adult hippocampal neurogenesis is regulated in response to environmental stimulation. Indeed, proliferation, survival and differentiation of progenitor cells and their progeny are each separately influenced by inheritable traits and are not uniformly

up-regulated in response to environmental stimulation.

Growth Factors

During development, growth factors provide important extracellular signals for regulating the proliferation and fate determination of stem and progenitor cells in the CNS (43). Several studies have been carried out to investigate progenitors in the adult brain respond to such growth factors. Intracerebroventricular infusion of EGF and FGF-2 in adult rats increased proliferation in the subventricular zone (44). Neither EGF nor FGF enhanced proliferation in the subgranular zone of the dentate gyrus. With regard to differentiation, EGF promoted glial differentiation, whereas FGF-2 did not influence phenotype distribution (44). In another series of experiments, FGF was administered systemically during the first postnatal weeks and in the adult rat. Cell proliferation was increased the dentate gyrus of infant rats but not in the adult hippocampus (45 ,46). Recent research has shown that intracerebral infusion of IGF increases both cell proliferation and neurogenesis in hypophysectomized rats (47). In songbirds, seasonal regulation of adult neurogenesis depends on testosterone levels that mediate their effect through BDNF (48). In addition, IGF-1, FGF mRNA, and BDNF mRNA are elevated in rodents by exercise (49 ,50 ,51 and 52). Expression of BDNF (53) and GDNF is increased by exposure to an enriched environment (54). Both running and enrichment increase net neurogenesis (17 ,55 ,56). The effects of intracerebral administration of trophins such as BDNF, NT-3, and GDNF remain to be determined; however, it appears that growth factors do play a role in *in vivo* regulation of proliferation and neurogenesis in the adult hippocampus. Better understanding of their mechanisms of action may lead to therapeutic application of these factors after brain injury or disease.

Neuroendocrine Factors and Stress

McEwen and Gould at Rockefeller University first investigated the effects of glucocorticoids or stress on adult hippocampal neurogenesis. The initial study reported that adrenalectomy, which leads to a reduction in serum glucocorticoid levels, elicits cell division in the dentate gyrus. This effect could be reversed by corticosterone replacement (57). Conversely, stress or increased levels of glucocorticoid hormones inhibit proliferative activity in the dentate gyrus (58 ,59). For example, administration of high levels of corticosterone diminishes cell division in the adult rat hippocampus (59). In addition, exposure of a rat to the odor of a natural predator (fox), causing stress and elevated corticosterone levels, transiently suppressed cell proliferation in the

adult rat dentate gyrus (60). Furthermore, exposure of marmoset monkeys to a resident intruder causes stress and results in a decrease in cell proliferation (61). In a recent study, rats that are highly reactive to novelty and exhibit a prolonged corticosterone secretion in response to novelty and stress were found have reduced dentate gyrus cell proliferation (62). Aging is accompanied by a reduction in neurogenesis (63), which may be caused in part by elevated glucocorticoid levels. Adrenalectomy in aged rodents has been shown to increase cell proliferation and neurogenesis (64). The effects of glucocorticoids on cell genesis appear to be mediated via a downstream effect on NMDA glutamate receptors (65). Thus, glucocorticoids and stress associated with increased corticosterone secretion inhibit cell genesis in the hippocampus. Enhanced stress or glucocorticoid levels therefore may impair hippocampal function, and lead to deficits in learning and memory. In contrast to the glucocorticoids, other steroid hormones, such as testosterone, enhance neurogenesis in birds (66), whereas estrogen results in a transient increase in proliferation in rats (67). Thyroid hormone can affect neuronal differentiation of hippocampal progenitor cells *in vitro* (27). *In vivo*, hypothyroidism interferes with cell migration (68), but does not affect postnatal cell proliferation (69).

Neurotransmitters

Neurotransmitters have also been suggested to play a role in adult dentate gyrus neurogenesis. Systemic injection of glutamate analogs inhibits birth of new cells, whereas an antagonist, such as MK801, enhances cell division (65 ,70). Recently, another class of neurotransmitters, the monoamines, has been suggested to be important as well. Prolonged administration of fluoxetine, as well as therapeutic agents acting on norepinephrine and dopamine receptors, and electroconvulsive shock enhance the number of BrdU-positive cells in rats (71 ,72 and 73). Acute administration of fluoxetine did not affect cell genesis (73). Grafting of fetal raphe neurons also stimulated granule proliferation in the hippocampus, whereas embryonic spinal tissue had no effect (74). Furthermore, depletion of serotonin reduces stem cell proliferation in the dentate gyrus (75). It is possible that these effects are mediated by the 1A receptor, because administration over 4 days of a specific 1A receptor antagonist (WAY) reduced basal rate of cell proliferation (Jacobs et al., unpublished observations). Taken together, these findings suggest that induction of cell proliferation is dependent on chronic administration of monoamines, consistent with the therapeutic time course for antidepressant treatments. Indeed, these studies have led to the hypothesis that therapeutic interventions that increase serotonergic transmission may act in part by augmenting dentate neurogenesis, promoting recovery from depression (76 ,77). It is of interest to note in this context that voluntary exercise increases cell proliferation (55), enhances monoamine levels and has an antidepressant effect (78). Thus, monoamines can affect cell genesis in the dentate gyrus. The receptors and mechanisms by which they exert their effects as well as possible interactions with other classes of neurotransmitters and/or growth factors remain to be determined.

Experience

As mentioned, stress (19) and depression may reduce the birth of new neurons. In addition, the aging process is accompanied by a decrease in neurogenesis (63); however, there are several environmental and behavioral interventions that can enhance neurogenesis (Fig. 8.2). In 1997 Kempermann and colleagues carried out the first of these studies comparing mice living under standard conditions with those housed in an enriched environment (17). Exposure to an enriched environment, consisting of larger housing; toys; and more opportunity for social stimulation, physical activity, and learning than standard laboratory conditions (79), resulted in a significant increase in neurogenesis, without affecting cell proliferation in mice and rats (17 ,80). Subsequent studies showed that the age-related decline in neurogenesis could be attenuated by enrichment (81). In addition, it was shown that enrichment inhibits cell death by apoptosis and prevents seizures (54). Moreover, it was determined that the most important components of enrichment are increased physical activity and possibly learning. Similar to enrichment, voluntary exercise in a running wheel increases net neurogenesis (55). In addition, running increases

cell proliferation in the dentate gyrus (55). It is interesting to note that enrichment and running had the same net effect on neurogenesis, but that running increased proliferation, whereas enrichment did not. Thus, not only the genetic factors mentioned, but also different environmental and behavioral factors can have differential effects on cell proliferation and neurogenesis. Others reported that hippocampus-dependent tasks, such as spatial learning in the Morris water maze (82), increases the number of surviving BrdU-positive cells (83 ,84); however, in our laboratory there was no effect of learning on proliferation or survival of newborn hippocampal cells (55).

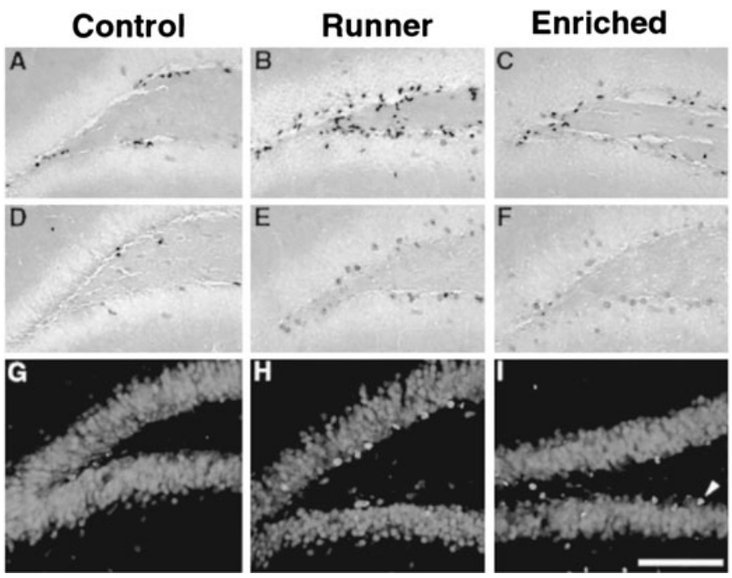


FIGURE 8.2. Proliferation and neurogenesis in the dentate gyrus. Photomicrographs of BrdU-positive cells 1 day *(a–c)* and 4 weeks *(d–f)* after the last injection in control *(a,d)*, running *(b,e)*, and enriched *(c,f)* mice. Confocal images of BrdU positive cells in control *(g)*, running *(h)*, and enriched *(i)* mice, 4 weeks after the last injection. Sections were immunofluorescent triple labeled for BrdU *(red)*, NeuN indicating neuronal phenotype *(green)*, and s100B selective for glial phenotype *(blue)*. Orange (*arrow*, newborn neuron) is red + green. Scale bar is 100 μm. See color version of figure.

Apart from these rather innocuous manipulations, there are several pathologic events that can affect granule cell number. Damage to the hippocampus by kindling (85 ,86), seizures (87 ,88 and 89), ischemia (90 ,91), or mechanical lesions (92) enhances proliferation. Thus, both normal and pathologic circumstances can affect cell genesis. Whether increased proliferation is beneficial for function or may represent compensation for lost cells and/or function remains to be determined.

FUNCTIONAL SIGNIFICANCE OF NEUROGENESIS

Part of "8 - Neurogenesis in Adult Brain "

Adult neurogenesis has been reported to exist over more than three decades, and to occur in a variety of species, including humans; however, the functional role of these new cells has yet to be determined. Given that the hippocampus is important for some forms of learning and memory and related mechanisms of neural plasticity such as long-term potentiation (LTP), much of the research has focused on finding a relationship between neurogenesis and memory function.

Behavior

It was found in several studies that animals living in an enriched environment not only had more new neurons, but also performed better on a spatial learning task (17 ,18 ,80). In addition, mice that were housed with a running wheel had more new neurons than their sedentary counterparts (55) and performed better on the water maze task (56). As mentioned, neurogenesis declines with aging (63); however, exposure to an enriched environment can restore some of the neurogenesis and improve performance on the water maze task in aging mice (81). Whether voluntary exercise in a running wheel would yield similar results in aged animals remains to be determined.

Electrophysiology

Although we can identify new granule cells using histologic and morphologic techniques, the question remains whether these cells are physiologic functional and if so, how similar or different are they from existing granule cells (Fig. 8.3). Moreover, what could be the functional significance of their (possible) neuronal activity? It has been shown that new hippocampal granule cells send axons along the mossy fiber tract to CA3 as do all other granule cells (93 ,94 and 95) and that they receive synaptic contacts (93). A recent study compared LTP, a physiologic model of certain forms of learning and memory (96) in the dentate gyrus, and CA1 in hippocampal slices from running and control mice. LTP amplitude was selectively enhanced in the dentate gyrus of running mice (56). It is possible that the newborn granule cell neurons play a role in increased dentate gyrus LTP because running increases learning and neurogenesis. Although the new cells are a small percentage of the granule cell layer, the possibility exists that they have greater plasticity than do mature cells. Indeed, dentate gyrus LTP last longer in immature rats than adults (97); however, in order to test this hypothesis it is necessary to record activity from individual new granule cells. In a recent study granule cells from the inner and outer layer of the dentate gyrus were compared. The inner layer cells were considered to be “young” cells and the outer layer “old” cells. The researchers found that the putative

young cells had a lower threshold for LTP and were unaffected by GABA-A inhibition, suggesting enhanced plasticity in the “young” cells (98); however, the problem with this study is that the definition of “young” cells is ambiguous. In our studies we have found newly generated cells throughout the granule cell layer. Moreover, the only way to be certain that a cell is a newborn neuron is by labeling the cell when it divides with a mitogenic marker. Ideally, recordings would be made from such a labeled cell and compared with preexisting neurons.

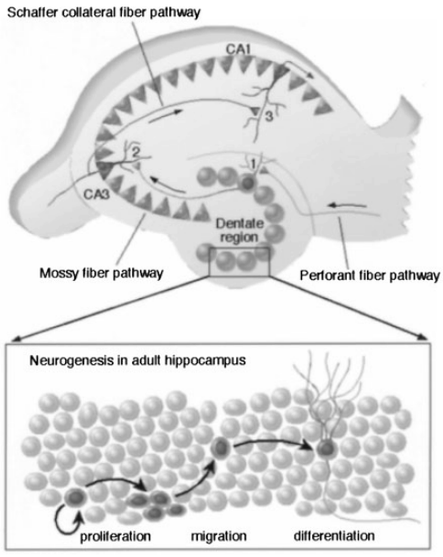


FIGURE 8.3. The three major areas of the hippocampus are the dentate gyrus, CA3, and CA1. The perforant pathway (1) from the subiculum forms excitatory connections with granule cells of the dentate gyrus. Both newborn and existing granule cells give rise to axons that form the mossy fiber pathway (2). This pathway projects to area CA3 pyramidal cells. The CA3 cells project to CA1 pyramidal cells by Schaffer collaterals. Within the hippocampal system, only the dentate gyrus gives rise to new neurons in the adult brain. These cells proliferate, migrate, and differentiate into mature neurons in the dentate gyrus (see box). See color version of figure.

In summary, adult hippocampal neurogenesis appears to be associated with memory function. Indeed, performance on a learning task is better in animals in which neurogenesis is stimulated, such as by running or enrichment, than in controls. This link may between neurogenesis and behavior may be causal rather than correlational, given that electrophysiologically measurable changes occur in the brain region where adult neurogenesis occurs.

POTENTIAL THERAPEUTIC IMPLICATIONS

Part of "8 - Neurogenesis in Adult Brain "

Several areas of interest emerge when considering the therapeutic potential for neurogenesis. First, a strategy that takes advantage of the ability of stem cells from the adult brain to be isolated and induced to divide in culture opens the opportunity for cellular transplantation to replace cells that have died because of injury or disease. Although evidence supports that ability of adult stem cells to survive grafting to the adult brain, the fate of the grafted cells appears to be dictated by the local environment. Thus, in order to accurately replace cells in damaged areas of the brain significant new information about the cellular and molecular mechanisms that control fate decisions is required in order to “train” the immature cells in culture to respond to the unique features of each of the environment to which they are grafted.

A second strategy emerges as a direct result of the fact that the adult brain retains stem cells *in situ* throughout life. By understanding the internal factors, molecules and mechanisms as well as the external stimuli and influences that control and regulate each of the steps in neurogenesis *in vivo*, eventually the potential of the endogenous cells could be harnessed. To this end, cell number could be amplified, direction of migration could be targeted, and finally terminal fate could be specified to the extent that a form of “self-repair” could be induced or orchestrated in the adult damaged brain. In all cases a more complete understanding the cellular and molecular events directing neurogenesis is a prerequisite for the use of this process in rational strategies for therapy.

ACKNOWLEDGMENTS

Part of "8 - Neurogenesis in Adult Brain "

This work was supported by NIA, NINDS, the Lookout Fund, Pasarow Foundation, Holfelder Foundation, and APA. We thank M.L. Gage for comments on the manuscript.

REFERENCES

1. Altman J, Das GD. Autoradiographic and histological evidence of postnatal neurogenesis in rats. *J Comp Neurol* 1965;124:319–335.

2. Altman J, Das GD. Autoradiographic and histological studies of postnatal neurogenesis. I. A longitudinal investigation of the kinetics, migration and transformation of cells incorporating tritiated thymidine in neonate rats, with special reference to postnatal neurogenesis in some brain regions. *J Comp Neurol* 1966;126:337–389.

3. Kaplan MS, Hinds JW. Neurogenesis in the adult rat: electron microscopic analysis of light radioautographs. *Science* 1997;197:1092–1094.

4. Goldman SA, Nottebohm F. Neuronal production, migration, and differentiation in a vocal control nucleus of the adult female canary brain. *Proc Natl Acad Sci USA* 1983;80:2390–2394.

5. Goldman SA. Adult neurogenesis: from canaries to the clinic. *J Neurobiol* 1998;36:267–286.

6. Richards LJ, Kilpatrick TJ, Bartlett PF. De novo generation of neuronal cells from the adult mouse brain. *Proc Natl Acad Sci USA* 1992;89:8591–8595.

7. Reynolds BA, Weiss S. Generation of neurons and astrocytes from isolated cells of the mammalian central nervous system. *Science* 1992;255:1707–1710.

8. Lois C, Alvarez Buylla A. Proliferating subventricular zone cells in the adult mammalian forebrain can differentiate into neurons and glia. *Proc Natl Acad Sci USA* 1993;90:2074–2077.

9. DeHamer MK, Guevara JL, Hannon K, et al. Genesis of olfactory receptor neurons *in vitro*: regulation of progenitor cell divisions by fibroblast growth factors. *Neuron* 1994;13:1083–1097.

10. Johansson CB, Momma S, Clarke DL, et al. Identification of a neural stem cell in the adult mammalian central nervous system. *Cell* 1999;96:25–34.

11. Doetsch F, Caille L, Lim DA, et al. Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell* 1999;97:703–716.

12. Lois C, Alvarez-Buylla A. Long-distance neuronal migration in the adult mammalian brain. *Science* 1994;264:1145–1148.

13. Horner PJ, Power AE, Kempermann G, et al. Proliferation and differentiation of progenitor cells throughout the intact adult rat spinal cord. *J Neurosci* 2000;20:2218–2228.

14. Wyss JM, Sripanidkulchai B. The development of Ammon’s horn and the fascia dentata in the cat: a [^{3}H]thymidine analysis. *Brain Res* 1985;1–2:185–198.

15. Gueneau G, Privat A, Drouet J, et al. Subgranular zone of the dentate gyrus of young rabbits as a secondary matrix. A high-resolution autoradiographic study. *Dev Neurosci* 1982;5:345–358.

16. Kempermann G, Kuhn HG, Gage FH. Genetic influence in the dentate gyrus of mice. *Proc Natl Acad Sci USA* 1997a;94:10409–10414.

17. Kempermann G, Kuhn HG, Gage FH. More hippocampal neurons in adult mice living in an enriched environment. *Nature* 1997b;386:493–495.

18. Kempermann G, Brandon EP, Gage FH. Environmental stimulation of 129/SvJ mice causes increased cell proliferation and neurogenesis in the adult dentate gyrus. *Curr Biol* 1998;8:939–942.

19. Gould E, Reeves AJ, Fallah M, et al. Hippocampal neurogenesis in adult Old World primates. *Proc Natl Acad Sci USA* 1999;96:5263–5267.

20. Kornack DR, Rakic P. Continuation of neurogenesis in the hippocampus of the adult macaque monkey. *Proc Natl Acad Sci USA* 1999;96:5768–5773.

21. Gould E, Reeves AJ, Graziano MS, et al. Neurogenesis in the neocortex of adult primates. *Science* 1999;286:548–552.

22. Eriksson PS, Perfilieva E, Bjork-Eriksson, et al. Neurogenesis in the adult human hippocampus. *Nat Med* 1998;11:1313–1317.

23. Vescovi AL, Reynolds BA, Fraser DD, et al. bFGF regulates the proliferative fate of unipotent (neuronal) and bipotent (neuronal/astroglial) EGF-generated CNS progenitor cells. *Neuron* 1993;11:951–966.

24. Kilpatrick TJ, Bartlett PF. Cloned multipotential precursors from the mouse cerebrum require FGF-2, whereas glial restricted precursors are stimulated with either FGF-2 or EGF. *J Neurosci* 1995;15:3563–3661.

25. Gage FH, Coates PW, Palmer TD, et al. Survival and differentiation of adult neuronal progenitor cells transplanted to the adult brain. *Proc Natl Acad Sci USA* 1995;92:11879–11883.

26. Palmer T, Takahashi J, Gage FH. The adult rat hippocampus contains primordial neural stem cells. *Mol Cell Neurosci* 1997;8:389–404.

27. Palmer TD, Ray J, Gage FH. FGF-2 responsive neuronal progenitors reside in proliferative and quiescent regions of the adult rodent brain. *Mol Cell Neurosci* 1995;6:474–486.

28. Shihabuddin LS, Ray J, Gage FH. FGF-2 is sufficient to isolate progenitors found in the adult spinal cord. *Exp Neurol* 1997;148:577–586.

29. Palmer TD, Markakis EA, Willhoite AR, et al. Fibroblast growth factor-2 activates a latent neurogenic program in neural stem cells from diverse regions of the adult CNS. *J Neurosci* 1999;19:8487–8497.

30. Tropepe V, Coles BLK, Chiasson BJ, et al. Retinal stem cells in the adult mammalian eye. *Science* 2000;287:2032–2036.

31. Morshead CM, Reynolds BA, Craig CG, et al. Neural stem cells in the adult mammalian forebrain: a relatively quiescent subpopulation of subependymal cells. *Neuron* 1994;13:1071–1082.

32. Gensburger C, Labourdette G, Sensenbrenner M. Brain basic fibroblast growth factor stimulates the proliferation of rat neuronal precursor cells *in vitro*. *FEBS Letts* 1987;217:1–5.

33. Arsenijvic Y, Weiss SJ. Insulin-like growth factor 1 is a differentiation factor for postmitotic CNS stem cell-derived neuronal precursors: distinct actions from those of brain-derived neurotrophic factor. *Neuroscience* 1998;18:2118–2128.

34. Ahmed S, Reynolds BA, Weiss S. BDNF enhances the differentiation but not the survival of CNS stem cell-derived neuronal precursors. *J Neurosci* 1995;15:5765–5778.

35. Ray J, Baird A, Gage FH. A 10-amino acid sequence of fibroblast growth factor 2 is sufficient for its mitogenic effect on progenitor cells. *Proc Natl Acad Sci USA* 1997;94:7047–7052.

36. Sakurada K, Ohshima-Sakurada M, Palmer TD, et al. Nurr1, an orphan nuclear receptor, is a transcriptional activator of endogenous tyrosine hydroxylase in neural progenitor cells derived from the adult brain. *Development* 1999;126:4017–4026.

37. Takahashi M, Palmer TD, Takahashi J, et al. Widespread integration and survival of adult-derived neural progenitor cells in the developing retina. *Mol Cell Neurosci* 1998;12:340–348.

38. Cameron HA, Hazel TG, McKay RD. Regulation of neurogenesis by growth factors and neurotransmitters. *J Neurobiol* 1998;2:287–306.

39. Bonni A, Sun Y, Nadal-Vicens M, et al. Regulation of gliogenesis in the central nervous system by the JAK-STAT signaling pathway. *Science* 1997;278:477–483.

40. Suhonen JO, Peterson DA, Ray J, et al. Differentiation of adult-derived hippocampal progenitor cells into olfactory bulb neurons. *Nature* 1996;382:624–627.

41. Shihabuddin LS, Horner PJ, Ray J, et al. Adult spinal cord stem cells generate neurons after transplantation in the adult dentate gyrus. *J Neurosci* 2000;20:8727–8735.

42. Bjornson CR, Rietze RL, Reynolds BA, et al. Turning brain into blood: a hematopoietic fate adopted by adult neural stem cells. *Science* 1999;283:3287–3297.

43. Calof AL. Intrinsic and extrinsic factors regulating vertebrate neurogenesis. *Curr Opin Neurobiol* 1995;5:19–27.

44. Kuhn HG, Winkler J, Kempermann G, et al. Epidermal growth factor and fibroblast growth factor-2 have different effects on neural progenitors in the adult rat brain. *J Neurosci* 1997;17:5820–5829.

45. Tao Y, Black IB, DiCicco-Bloom E. Neurogenesis in neonatal rat brain is regulated by peripheral injection of basic fibroblast growth factor (bFGF). *J Comp Neurol* 1996;376:653–663.

46. Wagner JP, Black IB, DiCicco-Bloom E. Stimulation of neonatal and adult brain neurogenesis by subcutaneous injection of basic fibroblast growth factor. *J Neurosci* 1999;19:6006–6016.

47. Aberg MAI, Aberg ND, Hedbacker H, et al. Peripheral infusion of IGF-1 selectively induces neurogenesis in the adult rat hippocampus. *J Neurosci* 2000;20:2896–2903.

48. Rasika S, Alvarez-Buylla, Nottebohm F. BDNF mediates the effects of testosterone on the survival of new neurons in an adult brain. *Neuron* 1999;22:53–62.

49. Neeper SA, Gomez-Pinilla F, Choi J, et al. Exercise and brain neurotrophins. *Nature* 1995;373:109.

50. Gomez-Pinilla F, Dao L, Vannarith S. Physical exercise induces FGF-2 and its mRNA in the hippocampus. *Brain Res* 1997;764:1–8.

51. Gomez-Pinilla F, So V, Kesslak JP. Spatial learning and physical activity contribute to the induction of fibroblast growth factor: neural substrates for increased cognition associated with exercise. *Neuroscience* 1998;85:53–61.

52. Carro E, Nunez A, Busiguina S, et al. Circulating insulin-like growth factor I mediates effects of exercise on the brain. *J Neurosci* 2000;20:2926–2933.

53. Falkenberg T, Mohammed AK, Henriksson B, et al. Increased expression of brain-derived neurotrophic factor mRNA in rat is associated with improved spatial memory and enriched environment. *Neuroscience Lett* 1992;138:153–156.

54. Young D, Lawlor PA, Leone P, et al. Environmental enrichment inhibits spontaneous apoptosis, prevents seizures and is neuroprotective. *Nat Med* 1999;5:448–453.

55. van Praag H, Kempermann G, Gage FH. Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus. *Nat Neurosci* 1999;2:266–270.

56. van Praag H, Christie BR, Sejnowski TJ, et al. Running enhances neurogenesis, learning and long-term potentiation in mice. *Proc Natl Acad Sci USA* 1999;96:13427–13431.

57. Gould E, Cameron HA, Daniels DC, et al. Adrenal hormones suppress cell division in the adult rat dentate gyrus. *J Neurosci* 1992;12:3642–3650.

58. McEwen BS. Gonadal and adrenal steroids regulate neurochemical and structural plasticity of the hippocampus via cellular mechanisms involving NMDA receptors. *Cell Mol Neurobiol* 1996;2:103–116.

59. Cameron HA, Gould E. Adult neurogenesis is regulated by adrenal steroids in the dentate gyrus. *Neuroscience* 1994;61:203–209.

60. Galea LAM, Tanapat P, Gould E. Exposure to predator odor suppresses cell proliferation in the dentate gyrus of adult rats via a cholinergic mechanism. *Soc Neurosci Abs* 1996;22:1196.

61. Gould E, Tanapat P, McEwen BS, et al. Proliferation of granule cell precursors in the dentate gyrus of adult monkeys is diminished by stress. *Proc Natl Acad Sci USA* 1998;95:3168–3171.

62. Lemaire V, Aurousseau C, Le Moal M, et al. Behavioural trait of reactivity to novelty is related to hippocampal neurogenesis. *Eur J Neurosci* 1999;11:4006–4014.

63. Kuhn HG, Dickinson-Anson H, Gage FH, Neurogenesis in the dentate gyrus of the adult rat: age-related decrease of neuronal progenitor proliferation. *J Neurosci* 1996;16:2027–2033.

64. Cameron HA, McKay RD. Restoring production of hippocampal neurons in old age. *Nat Neurosci* 1999;2:894–897.

65. Cameron HA, McEwen BS, Gould E. Regulation of adult neurogenesis by excitatory input and NMDA receptor activation in the dentate gyrus. *J Neurosci* 1995;15:4687–4692.

66. Rasika S, Nottebohm F, Alvarez-Buylla A. Testosterone increases the recruitment and/or survival of new high vocal center neurons in adult female canaries. *PNAS* 1994;91:7854–7858.

67. Tanapat P, Hastings NB, Reeves AJ, et al. Estrogen stimulates a transient increase in the number of new neurons in the dentate gyrus of the adult female rat. *J Neurosci* 1999;19:5792–5801.

68. Rami A, Rabie A, Patel AJ. Thyroid hormone and development of the rat hippocampus: morphological alterations in granule and pyramidal cells. *Neuroscience* 1986;19:1207–1216.

69. Seress L. The postnatal development of rat dentate gyrus and the effect of early thyroid hormone treatment. *Anat Embryol* 1977;151:335–339.

70. Gould E, Cameron HA, McEwen BS. Blockade of NMDA receptors increases cell death and birth in the developing dentate gyrus. *J Comp Neurol* 1994;340:551–565.

71. Dawirs RR, Hildebrandt K, Teuchert-Noodt G. Adult treatment with haloperidol increases dentate granule cell proliferation in the gerbil hippocampus. *J Neural Transm* 1998;105:317–327.

72. Fornal CA, Jacobs BL. Chronic fluoxetine treatment increases hippocampal neurogenesis in rats: a novel theory of depression. *Soc Neurosci Abs* 1999;25:714.

73. Malberg JE, Eisch AJ, Nestler EJ, et al. Chronic antidepressant administration increases granule cell genesis in the hippocampus of the adult male rat. *Soc Neurosci Abs* 1999;25:1029.

74. Brezun JM, Daszuta A. Serotonergic reinnervation reverses lesion-induced decreases in PSA-NCAM labeling and proliferation of hippocampal cells in adult rats. *Hippocampus* 2000;10:37–46.

75. Brezun JM, Daszuta A. Depletion in serotonin decreases neurogenesis in the dentate gyrus and the subventricular zone of adult rats. *Neuroscience* 1999;89:999–1002.

76. Duman RS, Malberg J, Thome J. Neural plasticity to stress and antidepressant treatment. *Biol Psychiatry* 1999;46:1181–1191.

77. Jacobs BL, van Praag H, Gage FH. Adult brain neurogenesis and psychiatry: a novel theory of depression. *Mol Psychiatry* 2000;5:262–269.

78. Chaouloff F. Physical exercise and brain monoamines: a review. *Acta Physiol Scand* 1989;137:1–13.

79. Rosenzweig MR, Krech D, Bennett EL, et al. Effects of environmental complexity and training on brain chemistry and anatomy. *J Comp Physiol Psychol* 1962;55:429–437.

80. Nilsson M, Perfilieva E, Johansson U, et al. Enriched environment increases neurogenesis in the adult rat dentate gyrus and improves spatial memory. *J Neurobiol* 1999;39:569–578.

81. Kempermann G, Kuhn HG, Gage FH. Experience-induced neurogenesis in the senescent dentate gyrus. *J Neurosci* 1998;18:3206–3212.

82. Morris RGM. Development of a water maze procedure for studying spatial learning in the rat. *J Neurosci Meth* 1984;11:47–60.

83. Gould E, Beylin A, Tanapat P, et al. Hippocampal-dependent learning enhances the survival of granule neurons generated in the dentate gyrus of adult rats. *Nat Neurosci* 1999;2:260–265.

84. Greenough WT, Cohen NJ, Juraska JM. New neurons in old brains: learning to survive? *Nat Neurosci* 1999;2:203–205.

85. Scott BW, Wang S, Burnham WM, et al. Kindling induced neurogenesis in the dentate gyrus of the rat. *Neuroscience Lett* 1998;248:73–76.

86. Parent JM, Janumpalli S, McNamara JO, et al. Increased dentate granule cell neurogenesis following amygdala kindling in the adult rat. *Neuroscience Lett* 1998;247:9–12.

87. Parent JM, Yu TW, Leibowitz RT, et al. Dentate granule neurogenesis is increased by seizures and contributes to aberrant network reorganization in the adult hippocampus. *J Neurosci* 1997;17:3727–3738.

88. Bengzon J, Kokaia Z, Elmer E, et al. Apoptosis and proliferation of dentate gyrus neurons after single and intermittent limbic seizures. *Proc Natl Acad Sci USA* 1997;94:10432–10437.

89. Nakagawa E, Aimi Y, Yasuhara O, et al. Enhancement of progenitor cell division in the dentate gyrus triggered by initial limbic seizures in rat models of epilepsy. *Epilepsia* 2000;41:10–18.

90. Liu J, Solway K, Messing RO, et al. Increased neurogenesis in the dentate gyrus after transient global ischemia in gerbils. *J Neurosci* 1998;18:7768–7778.

91. Takagi Y, Nozaki K, Takahashi J, et al. Proliferation of neuronal precursor cells in the dentate gyrus is accelerated after transient forebrain ischemia in mice. *Brain Res* 1999;831:283–287.

92. Gould E, Tanapat P. Lesion-induced proliferation of neuronal progenitors in the dentate gyrus of the adult rat. *Neuroscience* 1997;80:427–436.

93. Markakis EA, Gage FH. Adult-generated neurons in the dentate gyrus send axonal projections to field CA3 and are surrounded by synaptic vesicles. *J Comp Neurol* 1999;406:449–460.

94. Stanfield BB, Trice JE. Evidence that granule cells generated in the dentate gyrus of adult rats extend axonal projections. *Exp Brain Res* 1988;72:399–406.

95. Hastings NB, Gould E. Rapid extension of axons into the CA3 region by adult-generated granule cells. *J Comp Neurol* 1999;413:146–154.

96. Bliss TV, Collingridge GL. A synaptic model of memory: long-term potentiation ion the hippocampus. *Nature* 1993;361:31–39.

97. Bronzino JD, Abu-Hasaballah K, Austin-LaFrance RJ, et al. Maturation of long-term potentiation in the hippocampal dentate gyrus of the freely moving rat. *Hippocampus* 1994;4:439–446.

98. Wang S, Scott BW, Wojtowicz JM. Heterogenous properties of dentate granule neurons in adult rat. *J Neurobiol* 2000;42:248–257.

99. Ciaroni S, Cuppini R, Cecchini T. Neurogenesis in the adult rat dentate gyrus is enhanced by vitamin E deficiency. *J Comp Neurol* 1999;411:495–502.

100. Madsen TM, Treschow A, Bengzon J, et al. Increased neurogenesis in a model of electroconvulsive therapy. *Biol Psychiatry* 2000;47:1043–1049.

9

Dopamine

Anthony A. Grace

Anthony A. Grace: Departments of Neuroscience and Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania.

Studies into the regulation of the dopamine (DA) system and its postsynaptic actions are often stymied by the myriad of actions that this neurotransmitter can produce. Thus, DA has been found to exert actions on the neurons it innervates both directly and via G-protein–coupled receptors. Moreover, this transmitter can modulate afferent input within these target regions, as well as alter intercellular communication via its actions on gap junctions. Finally, DA can potently modulate its own dynamics, acting via autoreceptors on DA nerve terminals and on DA neuron somata. In fact, the DA system is under potent dynamic regulation in the short term by a multitude of feedback systems, and in response to prolonged alterations is subject to powerful homeostatic mechanisms that can compensate for dramatic changes in DA system function. Such homeostatic alterations can be compensatory in nature, such as those that occur in response to a partial DA system lesion, or pathologic, such as the sensitization that can occur with repeated psychostimulant administration. Nonetheless, the importance of this neurotransmitter system in a broad array of human disorders ranging from Parkinson's disease to schizophrenia has driven an intensive array of investigations oriented toward increasing our understanding of this complex system in normal conditions as well as disease states. This chapter attempts to summarize some of the major research findings that have occurred within the last 5 years, and place them into a functional framework. This is not meant to be inclusive: A search of Medline indicated that there were over 16,000 papers published on DA during the past 5 years! Because of the exceedingly broad range that this topic encompasses, the focus is primarily on a subset of the numerous improvements that are most related to advancing our understanding of psychiatric disorders in particular. Topics related to specific disorders, such as drug abuse, schizophrenia, and so on, are deferred to the appropriate chapters in this volume.

DA NEURON ANATOMY AND PHYSIOLOGY

Part of "9 - Dopamine "

Both *in vivo* and *in vitro* studies have demonstrated that DA-containing neurons in the midbrain exhibit spontaneous spike firing that is driven by an endogenous pacemaker conductance (1 ,2 and 3), with their activity modulated by afferent inputs. One of the prominent regulators of DA neuron activity is the DA autoreceptor. It has been known for some time that DA neurons are very sensitive to DA agonists, which inhibit spike firing as well as cause a presynaptic inhibition of DA synthesis and release. Studies indicate that DA neuron somatodendritic autoreceptors are stimulated by an extracellular pool of DA released from the dendrites of neighboring DA neurons rather than exclusively by autoinhibition back onto the releasing neuron. This is supported by data showing that partial lesions of the DA system result in DA autoreceptor supersensitivity in the remaining neurons, which would only occur if the remaining neurons were responding to the decrease in DA caused by the loss of neighboring neurons (4). The autoreceptors are believed to exert a tonic down-regulation of DA neuron activity, maintaining their firing within a stable range of activity (4 ,5). These autoreceptors appear to be primarily of the D2 type, because D2-deficient mice do not show autoreceptor-mediated inhibition of firing (6). Moreover, inhibition of monoamine oxidase potentiates this inhibition (7 ,8), whereas inhibition of catechol-o-methyl transferase does not alter this response (9).

Exogenous transmitters also potently regulate dopamine neurons. Thus, GABA afferents both from striatonigral neurons as well as from local circuit neurons in the midbrain cause inhibition of DA neuron activity (10) by both a GABA-A– and GABA-B–mediated action (11 ,12 and 13). Glutamate has also been shown to exert multiple actions on DA neuron activity. Glutamate applied *in vivo* increases burst firing (14). N-methyl-D-aspartate (NMDA) receptor activation mediates a slow excitatory postsynaptic potential (EPSP) in these neurons (8), whereas metabotropic glutamate agonists are reported to depress both excitatory and inhibitory afferent input to these neurons (15). This latter effect is apparently shared by muscarinic receptors, which also depress both excitatory and inhibitory afferents, presumably via a presynaptic action (16 ,17).

Burst Firing

Studies have shown that DA neuron discharge is an essential component of the DA release process (18). The firing pattern of DA neurons also is effective in modulating release, with burst firing in particular being an important regulator of DA transmission. Thus, studies have shown that burst firing in DA neurons is associated with induction of c-fos and NG1-A in postsynaptic sites (19 ,20), and this response demonstrates a spatial and temporal specificity with respect to brain region, genes activated, and cell phenotype. One factor that is thought to regulate burst firing is the glutamatergic system. Several studies have shown that iontophoresis of glutamate onto DA neurons *in vivo* lead to burst firing (14), as do stimulation of glutamatergic afferents to DA neurons (21 ,22); however, the evidence for glutamate acting alone to induce burst firing *in vitro* is equivocal (23). In contrast, evidence shows that burst firing can be induced *in vitro* by blockade of apamin-sensitive potassium channels that modulate a nifedipine-sensitive calcium conductance (24). One source of glutamatergic input to ventral tegmental area (VTA) DA neurons is proposed to arise from the prefrontal cortex (PFC) (25); however, recent studies (26,27) show that the PFC input to the VTA innervates only the small proportion of VTA DA neurons that project to the PFC, providing a direct feedback loop, whereas the VTA-accumbens neurons innervated from the PFC are exclusively GABAergic neurons; therefore, it is unlikely that activation of the PFC can induce increased DA levels in the accumbens by a direct projection to these neurons (28). On the other hand, there is evidence that activation of the subiculum by excitatory amino acids increases accumbens DA (29 ,30) via activation of DA neurons that involves a pathway through the nucleus accumbens (31).

Afferent Input

The feedback systems between DA neurons and their postsynaptic targets appear to be quite complex, particularly in the primate. By analyzing a large number of retrograde and anterograde tracings, Haber and associates (32) found that different striatal subdivisions are linked by overlapping feedback to DA neurons, in a manner that suggests an ascending spiral of regulation extending from the shell to the core to the central striatum and finally to the dorsolateral striatum (Fig. 9.1). As pointed out by Haber, such an anatomic arrangement could account for the parallel psychomotor, affective, and cognitive disturbances seen in a variety of psychiatric disorders.

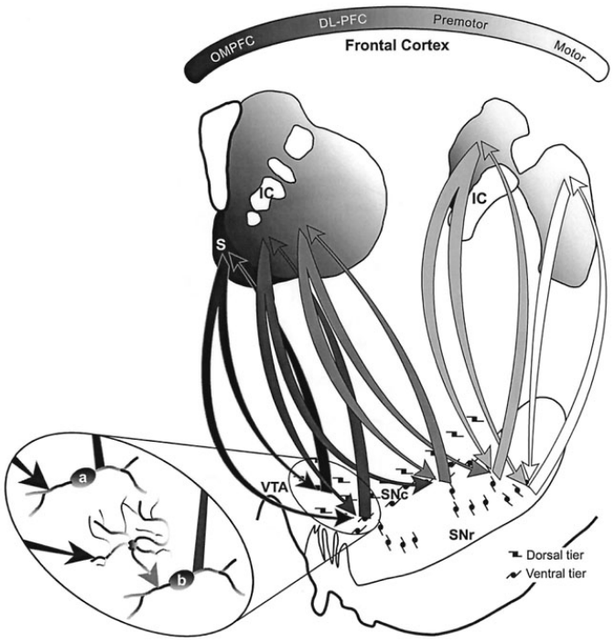


FIGURE 9.1. Studies using retrograde and anterograde tracers reveal that the feedback system between the midbrain DA neurons and their striatal targets contains both reciprocal and feed-forward components. The shell of the accumbens *(left)* receives inputs from hippocampus, amygdala, and limbic cortex, and projects to both the ventral tegmental area (VTA) and dorsomedial substantia nigra DA neurons. Projections from the VTA back to the shell form the "closed" portion of this loop. The core receives afferent input from the orbital and medial prefrontal cortex (OMPFC); the afferent projection to the core from the medial substantia nigra (SN) then forms the first part of the spiral. The core in turn projects to more ventrodorsal SN regions; therefore, ventral striatal regions can modulate the dopaminergic influence over more dorsal striatal regions via the spiraling midbrain-striatal-midbrain connections. The magnified insert shows a model of the reciprocal versus feed-forward loops. The reciprocal component is proposed to (a) directly inhibit the DA neuron, whereas the feed-forward, nonreciprocal component terminates on a GABAergic interneuron; or (b) indirectly excite the DA neuron by disinhibition. DL-PFC, dorsolateral prefrontal cortex; IC, internal capsule; S, shell; Snc, substantia nigra, zona compacta; Snr, substantia nigra, zona reticulata; VTA, ventral tegmental area. (From Haber et al., 2000; used with permission.)

The striatum provides a powerful feedback regulation of DA neuron firing. Thus, alterations in striatal activity potently affect DA cell activity states. Striatal neuron activation is known to cause an activation of DA neuron firing (10 ,33). Moreover, single-pulse stimulation of the striatum directly, or indirectly via activation of the PFC in rats, causes an inhibition/excitation response pattern (10 ,25). This relationship can be altered by manipulation of second messenger systems in the striatum. One system in particular that seems to affect striatal activation, leading to an alteration

in DA neuron activity, is the nitric oxide (NO) system. Increasing NO in the striatum by infusion of the substrate for the synthetic enzyme nitric oxide synthetase (NOS), coupled with striatal or cortical stimulation, was found to increase the firing rate of striatal neuron DA neurons. This effect was mimicked by infusion of the nitric oxide generator hydroxylamine. In contrast, NOS inhibitors failed to affect baseline DA cell firing but did increase their response to stimulation (34); therefore, NO signaling in the striatum facilitates DA neurotransmission by modulation of corticostriatal and striatonigral pathways. NO also appears to have a role in regulating terminal DA release (see the following).

Stress and DA Neuron Activity

Stress appears to affect many monoamine systems. Stress plays a role both in acute behavioral responses and adaptations to chronic stressful conditions. Although the noradrenergic system has played a major role in these processes, recent evidence supports a role for the DA system as well. Studies have shown that, on presentation of stress, there are differential increases in DA dynamics depending on the brain regions involved. Thus, stressful stimuli tend to cause the largest increase in DA levels in the PFC region, with markedly smaller changes in the limbic and dorsal striatal regions (35); however, this relationship is altered by lesions of different nuclei. Thus, stress causes release of DA in the amygdala (36), and lesions of the amygdala tend to block stress-induced increases in PFC DA levels (37). Lesions of the PFC also affect this response. Studies in which the PFC DA innervation is lesioned show that subsequent stressors cause a much larger increase in DA levels within the nucleus accumbens, particularly with respect to the duration of the response (38). This suggests that PFC DA released in response to stress actually blunts the responsiveness of the subcortical limbic DA system. In contrast, 6-OHDA lesions of PFC DA levels were found to decrease the basal electrophysiologic activity of VTA DA neurons (39). Given that basal DA levels in the accumbens are normal, one interpretation is that the DA release system has adapted to the diminished DA neuron drive, allowing normal levels of DA transmission to occur. However, if a stimulus then causes an increase in DA neuron firing, the compensated release mechanism would produce an augmented response. Thus, the magnitude of increase in action potential-dependent DA release into the accumbens that occurs in response to a challenge may be augmented when the PFC DA response is attenuated (39).

Repeated stress also has important clinical implications with regard to the DA system and exacerbation of schizophrenia. A recent study examined how chronic stress in the form of cold exposure affects the discharge of VTA DA neurons. Thus, after exposing rats to cold, there was a 64% decrease in the number of spontaneously active DA neurons, with no significant alteration in their average firing rate. Nonetheless, there was a subpopulation of neurons that exhibited excessive burst activity in the exposed rats (40). Therefore, unlike acute exposure to stressful or noxious stimuli, chronic stress actually attenuates DA neuron baseline activity. Such a decrease in baseline activity could enable the system to show a magnified response to activating stimuli, thereby producing a sensitized DA response.

REGULATION OF DA RELEASE

Part of "9 - Dopamine "

DA appears to be released by multiple factors within its postsynaptic target; moreover, once it is released, there are several mechanisms that can modulate its site of action. In general, the majority of evidence suggests that DA is released primarily in a spike-dependent manner, because inactivation of DA neuron firing virtually eliminates DA release within the striatum (18). Carbon fiber recordings, which allow rapid measurement of DA overflow, show that stimulation of DA axons causes rapid release of transmitter. Moreover, the release varies with tissue content, with PFC showing much lower levels of release compared to accumbens at a given stimulus frequency (41). DA released by impulse flow is then rapidly removed via the DA transporter, because mice with knockouts of this transporter exhibit 300 times longer clearance half-life compared to controls (42). The amount of DA released by impulses appears to depend on several factors. Previous volumes in the *Generations of Progress* series have detailed how DA release can be modulated by both synthesis- and release-modulating autoreceptors on DA terminals. It is becoming more evident that heteroceptors also play a significant role in modulating DA release (43).

DA release appears to occur via two functionally distinct components. One is the DA that is released in a high-amplitude, brief pulsatile manner by means of action potentials, and then is rapidly removed from the synaptic cleft via reuptake. This has been termed the phasic component of DA release (44), and is believed to underlie most of the behavioral indices of this transmitter. The other is the level of DA present in the extrasynaptic space. This tonic DA exists in very low concentrations; too low to stimulate intrasynaptic DA receptors, but of sufficient level to activate extrasynaptic receptors, including DA terminal autoreceptors (thereby causing feedback-inhibition of phasic DA release) and other extrasynaptic receptor sites. It is this tonic DA compartment that is sampled by slower measures of DA dynamics, such as microdialysis. Recently, evidence has been advanced to define what factors may contribute to the regulation of this tonic DA compartment.

Although studies suggest that neuronal impulse flow is necessary for DA overflow in the striatum, there is substantial evidence that the released DA can be controlled locally by a number of factors. For example, stimulation of cortical inputs increases DA release within the striatum, and evidence suggests that this can occur via afferents to DA cell bodies or presynaptically onto DA terminals, depending on the preparation and site of stimulation. Thus, infusion of excitatory amino acids into the hippocampus subiculum increases

DA neuronal activity (31) and DA levels in the striatum in a manner that is dependent on DA neuron impulse flow (29). It is proposed that this subicular-driven DA release may be involved in the modulation of investigatory response to novel and conditioned stimuli (45). Stimulation of the PFC also appears to result in impulse-dependent DA release in the striatum (28). On the other hand, there is evidence suggesting that DA can be released in a manner not dependent on DA neuron firing via stimulation of the hippocampal afferents (46), or amygdala afferents (47) to the accumbens, all of which use glutamate as a transmitter. This purported presynaptic action on DA terminals appears to occur via activation of either NMDA receptors on DA terminals (48) or by metabotropic glutamate receptors (49 ,50 and 51). There is also evidence that glutamate can release acetylcholine or serotonin in the striatum, which in turn can trigger DA release (43). Glutamate may also stimulate DA release via an action on other local systems, such as those producing NO. NO is known to be released from striatal interneurons containing the enzyme NOS, and exert actions on neuronal elements in the vicinity of the release site. Infusion of NOS substrates or NO generator compounds was found to facilitate the release of both glutamate and DA within the striatum in a calcium-dependent manner, and is dependent on vesicular stores (52 ,53). Moreover, the NO-induced efflux of striatal glutamate was found to indirectly enhance extracellular DA levels in the striatum in a manner dependent on NMDA and AMPA receptors (53 ,54). Therefore, it is likely that excitatory amino acids and NO interact with DA neuron firing to regulate DA release from presynaptic sites within the striatum.

The ability of cortical glutamate to release tonic DA in the striatum is supported by studies showing that lesions of the cortical input to the striatum cause a decrease in extracellular DA and glutamate within the striatum (55), which would thereby increase in the behavioral response to amphetamine (56). Thus, evidence indicates that alterations in tonic DA levels produced by cortical afferents can potently alter spike-dependent DA release, and thereby modulate DA-dependent behaviors (43 ,44 ,57). Such tonic down-modulation of spike-dependent DA release could play a particular role when the uptake system is inactivated by psychostimulants. Thus, although the DA transporter is normally highly effective at removing DA from the synaptic cleft before it can escape into the extracellular space, blockade of the DA transporter would allow substantially higher levels of DA to escape the cleft and contribute to the tonic extracellular DA pool (57). Such a condition is thought to underlie some of the therapeutic actions of psychostimulants in attention deficit/hyperactivity disorder (ADHD) (58).

One problem in attempting to examine the relationship between DA neuron firing rate and DA overflow is the potential disruption in the system caused by probe implantation. This was found to be a significant issue when testing the effects of chronic antipsychotic drug treatment-induced DA neuron depolarization block (59) on DA levels in the striatum. Thus, implantation of a microdialysis probe was found to disrupt DA neuron depolarization block when DA cell activity was assessed 24 hours following probe implantation. However, if the probe was inserted via a preimplanted guide cannula, depolarization block was maintained, and the DA levels were found to be approximately 50% less than in control conditions. Moreover, the relationship between DA neuron firing and release was altered. Thus, although there was no significant correlation between DA cell population burst firing and DA release in control rats, there was a significant correlation between burst firing in the remaining cells and DA levels following administration of chronic antipsychotic drug (60). Thus, correlations between cell firing patterns and DA levels postsynaptically appear to depend on the state of the system.

It is also possible that there may be local fluctuations in tonic DA stimulation that may be a consequence of increases in DA neuron firing. Indeed, studies using voltametric measures have shown that brief elevations in extracellular DA may occur as a consequence of rapid burst firing, overwhelming the DA uptake process (61). This relationship is particularly important during administrations of drugs that interfere with the uptake process, such as cocaine or amphetamine (57 ,58). Such drugs would cause phasic DA release to rapidly augment tonic DA levels, leading to high extracellular DA and abnormal levels of down-regulation of spike-dependent DA release. In a similar nature, in mice lacking the DA transporter, the extracellular DA is already elevated fivefold over control (62); therefore, there appears to be a tight dynamic interdependence on DA neuron activity levels and DA uptake that determines the contribution of phasic and tonic DA to activity within this system. This tonic/phasic balance has been proposed to underlie normal and dysfunctional DA regulation as it relates to the pathophysiology of schizophrenia, drug abuse, and the treatment of ADHD (44 ,57 ,58).

Given the importance of tonic DA system regulation, a literature has emerged regarding the functional relevance of extrasynaptic DA receptors. Indeed, studies have shown that in the PFC, the DA terminals located in the deep layers of cortex do not contain DA transporters (63). As a consequence, the DA released from these sites would be free to diffuse to a much greater extent than in areas such as the striatum and accumbens. This is further substantiated by evidence that a substantial portion of the DA that is released in the PFC is actually taken up and deaminated in norepinephrine (NE) terminals (64). This arrangement would have substantial functional implications. First, it would provide a mechanism for stimulation of the numerous extrasynaptically located D1 terminals on pyramidal neurons (65), which have been proposed to regulate information flow between compartments on pyramidal neurons (66 ,67 ,68 and 69). Moreover, such a condition could imply that NE uptake

blockers could serve to increase the functional actions of DA in the PFC by preventing its removal via NE terminals. This may also have implications regarding the clinical actions of NE-selective antidepressant drugs within this brain region.

POSTSYNAPTIC EFFECTS OF DA

Part of "9 - Dopamine "

DA exerts a myriad of actions on postsynaptic systems. These actions can occur at the level of individual cells in terms of direct postsynaptic actions, as well as altering cellular interactions (via presynaptic effects and network modulation). Moreover, the nature of these effects can vary depending on both the specific region examined and the time course of DA agonist administration.

Striatum

D1 stimulation decreases excitability of dorsal striatal and accumbens neurons (67 ,68 and 69), although others have reported excitation by this agonist (70). Within the dorsal striatum, D1 receptor stimulation decreases current-evoked action potential discharge in hyperpolarized neurons, although an enhancement in excitability can be obtained with longer duration or higher frequency current pulses (71). The decrease in spiking is believed to be owing to a reduction in the peak amplitude of the fast sodium conductance (72 ,73), and occurs through activation of protein kinase-A (PKA) (74). Studies show that the D1-mediated inhibition can act synergistically with D2 stimulation-induced inhibition when the agonists are applied simultaneously. However, the D1-mediated decrease in excitability can be reversed to facilitation if the D2 agonist is administered subsequently (75). This temporal dependence of D1 and D2 activation may have functional implications with regard to the tonic/phasic model of DA system regulation (44). For example, if the DA system exhibits sustained activation such as during a reward process, the large phasic DA release that results should stimulate both D1 and D2 receptors located within synapses. In addition, the large DA level released should be sufficient to escape the synaptic cleft, with the resultant elevated tonic DA levels stimulating the extrasynaptic D1 receptors (76 ,77). According to our data, this should produce synergistic inhibition. On the other hand, if the activity is maintained, there would be tonic stimulation of the extrasynaptic D1 receptors. Under this condition, subsequent stimulation of the D2 receptors preferentially located in the synaptic cleft (78) would be attenuated (75). Thus, the system appears to be oriented to provide a maximal initial response, whereas continuous activation would cause an attenuation of subsequent responses.

In addition to effects on sodium conductances, D1 stimulation also affects high voltage-activated calcium conductances. Thus, both D1 agonists and cAMP analogues reduce both N- and P-type calcium currents via a PKA-mediated process; however, these manipulations also enhance L-type calcium currents (79). In contrast, D2-receptor stimulation has been shown to modulate voltage-dependent potassium conductances in the striatum (80).

Evidence shows that a large part of the response to D1 stimulation requires the participation of a messenger cascade involving the phosphorylation of dopamine- and cAMP-regulated phosphoprotein (DARPP-32) (81). In particular, this phosphoprotein is a required component in the cascade mediating D1 function (Fig. 9.2). Moreover, mice with knockouts of DARPP-32 have been shown to lack D1 modulation of glutamate function, as well as other biochemical processes and behavioral responses known to involve D1 receptors (82). Recent studies have shown that DARPP-32 is also present in other, non–D1-containing neurons as well, including the enkephalin-containing striatal neurons (83). In this case, D2-receptor stimulation has been shown to cause a dephosphorylation of DARPP-32 via calcineurin activation by calcium influx. DARPP-32 is also present in striatal efferent projection areas, including the globus pallidus, entopeduncular nucleus, and substantia nigra (SN) (83). Thus, DARPP-32 is positioned to exert modulatory influences on DA function by affecting striatal outflow.

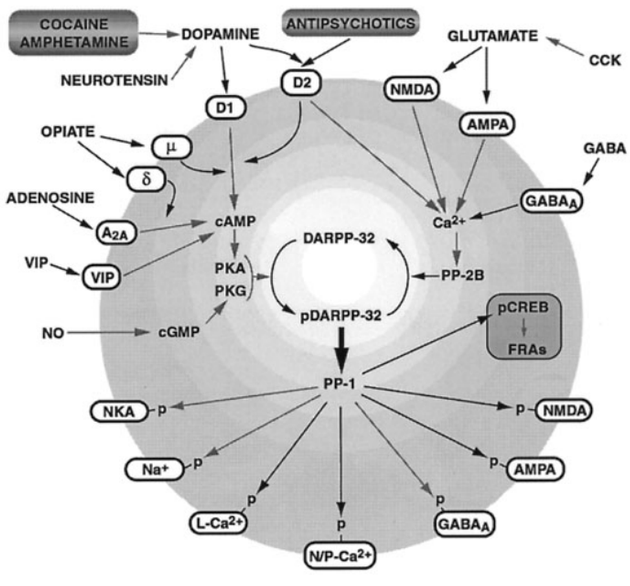


FIGURE 9.2. The DARPP-32 signaling pathway has a central role in mediating signal transduction within medium spiny neurons in the striatum. A variety of neurotransmitters act on systems regulating the phosphorylation of DARPP-32, which in turn modulates the activity of protein phosphatase-1 (PP-1). DA stimulation of D1 receptors acts via cAMP and PKA to phosphorylate DARPP-32, which in turn inhibits PP-1; this works in synergy with different protein kinases to increase the level of protein phosphorylation of their targets. In contrast, stimulation of D2 receptors attenuates D1 activation of adenylate cyclase as well as leading to calcium stimulation of protein phosphatase 2B; together, this decreases the phosphorylation state of DARPP-32. (From Greengard et al., 1999; used with permission.)

Modulation of Intercellular Coupling

In addition to its effects on single neurons, DA also is capable of affecting neuronal interactions on a network level. In

particular, substantial evidence has shown DA to have a potent effect over interactions among neighboring neurons in a region via its modulation of gap junction conductance. The DA system appears to regulate this coupling in two ways: (a) acutely, presumably by opening gap junctions that are already present between neurons in its target structures, and (b) as a compensatory change in response to a chronic compromise of the DA system.

Studies have shown that neurons within the dorsal and ventral striatum exhibit dye coupling, which is the morphologic correlate of gap junctions between neurons. In striatal slices recorded *in vitro*, application of the D2 agonist quinpirole causes a substantial increase in coupling, from nearly undetectable levels in the basal state to approximately 80% coupling after the agonist. D1 agonists, in contrast, do not affect coupling in a measurable way; however, in brain slices derived from a DARPP-32 knockout rat, the basal level of coupling is significantly higher than in control, and furthermore, the D2 agonist fails to increase coupling above this elevated baseline (75). These data suggest that coupling is normally suppressed by an action of DARPP-32, and that this suppression can be overcome by D2 agonist administration.

Dye coupling is also affected by maintained changes in DA system function. Changes in coupling are observed following lesions of the DA system with the neurotoxin 6-hydroxydopamine. Only the rats that exhibit severe loss of the DA innervation (i.e., >95%) also show a substantial increase in the level of dye coupling among striatal neurons (84). In all cases, the coupling was present only between cells of the same morphologic class; that is, between medium spiny neurons or between aspiny neurons. In addition, withdrawal from repeated drug treatment such as amphetamine (Fig. 9.3) (85) or antipsychotic drugs (86 ,87) cause a regionally selective increase in dye coupling. Amphetamine and antipsychotic drugs increase coupling in limbic striatum, whereas classic antipsychotic drugs also cause an increase in coupling in the motor-related dorsal striatum. These effects are only observed following withdrawal from the drug. Given that changes in gap junction composition are observed during repeated cocaine administration (88), it is possible that the system compensates for the presence of the drug by altering gap junctions to allow coupling to be maintained at its basal state. Under these conditions, the alteration is only observed when the adapted state is altered by withdrawal of the drug. Indeed, the observation that coupling is maintained for weeks following drug withdrawal suggests that the system may have reached a new stable steady state that could leave it more susceptible to destabilizing influences (85).

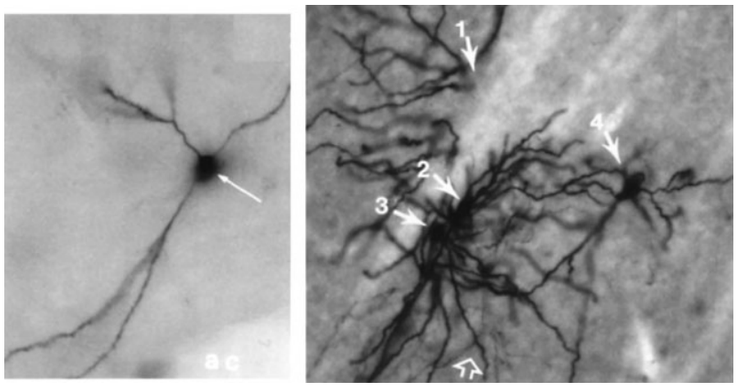


FIGURE 9.3. Long-term alterations in DA transmission lead to changes in dye coupling within the striatal complex. Medium spiny neurons in the nucleus accumbens were injected during *in vivo* intracellular recording with Lucifer yellow, which was then converted into a dense stain using antibodies. In a control rat, injection of Lucifer yellow typically labels only a single neuron *(left)*; overall, less than 15% of accumbens neurons injected in control rats exhibit labeling of more than a single neuron. In contrast, in rats that had been administered amphetamine for 2 to 4 weeks and then withdrawn for at least 7 days, the majority of injected neurons exhibited dye coupling (>60% of cells injected). In this case, four neurons were labeled after injecting a single neuron with Lucifer yellow. This increase in coupling persisted for at least 28 days following amphetamine withdrawal, but was not present if the rats were tested during the treatment phase. (From Onn and Grace, 2000; used with permission.)

Interactions with Other Neurotransmitters

DA has also been shown to affect the response of striatal neurons to other neurotransmitters. Thus, DA was found to

modulate the response of striatal neurons to glutamatergic excitation (89). Specifically, D1-receptor stimulation enhances NMDA-mediated currents (90), which may occur via a combination of two effects: (a) a facilitation of L-type calcium conductances on dendrites (90), and (b) activation of cAMP-PKA cascade (91). A similar D1-mediated cascade also attenuates responses to GABA in the striatum (92 ,93). In contrast, D2 stimulation appears to preferentially attenuate non–NMDA-mediated responses (89). There is also evidence that the activation of DA neuron firing by stimulation of DA axons (70 ,94) occurs via a D1-mediated facilitation of glutamate transmission (94). This response, which occurs in parallel with a D1-mediated increase in c-fos in striatonigral neurons (20), is more potent when the DA axons are stimulated in a burst-firing pattern (70). This suggests that, under physiologic conditions, D1-induced facilitation of glutamate transmission in the striatum is mediated by burst-firing–dependent phasic DA release (44).

In addition to its ability to modulate neurotransmitter actions on postsynaptic neurons in the striatum, DA also plays a significant modulatory role in the presynaptic regulation of neurotransmitter release. D2 stimulation is reported to presynaptically decrease GABA release from intrinsic neurons (95) and glutamate release from corticostriatal terminals. Several studies report that D2 agonists cause a down-regulation of glutamate-mediated EPSPs on neurons in the nucleus accumbens (96 ,97 ,98 and 99). This is consistent with biochemical studies showing D2-mediated down-regulation of stimulated glutamate release in striatal tissue (100 ,101) and the presence of D2 receptors on presynaptic terminals making asymmetric synapses in the striatum (78), which are presumed to be the glutamatergic corticostriatal afferents. Interestingly, D2 stimulation does not inhibit all corticostriatal EPSPs in normal preparations (97 ,99); however, after acute depletion of endogenous DA, all corticoaccumbens EPSPs are sensitive to DA (99). This suggests that under normal circumstances, the presynaptic DA receptors may already be saturated with DA, as suggested by the observation that sulpiride increase EPSP amplitude in a majority of cases when administered alone (99). This unusual pharmacology may reflect a contribution of presynaptic D4 receptors on the corticoaccumbens terminals to this response (102). Although another group has reported a D1-mediated presynaptic action EPSPs evoked by intrastriatal stimulation in slices, which was interpreted as a presynaptic effect on corticostriatal terminals (103), this study employed exceedingly high doses of the D1 agonist to achieve these effects (i.e., 100 μM, which is approximately two orders of magnitude higher than should be required for a selective D1 action). Moreover, anatomic studies have shown that D1 immunoreactive axons are exceedingly rare in the striatum (77). In contrast, recent studies suggest that DA acting on postsynaptic D1 receptors may actually cause a transsynaptic feed-forward inhibition of glutamate release. Both NMDA antagonists and adenosine antagonists can block this effect. These data suggested that dopamine depresses the excitatory postsynaptic conductance (EPSC) by causing an NMDA receptor-dependent increase in extracellular adenosine, which acts presynaptically to depress glutamate release (104). The D1–NMDA-R interaction appears to be postsynaptic and acts via PKC activation (105). It is of interest to note that there is other evidence of interdependence between DA and adenosine. Thus, a recent study by Ginés and colleagues (106) have shown that D1 and adenosine A1 receptors have the capacity to form heteromeric complexes, which appear to play a role in receptor desensitization and trafficking.

Consequently, there appears to be a complex, dynamic equilibrium between dopamine and glutamate transmission within the striatal complex, with glutamate contributing to DA release and DA causing a two-pronged inhibition of glutamate release, both directly via D2 presynaptic receptors and indirectly using adenosine as an intermediary. Finally, glutamate-released NO also appears to play a significant role in modulating DA systems and striatal neuron responsivity. The tight interdependence and coregulation between DA and glutamate suggest that the system is designed to maintain stable levels of transmission to the striatal neurons over the long term, whereas short-term changes in activity in either system in response to a signal are amplified by their coordinated effects on each of these interdependent processes.

LTP and LTD

DA also appears to have a role in short- and long-term synaptic plasticity within the striatum. Specifically, DA was found to influence two opposite types of synaptic plasticity within the striatum that depend on the history of synaptic input to this structure. In cases in which striatal excitatory amino acid afferents arising from the cortex are stimulated with high frequencies in the absence of magnesium (to enhance NMDA conductances), a long-term facilitation in synaptic transmission is induced, known as long-term potentiation. In contrast, if the stimulation is carried out at a low frequency, the opposite type of plasticity is induced; that is, long-term depression (LTD) (107). These forms of synaptic plasticity have been proposed to play a major role in learning and memory formation in other structures, such as the hippocampus. Such plasticity within the striatum may be involved in such phenomena as the acquisition of complex motor skills. Repetitive stimulation of corticostriatal fibers to release glutamate is required for the induction of LTP and LTD, which only occurs in the presence of DA afferent input (108). Thus, D1 and/or D2 antagonist pretreatment prevents the induction of LTD (107), suggesting that a synergistic interaction between these receptor subtypes is required for this process to occur. In contrast, cortical stimulation-induced LTP is blocked selectively by D1 antagonists, but is actually enhanced by D2 antagonists or in D2 receptor knockout mice (109).

Prefrontal Cortex

The effects of DA within the PFC have been controversial, in that several groups have failed to produce consistent results. Thus, although studies done *in vivo* have consistently shown that direct DA application inhibits PFC neuron firing, studies using *in vitro* slice preparations have found a DA-mediated increase (110 ,111) and a decrease (112 ,113) in neuronal excitability in this region. D1 stimulation has been shown to affect sodium conductances by increasing the sodium plateau potential and shifting the activation of sodium currents to more negative potentials (114). This increase in excitability was augmented by a D1-induced decrease in slow potassium conductances (110). D1 stimulation may also activate L-type calcium conductances located in proximal dendrites of pyramidal neurons to further increase excitability in these neurons (66). Such an interaction has been postulated to differentially modulate afferent input to these neurons (Fig. 9.4). Indeed, the highly organized DAergic input onto virtually every dendrite of PFC pyramidal neurons in the primate provides a means for this neurotransmitter to regulate nearly the entire complement of glutamatergic afferents to this cell type (115). In contrast, at least part of the inhibitory action of DA on PFC pyramidal neurons may occur by DA-induced excitation of GABAergic interneurons (116), which also receive a direct DA innervation (115 ,117).

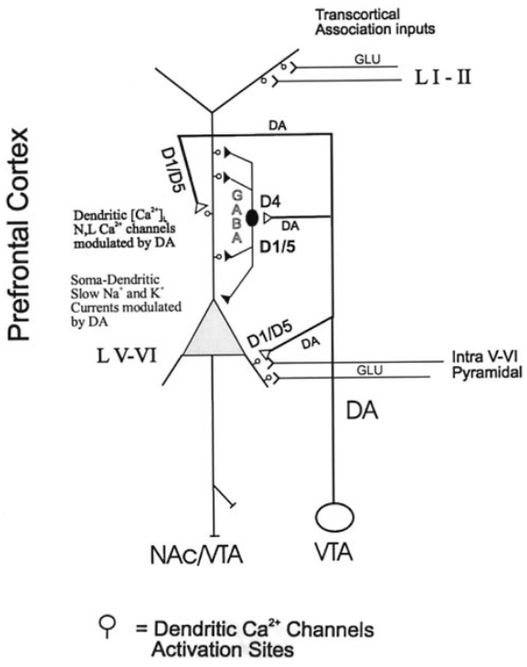


FIGURE 9.4. A simplified diagram illustrating the basic processing units within the prefrontal cortex. Each unit consists of a deep layer pyramidal neuron that projects to the nucleus accumbens or the ventral tegmental area (VTA), and a GABAergic interneuron. The apical and basal dendrites of the pyramidal neuron receive functionally segregated inputs from various cortical and subcortical regions, whereas the GABAergic interneuron, which is also modulated by DA, exerts inhibitory influences over both the apical dendrite and soma of the pyramidal neuron. By acting on both the interneuron and pyramidal neuron dendrites, the DA input has the capacity to modulate the integration of the functionally diverse array of inputs to this neuron. DA acting on D1/D5 receptors to modulate calcium channel subtypes on the apical dendrite are proposed to "sharpen" or, with greater stimulation, attenuate afferent signals arising from these distal regions. DA activation of the GABAergic interneuron can also serve to suppress information input from the apical dendrites. In contrast, DA modulation of conductances at the somatodendritic region amplifies low-level afferent inputs from neighboring pyramidal neurons. In this way, the DAergic input is proposed to change the pyramidal neuron from responding primarily to long-loop afferents to a state in which it responds primarily to local circuit interactions that may subserve working memory functions. (From Yang et al., 1999, with permission.)

It is known that PFC neurons *in vivo* exhibit bistable membrane potentials, which alternate between a hyperpolarized, nonfiring condition and a depolarized plateau state where they fire action potentials. Moreover, studies have shown that the effects of DA vary depending on the state of the membrane potential at which it is administered. In particular, DA and D1 agonists cause an increase in excitability of PFC neurons in the depolarized state but not at the hyperpolarized state (118). Furthermore, studies combining *in vivo* microdialysis administration of drugs with intracellular recording (119) found that DA could potentiate glutamate-driven bistable states of PFC neurons (Fig. 9.5). Therefore, the state of the membrane may significantly influence the response to DA observed.

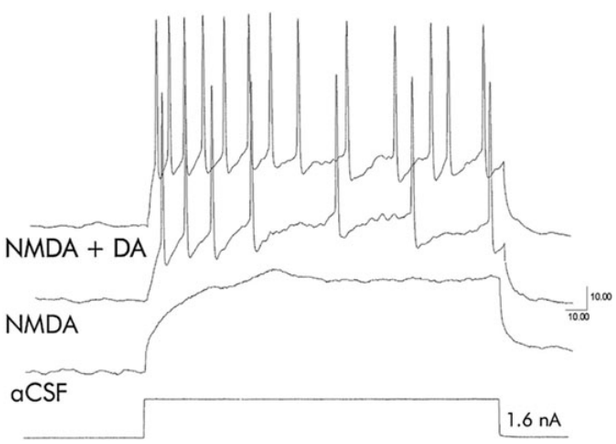


FIGURE 9.5. *In vivo* intracellular recordings from a pyramidal neuron in the frontal cortex of a chloral hydrate anesthetized rat is illustrated. The neuron was located near a microdialysis probe implanted to deliver the compounds to be tested by reverse dialysis to the environment of the cell. Administration of NMDA (20 μM) increases the number of spikes evoked by brief depolarizing pulses. Following washout, administration of NMDA + DA (30 μM) greatly increases the spikes evoked per unit current. This occurs despite the fact that DA did not appear to significantly affect current threshold. In the case of NMDA alone, the increase in the number of spikes per unit current occurs with (and may be secondary to) a decrease in current threshold; that is, the cell is simply more excitable.

In the case of DA added to NMDA, the cell fires more spikes during the current-induced depolarization, but the current threshold is not further decreased; therefore, DA allowed the cell to respond maximally to the NMDA input without altering the threshold for spiking. (From Moore et al., 1998b, with permission.)

Ventral Pallidum

The ventral pallidum (VP) receives a DA innervation from the midbrain (120), and is believed to play a significant role in several of the behavioral aspects of DA system function (121), particular related to drug sensitization (122). *In vivo* dopamine iontophoresis is known to: (a) increase and decrease VP neuronal firing (120 ,123); (b) potentiate or attenuate the excitatory effects of glutamate iontophoresis (124); and (c) modulate the firing rate enhancements produced in VP neurons by activating the amygdala (123) and attenuating the excitatory influences of the amygdala on pallidal cell firing at local concentration that are below that required to alter spontaneous firing (123). Dopamine receptors for the D1 and D2 class have been identified in the ventral pallidum (120), and electrophysiologic and behavioral evaluations have revealed that these two classes operate in opposition in this region (which contrasts the "enabling" effects reported for striatal regions and other pallidal regions). Local D1 activation induces a robust attenuation of cell firing (125) and enhances locomotor activity (126), whereas D2 activation slightly attenuates or has no effect on firing (125) and is largely without influence on motor activity (126).

Because the VP is positioned anatomically at the crossroads of the limbic and extrapyramidal system, DA modulation in this area has the ability to potently influence motivated behavior by its actions in this region (121).

Mediodorsal Thalamus

Anatomic studies have revealed the presence of a DA innervation of the mediodorsal (MD) thalamic nucleus arising from the midbrain. Using *in vitro* intracellular recordings, DA was found to alter MD neuron activity via a D2-mediated effect. In particular, quinpirole was found to increase membrane excitability and enhance the low threshold spike in these neurons (127). This was mediated at least in part via an alteration in potassium conductances. By increasing the low threshold spike, DA was found to facilitate oscillatory activity within the MD, which would potently impact thalamocortical information processing in this region.

Basolateral Amygdala

The basolateral nucleus of the amygdala (BLA) exhibits a substantial innervation from the midbrain DA neurons. The effects produced by DA in the BLA are dependent on the type of neuron recorded (128). DA causes an overall decrease in the firing rate of presumed projection neurons by two mechanisms: (a) a direct effect on the projection neuron, and (b) an activation of the firing of putative interneurons, which may be analogous to the interactions occurring in the PFC. In addition, DA produced effects on afferent drive of these neurons that was dependent on the origin of the projection system. Thus, DA attenuates afferents from limbic structures such as the PFC and MD thalamus, whereas afferent input from auditory association cortex (Te3) is potentiated (Fig. 9.6). Intracellular recordings revealed that this was a consequence of a D1-mediated decrease in PFC-evoked EPSP amplitude, combined with a D2-mediated increase in BLA input resistance that potentiated Te3 afferent drive (129). PFC stimulation also caused an excitation of BLA interneurons, which lead to a subsequent attenuation of input arising from Te3; however, in the presence of DA stimulation, the ability of the PFC stimulation to attenuate responses from Te3 was diminished (129). These data suggest that the PFC is normally capable of attenuating amygdala responses to sensory inputs, which could be a mechanism for decreasing emotional responses to familiar or nonthreatening stimuli. However, with excessive DA stimulation, the ability of the PFC to suppress amygdala-mediated emotional responses may be lost.

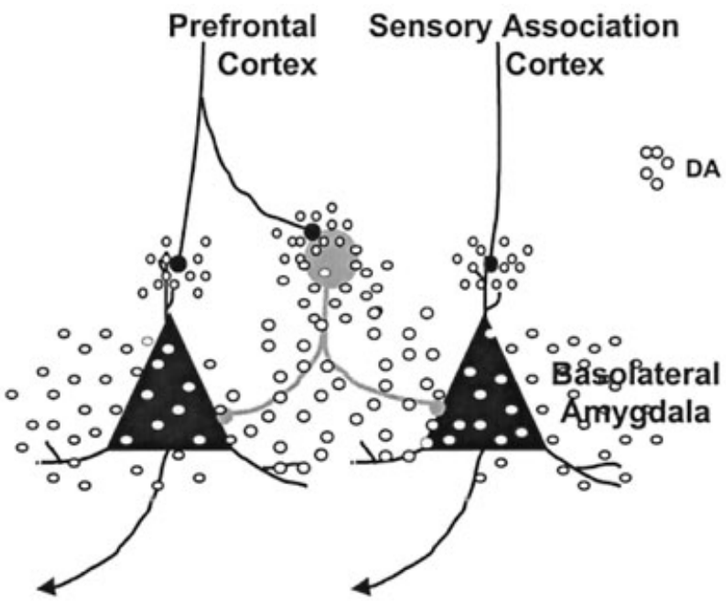


FIGURE 9.6. DA attenuates prefrontal cortex (PFC) modulation of basolateral amygdala (BLA) neuronal responses. The PFC provides a direct drive of BLA projection neurons and interneurons, whereas inputs from the sensory association cortex project only to the output neurons. As a result, the PFC inhibits the ability of the sensory association cortex to activate BLA neuron firing. However, the PFC inputs are attenuated by elevated DA levels in the BLA, removing a source of inhibition on BLA projection neurons. Furthermore, elevated DA levels in the BLA increase the input resistance of BLA projection neurons, leading to augmentation of nonsuppressed inputs to BLA neurons. Thus, DA receptor activation enables sensory-driven amygdala-mediated affective responses by removal of regulatory inputs and augmenting sensory inputs. (From Rosenkranz and Grace, 2001, with permission.)

Substantia Nigra

In addition to the effects of DA on DA neuron autoreceptors within the substantia nigra, there are also DA receptors

located on striatonigral afferents to this region. Locally evoked IPSCs in neurons of the substantia nigra zona reticulata (ZR) are GABAergic in nature, and are believed to arise from striatal afferents. These IPSCs are depressed by DA acting on D1 but not D2 receptors. The fact that this depression was accompanied by increased paired-pulse facilitation and not by a change in membrane potential or conductance indicates that the effect is likely presynaptic in origin (130). It is interesting to note that the striatonigral neurons that exhibit terminal D1 receptors do not exhibit D1 receptors on their local collaterals within the striatum (77). This suggests that these neurons can selectively traffic presynaptic receptors to long projection sites.

BEHAVIORAL CORRELATES OF DA SYSTEM FUNCTION

Part of "9 - Dopamine "

DA is known to play an important role in working memory and response sequencing in the PFC. In particular, DA acting on D1 receptors has been shown to exert dual actions on these types of behaviors. Thus, D1 agonist administration into the PFC of rats with poor performance on attentional function tasks significantly improved their performance, whereas impairing performance in rats that had higher baseline attentional skills (131). This is consistent with studies suggesting that optimal DA levels are required to maintain function in the PFC, with both too high or low D1 stimulation leading to impaired working memory function (132 ,133).

Several studies have shown that the DA system is activated by rewarding stimuli, such as food (134 ,135); however, it is becoming evident that DA is not the reward signal per se, but instead is necessary for the acquisition of reinforcing stimuli. In some cases, DA has been described as a type of error signal (136), in which the predicted occurrence of reward does not correlate with the behavioral response emitted to generate this reward. Thus, when a task is well learned, DA neuron firing no longer is a necessary correlate of the reward signal. But if reward is absent, DA neuron firing appears to decrease (137). Studies of DA overflow in the nucleus accumbens show that DA is released when the DA cell bodies are stimulated electrically. However, when the stimulation is contingent on a bar press by the rat, the DA overflow does not occur even in the presence of the electrical stimulus (138). These data suggest that the lack of DA system activation during a well-learned contingent reinforcement task is not simply a failure to activate DA neuron firing, but instead may represent an offsetting inhibitory influence over the DA system, either at the level of the DA cell body or the terminal. Indeed, the reports of an anticipatory increase in extracellular DA in the accumbens prior to self-administration of a DA drug such as cocaine (139) could potentially increase extracellular DA sufficiently to inhibit phasic DA release occurring via stimulation of the DA cell bodies (44).

Overall, studies support the suggestion that DA actions in the PFC may have a greater involvement in the regulation of novel circumstances, with the striatum involved more in expression of learned behaviors (140). This model is consistent with the physiologic studies cited that show that DA can selectively activate circuits within frontal cortex and striatal complex, potentially facilitating information flow along new pathways when a change occurs, but playing less of a role once a new stable steady state is achieved at which the internal representation is at equilibrium with the predicted external events.

SUMMARY OF DA ACTIONS

Part of "9 - Dopamine "

It is clear from the preceding that DA exerts multiple actions at each level of integration within the cortico-striato-pallido-thalamo-cortical loop. The actions exerted at each stage of this loop appear to have marked differences, however. For example, DA acting on primary inputs to this circuit (e.g., the amygdala and PFC) affect both primary neurons and interneurons, with the net effect being a selective potentiation of particular afferent drive sources. Within the striatum, DA exerts actions on presynaptic terminals containing glutamate, as well as affecting the actions of glutamate on postsynaptic neurons. Combined with the reciprocal feedback interactions between glutamate and DA terminals, this system appears to be designed to facilitate rapid changes in input states while attenuating any long-term alterations that may occur. Moreover, the effects of DA on cellular coupling provide a type of reversible hardwiring, which may facilitate performance of well-learned motor actions (141). Within the VP and MD, DA has effects that would alter the behavioral output by changing the state of neuronal activity within these structures. Therefore, DA could enable multiple state transitions within these regions, selecting among competing inputs, facilitating information transfer, and altering states that would ultimately feed back via the thalamus to reinforce cortical activity that is most pertinent to the task at hand (58 ,141).

The actions of DA may best be described not in terms of inhibition or excitation, but rather as related to the gating of inputs and modulation of states of neuronal elements. This modulation of information integration is then further influenced at the network level via the actions of DA on interneurons or cell coupling. Such a description is consistent with the behavioral actions of this transmitter as well, in that it does not directly produce a motor output or reward signal, but instead modulates inputs and adjusts the states of the organism in order to redirect the stimulus-response output to achieve the most effective behavioral strategy. Given these constraints, one could imagine how dysfunctions in such a system could produce the profound pathologic states that have been attributed to DA.

ACKNOWLEDGMENTS

Part of "9 - Dopamine "

Dr. Grace has received research support from Pharmacia-Upjohn and Warner-Lambert/Parke Davis, as well as travel support for presenters at a scientific panel he co-organized for Pharmacia-Upjohn.

REFERENCES

1. Grace AA, Bunney BS. Intracellular and extracellular electrophysiology of nigral dopaminergic neurons—1. Identification and characterization. *Neuroscience* 1983;10:301–315.

2. Grace AA, Onn SP. Morphology and electrophysiological properties of immunocytochemically identified rat dopamine neurons recorded *in vitro*. *J Neurosci* 1989;9:3463–3481.

3. Grace AA, Bunney BS. The control of firing pattern in nigral dopamine neurons: single spike firing. *J Neurosci* 1984;4:2866–2876.

4. Harden DG, Grace AA. Activation of dopamine cell firing by repeated L-DOPA administration to dopamine-depleted rats: its potential role in mediating the therapeutic response to L-DOPA treatment. *J Neurosci* 1995;15:6157–6166.

5. Pucak ML, Grace AA. Effects of haloperidol on the activity and membrane physiology of substantia nigra dopamine neurons recorded in vitro. *Brain Res* 1996;713:44–52.

6. Mercuri NB, Saiardi A, Bonci A, et al. Loss of autoreceptor function in dopaminergic neurons from dopamine D2 receptor deficient mice. *Neuroscience* 1997;79:323–327.

7. Mercuri NB, Scarponi M, Bonci A, et al. Monoamine oxidase inhibition causes a long-term prolongation of the dopamine-induced responses in rat midbrain dopaminergic cells. *J Neurosci* 1997;17:2267–2272.

8. Mercuri NB, Grillner P, Bernardi G. N-methyl-D-aspartate receptors mediate a slow excitatory postsynaptic potential in the rat midbrain dopaminergic neurons. *Neuroscience* 1996;74:785–792.

9. Mercuri NB, Federici M, Bernardi G. Inhibition of catechol-O-methyltransferase (COMT) in the brain does not affect the action of dopamine and levodopa: an in vitro electrophysiological evidence from rat mesencephalic dopamine neurons. *J Neural Transm* 1999;106:1135–1140.

10. Grace AA, Bunney BS. Opposing effects of striatonigral feedback pathways on midbrain dopamine cell activity. *Brain Res* 1985;333:271–284.

11. Grace AA, Bunney BS. Effects of baclofen on nigral dopaminergic cell activity following acute and chronic haloperidol treatment. *Brain Res Bull* 1980;5:537–543.

12. Erhardt S, Nissbrandt H, Engberg G. Activation of nigral dopamine neurons by the selective GABA(B)-receptor antagonist SCH 50911. *J Neural Transm* 1999;106:383–394.

13. Paladini CA, Celada P, Tepper JM. Striatal, pallidal, and pars reticulata evoked inhibition of nigrostriatal dopaminergic neurons is mediated by GABA(A) receptors in vivo. *Neuroscience* 1999;89:799–812.

14. Grace AA, Bunney BS. The control of firing pattern in nigral dopamine neurons: burst firing. *J Neurosci* 1984;4:2877–2890.

15. Bonci A, Grillner P, Siniscalchi A, et al. Glutamate metabotropic receptor agonists depress excitatory and inhibitory transmission on rat mesencephalic principal neurons. *Eur J Neurosci* 1997;9:2359–2369.

16. Grillner P, Berretta N, Bernardi G, et al. Muscarinic receptors depress GABAergic synaptic transmission in rat midbrain dopamine neurons. *Neuroscience* 2000;96:299–307.

17. Grillner P, Bonci A, Svensson TH, et al. Presynaptic muscarinic (M3) receptors reduce excitatory transmission in dopamine neurons of the rat mesencephalon. *Neuroscience* 1999;91:557–565.

18. Keefe KA, Zigmond MJ, Abercrombie, ED Extracellular dopamine in striatum: influence of nerve impulse activity in medial forebrain bundle and local glutamatergic input. *Neuroscience* 1992;47:325–332.

19. Chergui K, Nomikos GG, Mathe JM, et al. Burst stimulation of the medial forebrain bundle selectively increase Fos-like immunoreactivity in the limbic forebrain of the rat. *Neuroscience* 1996;72:141–156.

20. Chergui K, Svenningsson P, Nomikos GG, et al. Increased expression of NGFI-A mRNA in the rat striatum following burst stimulation of the medial forebrain bundle. *Eur J Neurosci* 1997;9:2370–2382.

21. Lokwan SJ, Overton PG, Berry MS, et al. Stimulation of the pedunculopontine tegmental nucleus in the rat produces burst firing in A9 dopaminergic neurons. *Neuroscience* 1999;92:245–254.

22. Smith ID, Grace AA. Role of the subthalamic nucleus in the regulation of nigral dopamine neuron activity. *Synapse* 1992;12:287–303.

23. Kitai ST, Shepard PD, Callaway JC, et al. Afferent modulation of dopamine neuron firing patterns. *Curr Opin Neurobiol* 1999;9:690–697.

24. Shepard PD, Stump D. Nifedipine blocks apamin-induced bursting activity in nigral dopamine-containing neurons. *Brain Res* 1999;817:104–109.

25. Tong ZY, Overton PG, Clark D. Stimulation of the prefrontal cortex in the rat induces patterns of activity in midbrain dopaminergic neurons, which resemble natural burst events. *Synapse* 1996;22:195–208.

26. Carr DB, O'Donnell P, Card JP, et al. Dopamine terminals in the rat prefrontal cortex synapse on pyramidal cells that project to the nucleus accumbens. *J Neurosci* 1999;19:11049–11060.

27. Carr DB, Sesack SR. Projections from the rat prefrontal cortex to the ventral tegmental area: target specificity in the synaptic associations with mesoaccumbens and mesocortical neurons. *J Neurosci* 2000;20:3864–3873.

28. Taber MT, Das S, Fibiger HC. Cortical regulation of subcortical dopamine release: mediation via the ventral tegmental area. *J Neurochem* 1995;65:1407–1410.

29. Legault M, Rompre PP, Wise RA. Chemical stimulation of the ventral hippocampus elevates nucleus accumbens dopamine by activating dopaminergic neurons of the ventral tegmental area. *J Neurosci* 2000;20:1635–1642.

30. Legault M, Wise RA. Injections of N-methyl-D-aspartate into the ventral hippocampus increase extracellular dopamine in the ventral tegmental area and nucleus accumbens. *Synapse* 1999;31:241–249.

31. Floresco SB, Todd CL, Grace AA. Glutamatergic afferents from the hippocampus to the nucleus accumbens regulate activity of ventral tegmental area dopamine neurons. *J Neurosci* 2001;21:4915–4922.

32. Haber SN, Fudge JL, McFarland NR. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci* 2000;20:2369–2382.

33. Braszko JJ, Bannon MJ, Bunney BS, et al. Intrastriatal kainic acid: acute effects on electrophysiological and biochemical measures of nigrostriatal dopaminergic activity. *J Pharmacol Exp Ther* 1981;216:289–293.

34. West AR, Grace AA. Striatal nitric oxide signaling regulates the neuronal activity of midbrain dopamine neurons in vivo. *J Neurophysiol* 2000;83:1796–1808.

35. Abercrombie ED, Keefe KA, DiFrischia DS, et al. Differential effect of stress on in vivo dopamine release in striatum, nucleus accumbens, and medial frontal cortex. *J Neurochem* 1989;52:1655–1658.

36. Inglis FM, Moghaddam B. Dopaminergic innervation of the amygdala is highly responsive to stress. *J Neurochem* 1999;72:1088–1094.

37. Davis M, Hitchcock JM, Bowers MB, et al. Stress-induced activation of prefrontal cortex dopamine turnover: blockade by lesions of the amygdala. *Brain Res* 1994;664:207–210.

38. King D, Zigmond MJ, Finlay JM. Effects of dopamine depletion in the medial prefrontal cortex on the stress-induced increase in extracellular dopamine in the nucleus accumbens core and shell. *Neuroscience* 1997;77:141–153.

39. Harden DG, King D, Finlay JM, et al. Depletion of dopamine in the prefrontal cortex decreases the basal electrophysiological activity of mesolimbic dopamine neurons. *Brain Res* 1998;794:96–102.

40. Moore H, Rose HJ, Grace AA. Chronic cold stress reduces the spontaneous activity of ventral tegmental dopamine neurons. *Neuropsychopharmacology* 2000;24:410–419.

41. Jones SR, Garris PA, Kilts CD, et al. Comparison of dopamine uptake in the basolateral amygdaloid nucleus, caudate-putamen, and nucleus accumbens of the rat. *J Neurochem* 1995;64:2581–2589.

42. Giros B, Jaber M, Jones SR, et al. Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. *Nature* 1996;379:606–612.

43. Moore H, West AR, Grace AA. The regulation of forebrain dopamine transmission: relevance to the pathophysiology and psychopathology of schizophrenia. *Biol Psychiatry* 1999;46:40–55.

44. Grace AA. Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* 1991;41:1–24.

45. Floresco SB, Phillips AG. Dopamine and hippocampal inputs to the nucleus accumbens play an essential role in the search for food in an unpredictable environment. *Psychobiology* 1999;27:277–286.

46. Blaha CD, Yang CR, Floresco SB, et al. Stimulation of the ventral subiculum of the hippocampus evokes glutamate receptor-mediated changes in dopamine efflux in the rat nucleus accumbens. *Eur J Neurosci* 1997;9:902–911.

47. Floresco SB, Yang CR, Phillips AG, et al. Basolateral amygdala stimulation evokes glutamate receptor-dependent dopamine efflux in the nucleus accumbens of the anaesthetized rat. *Eur J Neurosci* 1998;10:1241–1251.

48. Gracy KN, Pickel VM. Ultrastructural immunocytochemical localization of the N-methyl-D-aspartate receptor and tyrosine hydroxylase in the shell of the rat nucleus accumbens. *Brain Res* 1996;739:169–181.

49. Antonelli T, Govoni BM, Bianchi C, et al. Glutamate regulation of dopamine release in guinea pig striatal slices. *Neurochem Int* 1997;30:203–209.

50. Ohno M, Watanabe S. Persistent increase in dopamine release following activation of metabotropic glutamate receptors in the rat nucleus accumbens. *Neurosci Letts* 1995;200:113–116.

51. Verma A, Moghaddam B. Regulation of striatal dopamine release by metabotropic glutamate receptors. *Synapse* 1998;28:220–226.

52. West AR, Galloway MP. Intrastriatal infusion of (+/-)-S-nitroso-*N*-acetylpenicillamine releases vesicular dopamine via an ionotropic glutamate receptor-mediated mechanism: an in vivo microdialysis study in chloral hydrate-anesthetized rats. *J Neurochem* 1996;66:1971–1980.

53. West AR, Galloway MP. Endogenous nitric oxide facilitates striatal dopamine and glutamate efflux in vivo: role of ionotropic glutamate receptor-dependent mechanisms. *Neuropharmacology* 1997;36:1571–1581.

54. West AR, Galloway MP. Desensitization of 5-hydroxytryptamine-facilitated dopamine release in vivo. *Eur J Pharmacol* 1996;298:241–245.

55. Smolders I, De Klippel N, Sarre S, et al. Tonic GABAergic modulation of striatal dopamine release studied by *in vivo* microdialysis in the freely moving rat. *Eur J Pharmacol* 1995;284:83–91.

56. Wolf ME, Dahlin SL, Hu XT, et al. Effects of lesions of prefrontal cortex, amygdala, or fornix on behavioral sensitization to amphetamine: comparison with N-methyl-D-aspartate antagonists. *Neuroscience* 1995;69:417–439.

57. Grace AA. The tonic/phasic model of dopamine system regulation: its relevance for understanding how stimulant abuse can alter basal ganglia function. *Drug Alcohol Depend* 1995;37:111–129.

58. Grace AA. Psychostimulant action on dopamine and limbic system function: relevance to the pathophysiology and treatment of ADHD. In: *Stimulant drugs and ADHD: basic and clinical neuroscience*. New York: Oxford University Press, 2001:134–157.

59. Grace AA, Bunney BS, Moore H, et al. Dopamine-cell depolarization block as a model for the therapeutic actions of antipsychotic drugs. *Trends Neurosci* 1997;20:31–37.

60. Moore H, Todd CL, Grace AA. Striatal extracellular dopamine levels in rats with haloperidol-induced depolarization block of substantia nigra dopamine neurons. *J Neurosci* 1998;18:5068–5077.

61. Chergui K, Suaud-Chagny MF, Gonon F. Nonlinear relationship between impulse flow, dopamine release and dopamine elimination in the rat brain in vivo. *Neuroscience* 1994;62:641–645.

62. Jones SR, Gainetdinov RR, Jaber M, et al. Profound neuronal plasticity in response to inactivation of the dopamine transporter. *Proc Natl Acad Sci USA* 1998;95:4029–4034.

63. Sesack SR, Hawrylak VA, Matus C, et al. Dopamine axon varicosities in the prelimbic division of the rat prefrontal cortex exhibit sparse immunoreactivity for the dopamine transporter. *J Neurosci* 1998;18:2697–2708.

64. Gresch PJ, Sved AF, Zigmond MJ, et al. Local influence of endogenous norepinephrine on extracellular dopamine in rat medial prefrontal cortex. *J Neurochem* 1995;65:111–116.

65. Smiley JF, Levey AI, Ciliax BJ, et al. D1 dopamine receptor immunoreactivity in human and monkey cerebral cortex: predominant and extrasynaptic localization in dendritic spines. *Proc Natl Acad Sci USA* 1994;91:5720–5724.

66. Yang CR, Seamans JK, Gorelova N. Developing a neuronal model for the pathophysiology of schizophrenia based on the nature of electrophysiological actions of dopamine in the prefrontal cortex. *Neuropsychopharmacology* 1999;21:161–194.

67. Calabresi P, Mercuri N, Stanzione P, et al. Intracellular studies on the dopamine-induced firing inhibition of neostriatal neurons in vitro: evidence for D1 receptor involvement. *Neuroscience* 1987;20:757–771.

68. O'Donnell P, Grace AA. Dopaminergic reduction of excitability in nucleus accumbens neurons recorded in vitro. *Neuropsychopharmacology* 1996;15:87–97.

69. White FJ, Wang RY. Electrophysiological evidence for the existence of both D-1 and D-2 dopamine receptors in the rat nucleus accumbens. *J Neurosci* 1986;6:274–280.

70. Gonon F. Prolonged and extrasynaptic excitatory action of dopamine mediated by D1 receptors in the rat striatum in vivo. *J Neurosci* 1997;17:5972–5978.

71. Hernandez-Lopez S, Bargas J, Surmeier DJ, et al. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2+ conductance. *J Neurosci* 1997;17:3334–3342.

72. Cepeda C, Chandler SH, Shumate LW, et al. Persistent Na+ conductance in medium-sized neostriatal neurons: characterization using infrared videomicroscopy and whole cell patch-clamp recordings. *J Neurophysiol* 1995;74:1343–1348.

73. Surmeier DJ, Eberwine J, Wilson CJ, et al. Dopamine receptor subtypes colocalize in rat striatonigral neurons. *Proc Natl Acad Sci USA* 1992;89:10178–10182.

74. Schiffmann SN, Lledo PM, Vincent JD. Dopamine D1 receptor modulates the voltage-gated sodium current in rat striatal neurones through a protein kinase A. *J Physiol (Lond)* 1995;483:95–107.

75. Onn SP, West AR, Grace AA. Dopamine regulation of neuronal and network interactions within the striatum. *Trends Neurosci* 2000;23:S48–S56.

76. Jansson A, Goldstein M, Tinner B, et al. On the distribution patterns of D1, D2, tyrosine hydroxylase and dopamine transporter immunoreactivities in the ventral striatum of the rat. *Neuroscience* 1999;89:473–489.

77. Hersch SM, Ciliax BJ, Gutekunst CA, et al. Electron microscopic analysis of D1 and D2 dopamine receptor proteins in the dorsal striatum and their synaptic relationships with motor corticostriatal afferents. *J Neurosci* 1995;15:5222–5237.

78. Sesack SR, Aoki C, Pickel VM. Ultrastructural localization of D2 receptor-like immunoreactivity in midbrain dopamine neurons and their striatal targets. *J Neurosci* 1994;14:88–106.

79. Surmeier DJ, Bargas J, Hemmings HC Jr, et al. Modulation of calcium currents by a D1 dopaminergic protein kinase/phosphatase cascade in rat neostriatal neurons. *Neuron* 1995;14:385–397.

80. Lin YJ, Greif GJ, Freedman JE. Permeation and block of dopamine-modulated potassium channels on rat striatal neurons by cesium and barium ions. *J Neurophysiol* 1996;76:1413–1422.

81. Greengard P, Allen PB, Nairn AC. Beyond the dopamine receptor: the DARPP-32/protein phosphatase-1 cascade. *Neuron* 1999;23:435–447.

82. Fienberg AA, Hiroi N, Mermelstein PG, et al. DARPP-32: regulator of the efficacy of dopaminergic neurotransmission. *Science* 1998;281:838–842.

83. Langley KC, Bergson C, Greengard P, et al. Co-localization of the D1 dopamine receptor in a subset of DARPP-32-containing neurons in rat caudate-putamen. *Neuroscience* 1997;78:977–983.

84. Onn SP, Grace AA. Alterations in electrophysiological activity and dye coupling of striatal spiny and aspiny neurons in dopamine-denervated rat striatum recorded in vivo. *Synapse* 1999;33:1–15.

85. Onn SP, Grace AA. Amphetamine withdrawal alters bistable states and cellular coupling in rat prefrontal cortex and nucleus accumbens neurons recorded in vivo. *J Neurosci* 2000;20:2332–2345.

86. Onn SP, Grace AA. Repeated treatment with haloperidol and clozapine exerts differential effects on dye coupling between neurons in subregions of striatum and nucleus accumbens. *J Neurosci* 1995;15:7024–7036.

87. O'Donnell P, Grace AA. Different effects of subchronic clozapine and haloperidol on dye-coupling between neurons in the rat striatal complex. *Neuroscience* 1995;66:763–767.

88. Bennett SA, Arnold JM, Chen J, et al. Long-term changes in connexin32 gap junction protein and mRNA expression following cocaine self-administration in rats. *Eur J Neurosci* 1999;11:3329–3338.

89. Cepeda C, Levine MS. Dopamine and N-methyl-D-aspartate receptor interactions in the neostriatum. *Dev Neurosci* 1998;20:1–18.

90. Cepeda C, Colwell CS, Itri JN, et al. Dopaminergic modulation of NMDA-induced whole cell currents in neostriatal neurons in slices: contribution of calcium conductances. *J Neurophysiol* 1998;79:82–94.

91. Colwell CS, Levine MS. Excitatory synaptic transmission in neostriatal neurons: regulation by cAMP-dependent mechanisms. *J Neurosci* 1995;15:1704–1713.

92. Flores-Hernandez J, Hernandez S, Snyder GL, et al. D(1) dopamine receptor activation reduces GABA(A) receptor currents in neostriatal neurons through a PKA/DARPP-32/PP1 signaling cascade. *J Neurophysiol* 2000;83:2996–3004.

93. Nicola SM, Malenka RC. Modulation of synaptic transmission by dopamine and norepinephrine in ventral but not dorsal striatum. *J Neurophysiol* 1998;79:1768–1776.

94. Gonon F, Sundstrom L. Excitatory effects of dopamine released by impulse flow in the rat nucleus accumbens in vivo. *Neuroscience* 1996;75:13–18.

95. Delgado A, Sierra A, Querejeta E, et al. Inhibitory control of the GABAergic transmission in the rat neostriatum by D2 dopamine receptors. *Neuroscience* 2000;95:1043–1048.

96. Brown JR, Arbuthnott GW. The electrophysiology of dopamine (D2) receptors: a study of the actions of dopamine on corticostriatal transmission. *Neuroscience* 1983;10:349–355.

97. Flores-Hernandez J, Galarraga E, Bargas J. Dopamine selects glutamatergic inputs to neostriatal neurons. *Synapse* 1997;25:185–195.

98. Hsu KS, Huang CC, Yang CH, et al. Presynaptic D2 dopaminergic receptors mediate inhibition of excitatory synaptictransmission in rat neostriatum. *Brain Res* 1995;690:264–268.

99. O'Donnell P, Grace AA. Tonic D2-mediated attenuation of cortical excitation in nucleus accumbens neurons recorded in vitro. *Brain Res* 1994;634:105–112.

100. Maura G, Giardi A, Raiteri M. Release-regulating D-2 dopamine receptors are located on striatal glutamatergic nerve terminals. *J Pharmacol Exp Ther* 1988;247:680–684.

101. Yamamoto BK, Davy S. Dopaminergic modulation of glutamate release in striatum as measured by microdialysis. *J Neurochem* 1992;58:1736–1742.

102. Svingos AL, Periasamy S, Pickel VM. Presynaptic dopamine D(4) receptor localization in the rat nucleus accumbens shell. *Synapse* 2000;36:222–232.

103. Nicola SM, Kombian SB, Malenka RC. Psychostimulants depress excitatory synaptic transmission in the nucleus accumbens via presynaptic D1-like dopamine receptors. *J Neurosci* 1996;16:1591–604.

104. Harvey J, Lacey MG. A postsynaptic interaction between dopamine D1 and NMDA receptors promotes presynaptic inhibition in the rat nucleus accumbens via adenosine release. *J Neurosci* 1997;17:5271–5280.

105. Chergui K, Lacey MG. Modulation by dopamine D1-like receptors of synaptic transmission and NMDA receptors in rat nucleus accumbens is attenuated by the protein kinase C inhibitor Ro 32-0432. *Neuropharmacology* 1999;38:223–231.

106. Ginés S, Hillion J, Torvinen M, et al. Dopamine D1 and adenosine A1 receptors form functionally interacting heteromeric complexes. *Proc Natl Acad Sci USA* 2000;97:8606–8611.

107. Calabresi P, Centonze D, Gubellini P, et al. Synaptic transmission in the striatum: from plasticity to neurodegeneration. *Prog Neurobiol* 2000;61:231–265.

108. Centonze D, Gubellini P, Picconi B, et al. Unilateral dopamine denervation blocks corticostriatal LTP. *J Neurophysiol* 1999;82:3575–3579.

109. Calabresi P, Saiardi A, Pisani A, et al. Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. *J Neurosci* 1997;17:4536–4544.

110. Yang CR, Seamans JK. Dopamine D1 receptor actions in layers V–VI rat prefrontal cortex neurons in vitro: modulation of dendritic-somatic signal integration. *J Neurosci* 1996;16:1922–1935.

111. Shi WX, Zheng P, Liang XF, et al. Characterization of dopamine-induced depolarization of prefrontal cortical neurons. *Synapse* 1997;26:415–422.

112. Gulledge AT, Jaffe DB. Dopamine decreases the excitability of layer V pyramidal cells in the rat prefrontal cortex. *J Neurosci* 1998;18:9139–9151.

113. Geijo-Barrientos E, Pastore C. The effects of dopamine on the subthreshold electrophysiological responses of rat prefrontal cortex neurons in vitro. *Eur J Neurosci* 1995;7:358–366.

114. Gorelova N, Yang CR. Dopamine D1/D5 receptor activation modulates a persistent sodium current in rat prefrontal cortical neurons in vitro. *J Neurophysiol* 2000;84:75–87.

115. Krimer LS, Jakab RL, Goldman-Rakic PS. Quantitative three-dimensional analysis of the catecholaminergic innervation of identified neurons in the macaque prefrontal cortex. *J Neurosci* 1997;17:7450–7461.

116. Zhou FM, Hablitz JJ. Dopamine modulation of membrane and synaptic properties of interneurons in rat cerebral cortex. *J Neurophysiol* 1999;81:967–976.

117. Sesack SR, Snyder CL, Lewis DA. Axon terminals immunolabeled for dopamine or tyrosine hydroxylase synapse on GABA-immunoreactive dendrites in rat and monkey cortex. *J Comp Neurol* 1995;363:264–280.

118. Lavin A, Grace AA. Stimulation of D1-type dopamine receptors enhances excitation in PFC pyramidal neurons in a state-dependent manner. *Neuroscience* 2001;104:335–346.

119. Moore H, Lavin A, Grace AA. Interactions between dopamine and NMDA delivered locally by microdialysis during in vivo intracellular recording of rat prefrontal cortical neurons. *Soc Neurosci Abstr* 1998;24:2061.

120. Napier TC, Muench MB, Maslowsli, RJ, et al. Is dopamine a neurotransmitter within the ventral pallidum/substantia innominata? In: *The basal forebrain: anatomy to function*. New York: Plenum Press, 1991:183–195.

121. Napier TC. Transmitter actions and interactions on pallidal neuronal function. In: *Limbic motor circuits and neuropsychiatry*. Boca Raton, FL: CRC Press, 1993:125–153.

122. Johnson PI, Mitrovic I, Napier TC. Effect of chronic cocaine treatment on dopamine-induced neuronal activity in the ventral pallidum. *Soc Neurosci Abstr* 1997;23:526.

123. Maslowski-Cobuzzi RJ, Napier TC. Activation of dopaminergic neurons modulates ventral pallidal responses evoked by amygdala stimulation. *Neuroscience* 1994;62:1103–1119.

124. Johnson PI, Napier TC. GABA- and glutamate-evoked responses in the rat ventral pallidum are modulated by dopamine. *Eur J Neurosci* 1997;9:1397–1406.

125. Napier TC, Maslowski-Cobuzzi RJ. Electrophysiological verification of the presence of D1 and D2 dopamine receptors within the ventral pallidum. *Synapse* 1994;17:160–166.

126. Gong W, Neill DB, Lynn M, et al. Dopamine D1/D2 agonists injected into nucleus accumbens and ventral pallidum differentially affect locomotor activity depending on site. *Neuroscience* 1999;93:1349–1358.

127. Lavin A, Grace AA. Dopamine modulates the responsivity of mediodorsal thalamic cells recorded in vitro. *J Neurosci* 1998;18:10566–10578.

128. Rosenkranz JA, Grace AA. Modulation of basolateral amygdala neuronal firing and afferent drive by dopamine receptor activation in vivo. *J Neurosci* 1999;19:11027–11039.

129. Rosenkranz JA, Grace AA. Dopamine attenuates prefrontal cortical suppression of sensory inputs to the basolateral amygdala of rats. *J Neurosci* 2001;21:4090–4103.

130. Miyazaki T, Lacey MG. Presynaptic inhibition by dopamine of a discrete component of GABA release in rat substantia nigra pars reticulata. *J Physiol (Lond)* 1998;513:805–817.

131. Granon S, Passetti F, Thomas KL, et al. Enhanced and impaired attentional performance after infusion of D1 dopaminergic receptor agents into rat prefrontal cortex. *J Neurosci* 2000;20:1208–1215.

132. Murphy BL, Arnsten AF, Goldman-Rakic PS, et al. Increased dopamine turnover in the prefrontal cortex impairs spatial working memory performance in rats and monkeys. *Proc Natl Acad Sci USA* 1996;93:1325–1329.

133. Arnsten AF. Catecholamine regulation of the prefrontal cortex. *J Psychpharmacol* 1997;11:151–162.

134. Mirenowicz J, Schultz W. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 1996;379:449–451.

135. Taber MT, Fibiger HC. Feeding-evoked dopamine release in the nucleus, accumbens: regulation by glutamatergic mechanisms. *Neuroscience* 1997;76:1105–1112.

136. Hollerman JR, Schultz W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1998;1:304–309.

137. Schultz W. Predictive reward signal of dopamine neurons. *J Neurophysiol* 1998;80:1–27.

138. Garris PA, Kilpatrick M, Bunin MA, et al. Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. *Nature* 1999;398:67–69.

139. Gratton A, Wise RA. Drug- and behavior-associated changes in dopamine-related electrochemical signals during intravenous cocaine self-administration in rats. *J Neurosci* 1994;14:4130–4146.

140. Wilkinson LS, Humby T, Killcross AS, et al. Dissociations in dopamine release in medial prefrontal cortex and ventral striatum during the acquisition and extinction of classical aversive conditioning in the rat. *Eur J Neurosci* 1998;10:1019–1026.

141. Grace AA, Moore H. Regulation of information flow in the nucleus accumbens: A model for the pathophysiology of schizophrenia. In: *Origins and development of schizophrenia: Advances in experimental psychopathology*. Washington, DC: American Psychological Association Press, 1998:123–157.

10
Astrocytes

Pierre J. Magistretti

Bruce R. Ransom

Pierre J. Magistretti: Institut de Physiologie, University of Lausanne Medical School, Lausanne, Switzerland

Bruce R. Ransom: Department of Neurology, University of Washington Medical School, Seattle, Washington 98195

The astrocyte is a ubiquitous type of glial cell that is defined in part by what it lacks: axons, action potentials, and synaptic potentials. Astrocytes greatly outnumber neurons, often 10:1 and occupy 25% to 50% of brain volume (1 ,2 and 3). Although these cells are anatomically obvious, their functions have been difficult to determine. Discoveries in the last 25 years, however, have revealed some of their functions and established the essential nature of interactions between neurons and astrocytes for normal brain function. We briefly review several basic facts about astrocytes and then selectively survey some of their functions, particularly emphasizing recent findings about metabolic interactions between astrocytes and neurons. We also discuss features of astrocyte function as they relate to synaptic plasticity and emerging concepts in the pathophysiology of psychiatric disorders.

STRUCTURAL AND PHYSIOLOGIC PROPERTIES

Part of "10 - Astrocytes "

The form of astrocytes is important in thinking about their functions. Astrocytes are stellate cells (hence their name) with multiple fine processes. Astrocytes in white matter are complex cells with 50 to 60 long branching processes that radiate from the cell body and terminate in end-feet at the pial surface, on blood vessels, or freely among axons; white matter astrocytes are usually called fibrous astrocytes (4). Astrocytes in gray matter, called protoplasmic astrocytes, have profuse, short stubby processes that contact blood vessels and the pial surface, and surround neurons. Astrocytic end-feet cover the entire surface of intraparenchymal capillaries (5). These end-feet express glucose transporters of the GluT 1 type (6) and are a likely site of glucose uptake. In gray matter, astrocytic processes ensheath virtually every synapse; the ensheathing membranes constitute about 80% of total membrane surface and are devoid of organelles (7). Thus, astrocytes are polarized cells with some processes contacting cells of mesodermal origin (i.e., endothelial cells of the capillary or fibroblasts of the pia mater), whereas other processes are intimately intertwined with neuronal processes and synapses (4 ,7).

Astrocytes are the only cells in the brain that contain the energy storage molecule glycogen (8). The importance of this is discussed elsewhere in this chapter. They also contain distinctive 9-nm intermediate filaments composed of a unique protein called glial fibrillary acidic protein (GFAP). Fibrous astrocytes contain more of these filaments than protoplasmic astrocytes. Recent work has assessed the functional significance of this defining astrocytic protein using genetic knockout experiments (9 ,10 and 11). Astrocytes in GFAP knockout animals have disturbed neuronal plasticity manifest as a loss of long-term depression (10), late-onset dysmyelination (9), and increased susceptibility to ischemia (12). It is not known how GFAP deficiency causes these changes.

Astrocytes are strongly coupled to one another by gap junctions (13), aqueous pores that are permeable to ions and other molecules with a molecular weight less than 1,000. A broad range of biologically important molecules, including nucleotides, sugars, amino acids, small peptides, cAMP, Ca^{2+} and inositol triphosphate (IP_3) have access to this pathway. Such intercellular communication is believed to mediate the coordinated action of adjacent but individual cells in terms of electrical and biochemical activity (13), and equalizes their intracellular ion concentrations (14). Gap junction permeability is strongly reduced by intracellular acidification or large increases in intracellular [Ca^{2+}].

The membrane potential (V_m) of astrocytes is more negative than that of neurons. For example, astrocytes have a V_m of about -85 mV, whereas neuronal membrane potential is about -65 mV. Although glial cells express a variety of K^+ channels, inwardly rectifying K^+ channels seem to be important in setting the resting potential (15). These channels are voltage sensitive and are open at membrane potentials more negative than about -80 mV, close to the observed resting potential of astrocytes. Astrocytes express many other voltage-activated ion channels, previously thought to be restricted to neurons (15). The significance

of voltage-activated Na^+ and Ca^{2+} channels in glial cells is unknown. Because the ratio of Na^+ to K^+ channels is low in adult astrocytes, these cells are not capable of regenerative electrical responses like the action potential.

One consequence of the high K^+ selectivity of astrocytes, compared to neurons, is that the membrane voltage of astrocytes is more sensitive to changes in extracellular $[K^+]$ ($[K^+]_o$). For example, when $[K^+]_o$ is raised from 4 to 20 mM, astrocytes depolarize by ˜25 mV, compared to only ˜5 mV for neurons (16). This relative insensitivity of neuronal resting potential to changes in $[K^+]_o$ in the "physiologic" range may have emerged as an adaptive feature that stabilizes the resting potential of neurons in the face of the transient increases in $[K^+]_o$ that accompany neuronal activity. In contrast, natural stimulation, such as viewing visual targets of different shapes or orientations, can cause depolarizations of up to 10 mV in astrocytes of the visual cortex (17). The accumulation of extracellular K^+ that is secondary to neural activity may serve as a signal to glial cells that is proportional to the extent of the activity. For example, small increases in $[K^+]_o$ cause breakdown of glycogen (18), perhaps providing fuel for nearby active neurons (see later).

Neurons and glial cells do not make functional synaptic or gap junction contacts with one another; therefore, interactions between these cell types must occur via the narrow extracellular space (ECS) between them (16). There may be rare exceptions to this rule (19 ,20). In the mammalian central nervous system (CNS), the ECS is a uniform and very small compartment formed by adjacent cell membranes that are, on average, separated by approximately by 0.02 µm. Brain ECS is a dynamic compartment in terms of its ionic contents and even its dimensions (15 ,21). Because of the extreme narrowness of the ECS, molecules released from one cell diffuse almost instantly to adjacent cells. Glial cells interact with neurons by influencing the contents (e.g., ions, energy metabolites, neurotransmitters, etc.) of the ECS. It should be emphasized that nearly every neuron in the brain shares common ECS with adjacent astrocytes, and astrocytic processes entirely surround synapses. It has been surprising to discover that glial cells release and express receptors for a wide range of informational molecules, including neurotransmitters (22); this greatly expands the possibilities for glial–neuronal interactions. Indeed, astrocytes are in a position to sense and modulate synaptic transmission through the pervasive lamellar processes that surround synaptic contacts (7).

FUNCTIONS

Part of "10 - Astrocytes "

Ion Homeostasis

One of the best-established functions of astrocytes is regulation of brain $[K^+]_o$. Astrocytes are also likely to participate in the regulation of extracellular pH, but this aspect of astrocyte function is still evolving and is not considered further here (23 ,24). Neural activity can rapidly increase $[K^+]_o$, which is tightly regulated to a resting level of about 3 mM (25). A single action potential increases the instantaneous $[K^+]_o$ by ˜0.75 mM (26). The increase in $[K^+]_o$ is proportional to the intensity of neural activity but has a so-called "ceiling" level of accumulation of 10 to 12 mM (27 ,28), which is only exceeded under pathologic conditions (29). If diffusion alone were responsible for dissipating K^+ released from neurons, it is easily calculated that extracellular K^+ accumulation would exceed 10 mM during normal neural activity, whereas measured increases in $[K^+]_o$ are in the range of 1 to 3 mM indicating powerful control mechanisms (30). Homeostatic control of $[K^+]_o$ is needed because brain $[K^+]_o$ can influence transmitter release (31), cerebral blood flow (32), ECS volume (33 ,34), glucose metabolism (35), and neuronal activity (36). Unchecked increases in $[K^+]_o$ act as an unstable positive feedback loop increasing excitability.

Astrocytes expedite the removal of evoked increases in $[K^+]_o$ and limit its accumulation to a maximum level of 10 to 12 mM, the ceiling level seen with intense activity such as epileptic discharge (37 ,38). Neurons, and perhaps blood vessels, also participate in $[K^+]_o$ regulation, but glial mechanisms are probably most important. Two general mechanisms of astrocyte K^+ removal have been proposed (39): 1) net K^+ uptake into astrocytes (by transport mechanisms and/or Donnan forces) and 2) K^+ redistribution through astrocytes, which is known as K^+ spatial buffering. The relative importance of these two mechanisms of $[K^+]_o$ regulation remains an open question and may depend on the nature of the $[K^+]_o$ increase as well as brain region (38).

If glial cells take up K^+ during neural activity and release it thereafter, a transient increase in glial $[K^+]_i$ should result. Astrocyte $[K^+]_i$ does transiently increase during neural activity and has a similar time course to the K^+ lost from active neurons and the increase in $[K^+]_o$, indicating that the K^+ released from neurons is passing by way of the ECS into glial cells (40 ,41 and 42). Uptake of K^+ into glial cells depends on the glial Na^+ pump (38 ,42 ,43 and 44), an anion transporter that cotransports K^+ and Na^+ with Cl^- (43) and Donnan forces that propel KCl into glial cells in the face of elevated $[K^+]_o$ (42) (Fig 10.1). It has not been determined with certainty which of these mechanisms is quantitatively most important for K^+ uptake. The astrocyte Na^+ pump, however, is exquisitely sensitive to elevations of $[K^+]_o$. Even a 1 mM increase in $[K^+]_o$ activates the Na^+ pump in these cells indicating, perhaps, that this is the major mechanism of K^+ sequestration (44). Neurons, of course, must eventually reaccumulate K^+ lost during activity using their Na^+ pump, but only glial cells show net accumulation of K^+ (Fig. 10.1). It is interesting to note that the neuronal Na^+ pump is not sensitive to small increases in $[K^+]_o$ and is probably activated mainly by increases in intracellular $[Na^+]$ (45).

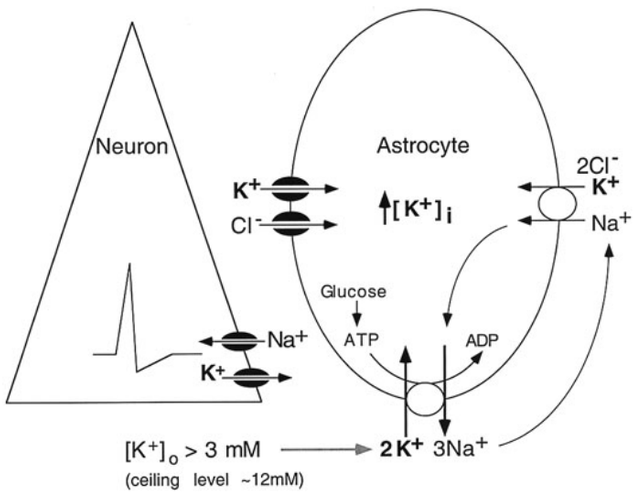


FIGURE 10.1. Schematic representation of mechanisms of K^+ uptake in astrocytes. K^+ released by firing neurons is actively accumulated by astrocytes in three ways. The sodium pump and an anion transporter both take up K^+; the sodium pump relies directly on the availability of ATP, whereas the anion transporter is indirectly powered by the energy stored in the transmembrane Na^+ gradient. The presence of channels for Cl^- and K^+, allow Donnan forces to produce KCl influx. These mechanisms along with K^+ spatial buffering (see text), prevent $[K^+]_o$ from exceeding ˜12 mM. Increases in $[K^+]_i$ are seen during neural activity as $[K^+]_o$ increases.

The idea that focal increases in $[K^+]_o$ could be redistributed by glial cells was introduced by Kuffler and colleagues (46). They realized that the selective K^+ permeability of glia coupled with their low-resistance intercellular connections (mediated by gap junctions), would permit them to transport K^+ from focal areas of high $[K^+]_o$, where a portion of the glial network would be depolarized, to areas of normal $[K^+]_o$, where the glial network would have a near normal membrane potential (46). Experiments suggest that under conditions of focal increases in $[K^+]_o$, five times as much K^+ moves by way of glial cells as through the ECS, except where only very localized K^+ gradients are involved (25). A further specialization that contributes to spatial buffering is a nonuniform distribution of K^+ channels on a single cell. The density of K^+ channels on the cell membrane of retinal Müller cells, which are specialized astrocytes, is highest on the cell's end-foot. Because the end-foot of the Müller cell, which abuts the vitreous humor of the eye, has the highest density of K^+ channels, accumulated $[K^+]_o$ is preferentially transported to the vitreous, which acts as a disposal site. It is not known if nonuniform K^+ channel distribution is a general feature of astrocytes.

Anoxia/ischemia causes rapid increases in $[K^+]_o$ in both gray matter (to ~60 to 80 mM) and white matter (to ~12 to 15 mM *in vitro*) of the brain (27 ,47). The increases in $[K^+]_o$ result because energy-dependent ion gradients can no longer be maintained and K^+ entering the ECS can no longer be taken up by glial cells, which also depend on ATP (48). In fact, under conditions of diminished energy supply, glial cells actually contribute K^+ to the ECS, rather than take it up (49).

Transmitter Synthesis

Glutamate is one of the most common amino acids in the brain, present at millimolar concentrations in brain tissue homogenate. It is also the predominant excitatory neurotransmitter (50). Only a small fraction of total brain glutamate is packaged for synaptic release and astrocytes are intimately involved in the synthesis of this crucial vesicular pool of glutamate.

Although glutamate can be derived from neuron glucose metabolism, carbon-labeling experiments reveal that astrocyte-derived glutamine is the principal precursor of synaptically released glutamate (51 ,52). The synthesis and release of glutamine by astrocytes is part of a biochemical shuttle mechanism called the glutamate-glutamine cycle (53) (Fig 10.2). After release from the presynaptic terminal, glutamate is taken up primarily by astrocytes (54 ,55). In the glial cell, glutamate is converted to glutamine through the ATP-dependent enzyme glutamine synthetase, located exclusively in astrocytes (56). In fact, glutamine synthetase is localized to astrocytic processes surrounding glutamatergic synapses (57).

Glutamine is released by the glial cells and taken up by the neurons through specific uptake carriers. In the presynaptic terminal, glutamine is converted to glutamate through glutaminase, a phosphate-dependent enzyme preferentially localized to synaptosomal mitochondria (58 ,59). The newly synthesized glutamate is then packed into vesicles and becomes available for release. The glutamate-glutamine cycle is a clear and important example of cooperativity between astrocytes and neurons (Fig. 10.2). It mediates removal of potentially toxic excess glutamate from the extracellular space and provides the neuron with a synaptically inert precursor for resynthesis of glutamate. (Glutamine does not bind to neurotransmitter receptors.)

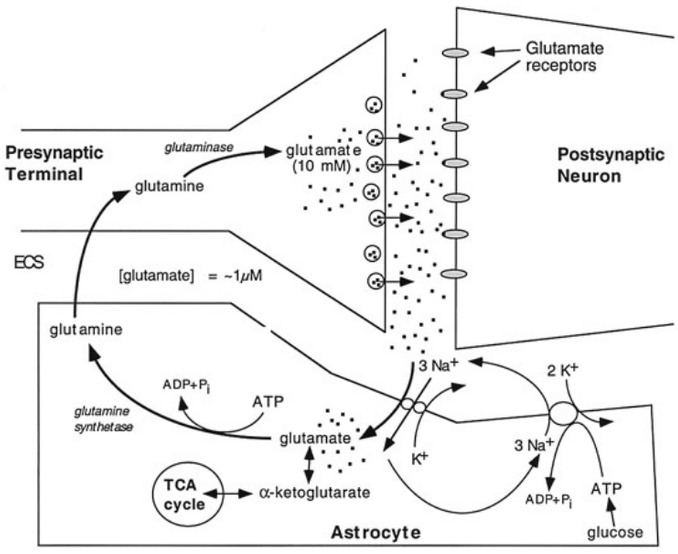


FIGURE 10.2. Scheme showing how astrocytes are involved in glutamate metabolism and uptake. Only astrocytes contain the enzyme glutamine synthetase, which converts glutamate to glutamine in an ATP-requiring reaction. Glutamine is transported to nearby presynaptic terminals where it is converted to glutamate for synaptic release. Finally, the released glutamate is recaptured by astrocytes via a high affinity glutamate uptake system. Although glutamate transporters are present in neurons, astrocytes are the most active in removing glutamate (see text). In the absence of the normal transmembrane Na^+ gradient, maintained by the ATP-dependent Na^+ pump, the glutamate transporter ceases to remove glutamate and can run in reverse so that it pumps glutamate into the extracellular space (ECS).

The cycle is surprisingly rapid. After a 6-min incubation of slices of rabbit hippocampus in [^{14}C] glutamine, half of the radioactivity was in the form of glutamate. Removal of glutamine from the bathing solution of the hippocampal slices decreased glutamate efflux by 60% to 80% after only 6 min (52).

Not all of the glutamate taken up by astrocytes is directly converted to glutamine. Glutamate can also enter the TCA cycle through its conversion to α-ketoglutarate (KG). Three enzymatic reactions can yield KG: one catalyzed by aspartate amino transferase and another by alanine aminotransferase, both reactions involving the transfer of an α-amino group. The third reaction is the direct conversion of glutamate to KG via the action of glutamate dehydrogenase (60) (Fig. 10.2). Theoretically, therefore, neurons *might not* get back in the form of glutamine (from astrocytes) all of the glutamate that they release for two reasons: (a) some of the glutamate diffuses away or is taken up by postsynaptic neurons, or (b) not all of the glutamate that enters astrocytes becomes glutamine. Two possibilities can be considered for stabilizing the pool of vesicular glutamate in neurons. First, contrary to the preceding premise, astrocytes might be able to compensate neurons for their loss of glutamate by appropriate adjustments in glutamine export. This would be possible because the pool of cytosolic glutamate in astrocytes is in equilibrium with TCA cycle intermediates, which in turn can be replenished by the carboxylation of pyruvate derived from glucose. The signal for more glutamine export could be extracellular (glutamine), which would fluctuate in response to the needs of glutamatergic neurons. Second, recent data have demonstrated that neurons can generate glutamate directly from pyruvate obtained from glucose or lactate (61 ,62). Indeed, lactate produced by astrocytes in response to synaptically released glutamate (see the following) appears to be taken up by neurons and could be a substrate for glutamate formation.

GABA, the most common inhibitory neurotransmitter in the brain, is synthesized from glutamate. Consequently, glutamine and α-ketoglutarate are used for GABA synthesis as well (63). Depolarization-released GABA is preferentially synthesized from glutamine supplied by astrocytes (64). Inhibition of astrocyte glutamine synthetase with methionine sulfoximine produces a significant decrease in GABA production both *in vivo* and in brain slices (65%); however, because a 90% decrease in glutamine did not fully suppress GABA synthesis, an additional metabolic source is considered to be likely (65).

Astrocytes, it would seem, are essential for normal glutamate- and GABA-mediated synaptic transmission. Indeed, selective inhibition of glial cells in the guinea pig hippocampus using the glial-specific metabolic blocker, fluoroacetate, decreases transmission at glutamate synapses (66). Intracellular recordings verified the integrity of neurons in fluoroacetate-treated slices and the persistence of normal responses to glutamate applied iontophoretically. A modulatory role of astrocytes in excitatory synaptic transmission is supported by this study.

Transmitter Removal

In addition to being the most important excitatory neurotransmitter in the brain, glutamate is also a potent neurotoxin and has been implicated in stroke, amyotrophic lateral sclerosis, and epilepsy. Highly efficient glutamate transporters remove synaptically released glutamate and also keep the extracellular concentration of this amino acid at about 2 μM (67). Glutamate transporters are expressed in oligodendrocytes, neurons, microglia, and astrocytes, but transporters in astrocytes are quantitatively the most important in regulating glutamate at synapses and in the extracellular space (Fig. 10.2). (See ref. 55 for review.)

Five main types of glutamate transporters have been described: GLAST (EAAT 1), GLT-1 (EAAT 2), EAAC 1 (EAAT 3), EAAT 4, and EAAT 5 (55). The latter two appear to be predominantly localized in cerebellum and retina, respectively. All have been cloned, functionally characterized, and their localization and distribution at the regional, cellular, and subcellular levels in the CNS are known (55 ,68). A detailed review of this fertile field of research is beyond the scope of this chapter. EAAC 1 transporters are neuronal, mostly localized on the cell body and dendrites, whereas GLAST and GLT-1 are predominantly glial (68). There are regional differences in the expression of GLAST and GLT-1; GLAST is more heavily expressed in the cerebellum and GLT-1 is more prevalent in the forebrain.

Glutamate uptake into astrocytes is driven by the electrochemical gradients of Na^+ and K^+, with a stoichiometry of 3 Na^+ and 1 H^+ in and 1 K^+ out with the uptake of each glutamate anion (55 ,69). The resulting increase in $[Na^+]_i$ must be corrected by a cycle of the Na^+ pump and ATP consumption (70). One advantage of a system where astrocytes take up most of the glutamate is that the metabolic burden of this work is offloaded from neurons (55).

Several lines of evidence support that astrocytes play an essential role in glutamate uptake in the brain (55): (a) Astrocytes preferentially accumulate glutamate transporter

substrate (71). (b) In the absence of astrocytes, neurons are 100-fold more vulnerable to glutamate toxicity (72). (c) Genetic down-regulation of GLAST or GLT-1, but not the neuronal subtype EAAC1, causes elevated extracellular levels of glutamate and neurotoxicity (73).

As emphasized, astrocytes form a ubiquitous part of all glutamatergic synapses. Astrocyte membrane facing a glutamate synapse expresses higher levels of GLAST than membrane facing other structures such as pia mater or capillaries (74). Most of the glutamate released at synapses appears to be taken up by the adjacent astrocytes (75), although there may be some regions in the brain where up to 20% of glutamate is transported into the postsynaptic neuron (55). The impact of astrocytic glutamate uptake at synapses is most emphatically detected when uptake is blocked. This increases both the amplitude and the duration of the excitatory postsynaptic current (76).

Metabolic Coupling with Neurons

Coupling of Synaptic Activity to Glucose Utilization

The cytoarchitecture of astrocytes is of particular relevance in a discussion about the coupling of synaptic activity to glucose utilization. As mentioned, astrocyte end-feet surround virtually all brain capillaries, whereas other astrocytic processes ensheath synaptic contacts (5 ,7) or end among axons (4). In addition, astrocytes possess receptors and reuptake sites for a variety of neurotransmitters, including the excitatory neurotransmitter glutamate (77), whereas astrocytic end-feet are enriched in the specific glucose transporter GLUT-1 (6). Thus, astrocytes possess the necessary features to sense synaptic activity, through receptors and reuptake sites for neurotransmitters, and to couple it with the entry of glucose into the brain parenchyma (78). Experimental evidence supporting this function is reviewed here.

Glutamate, the main excitatory neurotransmitter released by activated circuits, is a potent stimulator of glycolysis; that is, of glucose uptake and lactate production, in primary astrocyte cultures. (See refs. 60 and 79 for review.) The metabolic effect of glutamate is not mediated by receptors, but rather by glutamate transporters selectively expressed in astrocytes, in particular GLAST (80). These observations suggest a mechanism whereby astrocytes contribute to the uptake of glucose from the circulation into brain parenchyma in register with synaptic activity: The release of glutamate from synaptically active neurons stimulates glucose uptake in nearby astrocytes. The extent of glucose uptake would be proportional to the extent of activity, thereby "coupling" neuronal activity to glucose utilization (80). The Na-K-ATPase is critical for this coupling. Thus, ouabain, a specific inhibitor of the Na-K-ATPase, completely inhibits the glutamate-evoked glycolysis in astrocytes (80). Glutamate stimulates astrocytic Na-K-ATPase (81) by increasing intracellular Na^+ concentration ($[Na^+]_i$) via Na^+-dependent glutamate uptake by glutamate transporters (55 ,69) (Fig. 10.2 and Fig. 10.3). The overall stoichiometry of the molecular steps involved in the coupling between glutamate uptake and glucose utilization are as follows: one glutamate is taken up with 3 Na^+, whereas one glucose consumed through glycolysis produces two ATPs. One ATP is used by the Na^+/K^+ ATPase for the extrusion of 3 Na^+; the other ATP could be used for the synthesis of glutamine from glutamate by glutamine synthase (Fig. 10.3) (82). The glutamate-stimulated glycolytic processing of glucose results in approximately two lactate molecules produced per one glucose molecule; that is, an expected stoichiometric relationship between glucose and lactate. This scheme is, of course, primarily relevant for synaptic regions of the brain (i.e., gray matter). The mechanism that "couples" axonal activity to glucose utilization in white matter is not established. Because axonal activity produces proportional increases in $[K^+]_o$ (28) and increases in $[K^+]_o$ increase astrocyte glucose utilization (35 ,83), it is tempting to speculate that activity-dependent changes in $[K^+]_o$ in white matter play an analogous role to glutamate release in gray matter.

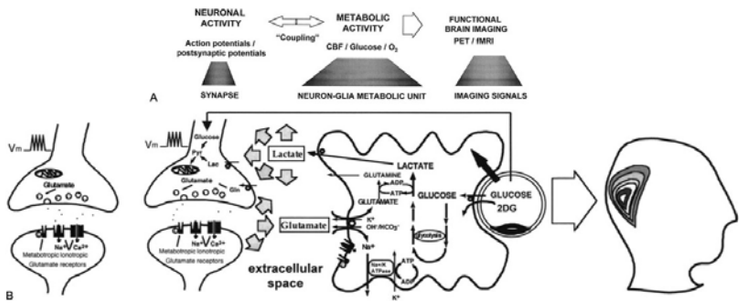


FIGURE 10.3. Schematic representation of the mechanism for glutamate-induced glycolysis in astrocytes during physiologic activation and its relevance for functional brain imaging. At glutamatergic synapses, presynaptically released glutamate depolarizes postsynaptic neurons by acting at specific receptor subtypes. The action of glutamate is terminated by an efficient glutamate uptake system located primarily in astrocytes. Glutamate is cotransported with Na^+, resulting in an increase in the intra-astrocytic concentration of Na^+, leading to an activation of the astrocyte Na^+/K^+ ATPase. Activation of the Na^+/K^+ ATPase stimulates glycolysis (i.e., glucose utilization and lactate production). Lactate, once released by astrocytes, can be taken up by neurons and serve them as an adequate energy supply. The proposed model of glutamate-induced glycolysis in astrocytes implies that the activity-linked uptake of 18FDG monitored with PET, reflects primarily an astrocyte-based signal. Because neuronally released glutamate triggers the cascade of events that leads to glucose uptake, the 18FDG-PET signal faithfully reflects activation of neuronal circuits. The model is also consistent with the transient lactate peak detected during activation by MRS in humans and by microdialysis and electrochemical detection in animals (see text for details).

Is the lactate released by astrocytes in this model used as fuel by neurons? A vast array of experimental data indicate that *in vitro*, lactate in the absence of glucose can adequately maintain synaptic (84 ,85 and 86) or axonal activity (87 ,88). If astrocytes transfer lactate to neurons as a fuel source (Fig. 10.3), several conditions must be met. There must be appropriate enzymes for the creation of lactate in astrocytes and its use in neurons, and appropriate transport mechanisms for the movement of lactate. Indeed one isoform of lactate dehydrogenase (LDH5), which is enriched in lactate-producing glycolytic tissues such as skeletal muscle, is predominantly localized in astrocytes in the human brain, whereas the other isoform, LDH1, expressed in highly oxidative tissues such the heart, which utilize lactate, is mostly found in neurons (89). Monocarboxylate transporters (MCTs) mediate the exchange of lactate between astrocytes and neurons. These transporters show a cell-specific distribution, with MCT1 predominantly present in astrocytes, whereas MCT2 is enriched in neurons (90 ,91).

Glutamate-mediated neuron–glia metabolic interactions provide an initial basis to better understand the cellular and molecular steps involved in neuro metabolic coupling (Fig. 10.3). In particular, the model proposed according to which the activity-dependent synaptic release of glutamate triggers glucose uptake into the brain parenchyma coupled with a transient production of lactate, provides a possible basis for functional brain imaging techniques such as positron emission tomography (PET) (78 ,82). Indeed, a host of PET human studies have indicated an activity-dependent partial uncoupling between glucose utilization (^{18}F-deoxyglucose PET) and oxygen consumption (^{15}O PET), whereas magnetic resonance spectroscopy (MRS) analyses show a transient

lactate production. (See ref. 79 for review.) In addition, recent MRS studies provide strong support for a tight coupling between glutamate-mediated synaptic activity and glucose utilization. Thus, the simultaneous measurements, over a range of synaptic activity, of glucose oxidation and the cycling of glutamate to glutamine (a process that occurs exclusively in astrocytes) using ^{15}C MRS has revealed a striking stoichiometric relationship of 1:1 between glutamate cycling (a reflection of synaptic activity) and glucose utilization (92). According to these data, for each glutamate released from active terminals and taken up by astrocytes one glucose would be oxidized.

Implications for Imaging

The model proposed on the basis of studies at the cellular level suggests an initial glycolytic processing of glucose occurring in astrocytes during neuronal activation, resulting in a transient lactate overproduction, followed by a recoupling phase during which lactate would be oxidized by neurons (Fig. 10.3).

Results obtained in a variety of *in vivo* paradigms both in laboratory animals and humans, support the existence of such a transient lactate production during activation. Thus, microdialysis studies in rats indicate a marked increase in the concentration of lactate in the dialysate in striatum and hippocampus during physiologic sensory stimulation (93). Interestingly, this activity-linked lactate peak is completely inhibited when the glutamate uptake inhibitor THA is present in the perfusate, thus providing further supporting evidence for the existence of glutamate stimulated glycolysis during activation (93). In humans, MRS reveals a transient lactate peak in primary visual cortex during physiologic activation of the visual system (94). Thus, microdialysis and MRS data *in vivo* support the notion of a transient glycolytic processing of glucose during activation. In addition, some PET studies have shown that oxygen consumption does not increase commensurately with blood flow and glucose utilization in activated brain areas (95), suggesting the occurrence of an activity-dependent glycolytic processing of glucose. In contrast, using ^{13}C-glucose MRS, recent data are consistent with a significant increase in oxygen utilization during activation (96). These contrasting views relevant to the degree of oxygen utilization during activation can be reconciled by the model proposed (80), which suggests that glucose imported into the brain parenchyma during activation undergoes a *transient* glycolysis in astrocytes, resulting in the production of lactate that is then oxidized by neurons. This latter process would imply a metabolic recoupling with increased oxygen consumption. The spatial and temporal “window” during which a transient glycolysis occurs and a lactate peak can be detected, may depend on the rapidity and degree of recoupling existing between astrocytic glycolysis and neuronal oxidative phosphorylation.

This operational model for coupling is consistent with the notion that the signals detected during physiologic activation in humans with FDG PET may result from signaling and metabolic exchanges between neurons (glutamate release) and astrocytes (glycolysis) (82). This conclusion does not question the validity of these techniques for monitoring neuronal function because the triggering event is neuronal glutamate release; rather, this conclusion provides a cellular and molecular basis for these functional brain-imaging techniques (Fig. 10.3).

Glycogen

Another facet of neuron–glia metabolic interactions concerns glycogen. Glycogen, the storage form of glucose, is the largest energy reserve of the brain and it is almost exclusively localized in astrocytes. Although the levels of glycogen are low compared to muscle or liver, they appear to vary in register with synaptic activity and are tightly regulated by a variety of neurotransmitters (97). For example, somatosensory stimulation readily mobilizes glycogen in the corresponding somatosensory cortex as well as in subcortical relays (98). In contrast, during anesthesia glycogen levels increase dramatically. Plastic adaptations in glycogen regulation appear to occur in astrocytes as a consequence of acute or slow neuronal loss; indeed, glycogen deposits are often observed in reactive astrocytes present in acutely lesioned areas, as well as at sites of slow neurodegeneration, such as those observed in Alzheimer’s disease (97). These latter observations suggest that impaired synaptic activity associated with neuronal loss results in a glycogen-sparing situation in astrocytes; active mechanisms driving glycogen metabolism toward increased resynthesis may also be operative under such conditions. The mechanisms that underlie the regulation of glycogen metabolism at the cellular and molecular levels have been partially characterized. (See ref. 97 for review.) Whereas the role of brain glycogen remains to be fully elucidated, recent results indicate that astrocytic glycogen in white matter is readily available to axons, mainly as lactate, and sustains their function during glucose withdrawal (99). It also seems likely that glycogen mobilization by certain neurotransmitters is an adaptation designed to provide additional energy substrate to synaptically active neurons (see the following).

It is appealing to think that astrocytic glycogen serves to provide fuel to the brain when glucose is in short supply. Indeed, astrocytic glycogen *in vitro* is rapidly degraded when glucose is withdrawn (100), and glycogen falls rapidly *in vivo* during ischemia, with a time course that is closely related to (87) ATP depletion and accumulation of lactate (101). Some experimental evidence supports this concept. Neurons grown in astrocyte-rich cultures are less severely injured by glucose withdrawal than are neurons in astrocyte-poor cultures (102). This benefit appeared to derive from the presence of greater amounts of glycogen in the astrocyte-rich cultures because depleting glycogen negated the benefit (102).

Two possible mechanisms for this outcome were suggested but not tested: (a) Astrocytes themselves utilize the energy from glycogen breakdown to prevent the accumulation of toxic levels of glutamate (removing it by a sodium-gradient-dependent transporter); or (b) Glycogen provides fuel to neurons to sustain their energy metabolism.

The role of astrocytic glycogen in CNS white matter has been analyzed using the rat optic nerve, a representative white matter tract. Optic nerve function, measured as the compound action potential (CAP), persists for about 40 min in the absence of glucose (87 ,103), suggesting the presence of an intrinsic energy reserve such as astrocytic glycogen. The theory was tested that during glucose withdrawal, astrocytic glycogen is converted to lactate and transported into axons to act as an energy source. Glycogen content of the optic nerve falls during glucose withdrawal, compatible with rapid use in the absence of glucose. Up-regulation of glycogen content increases latency to CAP failure and improves CAP recovery after glucose withdrawal, whereas down-regulation of glycogen content has the opposite effects (99). Lactate can replace glucose as an energy source and appears to be transported by MCT out of astrocytes and into axons in the absence of glucose. These results indicate that astrocytic glycogen in white matter is readily available to axons, mainly as lactate, during glucose withdrawal (Fig. 10.4). Conceptually, they provide “proof of principle” that astrocytic glycogen is an energy reserve that can be shared with, and thus can prolong the function of, neural elements in times of need.

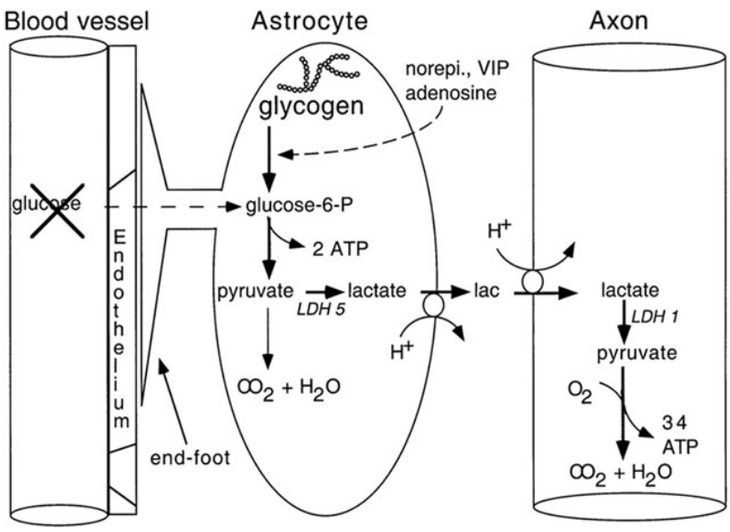


FIGURE 10.4. Schematic illustration of how astrocytic glycogen appears to fuel axons in the absence of glucose. Blood glucose first encounters astrocytic end-feet as it is transported into the brain. In the absence of glucose, astrocytic glycogen is broken down to lactate, which is transported to the extracellular space via a monocarboxylate transporter (MCT). It is then taken up by an MCT in axons and is oxidatively metabolized to produce energy needed to sustain excitability. LDH5 preferentially reduces pyruvate to lactate, whereas LDH1 preferentially oxidizes lactate to pyruvate. Neurotransmitters such as norepinephrine, vasoactive intestinal peptide, and adenosine, promote glycogenolysis. This scheme recognizes that astrocytes can subsist, at least transiently, on glycolytic energy metabolism, whereas axons require oxidative metabolism.

Glycogen levels in astrocytes are tightly regulated by various neurotransmitters. Several monoamine neurotransmitters, namely noradrenaline, serotonin, and histamine, are glycogenolytic in the brain, in addition to certain peptides such as vasoactive intestinal peptide (VIP) and pituitary adenylate cyclase activating peptide (PACAP), as well as adenosine and ATP (104 ,105). The effects of all these neurotransmitters are mediated by specific receptors coupled to second messenger pathways such as adenylate cyclase, for the ß-adrenergic, VIP/PACAP and adenosine A2 receptors, or phospholipase C for α-1 adrenergic receptors (106). The initial rate of glycogenolysis activated by VIP and noradrenaline is between 5 and 10 nmol/mg prot/min (106), a value that is remarkably close to glucose utilization of the gray matter as determined by the 2-DG autoradiographic method (107). This correlation raises the possibility that the glycosyl units mobilized in response to the glycogenolytic neurotransmitters are in register with the energy demands of the neuropil.

In addition to the rapid (within seconds) glycogenolysis, VIP, noradrenaline, and adenosine induce a long-lasting plastic response resulting in massive glycogen resynthesis (108). This effect is expressed after several hours and is transcriptionally regulated, involving the expression of new genes: two immediate-early and two late genes, all being induced in a cAMP-dependent manner in astrocytes. The immediate-early genes, C/EBP ß and δ are members of a family of transcription factors called CCAAT/enhancer binding protein. This family of transcription factors is predominantly involved in two types physiologic responses: inflammation, through the control of expression of several acute phase response genes, and energy metabolism, in particular through the regulation of cAMP-sensitive genes controlling glucose metabolism in peripheral tissues (109). The late genes regulated by VIP and NA are glycogen synthase (110) and protein targeting to glycogen (PTG) (111). The physiologic function of PTG appears to be that of a chaperone protein coordinating the activity and compartmentalization of glycogen-synthesizing enzymes (111).

Glycogen metabolism in astrocytes appears to be under the dynamic control of at least two neurotransmitters, VIP and NA, with the balance between short-term (glycogenolysis) and transcriptionally regulated long-term effects (glycogen resynthesis) setting the intracellular levels of glycogen.

Calcium Waves and Glial Signaling

The phenomenon of inducible waves of increased $[Ca^{2+}]_i$ moving through adjacent astrocytes was first reported in tissue culture by Cornell-Bell and associates (112). Such waves, stimulated by glutamate application or a mechanical

stimulus, are also known to occur in other cell types (113). It is of special interest that astrocyte Ca^{2+} waves may be elicited by activity in adjacent neurons (114). The mechanism of these waves, which move through cells at a rate of 10 to 20 μm/sec, is not entirely understood (115). It appears to involve the intracellular formation and intercellular transmission of inositol-1,4,5-trisphosphate (IP_3), and the release of an extracellular messenger substance, perhaps ATP (115). This phenomenon has become more interesting with the discovery that it is seen in intact brain tissue such as the retina (116).

The function of intracellular Ca^{2+} waves in astrocytes could be to coordinate the activity of these glial cells. A more intriguing possibility would be that the Ca^{2+} wave could influence neurons in its vicinity. Indeed, Ca^{2+} elevation in astrocytes can cause increases in neuronal $[Ca^{2+}]_i$ (117) and induce action potentials in hippocampal neurons (118 ,119). The most compelling demonstration to date of glia-to-neuron signaling mediated by Ca^{2+} waves, however, has come from studies in the rat retina (116). Mechanical stimulation of astrocyte Ca^{2+} waves led to changes in light-induced ganglion cell firing, usually inhibition, when the Ca^{2+} wave reached the neuron. The likely mechanism was the release of glutamate from the astrocytes with stimulation of inhibitory interneurons projecting to the ganglion cells. The importance of this fascinating observation for the normal operation of the nervous system remains unclear, but it suggests a unique form of glial modulation of neuronal activity.

Astrocytes and Synaptic Plasticity

Considerable attention has been given in recent years to the mechanisms of synaptic plasticity. It is now clear that plastic events outlast the early stages of nervous system development and are maintained, to different degrees, throughout adult life, providing the basis for the processes of learning and memory (120). Most experimental paradigms both *in vivo* and *in vitro*, such as the study of long-term potentiation in brain slices, have thoroughly explored the mechanisms of plasticity occurring as a consequence of intense activation of neuronal circuits purportedly associated with the learning process (12). For example, structural rearrangements at synapses characterized by an enhanced number of axodendritic synaptic contacts have been demonstrated (121), and the relative contribution of presynaptic adaptations (increased neurotransmitter release) versus postsynaptic ones (increased responsivity of the target elements) have been thoroughly investigated (122). The involvement of a variety of molecules, including glutamate and its receptors, nitric oxide, arachidonic acid, certain neurotrophins, and cell adhesion molecules, has been proposed (123).

However, few studies have explored the possible adaptations that may occur in glial elements, in particular at the astrocytic profiles that ensheath synapses, as a consequence of learning paradigms known to affect synaptic plasticity. The structural rearrangements occurring at synaptic contacts are likely to affect the morphology of the associated astrocytic profiles. Such an astrocytic structural plasticity has been well documented in the hypothalamus. Here, on physiologic stimulation (lactation, dehydration) the astrocytic profiles surrounding the soma and dendrites of oxytocin-containing neurons retract, allowing a marked increase in membrane surface available for synaptic contacts (124). This structural modification, in which a clear role for certain cell-adhesion molecules (e.g., PSA-NCAM, F3, and tenascin) has been shown (125), is reversible, being associated with the period of stimulation. A similar structural astrocytic plasticity has been shown in the arcuate nucleus during the estrous cycle (124). Evidence for activity-dependent astrocytic plasticity is beginning to be demonstrated also in extrahypothalamic areas of the brain. Striking structural modifications of astrocyte morphology surrounding synaptic contacts occur in parallel to neuronal plastic adaptations in brain areas activated by simple behavioral paradigms of learning. Thus, in animals reared in complex environments, the size and the number of GFAP-immunoreactive astrocytes was found to be increased in the visual cortex as compared to animals raised in normal laboratory cage environment (126 ,127 and 128). In addition, the extent of contact between astrocytic processes and synapses is increased under the same conditions (129). Similar astrocytic modifications are also observed in the cerebellum following synaptic plasticity induced by motor-skill learning (130). Following a spatial learning task (the Morris water maze test), an increase in the density of GFAP-immunoreactive astrocytes is observed in the hippocampus (131). This increase in GFAP is correlated with enhanced expression of basic-fibroblast growth factor. Considering the purported role of LTP in spatial learning (132), it is worth noting that induction of LTP *in vivo* by repeated high-frequency stimulation of the perforant pathway in the hippocampus causes a similar increase in numerical density, higher surface density, and closer apposition of astrocyte processes to the synaptic clefts, or dendritic spines in the potentiated synapses (133). The concept that glial cells, and in particular astrocytes, could contribute to plastic changes occurring after learning and memory or developmental paradigms has already been the subject of reviews (134).

In some of these learning paradigms involving structural plastic responses in astrocytes, long-lasting adaptations in local energy metabolism have been reported. Thus, increases in capillary formation, capillary branching, and surface area were reported in visual cortex following complex experience (127). Angiogenesis was also demonstrated in the cerebellum after motor skill learning (135), thus suggesting that a long-lasting enhancement in the supply of energy substrates (mostly glucose and O_2) is required in the activated area. Several studies have reported long-lasting changes in glucose utilization following various learning and memory tasks (136 ,137 ,138 and 139).

Spatial discrimination training was reported to cause persistent increases in glucose utilization, as measured with the 2-deoxyglucose autoradiographic technique in regions such as the hippocampus and various cortical areas (140).

GLIAL CELL FUNCTION: IMPLICATIONS FOR PSYCHIATRY

Part of "10 - Astrocytes "

From the overview presented in the preceding paragraphs, it clearly appears that all facets of intercellular communication in the brain are relevant to glial cell, and specifically astrocyte function. Thus, the presence on astrocytes of receptors for all chemical classes of neurotransmitters (amino acids, monoamines, and peptides) indicates that the receptor-mediated effects of psychoactive drugs can involve astrocytes. To take a simple example, increasing synaptic levels of noradrenaline with certain classes of antidepressants, by activating ß-adrenergic receptors on astrocytes, is likely to result in induction of genes that regulate glycogen metabolism. Plasticity is no longer exclusive to neurons, providing the conceptual background for considering the activity- and drug-induced adaptations observed in the function of specific circuits as also involving astrocyte-based mechanisms. As recently reviewed, alterations in excitatory neurotransmission, which appear to be involved in neuropsychiatric disorders such as schizophrenia or Alzheimer's disease, may well include an astroglial component (141). Recent neuroimaging studies have shown that the volume of the subgenual prefrontal cortex is reduced in familial forms of major depressive and in bipolar disorders. This decrease in cortical volume could be ascribed to a marked decrease in the number of glial cells (142). Functional imaging in patients suffering from both clinical conditions show a significant decrease in glucose utilization determined by ^{18}F-deoxyglucose PET (143). In addition to providing indirect evidence for the involvement of glia in certain psychiatric disorders, this set of observations is consistent with the proposed role of astrocytes in the ^{18}F-deoxyglucose PET signal (80). Chronic treatment with neuroleptics has been associated with increased glial density in the primate prefrontal cortex (144). Although these examples still need validation in terms of their relevance in the pathophysiology of certain neuropsychiatric disorders, they provide, along with advances in the understanding of glial cell biology, sufficient grounds for opening the scope of psychiatric neuroscience to the study of what Joseph Coyle and Robert Schwarcz have recently called "mind glue" (the term glia, meaning glue in German) (141). This is a major step indeed from "neuron-centric" dominated psychiatric theory and one that is justified by the available facts.

REFERENCES

1. O'Kusky J, Colonnier M. A laminar analysis of the number of neurons, glia and synapses in the visual cortex (area 17) of the adult macaque monkey. *J Comp Neurol* 1982;210:278–290.

2. Kimelberg HK, Norenberg MD. Astrocytes. *Sci Am* 1989;260:44–52.

3. Bignami A. Glial cells in the central nervous system. *Discussions in neuroscience*. Amsterdam: Elsevier, 1991:1–45.

4. Butt AM, Ransom BR. Visualization of oligodendrocytes and astrocytes in the intact rat optic nerve by intracellular injection of Lucifer yellow and horseradish peroxidase. *Glia* 1989;2:470–475.

5. Peters A, Palay SL, Webster HD. *The fine structure of the nervous system*. New York, Oxford Press, 1991.

6. Morgello S, Uson RR, Schwartz EJ, et al. The human blood-brain barrier glucose transporter (GLUT1) is a glucose transporter of gray matter astrocytes. *Glia* 1995;14:43–54.

7. Rohlmann A, Wolff JR. Subcellular topography and plasticity of gap junction distribution on astrocytes. In: *Gap junctions in the nervous system*. RG Landes Company, 1996:175–192.

8. Cataldo AM, Broadwell RD. Cytochemical identification of cerebral glycogen and glucose-6-phosphatase activity under normal and experimental conditions. II. Choroid plexus and ependymal epithelia, endothelia and pericytes. *J Neurocytol* 1986;15:511–524.

9. Liedtke W, Edelmann W, Bieri PL, et al. GFAP is necessary for the integrity of CNS white matter architecture and long-term maintenance of myelination. *Neuron* 1996;17:607–615.

10. McCall MA, Gregg RG, Behringer RR, et al. Targeted deletion in astrocyte intermediate filament (Gfap) alters neuronal physiology. *Proc Natl Acad Sci USA* 1996;93:6361–6366.

11. Gimenez YRM, Langa F, Menet V, et al. Comparative anatomy of the cerebellar cortex in mice lacking vimentin, GFAP, and both vimentin and GFAP. *Glia* 2000;31:69–83.

12. Nawashiro H, Brenner M, Fukui S, et al. High susceptibility to cerebral ischemia in GFAP-null mice. *J Cereb Blood Flow Metab* 2000;20:1040–1044.

13. Ransom BR. Gap junctions. In: *Neuroglia*. New York: Oxford University Press, 1995:299–318.

14. Rose CR, Ransom BR. Regulation of intracellular sodium in cultured rat hippocampal neurones. *J Physiol (Lond)* 1997a;499:573–587.

15. Ransom BR, Sontheimer H. The neurophysiology of glial cells. *J Clin Neurophysiol* 1992;9:224–251.

16. Kuffler SW, Nicholls JG. The physiology of neuroglial cells. *Ergeb Physiol* 1966;57:1–90.

17. Kelly JP, Van Essen DC. Cell structure and function in the visual cortex of the cat. *J Physiol (Lond)* 1974;238:515–547.

18. Hof PR, Pascale E, Magistretti PJ. K+ at concentrations reached in the extracellular space during neuronal activity promotes a Ca2+-dependent glycogen hydrolysis in mouse cerebral cortex. *J Neurosci* 1988;8:1922–1928.

19. Mudrick-Donnon LA, Williams PJ, Pittman QJ, et al. Postsynaptic potentials mediated by GABA and dopamine evoked in stellate glial cells of the pituitary pars intermedia. *J Neurosci* 1993;13:4660–4668.

20. Alvarez-Maubecin V, Garcia-Hernandez F, Williams JT, et al. Functional coupling between neurons and glia. *J Neurosci* 2000;20:4091–4098.

21. Nicholson C. Extracellular space as the pathway for neuron-glial cell interaction. In: *Neuroglia*. New York: Oxford University Press, 1995:387–397.

22. Kettenmann H, Ransom BR. *Neuroglia*. New York, Oxford University Press, 1995.

23. Ransom BR. Glial modulation of neural excitability mediated by extracellular pH: a hypothesis. *Prog Brain Res* 1992;94:37–46.

24. Kaila K, Ransom BR. *pH and brain function*. New York: Wiley-Liss, 1998.

25. Ransom BR, Carlini WG. Electrophysiological properties of astrocytes. In: *Astrocytes: biochemistry, physiology and pharmacology of astrocytes*. Orlando, FL: Academic Press, 1986:1–49.

26. Adelman WJ Jr, Fitzhugh R. Solutions of the Hodgkin-Huxley equations modified for potassium accumulation in a periaxonal space. *Fed Proc* 1975;34:1322–1329.

27. Heinemann U, Lux HD. Ceiling of stimulus induced rises in extracellular potassium concentration in the cerebral cortex of cat. *Brain Res* 1977;120:231–249.

28. Connors BW, Ransom BR, Kunis DM, et al. Activity-dependent K+ accumulation in the developing rat optic nerve. *Science* 1982;216:1341–1343.

29. Hansen AJ. Effect of anoxia on ion distribution in the brain. *Physiol Rev* 1985;65:101–148.

30. Somjen GG. Extracellular potassium in the mammalian central nervous system. *Annu Rev Physiol* 1979;41:159–177.

31. Balestrino M, Aitken PG, Somjen GG. The effects of moderate changes of extracellular K+ and Ca2+ on synaptic and neural function in the CA1 region of the hippocampal slice. *Brain Res* 1986;377:229–239.

32. Kontos HA. Regulation of the cerebral circulation. *Annu Rev Physiol* 1981;43:397–407.

33. Dietzel I, Heinemann U, Hofmeier G, et al. Transient changes in the size of the extracellular space in the sensorimotor cortex of cats in relation to stimulus-induced changes in potassium concentration. *Exp Brain Res* 1980;40:432–439.

34. Ransom BR, Yamate CL, Connors BW. Activity-dependent shrinkage of extracellular space in rat optic nerve: a developmental study. *J Neurosci* 1985;5:532–535.

35. Salem RD, Hammerschlag R, Brancho H, et al. Influence of potassium ions on accumulation and metabolism of (^{14}C)glucose by glial cells. *Brain Res* 1975;86:499–503.

36. Baylor DA, Nicholls JG. Changes in extracellular potassium concentration produced by neuronal activity in the central nervous system of the leech. *J Physiol (Lond)* 1969;203:555–569.

37. Newman EA. Glial cell regulation of extracellular potassium. *Neuroglia*. New York: Oxford University Press, 1995:717–731.

38. Ransom CB, Ransom BR, Sontheimer H. Activity-dependent extracellular K+ accumulation in rat optic nerve: the role of glial and axonal Na+ pumps. *J Physiol* 2000;522.3:427–442.

39. Orkand RK. Glial-interstitial fluid exchange. *Ann NY Acad Sci* 1986;481:269–272.

40. Kettenmann H, Sonnhof U, Schachner M. Exclusive potassium dependence of the membrane potential in cultured mouse oligodendrocytes. *J Neurosci* 1983;3:500–505.

41. Coles JA, Orkand RK, Yamate CL, et al. Free concentrations of Na, K, and Cl in the retina of the honeybee drone: stimulus-induced redistribution and homeostasis. *Ann NY Acad Sci* 1986;481:303–317.

42. Ballanyi K, Grafe P, ten Bruggencate G. Ion activities and potassium uptake mechanisms of glial cells in guinea-pig olfactory cortex slices. *J Physiol (Lond)* 1987;382:159–174.

43. Walz W, Hinks EC. Carrier-mediated KCl accumulation accompanied by water movements is involved in the control of physiological K+ levels by astrocytes. *Brain Res* 1985;343:44–51.

44. Rose CR, Ransom BR. Intracellular sodium homeostasis in rat hippocampal astrocytes. *J Physiol (Lond)* 1996;491:291–305.

45. Rose CR, Ransom BR. Gap junctions equalize intracellular Na+ concentration in astrocytes. *Glia* 1997b;20:299–307.

46. Orkand RK, Nicholls JG, Kuffler SW. Effect of nerve impulses on the membrane potential of glial cells in the central nervous system of amphibia. *J Neurophysiol* 1966;29:788–806.

47. Ransom BR, Walz W, Davis PK, et al. Anoxia-induced changes in extracellular K+ and pH in mammalian central white matter. *J Cereb Blood Flow Metab* 1992;12:593–602.

48. Rose CR, Waxman SG, Ransom BR. Effects of glucose deprivation, chemical hypoxia, and simulated ischemia on Na+ homeostasis in rat spinal cord astrocytes. *J Neurosci* 1998;18:3554–3362.

49. Ransom BR, Philbin DM Jr. Anoxia-induced extracellular ionic changes in CNS white matter: the role of glial cells. *Can J Physiol Pharmacol* 1992;70:S181–189.

50. Fornum F. Glutamate: a neurotransmitter in mammalian brain. *J Neurochem* 1984;42:1–11.

51. Bradford HF, Ward HK, Thomas AJ. Glutamine—a major substrate for nerve endings. *J Neurochem* 1978;30:1453–1459.

52. Hamberger AC, Chiang GH, Nylen ES, et al. Glutamate as a CNS transmitter. I. Evaluation of glucose and glutamine as precursors for the synthesis of preferentially released glutamate. *Brain Res* 1979;168:513–530.

53. Balazs R, Machiyama Y, Hammond BJ, et al. The operation of the gamma-aminobutyrate bypath of the tricarboxylic acid cycle in brain tissue in vitro. *Biochem J* 1970;116:445–461.

54. Schousboe A, Westergaard N. Transport of neuroactive amino acids in astrocytes. In: *Neuroglia*. New York: Oxford University Press, 1995:246–258.

55. Anderson CM, Swanson RA. Astrocyte glutamate transport: review of properties, regulation, and physiological functions. *Glia* 2000;32:1–14.

56. Norenberg MD, Martinez-Hernandez A. Fine structural localization of glutamine synthetase in astrocytes of rat brain. *Brain Res* 1979;161:303–310.

57. Derouiche A, Frotscher M. Astroglial processes around identified glutamatergic synapses contain glutamine synthetase: evidence for transmitter degradation. *Brain Res* 1991;552:346–350.

58. Yudkoff M, Zaleska MM, Nissim I, et al. Neuronal glutamine utilization: pathways of nitrogen transfer studied with [^{15}N]glutamine. *J Neurochem* 1989;53:632–640.

59. Westergaard N, Sonnewald U, Schousboe A. Metabolic trafficking between neurons and astrocytes: the glutamate/glutamine cycle revisited. *Dev Neurosci* 1995;17:203–211.

60. Magistretti PJ, Pellerin L. Cellular mechanisms of brain energy metabolism and their relevance to functional brain imaging. *Phil Trans R Soc Lond B* 1999;354:1155–1163.

61. Hassel B, Brathe A. Cerebral metabolism of lactate in vivo: evidence for neuronal pyruvate carboxylation. *J Cereb Blood Flow Metab* 2000;20:327–336.

62. Hassel B, Brathe A. Neuronal pyruvate carboxylation supports formation of transmitter glutamate. *J Neurosci* 2000;20:1342–1347.

63. Martin DL. Short-term control of GABA synthesis in brain. *Prog Biophys Mol Biol* 1993;60:17–28.

64. Sonnewald U, Westergaard N, Schousboe A, et al. Direct demonstration by [^{13}C]NMR spectroscopy that glutamine from astrocytes is a precursor for GABA synthesis in neurons. *Neurochem Int* 1993;22:19–29.

65. Paulsen RE, Odden E, Fonnum F. Importance of glutamine for gamma-aminobutyric acid synthesis in rat neostriatum in vivo. *J Neurochem* 1998;51:1294–1299.

66. Keyser DO, Pellmar TC. Synaptic transmission in the hippocampus: critical role for glial cells. *Glia* 1994;10:237–243.

67. Benveniste H, Drejer J, Schousboe A, et al. Elevation of the extracellular concentrations of glutamate and aspartate in rat hippocampus during transient cerebral ischemia monitored by intracerebral microdialysis. *J Neurochem* 1984;43:1369–1374.

68. Rothstein JD, Martin L, Levey AI, et al. Localization of neuronal and glial glutamate transporters. *Neuron* 1994;13:713–725.

69. Rose CR, Ransom BR. pH regulation in mammalian glia. In: *pH and brain function*. New York: Wiley-Liss, 1998:253–275.

70. Deitmer JW, Schneider HP. Enhancement of glutamate uptake transport by $CO_{(2)}$/bicarbonate in the leech giant glial cell. *Glia* 2000;30:392–400.

71. McLennan H. The autoradiographic localization of L-[^{3}H]glutamate in rat brain tissue. *Brain Res* 1976;115:139–144.

72. Rosenberg PA, Aizenman E. Hundred-fold increase in neuronal vulnerability to glutamate toxicity in astrocyte-poor cultures of rat cerebral cortex. *Neurosci Lett* 1989;103:162–168 [published erratum appears in *Neurosci Lett* 1990;116:399.]

73. Rothstein JD, Dykes-Hoberg M, Pardo CA, et al. Knockout of glutamate transporters reveals a major role for astroglial transport in excitotoxicity and clearance of glutamate. *Neuron* 1996;16:675–686.

74. Chaudhry FA, Lehre KP, van Lookeren Campagne M, et al. Glutamate transporters in glial plasma membranes: highly differentiated localizations revealed by quantitative ultrastructural immunocytochemistry. *Neuron* 1995;15:711–720.

75. Bergles DE, Jahr CE. Glial contribution to glutamate uptake at Schaffer collateral-commissural synapses in the hippocampus. *J Neurosci* 1998;18:7709–7716.

76. Tong G, Jahr CE. Block of glutamate transporters potentiates postsynaptic excitation. *Neuron* 1994;13:1195–1203.

77. Murphy S. *Astrocytes: pharmacology and function*. San Diego: Academic Press, 1993.

78. Tsacopoulos M, Magistretti PJ. Metabolic coupling between glia and neurons. *J Neurosci* 1996;16:877–885.

79. Magistretti PJ, Pellerin L. Cellular bases of brain energy metabolism and their relevance to functional brain imaging: evidence for a prominent role of astrocytes. *Cereb Cortex* 1996;6:50–61.

80. Pellerin L, Magistretti PJ. Glutamate uptake into astrocytes stimulates aerobic glycolysis: a mechanism coupling neuronal activity to glucose utilization. *Proc Natl Acad Sci USA* 1994;91:10625–10629.

81. Pellerin L, Magistretti PJ. Glutamate uptake stimulates Na+/K+-ATPase activity in astrocytes via activation of a distinct subunit highly sensitive to ouabain. *J Neurochem* 1997;69:2132–2137.

82. Magistretti PJ, Pellerin L, Rothman DL, et al. Energy on demand. *Science* 1999;283:496–497.

83. Walz W, Mukerji S. Lactate release from cultured astrocytes and neurons: a comparison. *Glia* 1988;1:366–370.

84. Larrabee MG. Lactate metabolism and its effects on glucose metabolism in an excised neural tissue. *J Neurochem* 1995;64:1734–1741.

85. Schurr A, West CA, et al. Lactate-supported synaptic function in the rat hippocampal slice preparation. *Science* 1988;240:1326–1328.

86. Poitry-Yamate CL, Poitry S, et al. Lactate released by Muller glial cells is metabolized by photoreceptors from mammalian retina. *J Neurosci* 1995;15:5179–5191.

87. Ransom BR, Fern R. Does astrocytic glycogen benefit axon function and survival in CNS white matter during glucose deprivation? *Glia* 1997;21:134–141.

88. Vega C, Poitry-Yamate CL, Jirounek P, et al. Lactate is released and taken up by isolated rabbit vagus nerve during aerobic metabolism. *J Neurochem* 1998;71:330–337.

89. Bittar PG, Charnay Y, Pellerin L, et al. Selective distribution of lactate dehydrogenase isoenzymes in neurons and astrocytes of human brain. *J Cereb Blood Flow Metab* 1996;16:1079–1089.

90. Bröer S, Rahman B, Pellegri G, et al. Comparison of lactate transport in astroglial cells and monocarbosylate transporter 1 (MCT1) expressing xenopus laevis oocytes. *J Biol Chem* 1997;272:30096–30102.

91. Pellerin L, Pellegri G, Martin J-L, et al. Expression of monocarboxylate transporter mRNA in mouse brain: support for a distinct role of lactate as an energy substrate for the neonatal vs adult brain. *Proc Natl Acad Sci USA* 1998;95:3990–3995.

92. Sibson NR, Dhankhar A, Mason GF, et al. Stoichiometric coupling of brain glucose metabolism and glutamatergic neuronal activity. *Proc Natl Acad Sci USA* 1998;95:316–321.

93. Fray AE, Forsyth RJ, Boutelle MG, et al. The mechanisms controlling physiologically stimulated changes in rat brain glucose and lactate: a microdialysis study. *J Physiol* 1996;496:49–57.

94. Prichard D, Rothman D, Novotny E, et al. Lactate rise detected by 1H NMR in human visual cortex during physiologic stimulation. *Proc Natl Acad Sci USA* 1991;88:5829–5831.

95. Fox PT, Raichle ME, Mintun MA, et al. Nonoxidative glucose consumption during focal physiologic neural activity. *Science* 1988;241:462–464.

96. Hyder F, Chase JR, Behar KL, et al. Increased tricarboxylic acid cycle flux in rat brain during forepaw stimulation detected with ^{1}H [^{13}C] NMR. *Proc Natl Acad Sci USA* 1996;93:7612–7617.

97. Magistretti PJ, Sorg O, Martin J-L. Regulation of glycogen metabolism in astrocytes: physiological, pharmacological, and pathological aspects. In: *Astrocytes: pharmacology and function*. Academic Press, 1993:243–265.

98. Swanson RA, Morton MM, Sagar SM, et al. Sensory stimulation induces local cerebral glycogenolysis: demonstration by autoradiography. *Neuroscience* 1992;51:451–461.

99. Wender R, Brown AM, Fern R, et al. Astrocytic glycogen influences axon function and survival during glucose deprivation in central white matter. [in process citation.] *J Neurosci* 2000;20:6804–6810.

100. Dringen R, Gebhardt R, Hamprecht B. Glycogen in astrocytes: possible function as lactate supply for neighboring cells. *Brain Res* 1993;623:208–214.

101. Swanson RA, Sagar SM, Sharp FR. Regional brain glycogen stores and metabolism during complete global ischaemia. *Neurol Res* 1989;11:24–28.

102. Swanson RA, Choi DW. Glial glycogen stores affect neuronal survival during glucose deprivation in vitro. *J Cereb Blood Flow Metab* 1993;13:162–169.

103. Fern R, Davis P, Waxman SG, et al. Axon conduction and survival in CNS white matter during energy deprivation: a developmental study. *J Neurophysiol* 1998;79:95–105.

104. Magistretti PJ, Morrison JH, Shoemaker WJ, et al. Vasoactive intestinal polypeptide induced glycogenolysis in mouse cortical slices: a possible regulatory mechanism for the local control of energy metabolism. *Proc Natl Acad Sci USA* 1981;78:6535–6539.

105. Magistretti PJ, Hof P, Martin JL. Adenosine stimulates glycogenolysis in mouse cerebral cortex: a possible coupling mechanism between neuronal activity and energy metabolism. *J Neurosci* 1986;6:2558–2562.

106. Sorg O, Magistretti PJ. Characterization of the glycogenolysis elicited by vasoactive intestinal peptide, noradrenaline and adenosine in primary cultures of mouse cerebral cortical astrocytes. *Brain Res* 1991;563:227–233.

107. Sokoloff L, Reivich M, Kennedy C, et al. The [^{14}C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *J Neurochem* 1977;28:897–916.

108. Sorg O, Magistretti PJ. Vasoactive intestinal peptide and noradrenaline exert long-term control on glycogen levels in astrocytes: blockade by protein synthesis inhibition. *J Neurosci* 1992;12:4923–4931.

109. Cardinaux J-R, Magistretti PJ. Vasoactive intestinal peptide, pituitary adenylate cyclase-activating peptide, and noradrenaline induce the transcription factors CCAAT/enhancer binding protein (C/EBP)-B and C/EBPd in mouse cortical astrocytes: involvement in cAMP-regulated glycogen metabolism. *J Neurosci* 1996;16:919–929.

110. Pellegri G, Rossier C, Magistretti PJ, et al. Cloning, localization and induction of mouse brain glycogen synthase. *Brain Res Mol Brain Res* 1996;38:191–199.

111. Allaman I, Pellerin L, Magistretti PJ. Protein targeting to glycogen (PTG) mRNA expression is stimulated by noradrenaline in mouse cortical astrocytes. *Glia* 2000;30:382–391.

112. Cornell-Bell AH, Finkbeiner SM, Cooper MS, et al. Glutamate induces calcium waves in cultured astrocytes: long-range glial signaling. *Science* 1990;247:470–473.

113. Sanderson MJ, Charles AC, Boitano S, et al. Mechanisms and function of intercellular calcium signaling. *Mol Cell Endocrinol* 1994;98:173–187.

114. Dani JW, Chernjavsky A, Smith SJ. Neuronal activity triggers calcium waves in hippocampal astrocyte networks. *Neuron* 1992;8:429–440.

115. Charles A. Intercellular calcium waves in glia. *Glia* 1998;24:39–49.

116. Newman EA, Zahs KR. Modulation of neuronal activity by glial cells in the retina. *J Neurosci* 1998;18:4022–4028.

117. Nedergaard M. Direct signaling from astrocytes to neurons in cultures of mammalian brain cells. *Science* 1994;263:1768–1771.

118. Hassinger TD, Atkinson PB, Strecker GJ, et al. Evidence for glutamate-mediated activation of hippocampal neurons by glial calcium waves. *J Neurobiol* 1995;28:159–170.

119. Parpura V, Basarsky TA, Liu F, et al. Glutamate-mediated astrocyte-neuron signalling [see comments]. *Nature* 1994;369:744–747.

120. Kimberley McAllister A, Katz LC, Lo DC. Neurotrophins and synaptic plasticity. *Annu Rev Neurosci* 1999;22:295–318.

121. Andersen P, Soleng AF. Long-term potentiation and spatial training are both associated with the generation of new excitatory synapses. *Brain Res Rev* 1998;26:353–359.

122. Stevens CF, Sullivan J. Synaptic plasticity. *Curr Biol* 1998;8:R151–R153.

123. Fitzsimonds RM, Poo MM. Retrograde signaling in the development and modification of synapses. *Physiol Rev* 1998;78:143–170.

124. Theodosis DT, El Majdoubi M, Pierre K, et al. Factors governing activity-dependent structural plasticity of the hypothalamoneurohypophysial system. *Cell Mol Neurobiol* 1998;18:285–298.

125. Pierre K, Rougon G, Allard M, et al. Regulated expression of the cell adhesion glycoprotein F3 in adult hypothalamic magnocellular neurons. *J Neurosci* 1998;18:5333–5343.

126. Sirevaag AM, Greenough WT. Differential rearing effects on rat visual cortex synapses. III. Neuronal and glial nuclei, boutons, dendrites, and capillaries. *Brain Res* 1987;424:320–332.

127. Sirevaag AM, Black JE, Shafron D, et al. Direct evidence that complex experience increases capillary branching and surface area in visual cortex of young rats. *DBR* 1988;43:299–304.

128. Jones TA, Hawrylak N, Greenough WT. Rapid laminar-dependent changes in GFAP immunoreactive astrocytes in the visual cortex of rats reared in a complex environment. *Psychoneuroendocrinology* 1996;21:189–201.

129. Jones TA, Greenough WT. Ultrastructural evidence for increased contact between astrocytes and synapses in rats reared in a complex environment. *Neurobiol Learn Memory* 1996;65:48–56.

130. Anderson BJ, Li X, Alcantara AA, et al. Glial hypertrophy is associated with synaptogenesis following motor-skill learning, but not with angiogenesis following exercise. *Glia* 1994;11:73–80.

131. Gomez-Pinilla F, So V, Kesslak JP. Spatial learning and physical activity contribute to the induction of fibroblast growth factor: neural substrates for increased cognition associated with exercise. *Neuroscience* 1998;85:53–61.

132. Silva AJ, Giese KP, Fedorov NB. Molecular, cellular, and neuroanatomical substrates of place learning. *Neurobiol Learn Memory* 1998;70:44–61.

133. Wenzel J, Lammert G, Meyer U, et al. The influence of long-term potentiation on the spatial relationship between astrocyte processes and potentiated synapses in the dentate gyrus neuropil of rat brain. *Brain Res* 1991;560:122–131.

134. Muller CM. Glial cell functions and activity-dependent plasticity of the mammalian visual cortex. *Perspect Dev Neurobiol* 1993;1:169–177.

135. Isaacs KR, Anderson BJ, Alcantara AA, et al. Exercise and the brain: angiogenesis in the adult rat cerebellum after vigorous physical activity and motor skill learning. *J Cereb Blood Flow Metab* 1992;12:110–119.

136. Kohsaka S-I, Takamatsu K, Aoki E, et al. Metabolic mapping of chick brain after imprinting using [^{14}C]2-deoxyglucose technique. *Brain Res* 1979;172:539–544.

137. Shimada M, Murakami TH, et al. Local cerebral alterations in [^{14}C-2]deoxyglucose uptake following memory formation. *J Anat* 1983;136:751–759.

138. Sif J, Messier C, et al. Time-dependent sequential increases in [^{14}C-2]deoxyglucose uptake in subcortical and cortical structures during memory consolidation of an operant training in mice. *Behav Neural Biol* 1991;56:43–61.

139. Siucinaka E, Kossut M. Short-lasting classical conditioning induces reversible changes of representational maps of vibrissae in mouse SI cortex: a 2DG study. *Cerebral Cortex* 1996;6:506–513.

140. Bontempi B, Jaffard R, Destrade C. Differential temporal evolution of post-training changes in regional brain glucose metabolism induced by repeated spatial discrimination training in mice: visualization of the memory consolidation process? *Eur J Neurosci* 1996;8:2348–2360.

141. Coyle JT, Schwarcz R. Mind glue. Implications of glial cell biology for psychiatry. *Arch Gen Psychiatry* 2000;57:90–93.

142. Ongür D, Drevets WC, Price JL. Glial reduction in the subgenual prefrontal cortex in mood disorders. *Proc Natl Acad Sci USA* 1998;95:13290–13295.

143. Drevets WC, Price JL, Simpson JRJ, et al. Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* 1997;386:824–827.

144. Salomon LD, Lidow MS, Goldman-Rakic P. Increased volume and glial density in primate prefrontal cortex associated with chronic antipsychotic drug exposure. *Biol Psychiatry* 1999;46:161–172.

11

Synaptic Plasticity

Robert C. Malenka

Robert A. Malenka: Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Palo Alto, California.

The most fascinating and important property of the mammalian brain is its remarkable plasticity, which can be thought of as the ability of experience to modify neural circuitry and thereby to modify future thought, behavior, and feeling. Thinking simplistically, neural activity can modify the behavior of neural circuits by one of three mechanisms: (a) by modifying the strength or efficacy of synaptic transmission at preexisting synapses, (b) by eliciting the growth of new synaptic connections or the pruning away of existing ones, or (c) by modulating the excitability properties of individual neurons. *Synaptic plasticity* refers to the first of these mechanisms, and for almost 100 years, activity-dependent changes in the efficacy of synaptic communication have been proposed to play an important role in the remarkable capacity of the brain to translate transient experiences into seemingly infinite numbers of memories that can last for decades. Because of its fundamental importance, there has been an enormous amount of work describing the many forms of synaptic plasticity and their underlying mechanisms.

Synaptic transmission can either be enhanced or depressed by activity, and these alterations span temporal domains ranging from milliseconds to enduring modifications that may persist for days or weeks and perhaps even longer. Transient forms of synaptic plasticity have been associated with short-term adaptations to sensory inputs, transient changes in behavioral states, and short-lasting forms of memory. More lasting changes are thought to play important roles in the construction of neural circuits during development and with long-term forms of memory in the mature nervous system. Given these diverse functions, it is not surprising that many forms and mechanisms of synaptic plasticity have been described. In this chapter, I provide a brief overview of some of the forms of synaptic plasticity found at excitatory synapses in the mammalian brain, focusing on long-term potentiation (LTP) and long-term depression (LTD).

SHORT-TERM SYNAPTIC PLASTICITY

Part of "11 - Synaptic Plasticity "

Virtually every synapse that has been examined in organisms ranging from simple invertebrates to mammals exhibits numerous different forms of short-term synaptic plasticity that last on the order of milliseconds to a few minutes (for detailed reviews, see 1 and 2). In general, these result from a short-lasting modulation of transmitter release that can occur by one of two general types of mechanisms. One involves a change in the amplitude of the transient rise in intracellular calcium concentration that occurs when an action potential invades a presynaptic terminal. This occurs because of some modification in the calcium influx before transmitter release or because the basal level of calcium in the presynaptic terminal has been elevated because of prior activity at the terminal. A second mechanism occurs downstream of calcium elevation in the presynaptic terminal and involves some modulation of the biochemical processes involved in synaptic vesicle exocytosis.

Paired-Pulse Facilitation and Depression

When two presynaptic stimuli are delivered within a short interval, the synaptic response to the second stimulus can be either enhanced or depressed relative to the first stimulus. Paired-pulse depression is commonly observed at all synapses at short (less than 20 milliseconds) interstimulus intervals. It may result from inactivation of voltage-dependent sodium or calcium channels or from a transient depletion of the synaptic vesicles that are “docked” adjacent to the presynaptic plasma membrane, waiting to be released. Many synapses at longer interstimulus intervals (20 to 500 milliseconds) exhibit paired-pulse facilitation that is thought to result from the influx of calcium that occurs in response to the first action potential. One simple idea is that the “residual” calcium left over from the first action potential combines with the calcium influx during the second action potential, and because the relationship between calcium concentration in the terminal and release is highly nonlinear, this small increase in resting calcium may cause substantial facilitation. However, with a single action potential, the increase in resting calcium concentration is very small, and

thus additional mechanisms are likely involved. Currently, there is much interest in the possibility that transient modulation, by activation of protein kinases, of some of the presynaptic phosphoproteins that are known to be involved in the control of transmitter release may play an important role in very short-term synaptic plasticity. For example, knockout mice lacking one or more of the synapsins (3 ,4), or lacking the small guanosine triphosphate–binding protein rab3A (5 ,6), exhibit abnormal short-term synaptic plasticity.

Whether a specific synapse displays paired-pulse facilitation or depression depends on the initial state of the synapse and its recent history of activation. Because these forms of plasticity largely result from changes in the probability of transmitter release, synapses that begin with a very high probability of release tend to show depression, whereas those with a low probability of release exhibit facilitation. Consistent with this idea, activation of presynaptic receptors that cause a decrease in transmitter release almost always causes an increase in the magnitude of paired-pulse facilitation (or even a conversion of paired-pulse depression to paired-pulse facilitation).

Facilitation and Depression Following Trains of Stimuli

Longer-lasting forms of plasticity are observed following repetitive or tetanic stimulation of synapses with prolonged (approximately 200-millisecond to 5-second) trains of stimuli applied at high frequencies (10 to 200 Hz). *Augmentation* and *posttetanic potentiation* refer to enhancements of transmitter release that can last anywhere from seconds (augmentation) to several minutes (posttetanic potentiation). They are thought to result in large part to the buildup of calcium concentration in the presynaptic terminal during the trains of stimuli. This residual calcium may both combine with the calcium influx elicited by subsequent single action potentials and lead to biochemical modifications of proteins in the presynaptic terminal. At some synapses, repetitive activation leads to depression that can last for several seconds or even minutes. As in paired-pulse depression, this generally occurs at synapses that exhibit a high probability of release and is thought to result, at least in part, from a transient depletion of the synaptic vesicles that are poised to be released by an action potential.

In large part because of these short-term forms of synaptic plasticity, the strength of communication between pairs of neurons can be modified even during short bursts of presynaptic activity (e.g., five to ten action potentials at 20 to 50 Hz) (7). The functional relevance of such short-term synaptic dynamics has received much less attention than long-lasting forms of synaptic plasticity and is just beginning to be explored (8). One potential role of these short-term forms of synaptic plasticity is to transform incoming information in the temporal domain into a spatially distributed code (9 ,10). Furthermore, given that presynaptic proteins that may be involved in short-term plasticity may be abnormal in neuropsychiatric disorders (11), it is not unreasonable to speculate that abnormal short-term synaptic dynamics in specific neural circuits may contribute to the pathophysiology of any number of mental illnesses.

LONG-TERM SYNAPTIC PLASTICITY

Part of "11 - Synaptic Plasticity "

During the last decade, there was enormous interest in elucidating the mechanisms responsible for activity-dependent long-lasting modifications in synaptic strength. The great interest in this topic is largely based on the simple idea that external and internal events are represented in the brain as complex spatiotemporal patterns of neuronal activity, the properties of which result from the pattern of synaptic weights at the connections made between the neurons that are contributing to this activity. The corollary to this hypothesis is that new information is stored (i.e., memories are generated) when activity in a circuit causes a long-lasting change in the pattern of synaptic weights. This simple idea was put forth by Ramon y Cajal almost 100 years ago, but experimental support for such a process was lacking until the early 1970s, when it was demonstrated that repetitive activation of excitatory synapses in the hippocampus caused an increase in synaptic strength that could last for hours or even days (12 ,13). This long-lasting synaptic enhancement, LTP, has been the object of intense investigation because it is widely believed that LTP provides an important key to understanding the molecular mechanisms by which memories are formed (14 ,15) and, more generally, by which experience modifies behavior. Furthermore, the activity- and experience-dependent refinement of neural circuitry that occurs during development shares features with learning, and thus a role for LTP in this process has been proposed (16 ,17 and 18).

Long-Term Potentiation

No form of synaptic plasticity has generated more interest and has been more extensively studied than LTP in the CA1 region of the hippocampus. The excitement surrounding this phenomenon derives mainly from four sources. First, there is compelling evidence from studies in rodents and higher primates, including humans, that the hippocampus is a critical component of a neural system involved in various forms of long-term memory (19). Second, several properties of LTP make it an attractive cellular mechanisms for information storage (20 ,21). Like memories, LTP can be generated rapidly and is prolonged and strengthened with repetition. It is also input specific in that it is elicited at the synapses activated by afferent activity and not at adjacent synapses on the same postsynaptic cell. This feature dramatically increases the storage capacity of individual neurons

that, because synapses can be modified independently, can participate in the encoding of many different bits of information. Third, LTP is readily generated in *in vitro* preparations of the hippocampus, thus making it accessible to rigorous experimental analysis. Indeed, much of what we know about the detailed mechanisms of LTP derives from studies of LTP at excitatory synapses on CA1 pyramidal cells in hippocampal slices. Fourth, LTP has been observed at virtually every excitatory synapse in the mammalian brain that has been studied. Table 11.1 gives a list of the brain regions in which LTP has been demonstrated, and it can be seen that regions thought to be particularly important for various forms of learning and memory are prominent. Although LTP is not a unitary phenomenon, most synapses appear to express a form of LTP that is identical or highly analogous to the LTP found at excitatory synapses on CA1 pyramidal cells. Thus, this form of LTP is the focus of the remainder of this section.

Hippocampus	Amygdala
Dentate gyrus	Cerebellum
CA1	Thalamus
CA3	Striatum
Cerebral cortex	Nucleus acumbens
Visual	Ventral tegmental area
Somatosensory	
Motor	
Prefrontal	

TABLE 11.1. AREAS OF BRAIN IN WHICH LTP HAS BEEN DEMONSTRATED

Triggering of LTP: A Critical Role for NMDA Receptors and Calcium

It is well established that the triggering of LTP requires synaptic activation of postsynaptic *N*-methyl-D-aspartate (NMDA) receptors, a subtype of ionotropic glutamate receptor (see Chapter 6) and postsynaptic depolarization, which is accomplished experimentally by repetitive tetanic stimulation of synapses or by directly depolarizing the cell while continuing low-frequency synaptic activation (a so-called “pairing protocol”). How do these requirements explain the properties of LTP? During basal low-frequency synaptic transmission, synaptically released glutamate binds to two different subtypes of ionotropic glutamate receptor, termed AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) and NMDA receptors, which are often, but not always (see later), co-localized on individual dendritic spines. The AMPA receptor has a channel that is permeable to monovalent cations (Na^+ and K^+), and activation of AMPA receptors provides most of the inward current that generates the excitatory synaptic response when the cell is close to its resting membrane potential (Fig. 11.1). In contrast, as described in Chapter 6 , the NMDA receptor exhibits a strong voltage dependence because of the block of its channel at negative membrane potentials by extracellular magnesium. As a result, NMDA receptors contribute little to the postsynaptic response during basal synaptic activity. However, when the cell is depolarized, magnesium dissociates from its binding site within the NMDA receptor channel and allows calcium as well as sodium to enter the dendritic spine (Fig. 11.1). The resultant rise in intracellular calcium is a necessary and perhaps sufficient trigger for LTP. This local source of calcium within the dendritic spine accounts for the input specificity of LTP.

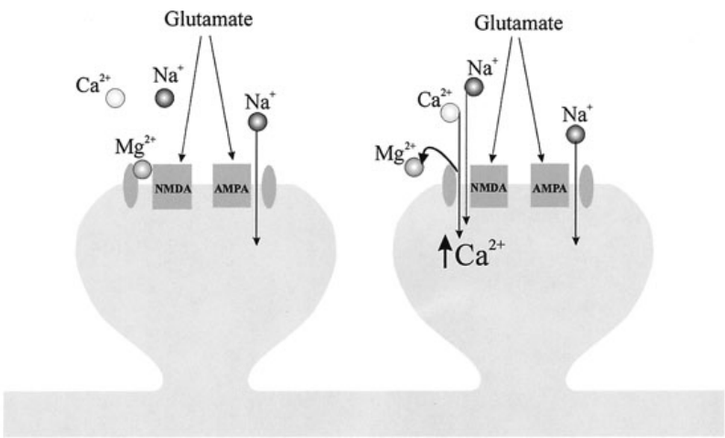


FIGURE 11.1. Model for the induction of long-term potentiation (LTP). During normal synaptic transmission (*left*), synaptically released glutamate acts on both NMDA and AMPA receptors. Na^+ flows through the AMPA receptor channel but not through the NMDA receptor channel because of the Mg^{2+} block of this channel. Depolarization of the postsynaptic cell (*right*) relieves the Mg^{2+} block of the NMDA receptor channel and allows Na^+ and Ca^{2+} to flow into the cell. The resultant rise in Ca^{2+} in the dendritic spine is a necessary trigger for the subsequent events leading to LTP.

The evidence in support of this model for the initial triggering of LTP is compelling. Specific NMDA receptor antagonists have minimal effects on basal synaptic transmission but block the generation of LTP (22 ,23). Preventing the rise in calcium by loading cells with calcium chelators blocks LTP (24 ,25), whereas directly raising intracellular calcium in the postsynaptic cell mimics LTP (25 ,26). Furthermore, imaging studies have demonstrated that NMDA receptor activation causes a large increase in calcium level within dendritic spines (see 23 for references). The exact properties of the calcium signal that is required to trigger LTP are unknown, but a transient signal lasting only 1 to 3 seconds appears to be sufficient (27). Whether additional sources of calcium, such as release from intracellular stores, are required for the generation of LTP is unclear. It is also uncertain whether additional factors provide by synaptic activity are required. Various neurotransmitters found in the hippocampus such as acetylcholine and norepinephrine can modulate the ability to trigger LTP, and such modulation

may be of great importance for the functional in vivo roles of LTP. However, there is no compelling evidence to suggest that any neurotransmitter other than glutamate is required to trigger LTP.

Signal Transduction Mechanisms in LTP

A bewildering array of signal transduction molecules has been suggested to play a role in translating the calcium signal that is required to trigger LTP into a long-lasting increase in synaptic strength (28). However, for only a few of these has compelling evidence of a mandatory role in LTP been presented. A major limitation of much of the work on this topic is that investigators have not adequately distinguished molecules that are key components of the signal transduction machinery absolutely required for LTP from biochemical processes that modulate the ability to generate LTP or play a permissive role. For example, any manipulation that modifies the activity of NMDA receptors may affect LTP. Therefore, several requirements must be met for a signaling molecule to be considered a key component of the biochemical machinery that triggers LTP. First, it must be activated or produced by stimuli that trigger LTP but not by stimuli that fail to do so. Second, inhibition of the pathway in which the molecule participates should block the generation of LTP. Third, activation of the pathway should lead to LTP.

Strong evidence indicates that calcium/calmodulin–dependent protein kinase II (CaMKII) fulfills these requirements and is a key component of the molecular machinery for LTP. Inhibiting its activity pharmacologically by directly loading postsynaptic cells with CaMKII inhibitors or genetic knockout of a critical CaMKII subunit blocks the ability to generate LTP (29 ,30 and 31). Conversely, acutely increasing the postsynaptic concentration of active CaMKII increases synaptic strength and occludes LTP (32 ,33). Furthermore, CaMKII undergoes autophosphorylation after the triggering of LTP (34 ,35). That this autophosphorylation is required for LTP was demonstrated by the finding that genetic replacement of endogenous CaMKII with a form lacking the autophosphorylation site prevented LTP (36).

Several other protein kinases have also been suggested to play roles in the triggering of LTP, but the experimental evidence supporting their role is considerably weaker than for CaMKII. Activation of the cyclic adenosine monophosphate–dependent protein kinase (PKA), perhaps by activation of a calmodulin-dependent adenylyl cyclase, has been suggested to boost the activity of CaMKII indirectly by decreasing competing protein phosphatase activity (37 ,38). This presumably happens by phosphorylation of inhibitor-1, an endogenous protein phosphatase inhibitor (see section on LTD later). Protein kinase C may play a role analogous to CaMKII, whereas the tyrosine kinases Fyn and Src may indirectly modulate LTP by affecting NMDA receptor function (see 23 for references). The mitogen-activated protein kinase (MAPK) has also been suggested to be important for LTP, albeit in unknown ways.

Expression Mechanisms and LTP

In the 1990s, tremendous confusion and controversy surrounded the seemingly simple issue of whether LTP is caused primarily by presynaptic or postsynaptic modifications. The great challenge to answering this question largely stemmed from the great technical difficulties inherent in examining the changes the occur at individual synapses that are embedded in a complex network in which each cell receives 10,000 or more synapses. Most neurobiologists studying this question agree that the simplest postsynaptic change that could cause LTP would be a change in AMPA receptor function or number, whereas the simplest presynaptic change would be an increase in the probability of neurotransmitter release.

Most studies examining this issue have used electrophysiologic assays, and most of these are inconsistent with the hypothesis that the release of glutamate increases significantly during LTP (23 ,39). For example, changes in transmitter release probability invariably influence various forms of short-term synaptic plasticity such as paired-pulse facilitation, yet these phenomena are not affected by LTP. To measure glutamate release more directly, two approaches were used. One took advantage of the finding that glial cells tightly ensheath synapses and respond to synaptically released glutamate by activation of electrogenic transporters that generate a current directly proportional to the amount of glutamate released (40 ,41). The other took advantage of use-dependent antagonists of the NMDA receptor or of a mutant AMPA receptor that lacks the GluR2 subunit. These antagonists decrease synaptic currents at a rate that is directly proportional to the probability of transmitter release (42 ,43). LTP had no discernible effect on these measures, even though they were affected in the predicted fashion by manipulations known to increase transmitter release.

In addition to these negative findings, certain electrophysiologic and biochemical measures were found to increase during LTP. An increase in the amplitude of miniature electrophysiologic synaptic currents (mEPSCs), which represent the postsynaptic response to the spontaneous release of a single quantum of neurotransmitter, normally indicates an increase in the number or function of postsynaptic neurotransmitter receptors. Such an increase occurs during LTP (44), as well as after manipulations that load dendritic spines with calcium (45 ,46). A more direct way of monitoring changes in AMPA receptors is to measure the postsynaptic response to direct application of agonist,

and such responses have also been reported to increase, albeit gradually (47).

That LTP is caused by a modification of AMPA receptors is supported by the finding that LTP causes an increase in the phosphorylation of the AMPA receptor subunit GluR1 at the site that is known to be phosphorylated by CaMKII (as well as PKC) (35 ,48 ,49). Using expression systems, this phosphorylation has been shown to increase the single-channel conductance of AMPA receptors (50). Because an increase in single-channel conductance of AMPA receptors has been reported to occur during LTP (51), one mechanism that seems likely to contribute to LTP is CaMKII-dependent phosphorylation of GluR1. Consistent with this idea, genetic knockout of GluR1 has been found to prevent the generation of LTP (52).

Silent Synapses and Quantal Synaptic Transmission

Although the evidence presented thus far makes a strong case for postsynaptic changes contributing to LTP, there remained one reproducible experimental result that was difficult to reconcile with this idea. This result derived from experiments that took advantage of the finding that the action potential-dependent release of quanta of neurotransmitter at individual synapses is probabilistic, and therefore release occurs only 10% to 40% of the time. Therefore, if a single synapse or a very small number of synapses is activated once every few seconds, on some of the trials no postsynaptic response is recorded, that is, a so-called failure occurs. An extensively replicated finding is that LTP causes a decrease in the proportion of failures that occur (see 53 for review). Because these failures were assumed to result from failures of neurotransmitter release, it was concluded that LTP involves an increase in the probability of transmitter release.

How can this result be reconciled with all the evidence suggesting that LTP is caused by modulation of AMPA receptors and is not accompanied by an increase in glutamate release? One straightforward idea to explain this apparent discrepancy is the *silent synapse hypothesis* (54), which predicts that some synapses express only NMDA receptors, whereas others express both AMPA and NMDA receptors (Fig. 11.2). Synapses with only NMDA receptors would be functionally silent at hyperpolarized membrane potentials, and thus, when transmitter is released, they would not yield a response. However, LTP at such silent synapses could occur by the rapid expression of AMPA receptors, and such a mechanism would account for the apparent change in failure rate.

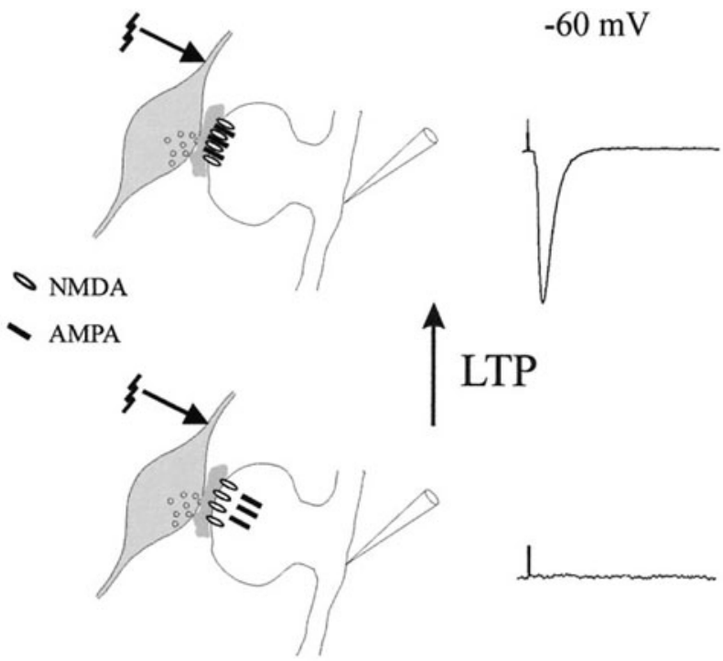


FIGURE 11.2. Diagram of the silent synapse hypothesis. A synapse is functionally silent when it expresses NMDA receptors but not AMPA receptors in its plasma membrane (bottom). **The induction of LTP causes the insertion of AMPA receptors (top)** from a putative cytosolic pool. To the right of each diagram are the synaptic currents (i.e., EPSCS) that would be recorded from the corresponding synapse.

There is now strong evidence to support this model of LTP. First, it is possible to record EPSCs that are mediated solely by NMDA receptors, and applying an LTP induction protocol at such synapses causes the rapid appearance of AMPA receptor-mediated EPSCs (55 ,56). Second, immunocytochemical analysis demonstrates that AMPA receptors are not found at a significant percentage of hippocampal synapses, whereas all synapses appear to contain NMDA receptors (see 23 for references). Third, LTP has been shown to cause the delivery of green fluorescent protein (GFP)-tagged AMPA receptors to dendritic spines and the insertion of recombinant AMPA receptors into the synaptic plasma membrane (57 ,58). Fourth, AMPA and NMDA receptors interact with different proteins at the synapse (59), a finding suggesting that they are regulated independently. Fifth, interference with membrane fusion and presumably exocytosis in the postsynaptic cells impairs LTP (60) and AMPA receptors can interact with proteins involved in membrane fusion (61). These findings are consistent with the idea that membrane fusion may be an important mechanism for the delivery of AMPA receptors to the synaptic plasma membrane.

Virtually all the data presented thus far are consistent with the simple model that LTP, at least initially, is caused by the phosphorylation of AMPA receptors and the delivery or clustering of AMPA receptors within the synaptic plasma membrane (23). These events will presumably occur both at synapses that already contain functional AMPA receptors and ones that are functionally silent. As discussed later, LTD appears to involve the converse process, that is, the removal or endocytosis of AMPA receptors. At the end of this chapter,

I discuss how these changes in the number of AMPA receptors at individual synapses may lead to more permanent structural changes, which, in turn, may mediate long-lasting forms of experience-dependent plasticity.

Long-Term Depression

Like LTP, LTD has been demonstrated in a large number of different brain regions and comes in a variety of different forms (62 ,63 and 64). This section focuses on the NMDA receptor-dependent form of LTD found at excitatory synapses on CA1 pyramidal cells and that appears to result, in large part, from a reversal of the processes that mediate LTP.

Triggering of LTD: A Critical Role for NMDA Receptors and Calcium

LTD is normally generated by prolonged (3- to 15-minute) low-frequency (1- to 5-Hz) afferent stimulation or by a pairing protocol during which the cell is held at approximately -50 mV. It shares many features with LTP including input specificity, and it can completely reverse LTP, a process often termed *depotentiation*. Surprisingly, the triggering of LTD requires NMDA receptor activation and an increase in postsynaptic calcium concentration (65 ,66). This can occur because at resting membrane potentials, the voltage-dependent block of the NMDA receptor channel by magnesium is not 100% effective, and thus, each stimulus will cause a very small amount of calcium entry. Current evidence suggests that the specific properties of the intracellular calcium signal dictate whether LTP or LTD is generated by a specific pattern of synaptic activity, with LTD requiring a modest rise in calcium (67) and LTP requiring a large rise beyond some critical threshold value (68). The temporal characteristics of this calcium signal may also be important.

Signal Transduction Mechanisms in LTD: A Role for Protein Phosphatases

If calcium is the critical triggering signal for LTD, it must be capable of activating biochemical processes that reverse LTP. Because LTP results, at least in part, from activation of protein kinases, a reasonable hypothesis is that LTD is caused by preferential activation of protein phosphatases, several of which are known to be found at excitatory synapses (69). This idea was first proposed in a theoretic article (70) that presented a specific model that accounted for the bidirectional control of synaptic strength by calcium (Fig. 11.3). It proposed that a balance between CaMKII activity and protein phosphatase 1 (PP1) influenced synaptic strength by controlling the phosphorylation state of unidentified synaptic phosphoproteins. Small rises in calcium favored activation of PP1, whereas large rises were required to increase CaMKII activity. Because PP1 is not directly influenced by calcium, a well-established calcium-dependent phosphatase cascade was invoked to translate the calcium signal into increased PP1 activity (69). This cascade (Fig. 11.3) begins with activation of the calcium/calmodulin–dependent phosphatase calcineurin (also known as protein phosphatase 2B or PP2B). PP2B then dephosphorylates inhibitor 1 (I1), a phosphoprotein that, in its phosphorylated state, is a potent inhibitor of PP1. Thus, activation

of PP2B causes an increase in PP1 activity through a mechanism of disinhibition. An attractive feature of this model is that the affinity of PP2B for calcium/calmodulin is significantly greater than that of CaMKII. Therefore, PP2B will be preferentially activated by small increases in synaptic calcium levels. Furthermore, a large rise in calcium will preferentially increase protein kinase activity not only by directly activating CaMKII but also by leading to the activation of PKA, which phosphorylates I1 and thereby further inhibits PP1.

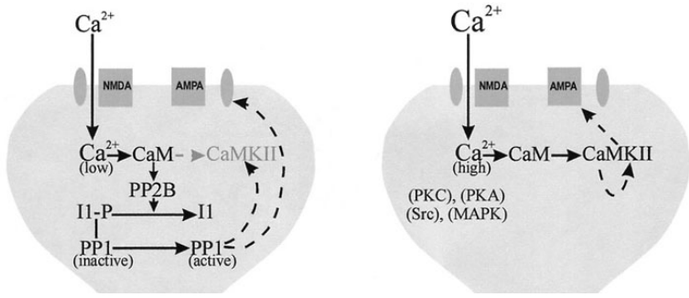


FIGURE 11.3. Model for the signaling cascades responsible for LTD and LTP. LTD is generated when a low rise in Ca^{2+} binds to calmodulin (cam) and activates calcineurin (PP2B). Calcineurin then dephosphorylates inhibitor 1 (I1), which therefore no longer inhibits protein phosphatase 1 (PP1). Active PP1 may act on any number of substrates including camkii or AMPA receptors (*left*). LTP is generated when a high rise in Ca^{2+} activates camkii. Other protein kinases that may also be involved in triggering LTP are protein kinase C (PKC), cAMP-dependent protein kinase (PKA), the tyrosine kinase src (Src), and MAP kinase (MAPK)(*right*).

Several experimental results are consistent with this model, which currently remains the leading hypothesis for the triggering of LTD. Pharmacologic inhibitors of PP1 or PP2B, when loaded directly into CA1 pyramidal cells, prevent the generation of LTD (71 ,72). Furthermore, loading cells with the phosphorylated form of I1 blocked LTD. However, although the results of such inhibitor studies are consistent with an important role for protein phosphatases in triggering LTD, other interpretations are possible, and more experimental work testing this hypothesis needs to be performed. Most notably, if PP1 plays a role in LTD analogous to that played by CaMKII in LTP, it should be possible to increase PP1 activity directly in postsynaptic cells and to mimic LTD.

Expression Mechanisms of LTD: A Role for Endocytosis of AMPA Receptors

The silent synapse hypothesis discussed earlier suggested that LTP involves the insertion of AMPA receptors into the synaptic plasma membrane. Consistent with the idea that LTD is a reversal of LTP, there is now considerable evidence that LTD involves removal (i.e., endocytosis) of synaptic AMPA receptors. The first direct evidence that the synaptic localization of AMPA receptors could be rapidly modified was the demonstration that, in cultured hippocampal neurons, short application of glutamate receptor agonists caused rapid a loss of synaptic AMPA receptors with no significant effect on the synaptic localization of NMDA receptors (73). This agonist-induced loss of synaptic AMPA receptors was subsequently shown to result from dynamin-dependent endocytosis (74). Perhaps more important, synaptically triggered LTD in the cultured neurons was accompanied by a decrease in the number of synaptic surface AMPA receptors with no discernible effect on the distribution of NMDA receptors (75). Consistent with these findings, loading CA1 pyramidal cells with inhibitors of endocytosis prevented the generation of LTD (76). These inhibitors also caused a gradual increase in the size of the synaptic responses, whereas inhibitors of exocytosis caused a gradual decrease (76). These results suggest that there is a pool of AMPA receptors that cycle into and out of the synaptic plasma membrane fairly rapidly and that LTP and LTD may involve a modification of the kinetics of these processes.

Structural Changes and Long-Term Synaptic Plasticity

How are the changes in synaptic strength that occur following the triggering of LTP or LTD maintained for periods lasting weeks or perhaps even years? Although the answer to this question is unknown, recent evidence suggests that the mechanisms described previously may be the initial steps in a more profound anatomic restructuring of synapses, including perhaps the growth of new synapses and the pruning away of preexisting ones. Dendritic spines, the postsynaptic sites that presynaptic boutons contact, have a complex ultrastructure and come in a large variety of shapes (77). With technical advances in microscopy and the use of recombinant fluorescent proteins such as GFP, it has become possible to image individual spines in living neurons. Such experiments have shown that spines are not static but can undergo rapid shape changes (78 ,79) that are influenced by activity (80). Furthermore, strong synaptic activation of the type that triggers LTP causes an NMDA receptor-dependent growth of spines as well as filopodia, which may be the precursors of spines (81 ,82). Prolonged application of NMDA to cultured neurons also can cause the loss of spines (83), a finding indicating that, like synaptic strength, the growth and loss of dendritic spines may be under the control of NMDA receptors.

Changes in synapse structure in response to activity also have been extensively explored using more standard electron microscopic techniques. One specific morphologic modification repeatedly associated with increased neuronal activity involves a reorganization of the postsynaptic density (PSD), the electron-dense thickening that contains synaptic glutamate receptors. Specifically, it has been suggested that LTP is associated with an increase in the fraction of synapses that contain discontinuities in their PSDs, termed *perforated synapses* (see 84 for references). This idea is strongly supported by studies in which the synapses activated by strong tetanic stimulation were identified in electron microscopic sections and were found to have larger total PSD surface areas and a larger proportion of perforated synapses (85 ,86). Several lines of evidence suggest that this growth of PSDs and their eventual perforation may be initiated by increasing the number of AMPA receptors in the PSD (see 84 for references).

The insertion of new AMPA receptors in the PSD and the generation of perforated synapses may also be early events in the generation of new synapses by a process of splitting or duplication of existing spines (87 ,88). Consistent with this hypothesis, LTP may be associated with an increase in spine density (84), as well as the frequency of multiple-spine synapses in which two adjacent spines arising from the same dendrite contact a single presynaptic bouton (86).

These types of observations have led to a model (Fig. 11.4) (84) that proposes a sequence of events by which the insertion of AMPA receptors leads to a growth and perforation of the PSD and eventually to multiple-spine synapses. Subsequently, retrograde communication perhaps involving cell adhesion molecules would cause appropriate presynaptic structural chances such that a completely new, independent synapse is formed. An attractive corollary to this hypothesis is that LTD involves a shrinkage of the PSD and eventually leads to a complete loss of the dendritic spine and its corresponding presynaptic bouton. However, minimal work on the structural changes associated with LTD has been performed.

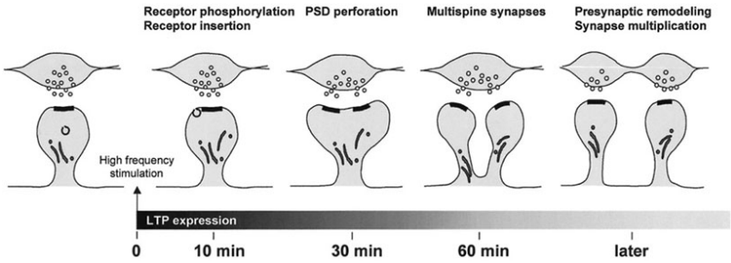


FIGURE 11.4. Model for sequence of events leading to structural changes following triggering of LTP. Within 10 minutes of LTP induction, AMPA receptors are phosphorylated and inserted into the postsynaptic membrane. This process leads to an increase in the size of the postsynaptic density (PSD) and the production of perforated synapses within 30 minutes. By 1 hour, some perforated synapses split and form multispine synapses. Eventually, retrograde communication, perhaps involving cell-adhesion molecules, leads to presynaptic structural changes and the production of new synapses.

An attractive feature of incorporating structural changes into the mechanisms of long-term synaptic plasticity is that it provides a straightforward means by which the activity generated by experience can cause very long-lasting modifications of neural circuitry. Structural changes also may explain the well-known requirement of long-lasting forms of synaptic plasticity for new protein synthesis and gene transcription (see 89 ,90 and 91 for reviews).

CONCLUSIONS

Part of "11 - Synaptic Plasticity "

This is a brief review of some of the most common forms of synaptic plasticity found at excitatory synapses throughout the mammalian brain. Although relatively little is known about the functional roles of these phenomena, such changes in synaptic function and structure remain the leading candidates for some of the fundamental mechanisms by which experiences of any type cause the reorganization of neural circuitry and thereby modify thoughts, feelings, and behavior. One hopes that it is also apparent why understanding the mechanisms of synaptic plasticity has important implications for many branches of clinical neuroscience. For example, the development of many pathologic behaviors, such as drug addiction, likely depends on the maladaptive use of neural mechanisms that normally are used for adaptive learning and memory (92). Similarly, the recovery of function following brain injury or the successful pharmacologic and behavioral treatment of mental illness also certainly must result from the reorganization of neural circuitry that is in part achieved by synaptic plasticity mechanisms. Thus further elucidation of the mechanisms of phenomena such as LTP and LTD will continue to have implications for all those interested in the neural basis of cognition and behavior.

ACKNOWLEDGMENTS

Part of "11 - Synaptic Plasticity "

I am currently supported by grants from the National Institute of Mental Health, the National Institute of Drug Abuse, and the National Institute on Aging, and I also acknowledge past support from the National Alliance for Research on Schizophrenia and Depression, the McKnight Endowment Fund for Neuroscience, and the Human Frontier Science Program. Thanks to Roger Nicoll, Steven Siegelbaum, Christian Lüscher and Dominique Muller,

with whom I have written previous reviews that provided much of the material for the present chapter.

REFERENCES

1. Zucker RS. Short-term synaptic plasticity. *Annu Rev Neurosci* 1989;12:13–31.

2. Zucker RS. Calcium- and activity-dependent synaptic plasticity. *Curr Opin Neurobiol* 1999;9:305–313.

3. Rosahl TW, Geppert M, Spillane D, et al. Short-term synaptic plasticity is altered in mice lacking synapsin I. *Cell* 1993;75:661–670.

4. Rosahl TW, Spillane D, Missler M, et al. Essential functions of synapsins I and II in synaptic vesicle regulation. *Nature* 1995;375:488–493.

5. Geppert M, Goda Y, Stevens CF, et al. The small GTP-binding protein Rab3A regulates a late step in synaptic vesicle fusion. *Nature* 1997;387:810–814.

6. Geppert M, Südhof TC. RAB3 and synaptotagmin: the yin and yang of synaptic membrane fusion. *Annu Rev Neurosci* 1998;21:75–95.

7. Markram H, Tsodyks M. Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* 1996;382:807–810.

8. Tsodyks M, Pawelzik K, Markram H. Neural networks with dynamic synapses. *Neural Comput* 1998;10: 821–835.

9. Buonomano DV, Merzenich MM. Temporal information transformed into a spatial code by a neural network with realistic properties. *Science* 1995;267:1028–1030.

10. Buonomano DV. Decoding temporal information: a model based on short-term synaptic plasticity. *J Neurosci* 2000;20:1129–1141.

11. Glantz LA, Lewis DA. Reduction of synaptophysin immunoreactivity in the prefrontal cortex of subjects with schizophrenia: regional and diagnostic specificity. *Arch Gen Psychiatry* 1997;54:660–669.

12. Bliss TV, Lomo T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol (Lond)* 1973;232:331–356.

13. Bliss TV, Gardner-Medwin AR. Long-lasting potentiation of synaptic transmission in the dentate area of the unanaestetized rabbit following stimulation of the perforant path. *J Physiol (Lond)* 1973;232:357–374.

14. Eichenbaum H. Spatial learning: the LTP-memory connection. *Nature* 1995;378: 131–132.

15. Eichenbaum H. Learning from LTP: a comment on recent attempts to identify cellular and molecular mechanisms of memory. *Learning Memory* 1996;3:61–73.

16. Crair MC, Malenka RC. A critical period for long-term potentiation at thalamocortical synapses. *Nature* 1995;375:325–328.

17. Singer W. Development and plasticity of cortical processing architectures. *Science* 1995;270:758–764.

18. Katz LC, Shatz CJ. Synaptic activity and the construction of cortical circuits. *Science* 1996;274:1133–1138.

19. Zola-Morgan S, Squire LR. Neuroanatomy of memory. *Annu Rev Neurosci* 1993;16:547–563.

20. Nicoll RA, Kauer JA, Malenka RC. The current excitement in long-term potentiation. *Neuron* 1988;1:97–103.

21. Bliss TV, Collingridge GL. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* 1993;361:31–39.

22. Collingridge GL, Kehl SJ, McLennan H. The antagonism of amino acid-induced excitations of rat hippocampal CA1 neurones *in vitro. J Physiol (Lond)* 1983;334:19–31.

23. Malenka RC, Nicoll RA. Long-term potentiation: a decade of progress? *Science* 1999;285:1870–1874.

24. Lynch G, Larson J, Kelso S, et al. Intracellular injections of EGTA block induction of hippocampal long-term potentiation. *Nature* 1983;305:719–721.

25. Malenka RC, Kauer JA, Zucker RS, et al. Postsynaptic calcium is sufficient for potentiation of hippocampal synaptic transmission. *Science* 1988;242:81–84.

26. Yang SN, Tang YG, Zucker RS. Selective induction of LTP and LTD by postsynaptic $[Ca^{2+}]i$ elevation. *J Neurophysiol* 1999;81:781–787.

27. Malenka RC, Lancaster B, Zucker RS. Temporal limits on the rise in postsynaptic calcium required for the induction of long-term potentiation. *Neuron* 1992;9:121–128.

28. Sanes JR, Lichtman JW. Can molecules explain long-term potentiation? *Nat Neurosci* 1999;2:597–604.

29. Malenka RC, Kauer JA, Perkel DJ, et al. An essential role for postsynaptic calmodulin and protein kinase activity in long-term potentiation. *Nature* 1989;340:554–557.

30. Malinow R, Schulman H, Tsien RW. Inhibition of postsynaptic PKC or CaMKII blocks induction but not expression of LTP. *Science* 1989;245:862–866.

31. Silva AJ, Wang Y, Paylor R, et al. Alpha calcium/calmodulin kinase II mutant mice: deficient long-term potentiation and impaired spatial learning. *Cold Spring Harb Symp Quant Biol* 1992;57:527–539.

32. Pettit DL, Perlman S, Malinow R. Potentiated transmission and prevention of further LTP by increased CaMKII activity in postsynaptic hippocampal slice neurons. *Science* 1994;266:1881–1885.

33. Lledo PM, Hjelmstad GO, Mukherji S, et al. Calcium/calmodulin-dependent kinase II and long-term potentiation enhance synaptic transmission by the same mechanism. *Proc Natl Acad Sci USA* 1995;92:11175–11179.

34. Fukunaga K, Muller D, Miyamoto E. Increased phosphorylation of Ca^{2+}/calmodulin-dependent protein kinase II and its endogenous substrates in the induction of long-term potentiation. *J Biol Chem* 1995;270:6119–6124.

35. Barria A, Muller D, Derkach V, et al. Regulatory phosphorylation of AMPA-type glutamate receptors by CaM-KII during long-term potentiation *Science* 1997;276:2042–2045.

36. Giese KP, Fedorov NB, Filipkowski RK, et al. Autophosphorylation at Thr286 of the alpha calcium-calmodulin kinase II in LTP and learning. *Science* 1998;279:870–873.

37. Blitzer RD, Connor JH, Brown GP, et al. Gating of CaMKII by cAMP-regulated protein phosphatase activity during LTP. *Science* 1998;280:1940–1942.

38. Makhinson M, Chotiner JK, Watson JB, et al. Adenylyl cyclase activation modulates activity-dependent changes in synaptic strength and Ca^{2+}/calmodulin-dependent kinase II autophosphorylation. *J Neurosci* 1999;19:2500–2510.

39. Nicoll RA, Malenka RC. Expression mechanisms underlying NMDA receptor-dependent long-term potentiation. *Ann NY Acad Sci* 1999;868:515–525.

40. Lüscher C, Malenka RC, Nicoll RA. Monitoring glutamate release during LTP with glial transporter currents. *Neuron* 1998;21:435–441.

41. Diamond JS, Bergles DE, Jahr CE. Glutamate release monitored with astrocyte transporter currents during LTP. *Neuron* 1998;21:425–433.

42. Manabe T, Nicoll RA. Long-term potentiation: evidence against an increase in transmitter release probability in the CA1 region of the hippocampus. *Science* 1994;265:1888–1892.

43. Mainen ZF, Jia Z, Roder J, et al. Use-dependent AMPA receptor block in mice lacking GluR2 suggests postsynaptic site for LTP expression. *Nat Neurosci* 1998;1:579–586.

44. Oliet SH, Malenka RC, Nicoll RA. Bidirectional control of quantal size by synaptic activity in the hippocampus. *Science* 1996;271:1294–1297.

45. Manabe T, Renner P, Nicoll RA. Postsynaptic contribution to long-term potentiation revealed by the analysis of miniature synaptic currents. *Nature* 1992;355:50–55.

46. Wyllie DJ, Manabe T, Nicoll RA. A rise in postsynaptic Ca^{2+} potentiates miniature excitatory postsynaptic currents and AMPA responses in hippocampal neurons. *Neuron* 1994;12:127–138.

47. Davies SN, Lester RA, Reymann KG, et al. Temporally distinct pre- and post-synaptic mechanisms maintain long-term potentiation. *Nature* 1989;338:500–503.

48. Barria A, Derkach V, Soderling T. Identification of the Ca^{2+}/calmodulin-dependent protein kinase II regulatory phosphorylation site in the alpha-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate-type glutamate receptor. *J Biol Chem* 1997;272:32727–32730.

49. Mammen AL, Kameyama K, Roche KW, et al. Phosphorylation of the alpha-amino-3-hydroxy-5-methylisoxazole4-propionic acid receptor GluR1 subunit by calcium/calmodulin-dependent kinase II. *J Biol Chem* 1997;272:32528–32533.

50. Derkach V, Barria A, Soderling TR. Ca^{2+}/calmodulin-kinase II enhances channel conductance of alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionate type glutamate receptors. *Proc Natl Acad Sci USA* 1999;96:3269–3274.

51. Benke TA, Luthi A, Isaac JT, et al. Modulation of AMPA receptor unitary conductance by synaptic activity. *Nature* 1998;393:793–797.

52. Zamanillo D, Sprengel R, Hvalby O, et al. Importance of AMPA receptors for hippocampal synaptic plasticity but not for spatial learning. *Science* 1999;284:1805–1811.

53. Kullmann DM, Siegelbaum SA. The site of expression of NMDA receptor-dependent LTP: new fuel for an old fire. *Neuron* 1995;15:997–1002.

54. Malenka RC, Nicoll RA. Silent synapses speak up. *Neuron* 1997;19:473–476.

55. Isaac JT, Nicoll RA, Malenka RC. Evidence for silent synapses: implications for the expression of LTP. *Neuron* 1995;15:427–434.

56. Liao D, Hessler NA, Malinow R. Activation of postsynaptically silent synapses during pairing-induced LTP in CA1 region of hippocampal slice. *Nature* 1995;375:400–404.

57. Shi SH, Hayashi Y, Petralia RS, et al. Rapid spine delivery and redistribution of AMPA receptors after synaptic NMDA receptor activation. *Science* 1999;284:1811–1816.

58. Hayashi Y, Shi SH, Esteban JA, et al. Driving AMPA receptors into synapses by LTP and CaMKII: requirement for GluR1 and PDZ domain interaction. *Science* 2000;287:2262–2267.

59. O'Brien RJ, Lau LF, Huganir RL. Molecular mechanisms of glutamate receptor clustering at excitatory synapses. *Curr Opin Neurobiol* 1998;8:364–369.

60. Lledo PM, Zhang X, Sudhof TC, et al. Postsynaptic membrane fusion and long-term potentiation. *Science* 1998;279:399–403.

61. Braithwaite SP, Meyer G, Henley JM. Interactions between AMPA receptors and intracellular proteins. *Neuropharmacology* 2000;39:919–930.

62. Linden DJ. Long-term synaptic depression in the mammalian brain. *Neuron* 1994;12:457–472.

63. Bear MF, Abraham WC. Long-term depression in hippocampus. *Annu Rev Neurosci* 1996;19:437–462.

64. Bear MF, Malenka RC. Synaptic plasticity: LTP and LTD. *Curr Opin Neurobiol* 1994;4:389–399.

65. Dudek SM, Bear MF. Homosynaptic long-term depression in area CA1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade. *Proc Natl Acad Sci USA* 1992;89:4363–4367.

66. Mulkey RM, Malenka RC. Mechanisms underlying induction of homosynaptic long-term depression in area CA1 of the hippocampus. *Neuron* 1992;9:967–975.

67. Cummings JA, Mulkey RM, Nicoll RA, et al. Ca^{2+} signaling requirements for long-term depression in the hippocampus. *Neuron* 1996;16:825–833.

68. Malenka RC. Postsynaptic factors control the duration of synaptic enhancement in area CA1 of the hippocampus. *Neuron* 1991;6:53–60.

69. Shenolikar S, Nairn AC. Protein phosphatases: recent progress. *Adv Second Messenger Phosphoprotein Res* 1991;23:1–121.

70. Lisman J. A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory. *Proc Natl Acad Sci USA* 1989;86:9574–9578.

71. Mulkey RM, Herron CE, Malenka RC. An essential role for protein phosphatases in hippocampal long-term depression. *Science* 1993;261:1051–1055.

72. Mulkey RM, Endo S, Shenolikar S, et al. Involvement of a calcineurin/inhibitor-1 phosphatase cascade in hippocampal long-term depression. *Nature* 1994;369:486–488.

73. Lissin DV, Carroll RC, Nicoll RA, et al. Rapid, activation-induced redistribution of ionotropic glutamate receptors in cultured hippocampal neurons. *J Neurosci* 1999;19:1263–1272.

74. Carroll RC, Beattie EC, Xia H, et al. Dynamin-dependent endocytosis of ionotropic glutamate receptors. *Proc Natl Acad Sci USA* 1999;96: 4112–4117.

75. Carroll RC, Lissin DV, von Zastrow M, et al. Rapid redistribution of glutamate receptors contributes to long-term depression in hippocampal cultures. *Nat Neurosci* 1999;2:454–460.

76. Lüscher C, Xia H, Beattie EC, et al. Role of AMPA receptor cycling in synaptic transmission and plasticity. *Neuron* 1999;24:649–658.

77. Harris KM, Kater SB. Dendritic spines: cellular specializations imparting both stability and flexibility to synaptic function. *Annu Rev Neurosci* 1994;17:341–371.

78. Kaech S, Brinkhaus H, Matus A. Volatile anesthetics block actin-based motility in dendritic spines. *Proc Natl Acad Sci USA* 1999;96:10433–10437.

79. Fischer M, Kaech S, Knutti D, et al. Rapid actin-based plasticity in dendritic spines. *Neuron* 1998;20:847–854.

80. Lendvai B, Stern EA, Chen B, et al. Experience-dependent plasticity of dendritic spines in the developing rat barrel cortex in vivo. *Nature* 2000;404:876–881.

81. Maletic-Savatic M, Malinow R, Svoboda K. Rapid dendritic morphogenesis in CA1 hippocampal dendrites induced by synaptic activity. *Science* 1999;283:1923–1927.

82. Engert F, Bonhoeffer T. Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *Nature* 1999;399:66–70.

83. Halpain S, Hipolito A, Saffer L. Regulation of F-actin stability in dendritic spines by glutamate receptors and calcineurin. *J Neurosci* 1998;18:9835–9844.

84. Lüscher C, Nicoll RA, Malenka RC, et al. Synaptic plasticity and dynamic modulation of the postsynaptic membrane. *Nat Neurosci* 2000;3:545–550.

85. Buchs PA, Muller D. Induction of long-term potentiation is associated with major ultrastructural changes of activated synapses. *Proc Natl Acad Sci USA* 1996;93:8040–8045.

86. Toni N, Buchs PA, Nikonenko I, et al. LTP promotes formation of multiple spine synapses between a single axon terminal and a dendrite. *Nature* 1999;402:421–425.

87. Edwards FA. Anatomy and electrophysiology of fast central synapses lead to a structural model for long-term potentiation. *Physiol Rev* 1995;75:759–787.

88. Carlin RK, Siekevitz P. Plasticity in the central nervous system: do synapses divide? *Proc Natl Acad Sci USA* 1983;80:3517–3521.

89. Frey U, Morris RG. Synaptic tagging: implications for late maintenance of hippocampal long-term potentiation. *Trends Neurosci* 1998;21:181–188.

90. Lisman JE, Fallon JR. What maintains memories? *Science* 1999;283:339–340.Kandel ER, Pittenger C. The past, the future and the biology of memory storage. *Philos Trans R Soc Lond B Biol Sci* 1999;354:2027–2052.

92. Berke JD, Hyman SE. Addiction, dopamine, and the molecular mechanisms of memory. *Neuron* 2000;25:515–532.

12

Gaba

Richard W. Olsen

Richard W. Olsen: Department of Molecular and Medical Pharmacology, University of California Los Angeles School of Medicine, Los Angeles, California.

GABA IS THE MAJOR INHIBITORY NEUROTRANSMITTER IN THE NERVOUS SYSTEM

Part of "12 - Gaba "

Several amino acids are found in high concentrations in brain, and some have been established as neurotransmitters. *L-Glutamic acid (glutamate)* is the major neurotransmitter for fast excitatory synaptic transmission, whereas *γ-aminobutyric acid (GABA)* is the major neurotransmitter for fast inhibitory synaptic transmission. Glycine is a secondary rapid inhibitory neurotransmitter, especially in the spinal cord (1 ,2). Because of the widespread presence and utilization of glutamate and GABA as transmitters, one could say that they are involved in all functions of the central nervous system (CNS), as well as in all diseases. At any point in the CNS, one is either at a cell that uses or responds to glutamate and GABA or no more than one cell removed. Many clinical conditions including psychiatric disorders appear to involve an imbalance in excitation and inhibition, and therapeutics thus involve attempts to restore the balance. The GABA system is the target of a wide range of drugs active on the CNS, including anxiolytics, sedative-hypnotics, general anesthetics, and anticonvulsants (3). See the chapters on GABA in previous editions of this book (1 ,4).

Since its discovery in the CNS in the early 1950s (5 ,6), GABA was shown to fulfill the criteria for establishment as a neurotransmitter (Fig. 12.1). It is synthesized by a specific enzyme, L-glutamic acid decarboxylase (GAD), in one step from L-glutamate. Thus, in addition to its role in protein synthesis, in cofactors such as folic acid and in hormones such as thyrotropin-releasing hormone, and its action as a neurotransmitter itself, glutamate must be available in certain nerve endings for biosynthesis of GABA. Much of the glutamate and GABA used as neurotransmitter is derived from glial storage pools of glutamine (2 ,6). Two genes for GAD have been cloned, and the two forms of the enzyme are proposed to differ in their affinity for the cofactor pyridoxal phosphate and the subcellular localization (7). GABA was shown to be released from electrically stimulated inhibitory nerve cells (8), and a mechanism of rapid removal from the synaptic release site was demonstrated by identification of high-affinity transporter proteins (9 ,10). The application of GABA and structural analogues to cells innervated by GABAergic neurons produces effects on that target cell identical to those produced by stimulating the inhibitory innervation (11).

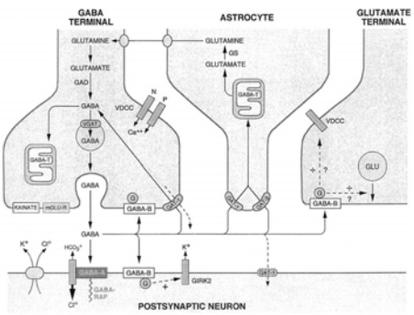


FIGURE 12.1. Schematic GABA synapse. Diagram showing the main features of the GABA synapse. Transporters are indicated by oval symbols, receptors and ion channels by rectangular symbols. **A:** Transporters: GAT-1, GAT-3, plasma membrane GABA transporters; VGAT, vesicular GABA transporter. **B:** Receptors: GABA-A, ionotropic GABA receptor; GABA-B, G-protein–coupled GABA receptor; KAINATE, presynaptic kainate receptor; MGLUR, metabotropic glutamate receptor. **C:** Ion channels: GIRK2, G-protein–coupled inwardly rectifying K^+ channel; VDCC: voltage-dependent calcium channel. **D:** Enzymes: GABA-T, GABA transaminase; GAD, glutamic acid decarboxylase; GS, glutamate synthetase. (Courtesy of O.P. Ottersen; design G. Lothe.)

PHYSIOLOGY AND PHARMACOLOGY OF $GABA_A$, $GABA_B$, AND $GABA_C$ RECEPTORS

Part of "12 - Gaba "

GABA-mediated synaptic inhibition involves rapid, less than 100-millisecond, inhibitory postsynaptic potentials and slower, more than 100-millisecond, inhibitory postsynaptic potentials. The former were shown by voltage clamp to involve increased chloride ion permeability and to be blocked by the plant convulsant drug picrotoxin, as seen with GABA action in invertebrates, such as crayfish muscle and nerve preparations (12). The rapid chloride current defined a physiologic receptor mechanism termed the *$GABA_A$receptor*, also pharmacologically defined by the antagonist bicuculline, as well as picrotoxin, and the agonist muscimol (Fig. 12.2). Thus, the $GABA_A$ receptor is a chloride channel regulated by GABA binding, and it is now grouped in the superfamily of ligand-gated ion channel receptors, which includes the well-characterized nicotinic acetylcholine receptor, present at the skeletal neuromuscular junction (13 ,14).

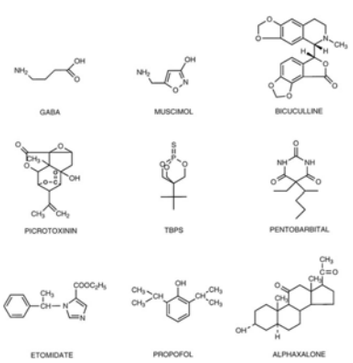


FIGURE 12.2. Chemical structures of $GABA_A$ receptor drugs.

Chloride channel gating is generally inhibitory on a neuron by virtue of stabilizing the membrane potential near the resting level. However, under conditions of high intracellular chloride, for example, in immature neurons with low capacity to maintain a chloride gradient, increasing chloride permeability can depolarize the membrane potential. This depolarization could be sufficient to fire the cell, and it would be likely to activate certain voltage-gated ion channels, including calcium, that can, in turn, regulate other cellular events. Variable permeability to bicarbonate ions for some subtypes of $GABA_A$ receptor (GABAR) could

also play a role in depolarization (15). Such depolarizing GABAR action has been proposed as an important excitatory system in developing brain (16), and it may explain the well-known trophic action of GABA to promote both survival and differentiation during development (17).

The slow inhibitory polysynaptic potentials were shown to be insensitive to $GABA_A$ drugs such as bicuculline, but to be activated by ß-chlorophenyl GABA (the antispastic drug baclofen) and to be mediated by a G-protein–coupled receptor that increases potassium conductance (18), now called the $GABA_B$ receptor. A further inhibitory GABA response was observed in some cells to be "non-A, non-B," neither bicuculline nor baclofen sensitive and sometimes called $GABA_C$ (19), and generally sensitive to the GABA analogue *cis*-aminocrotonic acid. $GABA_C$–type inhibition was shown to involve a rapid chloride conductance, as with $GABA_A$ receptors; however, it was not only insensitive to bicuculline, but also not modified by other $GABA_A$ drugs, such as benzodiazepines and anesthetics (19). The eventual cloning of a retinal-specific subunit cDNA ρ that produced bicuculline-insensitive GABA chloride channels appeared to account for $GABA_C$ receptors (20). However, because of the structural and functional homology with $GABA_A$ receptors, the International Union of Pharmacology subcommittee on nomenclature recommended that these ρ receptors not be called $GABA_C$ receptors, but rather a subtype of pharmacologically unique $GABA_A$ receptors (21).

$GABA_B$ receptors were shown to mediate presynaptic inhibition on some nerve endings and postsynaptic inhibition on some cell bodies or dendrites. The coupling mechanism depends on the cell location, because several G-protein–coupled effectors can be used, involving negative modulation of adenylate kinase and negative modulation of inositol tris phosphate production. These lead to activation of potassium channels or inhibition of voltage-gated calcium channels (22). Presynaptic inhibition of GABA release

by GABA involves $GABA_B$ autoreceptors, and their activation would be overall excitatory, as opposed to inhibition of glutamate release, which would be overall inhibitory. Considerable effort was therefore expended to determine whether different $GABA_B$ receptors could mediate these very different functions, possibly allowing the development of receptor subtype-specific drugs. Although some classic pharmacology studies supported this hypothesis (18, 22), it was the long-awaited cloning of the $GABA_B$ receptor (23) that established the true situation. The first receptor exists as two splice variants, and additional clones for $GABA_B$ receptor subtype genes have been isolated. Surprisingly, the $GABA_B$ receptors appear to exist as heterodimers, previously unknown for G-protein–coupled receptors. The dimers produce the diverse pharmacologic specificity for the GABA site and the diverse coupling mechanisms observed in nature (24). It seems that the pharmacology of $GABA_B$ receptors is in a very promising infancy.

STRUCTURE AND FUNCTION OF $GABA_A$ RECEPTORS

Part of "12 - Gaba "

The GABARs are the major players in CNS function and relevance to psychopharmacology. These receptors, defined by pharmacologists using electrophysiologic and other techniques (14 ,22), were identified in brain homogenates by radioligand binding (25), and are shown to have the correct specificity for GABA analogues expected from the neuropharmacology (26 ,27). The GABAR protein (Fig. 12.3) also contains binding sites for benzodiazepines, picrotoxin, barbiturates, and other anesthetics, all of which allosterically interact with each other (28). One or more polypeptides of 45 to 60 kd on sodium dodecylsulfate–polyacrylamide gel

electrophoresis were identified in brain homogenates as constituents of the GABAR by photoaffinity labeling with the radioactive benzodiazepine flunitrazepam (29 ,30), and monoclonal antibodies were developed to the partially purified bovine receptor, which recognized the photolabeled peptides using Western blotting (31).

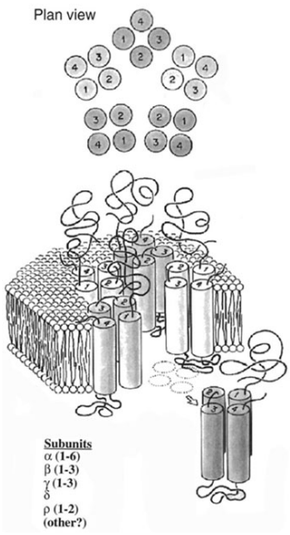


FIGURE 12.3. Schematic $GABA_a$ receptor structure. The chloride channel is shown as a pore in the center of five equivalent subunits, each with four membrane-spanning domains (see the isolated subunit at the **bottom**). Because of the existence of subunit families, many such heteropentamer combinations are possible, each with multiple drug sites. Ligand sites: GABA: agonists (muscimol), antagonists (bicuculline); Benzodiazepine: agonists (flunitrazepam), antagonists (flumazenil), inverse agonists (DMCM); Picrotoxin/Convulsant (TBPS); Barbituate (phenobarbitol); Steroid (alphaxalone, allopregnanolone); Volatile Anesthetic (halothane). (Modified from Olsen RW, Tobin AJ. Molecular biology of $GABA_A$ receptors. *FASEB J* 1990;4:1469–1480, with permission.)

The GABAR proteins were purified using benzodiazepine affinity chromatography (32), which allowed partial protein sequencing and expression cloning of two receptor genes (13). GABA-activated currents were demonstrated in *Xenopus* oocytes using cDNAs for two polypeptides that contained the partial sequences within their coded sequence, and these were designated α and ß. At first, these were thought (incorrectly) to correspond to the two bands seen in the purified protein (32). These two subunits were related to each other and also to the nicotinic acetylcholine receptor family of subunits, a finding indicating a superfamily of receptor polypeptide genes and a likely heteropentameric structure (Fig. 12.3) (13 ,14). These two cDNAs were used as probes to clone additional family members with more or less sequence homology to the first two. Those with high homology were named with the same Greek letter, whereas those with less homology were given other Greek letters. The current repertoire involves α1 to 6, ß1 to 3, γ1 to 3, δ, ε, θ, π, and ρ1 to 3 (21). There are also a few splice variants; for example, γ2 exists in two forms differing in an eight-amino acid insert in the intracellular loop that includes a substrate serine for protein kinase C (33). All the subunits are related to each other and have molecular weights of about 50 kd. The purified receptor protein thus actually contains about a dozen subunit polypeptides, of varying amount (6). Hydropathy plots show that they have a long extracellular N-terminal domain, which has glycosylation sites and is believed to carry the GABA binding site. They have four membrane-spanning domains (M1 to M4) of about 25 residues each, a long intracellular loop between M3 and M4, and a short extracellular C-terminal tail. These subunits are arranged as heteropentamers (Fig. 12.3), several of which are common in nature, but whose expression varies with both age and brain region. The different receptor subtypes have biological differences, such as location, affinity for GABA, and channel properties, as well as pharmacologic heterogeneity. Most receptors contain two copies of one type of α subunit, two copies of one type of ß subunit, and a γ subunit. Rarely, another subunit (δ, ε, θ) can substitute for γ (30 ,33). The presence of a γ subunit is needed for benzodiazepine sensitivity, and other subunits affect the detailed specificity. For example, the α subunits define the benzodiazepine pharmacology, and some subunits α4 and α6 do not bind classic benzodiazepine agonists; the detailed pharmacology depends on the small differences in polypeptide sequence for the various subunits (6 ,34 ,35 and 36). Because of the unique location of receptor subtypes, and thus unique functions of the circuits involved, great hope for new drugs of improved pharmacologic profile has been expressed. The GABAR strategy has certainly not been exhausted.

$GABA_A$ RECEPTORS ARE THE SITES OF ACTION OF BENZODIAZEPINES AND BARBITURATES

Part of "12 - Gaba "

The actions of several classes of CNS depressant drugs had for some time been suggested to involve enhancement of inhibitory synaptic transmission. In particular, the anxiolytic effects of benzodiazepines were shown probably to result from potentiation of GABA action (37 ,38). When the benzodiazepine receptors were discovered using radioligand binding to brain homogenates (1 ,4 ,39 ,40), it was quickly determined that the benzodiazepine binding sites were physically present on the $GABA_A$ receptor–chloride channel complex (28 ,41). The various types of drug binding site on the $GABA_A$ receptor allosterically interact with each other in the test tube. Barbiturates and related sedatives also enhance $GABA_A$ receptor–mediated inhibition, and their pharmacologic spectrum overlaps with that of the benzodiazepines and related substances, such as zolpidem, zopiclone, and abecarnil (Fig. 12.4). The selective actions of benzodiazepines not shown by barbiturates or vice versa are believed to arise from heterogeneity in GABA receptor sensitivity to the drugs, and corresponding heterogeneity in brain regions, circuits, and functions. Further, some GABARs are insensitive to benzodiazepines but not to barbiturates, as well as additional nonoverlapping, nonGABA actions of high doses, especially barbiturates. In addition, the two classes of drugs have a different mechanism of action at the molecular channel level; barbiturates prolong the lifetime of GABA currents, in addition to gating channels directly at high concentrations, whereas benzodiazepines increase the frequency of opening of GABAR channels and do not directly open channels in the absence of GABA (3 ,42).

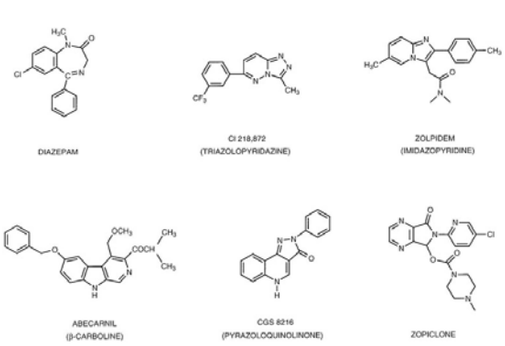


FIGURE 12.4. Chemical structures of drugs active at the benzodiazepine site on the $GABA_A$ receptor–chloride channel complex.

The classical benzodiazepines such as diazepam (Valium) have had a tremendous history in psychopharmacology, reaching tremendous sales, primarily for clinical anxiety (38 ,43 ,44 and 45). Other uses of benzodiazepines include sedation, muscle relaxation, and a significant utilization for treatment of panic (1 ,45). Various structural analogues were developed by numerous firms, with slight variations in pharmacokinetics and other details, and quite a few nonbenzodiazepine structures were discovered that act at the benzodiazepine site on the GABAR to enhance GABA-mediated inhibition (Fig. 12.4) (46). This group includes compounds called ß-carbolines, some of which were isolated from biological tissues (47). However, neither the ß-carbolines nor any peptides have been demonstrated to act as biological ligands at benzodiazepine receptor sites (45). Surprisingly, some ß-carbolines, and indeed, benzodiazepines and other types of chemical structures active on the benzodiazepine site, were found to have the opposite pharmacologic efficacy as classic benzodiazepine ligands such as diazepam; that is, they are anxiogenic and proconvulsant in animals and inhibit GABAR function in cells, while binding to the same sites as agonist benzodiazepine site ligands. These compounds were given the name *inverse agonists* (48 ,49).

Given this spectrum of efficacy, it would be expected that compounds with true antagonist efficacy would exist, and these were found, for example, Ro15-1788, or flumazenil (50). This compound does not affect GABAR function on its own, but it blocks the actions of both agonists to enhance GABAR function and inverse agonists to inhibit GABAR function. In animals, it also reverses the pharmacologic actions of both agonists and inverse agonists (50). Thus, an antagonist can be used to treat overdose of agonist, or inverse agonist, and it triggers withdrawal in individuals treated on a long-term basis with agonists (45 ,51). Flumazenil administration to rats after long-term administration of diazepam was found to reverse tolerance rapidly and permanently, and treated animals showed no long-lasting effects but resembled treatment-naive animals (52 ,53). This finding suggested that benzodiazepine antagonists may be useful in reversing benzodiazepine dependence and also potentially for other GABAR drugs, such as ethanol. This has not proved effective so far, however (43 ,45).

Certain benzodiazepines have considerable success in the treatment of some types of epilepsy (38). Every emergency medical cart contains injectable benzodiazepine (diazepam, clonazepam, lorazepam) for convulsions and status epilepticus. However, long-term therapy of epilepsy with benzodiazepines is often prevented by the development of tolerance to the anticonvulsant actions, without change in blood levels (54). The development of tolerance to long-term administration of benzodiazepines, and also of withdrawal signs (43 ,45 ,55), is consistent with the development of psychological and physical dependence with these drugs. The potential for abuse with CNS depressant drugs in general and benzodiazepines in particular is well known, as is the interaction with ethanol. This has led to a considerable drop in prescriptions of these agents for routine anxiety. Because the danger of fatal overdose with benzodiazepines is lower than that of ethanol and barbiturates, and because withdrawal symptoms are less dangerous for benzodiazepines than for alcohol, benzodiazepines reached considerable popularity in treatment of alcoholism. However, the two drugs show cross-tolerance and cross-dependence, so substitution of benzodiazepines for ethanol is merely substituting one addiction for another (55).

Conversely, an interesting observation was made with the benzodiazepine partial inverse agonist Ro15-4513. This compound was found to antagonize the behavioral effects of ethanol (49), as well as the *in vitro* action of ethanol to enhance GABAR function (56). (Ethanol and GABA are discussed further later.) Moreover, the action of Ro15-4513 to antagonize ethanol occurred under conditions of assay, such as behavior, tissue, or species, in which Ro15-4513 itself did not exhibit inverse agonist activity or inhibit GABAR function, nor did it reverse the actions of pentobarbital (56 ,57 ,58 and 59). Thus, this compound or one like it had potential as an “alcohol antidote” in humans, by reducing intoxication and perhaps withdrawal and craving. Unfortunately, the ethical decisions involved in prescribing such a drug were made moot by discovery that Ro15-4513 was tremorigenic and proconvulsant in nonhuman primates, as well as other animals (60).

Understanding the mechanism of tolerance development has been a research topic of high interest, especially for epilepsy treatment, but also because of the relevance to brain plasticity. Whereas long-term administration of benzodiazepines may produce tolerance in part by down-regulation of receptor levels, considerable evidence suggests that receptors are not removed, but rather are altered in some way to produce tolerance (61 ,62 ,63 and 64). Besides tolerance development to long-term use of agonist benzodiazepines, sensitization to the actions of inverse agonists is observed; that is, excitatory benzodiazepine receptor ligands become more efficacious (65). This may resemble the kindling process seen with long-term administration of inverse agonists; that is, repeated administration of nonconvulsant doses of inverse agonists eventually leads to convulsions to that dose. This resembles the electrical kindling model of epilepsy, in which repeated electrical stimuli with nonconvulsive amplitude eventually evoke a seizure (66). Thus, long-term administration of benzodiazepine agonists or inverse agonists may shift the set point of the GABAR toward the excitatory or lower functional end of the spectrum (65 ,67). Dependence on benzodiazepines and alcohol resulting from long-term administration (abuse) may be exacerbated by a kindling-like development of increased severity of withdrawal symptoms, with an increased risk of relapse (68). Another aspect of the tolerance model is the possibility of replacing one type of GABAR subunit with another that still responds to GABA

but not to the chronically administered modulatory drug (69 ,70).

$GABA_A$ RECEPTORS ARE THE TARGETS OF ALCOHOL, GENERAL ANESTHETICS, AND NEUROSTEROIDS

Part of "12 - Gaba "

Alcohols are CNS depressants with a pharmacologic spectrum of action overlapping those of the benzodiazepines and barbiturates, known to act by enhancement of GABAR. Long-chain alcohols have anesthetic activity, as does ethanol at high doses (greater than 100 mM), whereas the intoxicating effects at lower concentrations (10 to 100 mM) have been suggested to involve blockade of *N*-methyl-D-aspartate (NMDA)–type glutamate receptors (71) or enhancement of GABAR (72 ,73 and 74). Because the latter effect varies considerably among, for example, laboratories, preparations, assays, and brain regions, unique ethanol-sensitive subtypes of GABAR were suggested, but they have not been established. Alternatively, and most popular currently, is the hypothesis that ethanol acts on GABAR indirectly to produce important aspects of its pharmacologic actions in cells and in animals (75). For example, ethanol may interact with membrane signaling proteins that regulate GABAR and NMDA receptors.

$GABA_A$ receptor function appears to be modulated by an endogenous substance: not a benzodiazepine-like or a picrotoxin-like peptide, but a barbiturate-like steroid. The neurosteroids are endogenous steroid hormone metabolites that have direct and rapid actions on cells not involving steroid hormone receptors or regulation of gene expression. Progesterone was shown to produce rapid sedative activity, a finding that led to the development of the clinical intravenous steroid anesthetic, alphaxalone. Progesterone has anxiolytic and anticonvulsant activity; discontinuation after long-term administration leads to withdrawal signs that are clearly CNS mediated: these actions are mediated by the progesterone metabolite, produced primarily in the adrenals but to some extent in brain, 3α-hydroxy-5-α-pregnane-20-one (76 ,77 and 78). The neuroactive steroids act principally by binding directly to membrane $GABA_A$ receptors and enhancing their function in a manner resembling the barbiturates (79 ,80).

Many related steroid compounds have been developed as lead compounds for potential use as antiepileptics, anxiolytics, and sedative-hypnotics (81). Whether these compounds are biologically relevant is uncertain, but this is suggested by considerable evidence. Endogenous steroids reach levels sufficient to modulate $GABA_A$ receptors during conditions of stress and anxiety, and during pregnancy (82 ,83). These compounds are probably involved in CNS plasticity responses to chronic stress and possibly epileptogenesis, and even drug dependence (84 ,85). The progesterone metabolite is the endogenous steroid that appears to be the most likely to be biologically relevant, but metabolites of testosterone and cortisone are also active (77 ,81). Pregnenolone sulfate, a biosynthetic intermediate in the synthesis of all the steroid hormones, present in high levels in the CNS, has weak activity as an antagonist of GABA function, but this appears to involve another mechanism and is unlikely to be biological (85). Neurosteroid action apparently has relevance to alcohol action. GABA-active steroids can substitute for ethanol in discriminative stimulus testing in rats and monkeys, and neurosteroids are synthesized in brain in response to ethanol administration and may mediate some of the pharmacologic actions (86). The neurosteroid-GABA connection potentially may be fruitful for new applications in psychopharmacology. As the endogenous functions of neurosteroids in stress control, seizure protection, attention and learning, and possibly even sleep, become better delineated, additional therapeutic approaches may arise.

Enhancement of $GABA_A$ receptor-mediated inhibition is currently the major candidate molecular mechanism for a generalized theory of general anesthesia. Everyone agrees that the anesthetic action of the steroid alphaxalone occurs by enhancement of GABAR (84 ,85), and many investigators believe that the actions of high-dose ethanol and other alcohols as anesthetics probably do also (75 ,87). Further, the sedative-hypnotic effects, and possibly anesthetic effects, of barbiturates and related drugs are considered to act through GABAR (88). Anesthetics are now believed to have a greater effect on membrane ion channels than on many other biological systems and to affect synaptic transmission more potently than nerve conduction. Ligand-gated ion channels, especially receptors for glutamate, glycine, and GABA, are most sensitive (89). All general anesthetics enhance GABA function at anesthetic concentrations (36 ,75). The ketamine-phencyclidine category of dissociative anesthetics enhances some GABA synapses, but these agents probably inhibit NMDA receptors more potently; further, they produce a different sort of anesthesia (90).

The *Meyer-Overton hypothesis* shows a high correlation for many drugs with respect to potency as a general anesthetic and partition in an oil-water biphasic system. The Meyer-Overton correlation has been found wanting, because of the existence of compounds with identical lipid solubility (oil-water partition coefficient), boiling point, and dipole moment, such as halogenated cyclobutane isomers, that differ in anesthetic potency: only the anesthetic isomers enhance $GABA_A$ receptors (91). Volatile anesthetics and alcohols (87), as well as intravenous agents such as barbiturates, propofol, neuroactive steroids, and etomidate, are all able at anesthetic concentrations to modulate $GABA_A$ receptor binding assays *in vitro* as well as to enhance $GABA_A$ receptor function in cells (36 ,88).

$GABA_A$ RECEPTORS ARE THE TARGETS OF MANY CNS EXCITANTS

Part of "12 - Gaba "

Many naturally occurring and synthetic convulsive agents are blockers of GABA-mediated inhibition (46). The prototypic $GABA_A$ channel blocker picrotoxinin (Fig. 12.2) is isolated from plants of the moonseed family, Menispermaceae, and its close relatives tutin and coriamyrtin, from the New Zealand tutu plant *Coriaria arborea* (92), known as a loco weed, which causes occasional poisonings in cows and even in people. A major category of synthetic potent neurotoxic chemicals (93), comprising the cage convulsants, was discovered to consist of noncompetitive $GABA_A$ receptor antagonists acting at the picrotoxinin site (93 ,94 and 95). One of these drugs, *t*-butyl bicylcophosphorothionate (Fig. 12.2), is a major research tool used to assay GABA receptors by radioligand binding (96). Synthetic butyrolactones with depressant and excitatory actions have also been described for the picrotoxinin site (97). In addition, this drug target appears to be the site of action of the experimental convulsant pentylenetetrazol (PTZ) and numerous polychlorinated hydrocarbon insecticides, including dieldrin, α-endosulfan, and lindane (93). The monoterpenoid thujone is the active constituent of oil of wormwood, the major ingredient of the famous green liqueur, absinthe, outlawed in about 1910. Absinthe was reputed to have hallucinogenic action and to be an inspiration for *fin de siècle* French artists and poets (92). Oil of wormwood has a history as a medicinal herb for treating intestinal worms and killing insects, and thujone is known to cause convulsions in high doses; thujone was demonstrated to be a $GABA_A$ receptor channel blocker like picrotoxinin (98). It remains anecdotal whether thujone/wormwood/absinthe produces psychic actions additional to those of the ethanol in the liqueur.

GABA-blocking agents thus have potential pharmacologic utility as excitants. Although at one time listed in the *Merck Index* and in pharmacology textbooks as a "barbiturate overdose antidote," picrotoxin is too dangerous as a convulsant to attempt to find an appropriate dose in the clinic. PTZ and related agents are known to show anxiogenesis in low doses, but also proabsence seizures. An alerting, attention-activation mechanism may figure to promote learning and memory in certain tasks, that is, nootropism. Partial inverse agonists at the benzodiazepine site, such as Ro15-4513, have been considered as candidates for memory enhancement (38 ,99), as well as for actions as antagonists and possible anticraving, antiwithdrawal agents for the treatment of addiction to benzodiazepines, alcohol, and many other drugs of abuse, as discussed earlier (1 ,60 ,69).

GENETIC ENGINEERING AND PSYCHOPHARMACOLOGY

Part of "12 - Gaba "

Gene targeting and transgenic mice have demonstrated several important roles for GABA in the CNS. Knockouts of both GAD67 and $GABA_A$ receptor subunit ß3 lead to cleft palate and early neonatal lethality (100 ,101 and 102). GAD65 knockout mice show increased anxiety, increased sensitivity to benzodiazepines, and impaired developmental plasticity in the cortex (103 ,104). Epilepsy results from knockout of GAD65, GABAR ß3, and GABAR δ subunit. Other phenotypic deficits include motor incoordination, movement disorders, cognitive defects, and other CNS circuitry problems resulting from lack of inhibitory synaptic transmission. In particular, the GABAR ß3 subunit is implicated in the human genetic disease Angelman syndrome, associated with mutation in maternal chromosome 15q and typified by severe mental retardation, epilepsy, motor incoordination, and sleep disorder (105). Mice targeted for this subunit have a phenotype remarkably similar to Angelman syndrome, especially the epilepsy, but also including the cognitive, motor, and sleep impairment (106).

The γ2 subunit knockout shows early neonatal lethality (107), without cleft palate, involving impaired clustering of $GABA_A$ receptors at synapses (108). Even heterozygotes, with presumably a partial deficit of γ2-containing GABAR, have impaired synapses and overanxious and paranoid behavior (109). Because GABARs are important drug targets, some GABAR subunit knockout mice have impaired sensitivity to drugs, such as decreased response to benzodiazepines in γ2 homozygous knockouts (107). Increased response to benzodiazepines is seen in γ2 heterozygous knockouts or in γ2L null mutants (109, 110). Reduced sensitivity to anesthetics was seen in ß3 but not α6 knockouts (102), and reduced sensitivity to neuroactive steroids is observed in the δ subunit knockout (111). This finding may be interesting in light of the apparent biological role of the neurosteroids in normal CNS. Gene targeting in mice also has been employed to "knock in" a mutation of the α1 subunit H101N, which prevents benzodiazepine binding to GABAR containing this subunit (112). The resulting animals have greatly impaired sensitivity to the sedative but not the anxiolytic actions of the benzodiazepines, whereas anticonvulsant activity is partially reduced. This finding indicates that the subtypes of GABAR containing the α1 subunit and the brain circuits in which they function are the substrates for benzodiazepine-stimulated sedation, whereas other GABARs, containing α2, α3, and α5, with α2 the most abundant and the major candidate, subserve specifically the role of GABARs in anxiety pathways sensitive to benzodiazepine therapy. (The observations of Rudolph et al., 1999 (112) were verified by McKernan et al., 2000 (113) for the role of the α1 subunit in the sedative actions of benzodiazepines, and extended by Low et al., 2000 (114) for the role of the α2 subunit in the anxiolytic actions of benzodiazepines.) Thus, new biotechnology applied to drug development is continuing to make new advances in psychopharmacology based on this now relatively "old" or at least well-known neurotransmitter system, GABA.

ACKNOWLEDGMENTS

Part of "12 - Gaba "

This work was supported by National Institutes of Health grant NS35985.

REFERENCES

1. Paul SM. GABA and glycine. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven, 1995:87–94.

2. Olsen RW, DeLorey TM. GABA and glycine. In: Siegel GJ, Agranoff BW, Albers RW, et al., eds. *Basic neurochemistry*, sixth ed. New York: Lippincott Williams & Wilkins, 1999:335–346.

3. Macdonald RL, Olsen RW. $GABA_A$ receptor channels. *Annu Rev Neurosci* 1994;17: 569–602.

4. Enna SJ, Möhler H. γ-Aminobutyric acid (GABA) receptors and their association with benzodiazepine recognition sites. In: Meltzer HY, ed. *Psychopharmacology: the third generation of progress*. New York: Raven, 1987:265–272.

5. Roberts E. Disinhibition as an organizing principle in the nervous system: the role of the GABA system. In: Roberts E, Chase TN, Tower DB, eds. *GABA in nervous system function*. New York: Raven, 1976:515–539.

6. Martin DL, Olsen RW, eds. *GABA in the nervous system: the view at 50 years*. Philadelphia: Lippincott Williams & Wilkins, 2000.

7. Erlander MC, Tobin AJ. The structural and functional heterogeneity of glutamic acid decarboxylase: a review. *Neurochem Res* 1991;16:215–226.

8. Otsuka J, Iversen LL, Hall ZW, et al. Release of γ-aminobutyric acid from inhibitory nerves of lobster. *Proc Natl Acad Sci USA* 1966;56:1110–1115.

9. Guastella J, Nelson N, Nelson H, et al. Cloning and experession of a rat brain GABA transporter. *Science* 1990;249:1303–1306.

10. Borden LA. GABA transporter heterogeneity: pharmacology and cellular locations. *Neurochem Int* 1996;29:335–356.

11. Krnjevic K. Chemical nature of synaptic transmission in vertebrates. *Physiol Rev* 1974;54:418–540.

12. Takeuchi A. Studies of inhibitory effects of GABA in invertebrate nervous systems. In: Roberts E, Chase TN, Tower DB, eds. *GABA in nervous system function*. New York: Raven, 1976:255–267.

13. Schofield PR, Darlison MG, Fujita N, et al. Sequence and functional expression of the $GABA_A$ receptor shows a ligand gated receptor superfamily. *Nature* 1987;328:221–227.

14. Olsen RW, Tobin AJ. Molecular biology of $GABA_A$ receptors. *FASEB J* 1990;4:1469–1480.

15. Rivera C, Volpio J, Payne JA, et al. The K^+/Cl^- co-transporter KCC2 renders GABA hyperpolarizing during neuronal maturation. *Nature* 1999;397:251–254.

16. Ben-Ari Y, Khazipov R, Leinekugel X, et al. $GABA_A$, NMDA and AMPA receptors: a developmentally regulated “menage a trois.” *Trends Neurosci* 1997;20:523–529.

17. Belhage B, Hansen GH, Elster L, et al. Effects of γ-aminobutyric acid (GABA) on synaptogenesis and synaptic function. *Perspect Dev Neurobiol* 1998;5:235–246.

18. Bowery NG. $GABA_B$ receptor pharmacology. *Annu Rev Pharmacol Toxicol* 1993;33:109–117.

19. Johnston GAR. Molecular biology, pharmacology, and physiology of $GABA_C$ receptors. In: Enna SJ, Bowery NG, eds. *The GABA receptors*, second ed. Totowa, NJ: Humana, 1997:297–323.

20. Shimada S, Cutting G, Uhl GR. γ-Aminobutyric acid A or C receptors? GABA ρ1 receptor RNA induces bicuculline-, barbiturate-, and benzodiazepine-insensitive GABA responses in *Xenopus* oocytes. *Mol Pharmacol* 1992;41:683–687.

21. Barnard EA, Skolnick P, Olsen RW, et al. Sub-types of γ-aminobutyric $acid_A$ receptors: classification on the basis of subunit structure and receptor function. Int Union Pharmacol XV. *Pharmacol Rev* 1998;50:291–313.

22. Enna SJ, Bowery NG, eds. *The GABA receptors*, second ed. Totowa, NJ: Humana, 1997.

23. Kaupmann K, Huggel K, Heid J, et al. Expression cloning of $GABA_B$ receptors uncovers similarity to metabotropic glutamate receptors. *Nature* 1997;386:239–246.

24. Marshall FH, Jones KA, Kaupmann K, et al. $GABA_B$ receptors: the first 7TM heterodimers. *Trends Pharmacol Sci* 1999;20:396–399.

25. Zukin SR, Young AB, Snyder SH. γ-Aminobutyric acid binding to receptor sites in rat central nervous system. *Proc Natl Acad Sci USA* 1974;71:4802–4807.

26. Krogsgaard-Larsen P, Johnston GAR. Structure-activity studies on the inhibition of GABA binding to rat brain membranes by muscimol and related compounds. *J Neurochem* 1978;30:1377–1382.

27. Olsen RW, Venter JC, eds. *Benzodiazepine/GABA receptors and chloride channels: structural and functional properties, receptor biochemistry and methodology*, vol 5. New York: Alan R. Liss, 1986.

28. Olsen RW. Drug interactions at the GABA receptor-ionophore complex. *Annu Rev Pharmacol Toxicol* 1982;22:245–277.

29. Möhler H, Battersby MK, Richards JG. Benzodiazepine receptor protein identified and visualized in brain tissue by a photoaffinity label. *Proc Natl Acad Sci USA* 1980;77:1666–1670.

30. Sieghart W. Structure and pharmacology of $GABA_A$ receptor subtypes. *Pharmacol Rev* 1995;47:181–234.

31. Schoch P, Richards JG, Haring P, et al. Co-localization of $GABA_A$ receptor and benzodiazepine receptors in the brain shown by monoclonal antibodies. *Nature* 1985;314:168–171.

32. Sigel E, Stephenson FA, Mamalaki C, et al. A γ-aminobutryric acid/benzodiazepine receptor complex from bovine cerebral cortex: purification and partial characterization. *J Biol Chem* 1983;258:6965–6971.

33. McKernan RM, Whiting PJ. Which $GABA_A$ receptor subtypes really occur in the Brain? *Trends Neurosci* 1996;19:139–143.

34. Lüddens H, Korpi ER, Seeburg PH. $GABA_A$/benzodiazepine receptor heterogeneity: neurophysiological implications. *Neuropharmacology* 1995;34:245–254.

35. Sigel E, Buhr A. The benzodiazepine binding site of $GABA_A$ receptors. *Trends Pharmacol Sci* 1997;18:425–429.

36. Olsen RW. The molecular mechanism of action of general anesthetics: structural aspects of interactions with $GABA_A$ receptors. *Toxicol Lett* 1998;100–101:193–201.

37. Costa E, Greengard P, eds. Mechanism of action of benzodiazepines. *Adv Biochem Psychopharmacol* 1975;14:318.

38. Haefely W. Allosteric modulation of the $GABA_A$ receptor channel: a mechanism for interaction with a multitude of central nervous system functions. In: Möhler H, DaPrada M, eds. *The challenge of neuropharmacology*. Basel: Editiones Roche, 1994:15–39.

39. Squires RF, Braestrup C. Benzodiazepine receptors in rat brain. *Nature* 1977;266:732–734.

40. Möhler H, Okada T. Benzodiazepine receptor: demonstration in the central nervous system. *Science* 1977;198:849–851.

41. Tallman J, Thomas J, Gallager D. GABAergic modulation of benzodiazepine binding site sensitivity. *Nature* 1978;274:383–385.

42. Study RE, Barker JL. Diazepam and (-)-pentobarbital: fluctuation analysis reveals different mechanisms for potentiation of GABA responses in cultured central neurons. *Proc Natl Acad Sci USA* 1981;78:7180–7184.

43. Usdin E, Skolnick P, Tallman JF, et al., eds. *Pharmacology of benzodiazepines*. London: MacMillan, 1982.

44. Zorumski CF, Isenberg KE. Insights into the structure and function of GABA-benzodiazepine receptors: ion channels and psychiatry. *Am J Psychiatry* 1991;148:162–173.

45. Biggio G, Sanna E, Serra M, Costa E, eds. $GABA_A$ Receptors and anxiety. *Adv Biochem Psychopharmacol* 1995;48:459.

46. Olsen RW, Gordey M. $GABA_A$ receptor chloride ion channels. In: Endo M, Kurachi Y, Mishina M, eds. *Handbook of experimental pharmacology*, vol 147. Heidelberg: Springer, 2000:497–515.

47. Braestrup C, Nielsen M, Olsen CE. Urinary and brain beta-carboline-3-carboxylates as potent inhibitors of brain benzodiazepine receptors. *Proc Natl Acad Sci USA* 1980;77:2288–2292.

48. Braestrup C, Schmiechen R, Neff G, et al. Interaction of convulsive ligands with benzodiazepine receptors. *Science* 1982;216:1241–1243.

49. Bonetti EP, Burkhard WP, Gabl M, et al. Ro15-4513:Partial inverse agonist at the BZR and interaction with ethanol. *Pharmacol Biochem Behav* 1989;31:733–749.

50. Hunkeler W, Möhler H, Pieri L, et al. Selective antagonists of benzodiazepines. *Nature* 1981;290:414–516.

51. File SE, Pellow S. Do the intrinsic actions of benzodiazepine receptor antagonists imply the existence of an endogenous ligand for benzodiazepine receptors? *Adv Biochem Psychopharmacol* 1986;41:187–202.

52. Gonsalves SF, Gallager DW. Persistent reversal of tolerance to anticonvulsant effects and GABAergic subsensitivity by a single exposure to benzodiazepine antagonist during chronic benzodiazepine administration. *J Pharmacol Exp Ther* 1987;244:79–83.

53. Tietz EI, Zeng XJ, Chen S, et al. Antagonist-induced reversal of functional and structural measures of hippocampal benzodiazepine tolerance. *J Pharmacol Exp Ther* 1999;291:932–942.

54. Olsen RW. Antiepileptic actions of benzodiazepines. In: Faingold CL, Fromm GH, eds. Drugs for control of epilepsy: actions on neuronal networks involved in seizure disorders. Boca Raton, FL: CRC, 1991:463–476.

55. Trimble MR, ed. *Benzodiazepines divided: a multidisciplinary review*. New York: Wiley, 1983.

56. Suzdak PD, Glowa JR, Crawley JN, et al. A selective imidazobenzodiazepine antagonist of ethanol in the rat. *Science* 1986;234:1243–1247.

57. Hoffman PL, Tabakoff B, Szabo G, et al. Effect of an imidazobenzodiazepine, Ro15-4513, on the incoordination and hypothermia produced by ethanol and pentobarbital. *Life Sci* 1987;41:611–619.

58. Kulkarni SK, Ticku MK. Ro15-4513 but not FG7142 preferentially reverses anticonvulsant effects of ethanol against bicuculline and picrotoxin-induced convulsions in rats. *Pharmacol Biochem Behav* 1989;32:233–240.

59. Nutt DJ, Lister RG. Antagonizing the anticonvulsant effect of ethanol using drugs acting at the benzodiazepine/GABA receptor complex. *Pharmacol Biochem Behav* 1989;31:751–755.

60. Lister RG, Nutt DJ. Is Ro 15-4513 a specific alcohol antagonist? *Trends Neurosci* 1987;10:223–225.

61. Gallager DW, Laskoski JM, Gonsalves SF, et al. Chronic benzodiazepine treatment decreases postsynaptic GABA sensitivity. *Nature* 1984;308:74–77.

62. Tietz EI, Chiu TH, Rosenberg HC. Regional GABA/benzodiazepine/chloride channel coupling after acute and chronic benzodiazepine treatment. *Eur J Pharmacol* 1989;167:57–65.

63. Miller LG, Greenblatt DJ, Barnhill JG, et al. Chronic benzodiazepine administration. I. Tolerance is associated with benzodiazepine receptor down-regulation and decreased GABA receptor function. *J Pharmacol Exp Ther* 1988;246:170–177.

64. Poisbeau P, Williams SR, Mody I. Silent $GABA_A$ synapses during flurazepam withdrawal are region-specific in the hippocampal formation. *J Neurosci* 1997;17:3467–3475.

65. Little HJ, Nutt DJ, Taylor SC. Bidirectional effects of chronic treatment with agonists and inverse agonists at the benzodiazepine receptor. *Brain Res Bull* 1987;19:371–378.

66. Goddard GV. Development of epileptic seizures through brain stimulation at low intensity. *Nature* 1967;214:1020–1021.

67. Stephens DN, Schneider HH, Weidmann R, et al. Decreased sensitivity to benzodiazepine receptor agonists and increased sensitivity to inverse agonists following chronic treatments: evidence for separate mechanisms. *Adv Biochem Psychopharmacol* 1988;45:337–345.

68. Kokka N, Sapp DW, Taylor AN, et al. The kindling model of alcohol dependence: similar persistent reduction in seizure threshold to pentylenetetrazol in animals receiving chronic ethanol or chronic pentylenetetrazol. *Alcohol Clin Exp Res* 1993;17:525–531.

69. Costa E, Auta J, Caruncho H, et al. A search for a new anticonvulsant and anxiolytic benzodiazepine devoid of side effects and tolerance liability. *Adv Biochem Psychopharmacol* 1995;48:75–92.

70. Holt RA, Bateson AN, Martin IL. Chronic treatment with diazepam or abecarnil differentially affects the expression of $GABA_A$ receptor subunit mRNAs in the rat cortex. *Neuropharmacology* 1996;35:1457–1463.

71. Lovinger DM, White G, Weight FF. Ethanol inhibits NMDA-activated ion current in hippocampal neurons. *Science* 1989;243:1721–1724.

72. Suzdak PD, Schwartz RD, Skolnick P, et al. Ethanol stimulates GABA receptor-mediated chloride transport in rat brain synaptoneurosomes. *Proc Natl Acad Sci USA* 1986;83:4071–4075.

73. Allan AM, Harris RA. Acute and chronic ethanol treatments alter GABA receptor-operated chloride channels. *Pharmacol Biochem Behav* 1987;27:665–670.

74. Mehta AK, Ticku MK. Ethanol potentiation of GABAergic transmission in cultured spinal cord neurons involves $GABA_A$-gated chloride channels. *J Pharmacol Exp Ther* 1988;246:558–564.

75. Mihic SJ, Harris RA. Alcohol actions at the $GABA_A$ receptor/chloride channel complex. In: Dietrich RA, Erwin VG, eds. *Pharmacological effects of ethanol on the nervous system*. Boca Raton, FL: CRC, 1996:51–72.

76. Majewska MD, Harrison NL, Schwartz RD, et al. Steroid hormone metabolites are barbiturate-like modulators of the GABA receptor. *Science* 1986;232:1004–1007.

77. Simmonds MA. Modulation of the $GABA_A$ receptor by steroids. *Semin Neurosci* 1991;3:231–239.

78. Smith SS, Gong QH, Hsu FC, et al. $GABA_A$ receptor α4 subunit suppression prevents withdrawal properties of an endogenous steroid. *Nature* 1998;392:926–930.

79. Harrison NL, Vicini S, Barker JL. A steroid anesthetic prolongs inhibitory postsynaptic currents in cultured rat hippocampal neurons. *J Neurosci* 1987;7:604–609.

80. Lambert JJ, Belelli D, Hill-Venning C, et al. Neurosteroids and $GABA_A$ receptor function. *Trends Pharmacol Sci* 1995;16:295–303.

81. Gee KW, McCauley L, Lan NC. A putative receptor for neurosteroids on the $GABA_A$ receptor complex: the pharmacological properties and therapeutic potential of epalons. *Crit Rev Neurobiol* 1995;9:207–277.

82. Barbaccia ML, Roscetti G, Trabucchi M, et al. Time-dependent changes in rat brain neuroactive steroid concentrations and $GABA_A$ receptor function after acute stress. *Neuroendocrinology* 1996;63:166–172.

83. Concas A, Mostallino MC, Porcu P, et al. Role of brain allopregnanolone in the plasticity of $GABA_A$ receptor in rat brain during pregnancy and after delivery. *Proc Natl Acad Sci USA* 1998;95:13284–13289.

84. Paul SM, Purdy RH. Neuroactive steroids. *FASEB J* 1992;6:2311–2322.

85. Olsen RW, Sapp DW. Neuroactive steroid modulation of $GABA_A$ receptors. *Adv Biochem Psychopharmacol* 1995;48:57–74.

86. Morrow AL, Janis GC, VanDoren MJ, et al. Neurosteroids mediate pharmacological actions of ethanol: a new mechanism of ethanol action? *Alcohol Clin Exp Res* 1999;23:1933–1940.

87. Narahashi T, Arakawa O, Nakahiro M, et al. Effects of alcohols on ion channels of cultured neurons. *Ann NY Acad Sci* 1991;625:26–36.

88. Olsen RW, Fischer JB, Dunwiddie TV. Barbiturate enhancement of GABA receptor binding and function as a mechanism of anesthesia. In: Roth S, Miller K, eds. *Molecular and cellular mechanisms of anaesthetics*. New York: Plenum, 1986:165–177.

89. Franks NP, Lieb WR. Which molecular targets are most relevant to general anesthesia? *Toxicol Lett* 1998;100–101:1–8.

90. Murray TF. Basic pharmacology of ketamine. In: Bowdle TA, Horita A, Kharasch ED, eds. *The Pharmacologic basis of anesthesiology*. New York: Churchill Livingstone, 1994:337–355.

91. Mihic SJ, McQuilken SJ, Eger EI, et al. Potentiation of γ-aminobutyric acid type A receptor-mediated chloride currents by novel halogenated compounds correlates with their abilities to induce general anesthesia. *Mol Pharmacol* 1994;46:851–857.

92. Olsen RW. Absinthe and GABA receptors. *Proc Natl Acad Sci USA* 2000;97:4417–4418.

93. Casida JE, Palmer CJ. 2,6,7-Trioxabicyclo[2.2.2]octanes: chemistry, toxicology, and action at the GABA-gated chloride channel. *Adv Biochem Psychopharmacol* 1988;45:109–123.

94. Bowery NG, Collins JF, Hill RG. Bicyclic phosphorus esters that are potent convulsants and GABA antagonists. *Nature* 1976;261:601–603.

95. Ticku MK, Olsen RW. Cage convulsants inhibit picrotoxinin binding. *Neuropharmacology* 1979;18:315–318.

96. Squires RF, Casida JE, Richardson M, et al. [^{35}S]*t*-Butyl bicyclophosphorothionate binds with high affinity to brain-specific sites coupled to $GABA_A$ and ion recognition sites. *Mol Pharmacol* 1983;23:326–336.

97. Holland KD, Ferrendelli JA, Covey DF, et al. Physiological regulation of the picrotoxin receptor by γ-butyrolactone and γ-thiobutyrolactone in cultured hippocampal neurons. *J Neurosci* 1990;10:1719–1727.

98. Höld KM, Sirisoma SI, Ikeda T, et al. α-Thujone (the active component of absinthe): $GABA_A$ receptor modulation and metabolic detoxification. *Proc Natl Acad Sci USA* 2000;97:3826–3831.

99. Venault P, Chapouthier G, Prado de Carvalho L, et al. Benzodiazepine impairs and ß-carboline enhances performance in learning and memory tasks. *Nature* 1986;321:864–866.

100. Asada H, Kawamura Y, Maruyama K, et al. Cleft palate and decreased brain GABA in mice lacking the 67-kDa isoform of glutamic acid decarboxylase. *Proc Natl Acad Sci USA* 1997;94:6496–6499.

101. Kash SF, Johnson RS, Tecott LH, et al. Epilepsy in mice deficient in the 65-kDa isoform of glutamic acid decarboxylase. *Proc Natl Acad Sci USA* 1997;94:14060–14065.

102. Olsen RW, Homanics GE. Function of $GABA_A$ receptors: insights from mutant and knockout mice. In: Martin DL, Olsen RW, eds. GABA in the nervous system: the view at 50 years. Philadelphia: Lippincott Williams & Wilkins, 2000:81–96.

103. Hensch TK, Fagiolini M, Mataga N, et al. Local GABA circuit control of experience-dependent plasticity in developing visual cortex. *Science* 1998;282:1504–1508.

104. Kash SF, Tecott LH., Hodge C, Baekkeskov S. Increased anxiety and altered responses to anxiolytics in mice deficient in the 65-kDa isoform of glutamic acid decarboxylase. *Proc Natl Acad Sci USA* 1999;96:1698–1703.

105. Williams CA, Angelman H, Clayton-Smith J, et al. Angelman syndrome: consensus for diagnostic criteria. *Am J Med Genet* 1995;56:237–238.

106. DeLorey TM, Handforth A, Homanics GE, et al. Mice lacking the ß3 subunit of the $GABA_A$ receptor have the epilepsy phenotype and many of the behavioral characteristics of Angelman syndrome. *J Neurosci* 1998;18:8505–8514.

107. Günther U, Benson J, Benke D, et al. Benzodiazepine-insensitive mice generated by targeted disruption of the γ2 subunit gene of $GABA_A$ receptors. *Proc Natl Acad Sci USA* 1995;92:7749–7753.

108. Essrich C, Fritschy JM, Lorez M, et al. Postsynaptic clustering of major $GABA_A$ receptor subtypes requires the γ2 subunit and gephyrin. *Nat Neurosci* 1998;1:563–571.

109. Crestani F, Lorez M, Baer K, et al. Decreased $GABA_A$ receptor clustering results in enhanced anxiety and a bias for threat cues. *Nat Neurosci* 1999;2:833–839.

110. Quinlan JJ, Firestone LL, Homanics GE. Mice lacking the long splice variant of the γ2 subunit of the $GABA_A$ receptor are more sensitive to benzodiazepines. *Pharmacol Biochem Behav* 2000;66:371–374.

111. Mihalek RM, Banerjee PK, Korpi ER, et al. Attenuated sensitivity to neuroactive steroids in $GABA_A$ receptor δ subunit knockout mice. *Proc Natl Acad Sci USA* 1999;96:12905–12910.

112. Rudolph U, Crestani F, Benke D, et al. Benzodiazepine actions mediated by specific $GABA_A$ receptor subtypes. *Nature* 1999;401:796–800.

113. McKernan RM, Rosahl TW, Reynolds DS, et al. Sedative but not anxiolytic properties of benzodiazepines are deiated by the $GABA_A$ receptor α1 subtype. *Nature Neurosci* 2000;3:587–592.

114. Low K, Crestani F, Keist R, et al. Molecular and neuronal substrate for the selective attenuation of anxiety. *Science* 2000;290:131–134.

13

Substance P and Related Tachykinins

Nadia M.J. Rupniak

Mark S. Kramer

Nadia M.J. Rupniak: Department of Pharmacology, Merck Sharp & Dohme, Harlow, Essex, United Kingdom.

M.S. Kramer: Department of Clinical Neuroscience, Merck & Co., West Point, PA; University of Pennsylvania, Philadelphia, Pennsylvania.

Substance P belongs to a family of neuropeptides known as tachykinins that share the common C-terminal sequence: Phe-X-Gly-Leu-Met-NH2. The three most common tachykinins are substance P, neurokinin A (NKA), and neurokinin B (NKB); their biologic actions are mediated through specific cell-surface receptors designated NK_1, NK_2, and NK_3, with substance P the preferred agonist for NK_1 receptors, NKA for NK_2 receptors, and NKB for NK_3 receptors.

Preclinical studies with substance P antagonists have been complicated not only by phylogenetic differences in central nervous system (CNS) localization of tachykinin receptors, but also by species variants in NK_1 receptor pharmacology. This situation greatly complicates preclinical evaluation of selective substance P receptor antagonists because most of these have only low affinity for the rat receptor, which is the most commonly used preclinical species. Substance P and the NK_1 receptor have a widespread distribution in the brain and are found in brain regions that regulate emotion (e.g., amygdala, periaqueductal gray, hypothalamus). They are also found in close association with 5-hydroxytryptamine (5-HT) and norepinephrine-containing neurons that are targeted by the currently used antidepressant drugs.

The effects of substance P antagonists in preclinical assays for analgesic, antiemetic, antipsychotic, anxiolytic, and antidepressant drugs is reviewed. The process of elucidating the clinical uses of substance P antagonists raises certain fundamental issues that will apply to other novel neurotransmitter ligands in future. The difficulty of predicting clinical efficacy from preclinical data, and of testing novel therapeutic drugs in patients with psychiatric disorders, is discussed.

Substance P, NKA, and NKB are related neuropeptides that are widely distributed in the peripheral nervous system and the CNS. With the development of selective nonpeptide receptor antagonists, it has become possible to investigate the physiologic roles of these peptides and to explore their use as novel treatments for neurologic and psychiatric disorders. Because the substance P–preferring NK_1 receptor is the predominant tachykinin receptor expressed in the human brain, most compounds that have been developed for clinical use are substance P–preferring (NK_1) receptor antagonists.

TACHYKININ FAMILY OF PEPTIDES

Part of "13 - Substance P and Related Tachykinins "

Substance P belongs to a family of neuropeptides known as tachykinins that share the common C-terminal sequence: Phe-X-Gly-Leu-Met-NH_2. Two other mammalian tachykinins are NKA and NKB (Table 13.1). Their biologic actions are mediated through specific G-protein–coupled neurokinin receptors designated NK_1, NK_2, and NK_3, with substance P the preferred agonist for NK_1 receptors, NKA for NK_2 receptors, and NKB for NK_3 receptors. However, the receptor selectivity of these peptides is relatively poor, and it is possible that their actions could be mediated by activation of their less preferred receptors. Indeed, this possibility is suggested by the mismatch between tachykinin-containing neurons and fibers and their corresponding receptor that is seen in certain brain regions. This is particularly apparent in the case of NKA, because NK_2 receptor expression appears to be extremely low in the adult mammalian brain (1).

Substance P	Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-MetNH_2
Neurokinin A	His-Lys-Thr-Asp-Ser-Phe-Val-Gly-Leu-MetNH_2
Neurokinin B	Asp-Met-His-Asp-Phe-Phe-Val-Gly-Leu-MetNH_2

TABLE 13.1. MAMMALIAN TACHYKININS

Substance P was originally discovered in 1931 by von Euler and Gaddum as a tissue extract that caused intestinal contraction *in vitro;* its biologic actions and tissue distribution were further investigated over subsequent decades. NKA (previously known as substance K or neuromedin L) and NKB (originally known as neuromedin K), were isolated from porcine spinal cord in 1983 and were also found to stimulate intestinal contraction (2). Substance P and NKA are produced from a polyprotein precursor after differential splicing of a single precursor gene, preprotachykinin

A (3). NKB is produced from a distinct precursor protein encoded by preprotachykinin B (4).

The existence of several neurokinin receptors was originally suggested by the differential contractile responses elicited in various tissues by mammalian and nonmammalian tachykinins (5). Subsequently, specific binding sites labeled by Bolton Hunter substance P, NKA, and eledoisin were identified in the CNS (6), a finding suggesting that at least three receptors mediated the actions of tachykinins. This was confirmed by cloning of three distinct functional cDNA constructs corresponding to NK_1, NK_2, and NK_3 receptor, which preferentially bound substance P, NKA, and NKB, respectively (7 ,8 and 9). However, the endogenous neurokinins exhibit a high degree of cross-reactivity with these tachykinin receptors.

SPECIES DIFFERENCES IN THE DISTRIBUTION OF NEUROKININS AND THEIR RECEPTORS IN THE NERVOUS SYSTEM

Part of "13 - Substance P and Related Tachykinins "

The substance P–preferring NK_1 receptor has attracted most interest as a CNS drug target because it is the predominant tachykinin receptor expressed in the human brain, whereas NK_2 and NK_3 receptor expression is extremely low or absent (10 ,11 and 12). Therefore, it appears that the central actions of all tachykinins may be mediated predominantly through the NK_1 receptor in humans. However, understanding the role of substance P in the brain has been greatly complicated by marked differences in the distribution of tachykinin receptor subtypes in rodent species that are normally used for such studies. For example, in the rat and guinea pig brain, both NK_1 and NK_3 receptors are expressed (10), findings suggesting that the CNS functions mediated by NK_1 receptors in the human brain may be subserved by NK_1 and/or NK_3 receptors in rodents. NK_2 receptors appear to be absent in the adult mammalian brain of all species examined (10). For these reasons, interpretation of the effects of selective tachykinin receptor antagonists in preclinical assays requires great caution. If such compounds either succeed or fail to exhibit activity in rodent assays for psychiatric and neurologic disorders, this may merely reflect different roles of tachykinin receptors in rodent versus human brain. Hence there is a risk of both false-positive and false-negative extrapolations from preclinical species to humans.

Substance P is widely distributed throughout the CNS and in primary sensory neurons. The demonstration of substance P immunoreactivity in the cell bodies of dorsal root ganglia, in sensory nerve fibers, and in the dorsal horn of the spinal cord led to early speculation that substance P is involved in pain perception (13). Substance P and the NK_1 receptor have a widespread distribution in the brain and are found in brain regions that regulate emotion (e.g., amygdala, periaqueductal gray, hypothalamus) (14 ,15). They are also found in close association with major catecholamine-containing nuclei, including the substantia nigra and the nucleus tractus solitarius (16), as well as with 5-HT- and norepinephrine-containing neurons that are targeted by currently used antidepressant drugs. NKA and NKB are also expressed in varying ratios in the CNS and spinal cord (17 ,18) and in the rodent (but not human) brain, and NK_3 receptors and mRNA have also been demonstrated in various regions, including the substantia nigra, raphe nuclei, and locus ceruleus (19 ,20 and 21).

An interesting aspect of the neuroanatomic localization of substance P is that that it is coexpressed with 5-HT in approximately 50% of ascending dorsal raphe neurons in the primate brain (22 ,23). In contrast, coexpression of substance P and 5-HT in ascending raphe neurons is not seen in the rat brain (24). These findings provide further illustrations of the marked species differences in the neuroanatomy, and possibly physiology, of neurokinin systems. The functional significance of substance P and 5-HT coexpression in the human brain is not known, but it suggests that both neurotransmitters may be coreleased in certain brain regions receiving terminal innervation.

Other evidence suggests that substance P and NKB may also modulate ascending norepinephrine systems. NK_1 receptors (25) have been shown to be expressed on tyrosine hydroxylase–positive cell bodies in the rat locus ceruleus, and both substance P and senktide (a selective NK_3 receptor agonist) excite the firing of locus ceruleus neurons in rats and guinea pigs (26 ,27).

PHYLOGENETIC DIFFERENCES IN TACHYKININ RECEPTOR PHARMACOLOGY

Part of "13 - Substance P and Related Tachykinins "

Preclinical studies with NK_1 receptor antagonists have also been complicated by species variants in NK_1 receptor pharmacology (28 ,29). Compounds such as CP-96,345 were found have high (nM) affinity for the NK_1 receptor expressed in human, gerbil, rabbit, guinea pig, cat, and monkey brain, but they had considerably lower affinity for the mouse and rat NK_1 receptor. Subsequent mutation analysis revealed that subtle differences in the amino acid sequence between the human and the rat NK_1 receptor dramatically alter antagonist binding affinity (30). This feature has greatly hindered preclinical evaluation of high-affinity human NK_1 receptor antagonists because most of these have considerably lower affinity for the rat receptor, the most

commonly used preclinical species (Table 13.2). A few compounds have high affinity for the rat receptor (e.g., SR140333), but their utility for *in vivo* studies may be severely limited by poor brain penetration (31). Although these difficulties may be overcome by administering high doses of NK_1 receptor antagonists to rats, unspecific pharmacologic effects are then frequently encountered, mostly attributable to ion channel blockade. It has therefore been necessary to examine the preclinical pharmacology of these compounds in species with humanlike NK_1 receptor pharmacology (gerbils, guinea pigs, ferrets, hamsters) whenever possible. Pharmacologic differences among human, guinea pig, and rat NK_3 receptors also exist (32).

TABLE 13.2. SPECIES VARIANTS IN NK, RECEPTOR PHARMACOLOGY

	IC_{50} For Inhibition of [^{125}I]SP Binding (nM)			
Compound	Human	Gerbil	Guinea Pig	Rat
L-733060	0.87	0.36	0.3	550
L-760735	0.3	0.5	0.34	10
SR140333	0.04	—	—	0.2
GR205171	0.08	0.06	0.09	1.4

From G. Chicchi and M.A. Cascieri, unpublished observations.

POTENTIAL FOR USE OF TACHYKININ RECEPTOR ANTAGONISTS TO TREAT PSYCHIATRIC AND NEUROLOGIC DISORDERS

Part of "13 - Substance P and Related Tachykinins "

The distribution of neurokinins in the central and peripheral nervous system has generated much speculation about the potential therapeutic uses of selective tachykinin receptor antagonists. The major hypotheses that are supported by preclinical data and have been investigated in clinical trials are considered here. Numerous clinical trials have now been conducted with NK_1 receptor antagonists to define their therapeutic potential in psychiatric and neurologic disorders. In all these studies, the compounds have been extremely well tolerated, with no significant side effects. There are as yet no reports of clinical trials with NK_2 or NK_3 receptor antagonists in patients with CNS disorders.

Pain

Radioligand-binding studies confirm the expression of tachykinin NK_1 and NK_3 (but not NK_2) receptors in the dorsal horn of the spinal cord (33 ,34 and 35). A role of spinal substance P and NKA in nociception is suggested by the reduction in response thresholds to noxious stimuli by central administration of NK_1 and NK_2 (but not NK_3) agonists (36 ,37 and 38). Based on these neuroanatomic and functional studies, it was anticipated that NK_1, and possibly NK_2, receptor antagonists could be developed as analgesic drugs.

Electrophysiologic studies on anesthetized or decerebrate animals provide evidence of potent and selective inhibition of facilitated nociceptive spinal reflexes by NK_1 receptor antagonists. Responses of dorsal horn neurons to noxious or repetitive electrical stimulation of a peripheral nerve was blocked by CP-96,345 (39); NK_1 receptor antagonists also blocked the flexor reflex facilitation produced by C-fiber–conditioning stimulation, but they did not affect protective nociceptive reflexes (40 ,41). NK_1 receptor antagonists have also been shown to inhibit the late-phase response to formalin in gerbils (42), to inhibit carrageenan and Freund adjuvant–induced hyperalgesia in guinea pigs (J. Webb, S. Boyce, and N. Rupniak, unpublished observations; 43), and to attenuate peripheral neuropathy in rats and guinea pigs (43 ,44). Overall, the profile of activity of NK_1 receptor antagonists in a range of assays is comparable to that seen with clinically used analgesic agents such as indomethacin (Table 13.3).

TABLE 13.3. PRECLINICAL EVIDENCE OF AN ANALGESIC PROFILE OF NK_1 RECEPTOR ANTAGONISTS

Assay	Morphine	Indomethacin	NK_1 Antagonist
Tail flick/hot plate	√	X	X
Paw pressure	√	X	X
Writhing	√	√	√
Formalin paw	√	√	√
Carrageenan paw	√	√	√
Nerve injury	√	X	√
CFA arthritis	√	√	√
Facilitated spinal reflex	√	√	√

The first clinical trials with NK_1 receptor antagonists were conducted in patients with various pain conditions. These trials uniformly failed to confirm the analgesic efficacy of these compounds in humans and are reviewed in detail elsewhere (45 ,46). The patient populations and compounds examined included the following: peripheral neuropathy, in which CP-99,994 had no analgesic effect (47); molar extraction, in which MK-869 was ineffective (48); and postherpetic neuralgia, in which MK-869 was ineffective (49). Further unpublished studies with other compounds support these conclusions. Thus, clinical studies to date indicate that NK_1 receptor antagonists do not have major potential as analgesics.

Less is known about the profile of NK_2 receptor antagonists in nociception assays. The NK_2 antagonist MEN 10207 completely blocked both facilitation and protective nociceptive reflex responses (40), and SR48968 reduced responses to both noxious and innocuous pressure applied to

the knee joint (50). In conscious rats, Sluka et al. found that pretreatment with SR48968 prevented the induction of hyperalgesia induced by intraarticular injection of kaolin and carrageenan (51), but it was not effective after hyperalgesia had been established.

Migraine

The vasculature of meningeal tissues such as the dura mater is densely innervated by nociceptive sensory afferents that run in the trigeminal nerve and contain substance P and other neuropeptides. The release of neuropeptides from these sensory fibers during a migraine attack is thought to cause neurogenic inflammation within the meninges and activation of nociceptive afferents projecting to the trigeminal nucleus caudalis (52). In rats, antidromic stimulation of the trigeminal nerve increases vascular permeability and causes plasma protein extravasation in the meninges that is inhibited by NK_1 receptor antagonists (53). These findings suggest that if meningeal plasma extravasation and inflammation of the meninges is involved in the pathogenesis of migraine, then NK_1 receptor antagonists should provide an effective antimigraine therapy. In addition, because of their potential analgesic activity, CNS-penetrant NK_1 antagonists may also be able to alleviate headache by preventing activation of sensory neurons in the trigeminal nucleus caudalis. However, this hypothesis was not confirmed in clinical trials in patients with migraine, in whom neither LY 303870 (54) nor GR205171 (55) gave headache relief.

Emesis

Substance P is present in the nucleus tractus solitarius and the area postrema (56), regions implicated in the control of emesis. Local application of substance P in the area postrema causes retching in ferrets (57), a finding suggesting that NK_1 receptor antagonists may be antiemetic. Consistent with this proposal, these compounds have emerged as an important new class of antiemetics in preclinical studies using ferrets. CP-99,994 completely abolished cisplatin-induced retching and vomiting and exhibited broad-spectrum activity against peripheral and centrally acting emetogens (58 ,59 and 60). Importantly, CP-99,994 markedly attenuated both acute and delayed emesis induced by cisplatin, a profile that distinguishes NK_1 receptor antagonists from established antiemetics (61 ,62). The ability of CP-99,994 to block both peripherally and centrally acting emetogens and the demonstration that direct injection of CP-99,994 into the region of the nucleus tractus solitarius inhibited cisplatin-induced emesis in ferrets (63) suggest that the antiemetic activity of NK_1 antagonists is centrally mediated. This proposal was confirmed by the use of a poorly brain-penetrant quaternary NK_1 receptor antagonist, L-743,310, which prevented cisplatin-induced retching in ferrets when it was infused directly into the CNS, but not systemically (64).

Evaluation of NK_1 receptor antagonists as antiemetics in patients has produced encouraging results. Three independent trials have confirmed that CP-122,721 (65), CJ-11974 (66), and MK-869 (67) are extremely effective in the prevention of acute and delayed emesis after cisplatin chemotherapy. CP-122,721 was also effective in preventing postoperative nausea and vomiting after gynecologic surgery (68), a finding suggesting the utility of NK_1 receptor antagonists as broad-spectrum antiemetics in humans. There are no published studies examining the effects of selective NK_2 and NK_3 receptor agonists and antagonists on emesis.

Schizophrenia

A rationale that NK_1 receptor antagonists may be useful as antipsychotic drugs has been built on evidence that substance P modulates the activity of the mesolimbic dopamine system through which established antipsychotic drugs are thought to act. Substance P–containing fibers have been shown to make synaptic contact with tyrosine hydroxylase–positive neurons in the ventral tegmental area (VTA) from which the mesolimbic dopamine projection arises (69). Infusion of substance P agonists into the VTA stimulates locomotor activity in rats, an effect attributed to the activation of dopamine neurons because this is accompanied by an increase in dopamine turnover in the terminal projection area (nucleus accumbens) (70). Consistent with this interpretation, the locomotor hyperactivity and changes in accumbens cell firing induced by intra-VTA infusion of substance P were blocked by the dopamine receptor antagonist haloperidol, an antipsychotic drug (71).

The ability of a monoclonal antibody to substance P, injected into the nucleus accumbens, to attenuate the locomotor response to amphetamine (72) was consistent with the proposal that endogenous substance P modulates the release of dopamine in the mesolimbic system. A subsequent study appeared to support this interpretation because the NK_1 receptor antagonist CP-96,345 reduced the firing of cells in the VTA in rats (73). However, other studies with NK_1 receptor antagonists are not consistent with these findings. Surprisingly, intra-VTA coinfusion of CP-96,345 was unable to block substance P agonist–induced locomotor activation in rats (71), and amphetamine-induced hyperactivity in guinea pigs was not selectively inhibited by CP-99,994.

A possible explanation for the lack of effect of NK_1 receptor antagonists in these studies is that the effects of substance P in the rodent VTA may be mediated by stimulation of NK_3, rather than NK_1, receptors, as is suggested by anatomic (19), electrophysiologic (74), and behavioral (75) evidence. Intra-VTA application of the NK_3 receptor agonist senktide was shown to enhance markedly the extracellular concentration of dopamine in the nucleus accumbens and

prefrontal cortex of anesthetized guinea pigs, and this was blocked by the selective NK_3 receptor antagonist SR142801 (76). SR142801 (but not the NK_1 receptor antagonist GR205171 or the NK_2 antagonist SR144190) was able to antagonize the increase in neuronal activity caused by acute administration of haloperidol in guinea pigs (77), a finding suggesting that NK_3 receptors play a key role in regulating midbrain dopamine neurons in this species.

Preliminary findings from an exploratory trial with MK-869 in patients with schizophrenia indicated that this compound did not ameliorate the core symptoms of acute psychosis (46).

Anxiety and Depression

Substance P and its preferred NK_1 receptor are highly expressed in brain regions that are critical for the regulation of emotion and neurochemical responses to stress (14 ,15 ,24). Direct central injection of substance P agonists produces a range of fear-related behaviors and defensive cardiovascular changes in animals (78 ,79 ,80 and 81). Neurochemical studies have revealed rapid reductions in substance P content in the mesolimbic system, hippocampus, septum, periaqueductal gray, and hypothalamus of rats after inescapable footshock (82 ,83) and immobilization stress (84). These findings indicate that activation of central substance P pathways occurs in response to noxious or aversive stimulation and suggest that NK_1 receptor antagonists may have anxiolytic or antidepressant-like properties.

Substance P antagonists are capable of attenuating psychological stress responses in paradigms using neurochemical and behavioral endpoints. This was first suggested by the demonstration that intra-VTA injection of a monoclonal antibody to substance P prevented stress-induced activation of mesocortical dopamine neurons (85). More recently, the NK_1 receptor antagonist GR205171 was shown to inhibit the stress-induced elevation in the dopamine metabolite DOPAC in the frontal cortex (86). Certain chemically diverse NK_1 receptor antagonists have also shown activity in a range of assays for anxiolytic and antidepressant drugs after intracerebral or systemic administration. One of the earliest reported studies demonstrated a direct substance P–ergic projection from the medial amygdala to the medial hypothalamus that regulates the expression of defensive rage in cats. Either systemic or intrahypothalamic injection of CP-96,345 inhibited amygdaloid facilitation of defensive rage (87). A second study examined the role of NK_1 receptors in the caudal pontine reticular nucleus and showed that injection of CP-96,345 or CP-99,994 into this region blocked potentiation of the acoustic startle response by footshock in rats (88). In the resident-intruder paradigm, L-760,735 reduced aggression in singly housed hamsters in a dose-dependent manner resembling the effect of fluoxetine (J. Webb, E. Carlson, N. Rupniak, unpublished observations) (Fig. 13.1). CGP 49823 has been reported to be active in the rat social interaction test for anxiolytic activity (89 ,90) and the forced swim test for antidepressant drugs (89). In guinea pig pups, the vocalization response elicited by maternal separation is inhibited by brain-penetrant NK_1 receptor antagonists (L-773,060, L-760,735, GR205171), a property also seen with clinically used antidepressant and anxiolytic drugs (91 ,92). The amygdala is a potential site of action for this effect of NK_1 receptor antagonists because separation stress caused internalization of NK_1 receptors (reflecting the release of substance P) in this brain region (91 ,93), and intraamygdala injection of L-760735 attenuated the neonatal vocalizations (93). Further evidence for an antidepressant-like preclinical profile of substance P antagonists is suggested by preliminary findings with L-733,060, which was active in the learned helplessness paradigm in rats (94), despite having only low affinity for the rat NK_1 receptor. These findings are summarized in Table 13.4 .

TABLE 13.4. PRECLINICAL EVIDENCE OF AN ANTIDEPRESSANT AND ANXIOLYTIC-LIKE PROFILE OF NK_1 RECEPTOR ANTAGONISTS

Assay	Species	BZ Anxiolytics	SSRI/TCA	NK_1 Antagonist
Neonatal vocalization	Guinea pig	√	√	√
Aggression	Hamster	√	√	√
Learned helpessness	Rat		√	√
Forced swim	Rat	X	√	√
Shock-potentiated startle	Rat	√		√
Social interaction	Rat	√		√

BZ, benzodiazepine; SSRI,-selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant.

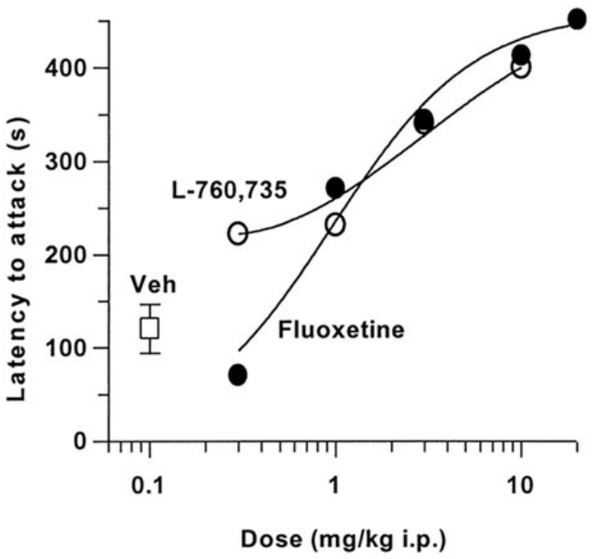


FIGURE 13.1. Activity of L-760,735 in the hamster residentintruder test. The subjects were individually housed adult male hamsters. On test days, both resident and intruder hamsters were dosed with the same drug and were returned to their own cages for 30 minutes before testing. Pretreatment with either fluoxetine (0.3 to 30 mg/kg intraperitoneally) or the selective NK_1 receptor antagonist L-760735 (0.3 to 10 mg/kg intraperitoneally) caused a dose-dependent increase in the latency to initiate an aggressive encounter.

The NK_2 receptor antagonists SR48968, GR100679, and GR159897 have been reported to exhibit anxiolytic-like effects in several preclinical assays (mouse light-dark box, rat social interaction test, rat elevated plus maze, and marmoset threat test) (95 ,96 and 97). However, these compounds were reported to be extremely potent, and the micrograms per kilogram anxiolytic dose range was considerably lower

than that required to block NK_2 agonist–mediated effects in peripheral tissues (mg/kg dose range) (98 ,99). A second difficulty concerns the failure to establish convincing expression of NK_2 receptors in the adult rat brain (100).

In rodents, there is evidence that NK_3 receptors are able to modulate monoaminergic neurotransmission. Because the clinical efficacy of currently used antidepressant drugs is ascribed to their ability to increase the synaptic availability of 5-HT and norepinephrine, modulation of these systems by NK_3 receptor ligands may suggest an antidepressant-like profile. The ability of central infusion of senktide to elicit a 5-HT behavioral syndrome (101) and to increase the release of norepinephrine in brain (27) indicates that monoamine systems can be activated by NK_3 receptor agonists. The ability of senktide to increase locus ceruleus firing, to increase norepinephrine release, and to decrease locomotor activity in animals was blocked by the selective NK_3 receptor antagonist SR142801 (102). These actions are not clearly indicative of an antidepressant-like profile of NK_3 receptor antagonists, and the low abundance of these receptors in human brain suggests that, like NK_2 receptor antagonists, NK_3 antagonists are less attractive candidates for clinical development in psychiatry than NK_1 receptor antagonists.

There is currently only one published study in which a tachykinin antagonist has been examined in patients with depression. The clinical efficacy of the NK_1 receptor antagonist MK-869 was comparable to that of paroxetine in outpatients with major depressive disorder and moderately high anxiety. As in other clinical trials, MK-869 was extremely well tolerated (94). Further studies are currently in progress with this and other NK_1 receptor antagonists in patients with depression and anxiety disorders.

CONCLUSIONS AND IMPLICATIONS FOR FUTURE STUDIES OF NEUROKININ ANTAGONISTS IN PSYCHIATRIC AND NEUROLOGIC DISORDERS

Part of "13 - Substance P and Related Tachykinins "

The process of elucidating the potential clinical uses of tachykinin receptor antagonists raises several fundamental issues that will apply to other novel neurotransmitter ligands in the future. Preclinical studies have suggested therapeutic potential of neurokinin antagonists in certain neurologic and psychiatric disorders, including migraine, pain, schizophrenia, anxiety, and depression. Of these, antagonists of tachykinin NK_1 receptors are the most attractive agents because this is the predominant receptor expressed in the human brain. However, expectations have been only partially fulfilled in clinical trials, and although preliminary findings suggest efficacy of NK_1 receptor antagonists in the control of emesis and depression, these compounds do not appear to possess analgesic or antipsychotic activity. It was not possible to predict this outcome from preclinical evidence, in which interpretation was complicated by species variants in tachykinin receptor pharmacology and possibly physiology, and this was coupled with uncertainty about whether relevant aspects of human disease can be accurately modeled in animals.

This chapter has focused on the intricacies of prioritizing efforts to identify the therapeutic uses of neurokinin antagonists for CNS disorders. However, there are many other potential uses for NK_1, NK_2, and NK_3 antagonists that have not yet been fully explored. These include inflammatory diseases such as cystitis and inflammatory bowel disease, asthma, cancer, glaucoma, ocular hypotension, cardiac disorders, and psoriasis.

REFERENCES

1. Saffroy M, Beaujouan JC, Torrens Y, et al. Localization of tachykinin binding sites (NK_1, NK_2, NK_3 ligands) in the rat brain. *Peptides* 1987;9:227–241.

2. Kangawa K, Minamino N, Fukuda A, et al. Neuromedin K: a novel mammalian tachykinin identified in porcine spinal cord. *Biochem Biophys Res Commun* 1983;114:533–540.

3. Krause JE, MacDonald MR, Takeda Y. The polyprotein nature of substance P precursors. *Bioessays* 1989;10:62–69.

4. Kotani H, Hoshimaru M, Nawa H, et al. Structure and gene organization of bovine neuromedin K precursor. *Proc Natl Acad Sci USA* 1986;83:7074–7078.

5. Regoli D, Mizrahi J, D'Orleans-Juste P, et al. Receptors for substance P. II. Classification by agonist fragments and homologues. *Eur J Pharmacol* 1984;97:171–177.

6. Cascieri MA, Chicchi GG, Liang T. Demonstration of two distinct tachykinin receptors in rat brain cortex. *J Biol Chem* 1985;260:1401–1507.

7. Masu Y, Nakayama K, Tamaki H, et al. cDNA cloning of bovine substance-K receptor through oocyte expression system. *Nature* 1987;329:836–838.

8. Yokota Y, Sasai Y, Tanaka K, et al. Molecular characterization of a functional cDNA for rat substance P receptor. *J Biol Chem* 1989;264:17649–17652.

9. Shigemoto R, Yokota Y, Tsuchida K, et al. Cloning and expression of a rat neuromedin K receptor cDNA. *J Biol Chem* 1990;265:623–628.

10. Dietl MM, Palacios JM. Phylogeny of tachykinin receptor localization in the vertebrate central nervous system: apparent absence of neurokinin-2 and neurokinin-3 binding sites in the human brain. *Brain Res* 1991;539:211–222.

11. Buell G, Schultz MF, Arkinstall SJ. Molecular characterisation, expression and localization of human neurokinin-3 receptor. *FEBS Lett* 1992;299:90–95.

12. Mileusnic D, Magnuson DJ, Hejna MJ, et al. Age and species-dependent differences in the neurokinin B system in rat and human brain. *Neurobiol Aging* 1999;20:19–35.

13. Ljungdahl A, Hokfelt T, Nilsson G. Distribution of substance P-like immunoreactivity in the central nervous system of the rat. I. Cell bodies and nerve terminals. *Neuroscience* 1978;3:861–943.

14. Mantyh PW, Hunt SP, Maggio JE. Substance P receptors: localization by light microscopic autoradiography in rat brain using [^{3}H]SP as the radioligand. *Brain Res* 1984;307:147–165.

15. Arai H, Emson PC. Regional distribution of neuropeptide K and other tachykinins (neurokinin A, neurokinin B and substance P) in rat central nervous system. *Brain Res* 1986;399:240–249.

16. Ljungdahl A, Hokfelt T, Nilsson G, et al. Distribution of substance P-like immunoreactivity in the central nervous system of the rat. II. Light microscopic localization in relation to catecholamine-containing neurons. *Neuroscience* 1978;3:945–976.

17. Minamino N, Masuda H, Kangawa K, et al. Regional distribution of neuromedin K and neuromedin L in rat brain and spinal cord. *Biochem Biophys Res Commun* 1984;124:731–738.

18. Lindefors N, Brodin E, Theodorsson-Norheim E, et al. Regional distribution and *in vivo* release of tachykinin-like immunoreactivities in rat brain: evidence for regional differences in relative proportions of tachykinins. *Regul Pept* 1985;10:217–230.

19. Saffroy M, Beaujouan JC, Torrens Y, et al. Localization of tachykinin binding sites (NK_1, NK_2, NK_3 ligands) in the rat brain. *Peptides* 1988;9:227–241.

20. Dam TV, Escher E, Quirion R. Visualisation of neurokinin-3 receptor sites in rat brain using the highly selective ligand [^{3}H]senktide. *Brain Res* 1990;506:175–179.

21. Whitty CJ, Walker PD, Goebel DJ, et al. Quantitation, cellular localization and regulation of neurokinin receptor gene expression within the rat substantia nigra. *Neuroscience* 1995;64:419–425.

22. Baker KG, Halliday GM, Hornung JP, et al. Distribution, morphology and number of monoamine-synthesizing and substance P-containing neurons in the human dorsal raphe nucleus. *Neuroscience* 1991;42:757–775.

23. Sergeyev V, Hokfelt T, Hurd Y. Serotonin and substance P co-exist in dorsal raphe neurons of the human brain. *Neuroreport* 1999;10:3967–3970.

24. Hokfelt T, Johansson O, Holets V, et al. Distribution of neuropeptides with special reference to their coexistence with classical transmitters. In: Meltzer HY, ed. *Psychopharmacology: the third generation of progress*. New York: Raven, 1987:401–416.

25. Hahn MK, Bannon MJ. Stress-induced c-fos expression in the rat locus coeruleus is dependent on neurokinin-1 receptor activation. *Neuroscience* 1999;94:1183–1188.

26. Guyenet PG, Aghajanian GK. Excitation of neurons in the nucleus locus coeruleus by substance P and related peptides. *Brain Res* 1977;136:178–184.

27. Jung M, Michaud JC, Steinberg R, et al. Electrophysiological, behavioural and biochemical evidence for activation of brain noradrenergic systems following neurokinin NK_3 receptor stimulation. *Neuroscience* 1996;74:403–414.

28. Beresford IJM, Birch PJ, Hagan RM, et al. Investigation into species variants in tachykinin NK_1 receptors by use of the non-peptide antagonist, CP-96,345. *Br J Pharmacol* 1991;104:292–293.

29. Gitter BD, Bruns RF, Howbert JJ, et al. Pharmacological characterization of LY303870: a novel, potent and selective nonpeptide substance P (neurokinin-1) receptor antagonist. *J Pharmacol Exp Ther* 1995;275:737–744.

30. Fong TM, Yu H, Strader CD. Molecular basis for the species selectivity of the neurokinin-1 receptor antagonists CP-96,345 and RP67580. *J Biol Chem* 1992;267:25668–25671.

31. Rupniak NMJ, Tattersall FD, Williams AR, et al. *In vitro* and *in vivo* predictors of the anti-emetic activity of tachykinin NK_1 receptor antagonists. *Eur J Pharmacol* 1997;326:201–209.

32. Suman-Chauhan N, Grimson P, Guard S, et al. Characterisation of [^{125}I][MePhe^7]neurokinin B binding to tachykinin NK_3 receptors: evidence for interspecies variance. *Eur J Pharmacol* 1994;269:65–72.

33. Shults CW, Quirion R, Chronwall B, et al. A comparison of the anatomical distribution of substance P and substance P receptors in the rat central nervous system. *Peptides* 1984;5:1097–1128.

34. Beresford IJM, Ireland SJ, Stables J, et al. Ontogeny and characterisation of ^{125}I-Bolton Hunter-eledoisin binding sites in rat spinal cord by quantitative autoradiography. *Neuroscience* 1992;46:225–232.

35. Humpel C, Saria A. Characterisation of neurokinin binding sites in rat brain membranes using highly selective ligands. *Neuropeptides* 1993;25:65–71.

36. Cridland RA, Henry JL. Comparison of the effects of substance P, neurokinin A, physalaemin and eledoisin in facilitating a nociceptive reflex in the rat. *Brain Res* 1986;381:93–99.

37. Laneuville O, Dorais J, Couture R. Characterization of the effects produced by neurokinins and three antagonists selective for neurokinin receptor subtypes in a spinal nociceptive reflex of the rat. *Life Sci* 1988;42:1295–1305.

38. Picard P, Boucher S, Regoli D, et al. Use of non-peptide tachykinin receptor antagonists to substantiate the involvement of NK_1 and NK_2 receptors in a spinal nociceptive reflex in the rats. *Eur J Pharmacol* 1993;232:255–261.

39. de Koninck Y, Henry JL. Substance P-mediated slow excitatory postsynaptic potential elicited in dorsal horn neurons *in vivo* by noxious stimulation. *Proc Natl Acad Sci USA* 1991;88:11344–11348.

40. Xu XJ, Dalsgaard CJ, Wiesenfeld-Hallin Z. Intrathecal CP-96,345 blocks reflex facilitation induced in rats by substance P and C-fiber-conditioning stimulation. *Eur J Pharmacol* 1992;216:337–344.

41. Laird JMA, Hargreaves RJ, Hill RG. Effect of RP67580, a non-peptide neurokinin-1 receptor antagonist, on facilitation of a nociceptive spinal flexion reflex in the rat. *Br J Pharmacol* 1993;109:713–718.

42. Rupniak NMJ, Carlson EJ, Boyce S, et al. Enantioselective inhibition of the formalin paw late phase by the NK_1 receptor antagonist L-733,060 in gerbils. *Pain* 1996;67:189–195.

43. Walpole C, Ko SY, Brown M, et al. 2-Nitrophenylcarbamoyl-(S)-prolyl-(S)-3-(2-naphthyl)alanyl-N-benzyl-N-methylamide (SDZ NKT 343), a potent human NK_1 tachykinin receptor antagonist with good oral analgesic activity in chronic pain models. *J Med Chem* 1998;41:3159–3173.

44. Cumberbatch MJ, Wyatt A, Boyce S, et al. Reversal of behavioural and electrophysiological correlates of experimental peripheral neuropathy by the NK_1 receptor antagonist GR205171 in rats. *Neuropharmacology* 1998;37:1535–1543.

45. Rupniak NMJ, Hill RG. Neurokinin antagonists. In: Sawynok J, Cowan A, eds. *Novel aspects of pain management: opioids and beyond*. New York: John Wiley, 1999:135–155.

46. Rupniak NMJ, Kramer MS. Discovery of the antidepressant anti-emetic efficacy of substance P antagonists. *Trends Pharmacol Sci* 1999;20, 485–490.

47. Suarez GA, Opfer-Gehrking TL, MacLean DB, et al. Double-blind, placebo-controlled study of the efficacy of a substance P (NK_1) receptor antagonist in painful peripheral neuropathy. *Neurology* 1994;44:373P.

48. Reinhardt RR, Laub JB, Fricke JR, et al. Comparison of a neurokinin-1 antagonist, L-754,030, to placebo, acetaminophen and ibuprofen in the dental pain model. *Clin Pharmacol Ther* 1998;63:168.

49. Block GA, Rue D, Panebianco D, et al. The substance P receptor antagonist L-754,030 (MK-0869) is ineffective in the treatment of postherpetic neuralgia. *Neurology* 1998;50:A225.

50. Neugebauer V, Rumenapp P, Schaible HG. The role of spinal neurokinin-2 receptors in the processing of nociceptive information from the joint and in the generation and maintenance of inflammation-evoked hyperexcitability of dorsal horn neurons in the rat. *Eur J Neurosci* 1996;8:249–260.

51. Sluka KA, Milton MA, Willis WD, et al. Differential roles of neurokinin 1 and neurokinin 2 receptors in the development and maintenance of heat hyperalgesia induced by acute inflammation. *Br J Pharmacol* 1997;120:1263–1273.

52. Uddman R, Edvinsson L. Neuropeptides in the cerebral circulation. *Cerebrovasc Brain Metab Rev* 1989;1:230–252.

53. Shepheard SL, Williamson DJ, Williams J, et al. Comparison of the effects of sumatriptan and the NK_1 antagonist CP-99,994 on plasma extravasation in dura mater and c-fos mRNA expression in trigeminal nucleus caudalis of rats. *Neuropharmacology* 1995;34:255–261.

54. Goldstein D, Wang O, Saper JR, et al. Ineffectiveness of neurokinin-1 antagonist in acute migraine: a crossover study. *Cephalalgia* 1997;17:785–790.

55. Connor HE, Bertin L, Gillies, et al. Clinical evaluation of a novel, potent, CNS penetrating NK_1 receptor antagonist in the acute treatment of migraine. *Cephalalgia* 1998;18:392.

56. Armstrong DM, Pickel VM, Joh TH, et al. Immunocytochemical localization of catecholamine synthesizing enzymes and neuropeptides in the area postrema and medial nucleus tractus solitarius of rat brain. *J Comp Neurol* 1981;196:505–517.

57. Andrews PLR. 5-HT_3 receptor antagonists and anti-emesis. In: King FD and Sanger GJ, eds. *5-HT_3 receptor antagonists*. Boca Raton, FL: CRC, 1994:255–317.

58. Bountra C, Bunce K, Dale K, et al. Anti-emetic profile of a non-peptide neurokinin NK_1 receptor antagonist, CP-99,994, in ferrets. *Eur J Pharmacol* 1993;249:R3–R4.

59. Tattersall FD, Rycroft W, Hill RG, et al. Enantioselective inhibition of apomorphine-induced emesis in the ferret by the neurokinin-1 receptor antagonist CP-99,994. *Neuropharmacology* 1994;33:259–260.

60. Watson JW, Gonsalves SF, Fossa AA, et al. The anti-emetic effects of CP-99,994 in the ferret and the dog: role of the NK_1 receptor. *Br J Pharmacol* 1995;115:84–94.

61. Watson JW, Nagahisa A, Lucot JB, et al. The tachykinins and emesis: towards complete control? In: DJM Reynolds, PLR Andrews, CJ Davis, eds. *Serotonin and the scientific basis of anti-emetic therapy*. Oxford: Oxford Clinical Communications, 1995:233–238.

62. Rudd JA, Jordan CC, Naylor RJ. The action of the NK_1 tachykinin receptor antagonist, CP 99,994, in antagonizing the acute and delayed emesis induced by cisplatin in the ferret. *Br J Pharmacol* 1996;119:931–936.

63. Gardner CJ, Bountra C, Bunce KT, et al. Anti-emetic activity of neurokinin NK_1 receptor antagonists is mediated centrally in the ferret. *Br J Pharmacol* 1994;112:516P.

64. Tattersall FD, Rycroft W, Francis B, et al. Tachykinin NK_1 receptor antagonists act centrally to inhibit emesis induced by the chemotherapeutic agent cisplatin in ferrets. *Neuropharmacology* 1996;35:1121–1129.

65. Kris MG, Radford JE, Pizzo BA, et al. Use of an NK_1 receptor antagonist to prevent delayed emesis after cisplatin. *J Natl Cancer Inst* 1997;89:817–818.

66. Hesketh PJ, Gralla RJ, Webb RT, et al. Randomized phase II study of the neurokinin-1 receptor antagonist CJ-11,974 in the control of cisplatin-induced emesis. *J Clin Oncol* 1999;17:338–343.

67. Navari RM, Reinhardt RR, Gralla RJ, et al. Reduction of cisplatin-induced emesis by a selective neurokinin-1-receptor antagonist, L-754,030. *N Engl J Med* 1999;340:190–195.

68. Gesztesi ZS, Song D, White PF. Comparison of a new NK-1 antagonist (CP-122,721) to ondansetron in the prevention of postoperative nausea and vomiting. *Anesth Analg* 1998;86:S32.

69. Tamiya R, Hanada M, Kawai Y, et al. Substance P afferents have synaptic contacts with dopaminergic neurons in the ventral tegmental area of the rat. *Neurosci Lett* 1990;110:11–15.

70. Elliott PJ, Mason GS, Stephens-Smith M, et al. Behavioural and biochemical responses following activation of midbrain dopamine pathways by receptor selective neurokinin agonists. *Neuropeptides* 1991;19:119–123.

71. Elliott PJ, Mason GS, Graham EA, et al. Modulation of the rat mesolimbic dopamine pathway by neurokinins. *Behav Brain Res* 1992;51:77–82.

72. Elliott PJ, Nemeroff CB, Kilts CD. Evidence for a tonic facilitatory influence of substance P on dopamine release in the nucleus accumbens. *Brain Res* 1986;385:379–382.

73. Minabe Y, Emori K, Toor A, et al. The effect of the acute and chronic administration of CP-96,345, a selective neurokinin-1 receptor antagonist, on midbrain dopamine neurons in the rat: a single unit, extracellular recording study. *Synapse* 1996;22:35–45.

74. Seabrook GR, Bowery BJ, Hill RG. Pharmacology of tachykinin receptors on neurones in the ventral tegmental area of rat brain slices. *Eur J Pharmacol* 1995;273:113–119.

75. Stoessl AJ, Szczutkowski E, Glenn B, et al. Behavioural effects of selective tachykinin agonists in midbrain dopamine regions. *Brain Res* 1991;565:254–262.

76. Marco N, Thirion A, Mons G, et al. Activation of dopaminergic and cholinergic neurotransmission by tachykinin NK_3 receptor stimulation: an *in vivo* microdialysis approach in guinea pig. *Neuropeptides* 1998;32:481–488.

77. Gueudet C, Santucci V, Soubrie P, et al. Blockade of neurokinin 3 receptors antagonizes drug-induced population response and depolarization block of midbrain dopamine neurons in guinea pigs. *Synapse* 1999;33:71–79.

78. Elliott P. Place aversion induced by the substance P analogue, dimethyl-C7, is not state dependent: implication of substance P in aversion. *Exp Brain Res* 1988;381:354–356.

79. Aguiar MS, Brandao ML. Conditioned place aversion produced by microinjections of substance P into the periaqueductal gray of rats. *Behav Pharmacol* 1994;5:369–373.

80. Teixeira RM, Santos ARS, Ribeiro SJ, et al. Effects of central administration of tachykinin receptor agonists and antagonists on plus-maze behaviour in mice. *Eur J Pharmacol* 1996;311:7–14.

81. Unger T, Carolus S, Demmert G, et al. Substance P induces a cardiovascular defense reaction in the rat: pharmacological characterization. *Circ Res* 1988;63:812–820.

82. Bannon MJ, Deutch AY, Tam SY, et al. Mild footshock stress dissociates substance P from substance K and dynorphin from Met- and Leu-enkephalin. *Brain Res* 1986;381:393–396.

83. Siegel RA, Duker E-M, Fuchs E, et al. Responsiveness of mesolimbic, mesocortical, septal and hippocampal cholecystokinin and substance P neuronal systems to stress in the male rat. *Neurochem Int* 1984;6:783–789.

84. Takamaya H, Ota Z, Ogawa N. Effect of immobilization stress on neuropeptides and their receptors in rat central nervous system. *Regul Pept* 1986;15:239–248.

85. Bannon MJ, Elliott PJ, Alpert JE, et al. Role of endogenous substance P in stress-induced activation of mesocortical dopamine neurones. *Nature* 1986;306:791–792.

86. Barton CL, Jay MT, Meurer L, et al. GR205171, a selective NK_1 receptor antagonist, attenuates stress-induced increase of dopamine metabolism in rat medial prefrontal cortex. *Br J Pharmacol* 1999;126:284P.

87. Shaikh MB, Steinberg A, Siegel A. Evidence that substance P is utilised in the medial amygdaloid facilitation of defensive rage behavior in the cat. *Brain Res* 1993;625:283–294.

88. Krase W, Koch M, Schnizler HU. Substance P is involved in the sensitization of the acoustic startle response by footshock in rats. *Behav Brain Res* 1994;63:81–88.

89. Vassout A, Schaub M, Gentsch C, et al. CGP 49823, a novel NK_1 receptor antagonist: behavioural effects. *Neuropeptides* 1994;26:S38.

90. File SE. Anxiolytic action of a neurokinin-1 receptor antagonist in the social interaction test. *Pharmacol Biochem Behav* 1997;58:747–752.

91. Kramer MS, Cutler N, Feighner J, et al. Distinct mechanism for antidepressant activity by blockade of central substance P receptors. *Science* 1998;281:1640–1645.

92. Rupniak NMJ, Carlson EC, Harrison T, et al. Pharmacological blockade or genetic deletion of substance P (NK_1) receptors attenuates neonatal vocalisation in guinea-pigs and mice. *Neuropharmacology* 2000;39:1413–1421.

93. Boyce S, Smith D, Carlson E, et al. Intra-amygdala Injection of the substance P (NK_1 receptor) antagonist L-760735 inhibits neonatal vocalisations in guinea-pigs. *Neuropharmacology* 2001;41:130–137.

94. McElroy JF, Weidemann KA, Zeller KL, et al. Acute efficacy of the substance P (NK_1) antagonist L-733,060 in rat learned helplessness, a chronic animal model of depression. *Soc Neurosci Abs* 1999;25:31.15.

95. Stratton SC, Beresford IJM, Hagan RM. Anxiolytic activity of tachykinin NK_2 receptor antagonists in the mouse light-dark box. *Eur J Pharmacol* 1993;250:R11–R12.

96. Stratton SC, Beresford IJM, Hagan RM. GR159897, a potent non-peptide tachykinin NK_2 receptor antagonist, releases suppressed behaviours in a novel aversive environment. *Br J Pharmacol* 1994;112:49P.

97. Walsh DM, Stratton SC, Harvey FJ, et al. The anxiolytic-like activity of GR159897, a non-peptide NK_2 receptor antagonist, in rodent and primate models of anxiety. *Psychopharmacology* 1995;121:186–191.

98. Tousignant C, Chan C-C, Guevremont D, et al. NK_2 receptors mediate plasma extravasation in guinea-pig lower airways. *Br J Pharmacol* 1993;108:383–386.

99. Ball DI, Beresford IJM, Wren GPA, et al. *In vitro* and *in vivo* pharmacology of the non-peptide antagonist at tachykinin NK_2 receptors, GR159897. *Br J Pharmacol* 1994;112:48P.

100. Hagan RM, Beresford IJ, Stables J, et al. Characterisation, CNS distribution and function of NK_2 receptors studied using potent NK_2 receptor antagonists. *Regul Pept* 1993;46:9–19.

101. Stoessl AJ, Dourish CT, Young SC, et al. Senktide, a selective neurokinin B–like agonist, elicits serotonin-mediated behaviour following intracisternal administration in the mouse. *Neurosci Lett* 1987;80:321–326.

102. Jung M, Michaud JC, Steinberg R, et al. Electrophysiological, behavioral and biochemical evidence for activation of brain noradrenergic systems following neurokinin NK3 receptor stimulation. *Neuroscience* 1996;74:403–414.

14
Histamine

Jean-Charles Schwartz

Jean-Michel Arrang

Jean-Charles Schwartz and Jean-Michel Arrang: Unité de Neurobiologie et Pharmacologie Moléculaire de l'INSERM, Centre Paul Broca, Paris, France.

In a certain way, histaminergic systems have had a great, although indirect, historical importance in the development of neuropsychopharmacology. Indeed, the discovery of both the neuroleptic agents and the tricyclic antidepressant drugs in the 1950s was derived from the clinical study of behavioral actions of "antihistamines," a class of antiallergic drugs now designated *H_1-receptor antagonists*.

Nevertheless, the histaminergic neuronal system in the brain, although already understood by the mid-1970s, has remained largely unexploited in drug design. Thus, only the traditional brain-penetrating H_1-receptor antagonists, used as over-the-counter sleeping pills, are known to interfere with histaminergic transmissions in the central nervous system (CNS). This situation contrasts with the emergence, in the 1990s, of detailed knowledge of the system that revealed that it shares many biological and functional properties with other aminergic systems overexploited in CNS drug design.

Histamine and its receptors in the brain have been the subject of two comprehensive reviews (1 ,2). Therefore, to limit the length of the present chapter, we have deliberately elected to summarize the detailed information that can be found in these reviews and have added only more recent information and major references.

ORGANIZATION OF THE HISTAMINERGIC NEURONAL SYSTEM

Part of "14 - Histamine "

One decade after the first evidence by Garbarg et al. of an ascending histaminergic pathway obtained by lesions of the medial forebrain bundle (3), the exact localization of corresponding perikarya in the posterior hypothalamus was revealed immunohistochemically, and the distribution, morphology, and connections of histamine and histidine decarboxylase-immunoreactive neurons were determined. Data were comprehensively reviewed (4 , 5 , 6 and 7), and they are summarized only briefly here.

All known histaminergic perikarya constitute a continuous group of mainly magnocellular neurons (about 2,000 in the rat), located in the posterior hypothalamus and collectively named the *tuberomammillary nucleus* (Fig.14.1). It can be subdivided into medial, ventral, and diffuse subgroups extending longitudinally from the caudal end of the hypothalamus to the midportion of the third ventricle. A similar organization was described in humans, except histaminergic neurons are more numerous (approximately 64), and occupy a larger proportion of the hypothalamus (8). Besides their large size (25 to 35 µm), tuberomammillary neurons are characterized by few thick primary dendrites, with overlapping trees, displaying few axodendritic synaptic contacts. Another characteristic feature is the close contact of dendrites with glial elements in a way suggesting that they penetrate the ependyma and come in close contact with the cerebrospinal fluid, perhaps to secrete or receive still unidentified messengers. Neurons expressing mRNAs for histidine decarboxylase (EC 4.1.1.22), the enzyme responsible for the one-step histamine formation in the brain (2), were found by *in situ* hybridization in the tuberomammillary nucleus, but not in any other brain area (9). Tuberomammillary neurons possess the vesicular monoamine transporter 2 (10), which accounts for the histamine-releasing effect of reserpine (2).

The histaminergic neurons are characterized by the presence of an unusually large variety of markers for other neurotransmitter systems: glutamic acid decarboxylase, the γ-aminobutyric acid (GABA–synthesizing enzyme; adenosine deaminase, a cytoplasmic enzyme possibly involved in adenosine inactivation; galanin, a peptide co-localized with all other monoamines; (Met^5)enkephalyl-Arg^{6}Phe^7, a product of the proenkephalin A gene; and other neuropeptides such as substance P, thyroliberin, or brain natriuretic peptide. Tuberomammillary neurons also contain monoamine oxidase B, an enzyme responsible for deamination of telemethylhistamine, a major histamine metabolite in brain. Finally, a subpopulation of histaminergic neurons is able to take up and decarboxylate exogenous 5-hydroxytryptophan,

a compound that they do not synthesize, however (5). Discovering the functions of such a high number of putative cotransmitters in the same neurons remains an exciting challenge.

Like other monoaminergic neurons, histaminergic neurons constitute long and highly divergent systems projecting in a diffuse manner to many cerebral areas (Fig.14.1). Immunoreactive, mostly unmyelinated, varicose or nonvaricose fibers are detected in almost all cerebral regions, particularly limbic structures, and it was confirmed that individual neurons project to widely divergent areas. Ultrastructural studies suggest that these fibers make few typical synaptic contacts (6).

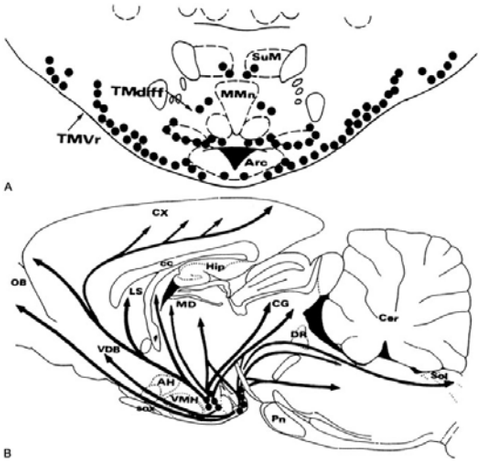


FIGURE 14.1. Localization of histaminergic perikarya *(closed circles)* in tuberomammillary nucleus and disposition of main histaminergic pathways *(arrows)* in rat brain. **A:** Frontal section of the caudal hypothalamus. **B:** Sagittal section of the brain. AH, anterior hypothalamic area; Arc, arcuate nucleus; cc, corpus callosum; Cer, cerebellum; CG, central gray; CX, cerebral cortex; DR, dorsal raphe nucleus; f, fornix; Hip, hippocampus; LS, lateral septum; MD, mediodorsal thalamus; MMn, medial mammillary nucleus median part; OB, olfactory bulb; Pn, pontine nuclei; Sol, nucleus of solitary tract; Sox, supraoptic decussation; sum, supramammillary nucleus; TMdiff, tuberomammillary nucleus diffuse part; TMVr, ventral tuberomamillary subgroup rostral part; VDB, nucleus of vertical limb of diagonal band; VMH, ventromedial hypothalamic nucleus.

Fibers arising from the tuberomammillary nucleus constitute two ascending pathways: one laterally, through the medial forebrain bundle, and the other periventricularly. These two pathways combine in the diagonal band of Broca to project, mainly in an ipsilateral fashion, to many telencephalic areas, for example, in all areas and layers of the cerebral cortex, the most abundant projections being to the external layers. Other major areas of termination of these long ascending connections are the olfactory bulb, the hippocampus, the caudate putamen, the nucleus accumbens, the globus pallidus, and the amygdaloid complex. Many hypothalamic nuclei exhibit a very dense innervation, for example, the suprachiasmatic, supraoptic, arcuate, and ventromedial nuclei.

Finally, a long descending histaminergic subsystem also arises from the tuberomammillary nucleus to project to various mesencephalic and brainstem structures such as the cranial nerve nuclei (e.g., the trigeminal nerve nucleus), the central gray, the colliculi, the substantia nigra, the locus ceruleus, the mesopontine tegmentum, the dorsal raphe nucleus, the cerebellum (sparse innervation), and the spinal cord.

Several anterograde and retrograde tracing studies established the existence of afferent connections to the histaminergic perikarya, namely, from the infralimbic cortex, the septum-diagonal band complex, the preoptic region, the hypothalamus, and the hippocampal area (subiculum) (7 , 11). Sleep-active GABAergic neurons in the ventrolateral preoptic nucleus provide a major input to the tuberomammillary nucleus (12 , 13). Histaminergic neurons also receive very dense orexin innervation originating from the lateral hypothalamus (14). Electrophysiologic studies provided evidence of inhibitory and excitatory synaptic control of tuberomammillary neuron activity by afferents from the diagonal band of Broca, the lateral preoptic area and the anterior lateral hypothalamic area (15). Projections from the brainstem to the tuberomammillary nucleus have also been demonstrated. Retrograde tracing studies combined with immunohistochemistry showed that monoaminergic inputs to the tuberomammillary nucleus originate mainly from the ventrolateral and dorsomedial medulla oblongata and from the raphe nuclei, with a low innervation originating from the locus ceruleus, the ventral tegmental area, and the substantia nigra (16).

MOLECULAR PHARMACOLOGY AND LOCALIZATION OF HISTAMINE RECEPTOR SUBTYPES

Part of "14 - Histamine "

Three histamine receptor subtypes (H_1, H_2 and H_3) have been defined by means of functional assays, followed by design of selective agonists and antagonists and, more recently, cloning of their genes (1). All three belong to the superfamily of receptors with seven transmembrane domains (TMs) and coupled to guanylnucleotide-sensitive G proteins (Table 14.1). In addition, histamine affects the glutamatergic *N*-methyl-D-aspartate (NMDA) receptor (17 ,18).

TABLE 14.1. PROPERTIES OF THREE HISTAMINE RECEPTOR SUBTYPES

	H_1	H_2	H_3
Coding sequence	491 a.a. (bovine) 488 a.a. (guinea pig) 486 a.a. (rat)	358 a.a. (rat) 359 a.a. (dog, human, guinea pig)	445 a.a. (human) H_{3L} 445 a.a., H_{3S} 415 a.a. (guinea pig) H_{3L} 445 a.a., H_{3S} 413 a.a. (rat)
Chromosome localization	3p25	5	20qTEL
Highest brain densities	Thalamus Cerebellum Hippocampus	Striatum Cerebral cortex Amydgala	Striatum Frontal cortex Substantia nigra
Autoreceptor	No	No	Yes
Affinity for histamine	Micromolar	Micromolar	Nanomolar
Characteristic agonists	2-(3-Trifluoromethyl) histamine	Impromidine Sopromidine	(R)α-Methylhistamine Imetit
Characteristic antagonists	Mepyramine	Cimetidine	Thioperamide Ciproxifan
Radioligands	[^{3}H]Mepyramine [^{125}I]Iodobolpyramine	[^{3}H]Tiotidine [^{125}I]Iodoamino-potentidine	[^{3}H](R)α-Methylhistamine [^{125}I]Iodoproxyfan
Second messengers	Inositol phosphates (+) Arachidonic acid (+) cAMP (potentiation)	cAMP (+) Arachidonic acid (–) Ca^{++} (+)	cAMP (–) Inositol phosphates (–) Arachidonic acid (+) Ca^{++} (–)

Histamine H_1 Receptor

The H_1 receptor was initially defined in functional assays (e.g., smooth muscle contraction) and in the design of potent

antagonists, the so-called *antihistamines* (e.g., mepyramine), most of which display prominent sedative properties. Biochemical and localization studies of the H_1 receptor were made feasible with the design of reversible and irreversible radiolabeled probes such as [^{3}H] mepyramine, [^{125}I]iodobolpyramine, and [^{125}I]iodoazidophenpyramine (19 ,20).

Various intracellular responses were found to be associated with H_1-receptor stimulation: inositol phosphate release, increase in Ca^{2+} fluxes, cyclic adenosine monophosphate (cAMP) or cyclic guanosine monophosphate accumulation in whole cells and arachidonic acid release (1).

The deduced amino acid sequence of a bovine H_1 receptor was disclosed after expression cloning of a corresponding cDNA. The latter was based on the detection of a Ca^{2+}-dependent Cl^- influx into microinjected *Xenopus* oocytes. After the transient expression of the cloned cDNA into COS-7 cells, the identity of the protein as an H_1 receptor was confirmed by binding studies (21).

Starting from the bovine sequence, the H_1 receptor DNA was also cloned in the guinea pig (22), a species in which the pharmacology of the receptor is better established, as well as from several other species including humans (1). Although marked species differences in H_1-receptor pharmacology had been reported (2), the sequence homology between the putative TMs of the proteins is high (90%).

The “anatomy” of the H_1 receptor, with a long third intracellular domain and a short C-terminal tail, is similar to that of other receptors positively coupled to phospholipases A_2 and C. Amino acid sequence homology between the TMs of the H_1 and those of the muscarinic receptors (approximately 45%) is higher than between those of H_1 and H_2 receptors (approximately 40%). H_1-receptor antagonists often display significant antimuscarinic activity but only limited H_2-receptor–antagonist properties.

A single gene seems to encode the guinea pig H_1 receptor, and mRNAs of similar size were detected in brain areas and peripheral tissues (22).The structure of the human gene was disclosed (23). Like other receptors of this superfamily, it contains an intron in the 5′ flanking untranslated region, close to the translation initiation codon, but the translated region is intronless.

When stably expressed in transfected fibroblasts, the guinea pig H_1 receptor was found to trigger a large variety of intracellular signals involving or not coupling to pertussis toxin-sensitive G proteins (G_i or G_o), namely, Ca^{2+} transients, inositol phosphates, or arachidonate release (24). H_1 receptor stimulation potentiates cAMP accumulation induced by forskolin in the same transfected fibroblasts, a response that resembles the H_1 potentiation of histamine H_2- or adenosine A_2-receptor–induced accumulation of cAMP in brain slices. All these responses mediated by a single H_1 receptor were known to occur in distinct cell lines or brain slices, but they could have resulted from stimulation of isoreceptors.

Constitutive activity of the recombinant human H_1 receptor

consisting in an agonist-independent increase in inositol phosphates accumulation in COS-7 cells was evidenced. Several H_1-receptor antagonists behaved as inverse agonists (i.e., reduced this constitutive activity), but the physiologic relevance of the process, such as in brain, was not established (25).

The H_1 receptor mediates various excitatory responses in brain (26). A reduction of a background leakage K^+ current was implicated in these responses, in cortical, striatal, and lateral geniculate relay neurons (27 ,28).

H_1-receptor distribution in the guinea pig brain was established autoradiographically using [^{3}H]mepyramine or the more sensitive probe [^{125}I]iodobolpyramine (20), and the information was complemented by *in situ* hybridization of the mRNA (22). For instance, the high density of H_1 receptors in the molecular layers of cerebellum and hippocampus seems to correspond to dendrites of Purkinje and pyramidal cells, respectively, in which the mRNA is highly expressed. H_1 receptors are also abundant in guinea pig thalamus, hypothalamic nuclei (e.g., ventromedial nuclei), nucleus accumbens, amygdaloid nuclei, and frontal cortex but not in neostriatum (20), whereas they are more abundant in the human neostriatum (29). The H_1 receptor was visualized in the primate and human brain *in vivo* by positron emission tomography using [^{11}C]mepyramine (30).

Blockade of H_1 receptors located in cerebral areas involved in wakefulness and cognition, and including those mediating excitation of thalamic relay neurons (31), neocortical pyramidal neurons (28) and ascending cholinergic neurons (32 ,33), presumably accounts for the sedative properties of “antihistamines” of the first generation.

Histamine H_2 Receptor

The molecular properties of the H_2 receptor remained largely unknown for a long time. Reversible labeling of the H_2 receptor was achieved using [^{3}H]tiotidine or, more reliably, [^{125}I]iodoaminopotentidine (2).

By screening cDNA or genomic libraries with homologous probes, the intronless gene encoding the H_2 receptor was first identified in dogs (34) and, subsequently, in other species including humans (1). The H_2 receptor is organized like other receptors positively coupled to adenylyl cyclase: it displays a short third intracellular loop and a long C-terminal cytoplasmic tail.

Using transfected cell lines, positive linkage of the H_2 receptor with adenylyl cyclase was confirmed, and unexpected inhibition of arachidonate release and stimulation of Ca^{2+} transients was evidenced (1). Hence H_2 receptor stimulation can trigger intracellular signals either opposite or similar to those evoked by H_1 receptor stimulation. Parallel observations were made for a variety of biological responses mediated by the two receptors in peripheral tissues.

Helmut Haas and colleagues showed that, in hippocampal pyramidal neurons, H_2-receptor stimulation potentiates excitatory signals by decreasing a Ca^{2+}-activated K^+ conductance, presumably by cAMP production (26). H_2-receptor activation depolarizes thalamic relay neurons slightly and increases apparent membrane conductance markedly, responses caused by enhancement of the hyperpolarization-activated cation current I_h (27). In addition to these short-lasting effects, histamine also induces very long-lasting increases in excitability in the CA1 region of the hippocampus through activation of H_2 receptors and the cAMP/cAMP-dependent protein kinase signal transduction cascade. This process is modulated by other receptors such as the H_1 receptor (35).

The sole selective H_2-receptor antagonist known to enter the brain is zolantidine, a compound used sometimes in animal behavioral studies but not introduced in therapeutics (36). However, some tricyclic antidepressants are known to block H_2-receptor–linked adenylyl cyclase potently and interact with [^{125}I]iodoaminopotentidine binding in a complex manner (37).

Autoradiographic localization of the H_2 receptor in guinea pig (20) and human brain (29) shows it distributed heterogeneously. The H_2 receptor is found in most areas of the cerebral cortex, with the highest density in the superficial layers, the piriform, and the occipital cortices, which contain low H_1-receptor density. The caudate putamen, the ventral striatal complex, and the amygdaloid nuclei (bed nucleus of the stria terminalis) are among the richest brain areas. In the hippocampal formation, the relative localizations of the H_2 receptor and its gene transcripts are similar to those observed for the H_1 receptor: the gene transcripts are expressed in all pyramidal cells of the Ammon horn and in granule cells of the dentate gyrus (38), whereas the H_2 receptor is expressed in the molecular layers of these areas, which contain the dendritic trees of the mRNA-containing neurons. The partial overlap with the H_1 receptor may account for their synergistic interaction in cAMP accumulation.

Histamine H_3 Receptor

The H_3 receptor was initially detected as an autoreceptor controlling histamine synthesis and release in brain. Thereafter, it was shown to inhibit presynaptically the release of other monoamines in brain and peripheral tissues as well as of neuropeptides from unmyelinated C fibers (39 ,40).

Reversible labeling of this receptor was first achieved using the highly selective agonist [^{3}H](R)α-methylhistamine (2), then [^{3}H]N^{α}-methylhistamine, a less selective agonist, was also proposed (19), as well as, more recently, [^{125}I]iodophenpropit and [^{125}I]iodoproxyfan, two antagonists (41).

The regulation of agonist binding by guanylnucleotides (39), and the sensitivity of several H_3-receptor–mediated

responses to pertussis toxin (42 ,43), suggested that the H_3 receptor was G_i/G_o protein coupled, a suggestion confirmed by the cloning of the corresponding human (44) and rodent (45) cDNAs. The H_3 receptor gene contains two introns in its coding sequence and several splice variants such as H_{3L} and H_{3S} differing by a stretch of 30 amino acids in the third intracellular loop, were identified (45). The existence of these variants may partly account for the apparent H_3-receptor heterogeneity in binding or functional studies (46).

Significant differences in the pharmacology of the human and rodent H_3 receptor (47) could be assigned to differences in only two amino acid residues in the third TM (48). In various cell lines, stimulation of the H_3 receptor, like that of other G_i-protein–coupled receptors, inhibits adenylate cyclase (44) or phospholipase C (42) and activates phospholipase A_2 (48a).

On neurons, the H_3 receptor mediates presynaptic inhibitions of release of several neurotransmitters, including histamine itself (2 ,39), norepinephrine, serotonin, dopamine, glutamate, GABA, and tachykinins (40), presumably by inhibiting voltage-dependent calcium channels (39 ,43).

Several H_3-receptor antagonists, such as thioperamide and ciproxifan, potently enhance histamine release *in vitro* and *in vivo* (2 ,39 ,49).This response was originally attributed to blockade of the inhibitory effects of endogenous histamine and was therefore used in many studies, such as behavioral studies, to delineate the functions of histaminergic neurons. However, these drugs were shown to act, in fact, as inverse agonists, and the native H_3 receptor in brain display high constitutive activity including *in vivo* (48a).

Autoradiography of the H_3 receptor in rat (50 ,51) and monkey brain shows it highly concentrated in the neostriatum, the nucleus accumbens, the cingulate and infralimbic cortices, the bed nucleus of the stria terminalis, and the substantia nigra pars lateralis. In contrast, its density is relatively low in the hypothalamus (including the tuberomammillary nucleus), which contains the highest density of histaminergic axons (and perikarya), a finding indicating that most H_3 receptors are not autoreceptors. In agreement with this concept, intrastriatal kainate strongly decreases H_3 binding sites in the forebrain (as well as in the substantia nigra, consistent with their presence in striatonigral neurons) (50 ,51). In the human brain, the high densities of H_3 receptors found in the striatum and globus pallidus (29) were lower in patients with Huntington disease, a finding suggesting that the H_3 receptor is also located on striatonigral projection neurons of the direct and indirect pathways (52). Consistent with the proposal that most H_3 receptors are not autoreceptors, a strong expression of H_3-receptor mRNAs was observed not only within the tuberomammillary nucleus, but also in various regions of the rat (44) and guinea pig (45) brain, including the cerebral cortex, the basal ganglia, and the thalamus.

Interaction with NMDA Receptors

Histamine potentiates NMDA-evoked currents in acutely dissociated and cultured hippocampal and cortical neurons, an effect that could not be ascribed to activation of the known histamine receptors (17 ,18), but rather of a novel recognition site on NMDA receptors containing the subunits NR1A/NR2B (53).

Histamine may play a role in modulating the functions of NMDA receptors *in vivo*. It facilitates the NMDA-induced depolarization of projection neurons in cortical slices (54) and phase shifts the circadian clock by a direct potentiation of NMDA currents in the suprachiasmatic nucleus (55). Histamine, presumably acting through NMDA receptors, facilitates the induction of long-term potentiation and causes long-lasting increases of excitability in the CA1 region of rat hippocampal slices (35).

The histamine-induced modulation of NMDA responses is higher under slightly acidic conditions (56), which occur during hypoxia or epileptiform activity. This may lead to enhancement of neurotransmission or histamine-mediated neuronal death such as that observed in a rat model of Wernicke encephalopathy (57).

HISTAMINERGIC NEURON ACTIVITY AND THEIR CONTROL

Part of "14 - Histamine "

Electrophysiologic Properties

Cortically projecting histaminergic neurons share with other aminergic neurons certain electrophysiologic properties evidenced by extracellular recording. They fire spontaneously slowly and regularly, and their action potentials are of long duration (26). Among the pacing events that may contribute to their spontaneous firing, tuberomammillary neurons exhibit a tetrodotoxin-sensitive persistent Na^+ current (58), a Ca^{2+} current probably of the low-threshold type (59), and multiple high-voltage–activated Ca^{2+} currents (43). In addition, they exhibit inward rectification attributed to an I_h current that may increase whole-cell conductance and may decrease the efficacy of synaptic inputs during periods of prolonged hyperpolarization, that is, when histaminergic neurons fall silent (60).

Modulation of Histamine Synthesis and Release In Vitro

The autoreceptor-regulated modulation of histamine synthesis in, and release from, brain neurons is well documented (2). It was initially evidenced in brain slices or synaptosomes after labeling the endogenous pool of histamine using the [^{3}H]histamine precursor. Exogenous histamine decreases the release and formation of [^{3}H]histamine induced by depolarization, and analysis of these responses led to the pharmacologic definition of H_3 receptors. The autoregulation

was found in various brain regions known to contain histamine nerve endings, a finding suggesting that all terminals are endowed with H_3 autoreceptors.

Regulation of histamine synthesis was also observed in the posterior hypothalamus (39), and somatodendritic H_3 autoreceptors inhibit the firing of tuberomammillary neurons (26) by modulating high-voltage–activated calcium channels (43).

Galanin, a putative cotransmitter of a subpopulation of histaminergic neurons, regulates histamine release only in regions known to contain efferents of this subpopulation, that is, in hypothalamus and hippocampus but not in cerebral cortex or striatum (61). In brain slices, galanin also hyperpolarizes and decreases the firing rate of tuberomammillary neurons (26). It is not known, however, whether these galanin receptors behave as “autoreceptors” modulating galanin release from histaminergic nerve terminals, inasmuch as the tuberomammillary nucleus receives a strong galaninergic innervation from the ventrolateral preoptic area (12 ,13). Other putative cotransmitters of histaminergic neurons failed to affect [^{3}H]histamine release from slices of rat cerebral cortex (62). However, GABAergic inhibitory postsynaptic potentials are mediated by $GABA_A$ receptors located on histaminergic neurons (63). To what extent these receptors play an autoinhibitory role is unclear. A subpopulation of histaminergic neurons contains GABA (5), but the tuberomammillary nucleus also receives dense GABAergic innervation (12 ,13 ,15).

[^{3}H]Histamine synthesis and release are inhibited in various brain regions by stimulation of not only autoreceptors but also α_2-adrenergic receptors, M_1-muscarinic receptors, and κ-opioid receptors (2). Because these regulations are also observed with synaptosomes (62), all these receptors presumably represent true presynaptic heteroreceptors. In contrast, histamine release is enhanced by stimulation of nicotinic receptors in rat hypothalamus (64) and by μ-opioid receptors in mouse cerebral cortex (2).

Some molecular mechanisms regulating neuronal histamine dynamics remain unclear. *N*-methylation catalyzed by histamine *N*-methyltransferase is the major process responsible for termination of histamine actions in the brain (2), and genetic polymorphisms for the enzyme have been associated with altered levels of its activity (65). No histamine transporter could be evidenced, and direct feedback inhibition of histidine decarboxylase by histamine has been excluded (2).

Changes in Histaminergic Neuron Activity In Vivo

Both neurochemical and electrophysiologic studies indicate that the activity of histaminergic neurons is high during arousal. In rat hypothalamus, histamine levels are low, whereas synthesis is high during the dark period, a finding suggesting that neuronal activity is enhanced during the active phase (2). In mouse cerebral cortex, striatum, and hypothalamus, telemethylhistamine levels are doubled at the end of the dark phase of the cycle as compared with the beginning of the light phase (66). Histamine release from the anterior hypothalamus of freely moving rats, evaluated by *in vivo* microdialysis, gradually increases in the second half of the light period and is maintained at a maximal level during the active phase (67). Such state-related changes are also found in single-unit extracellular recordings performed in the ventrolateral posterior hypothalamus of freely moving cats. Neurons with properties consistent with those of histaminergic neurons exhibited a circadian rhythm of their firing rate, falling silent during deep slow-wave or paradoxical sleep (2). An important determinant of this circadian rhythm of tuberomammillary histaminergic neuron activity is a GABAergic inhibitory pathway originating in the ventrolateral preoptic area and activated during sleep (12 ,15).

A feeding-induced increase in the activity of histaminergic neurons has also been shown by microdialysis performed in the hypothalamus of conscious rats (68). Histaminergic neurons are a target for leptin in its control of feeding. An enhancement of histamine turnover was observed after intracerebroventricular infusion of leptin (70). Changes in the metabolism and release of histamine observed *in vivo* after occlusion of the middle cerebral artery in rats suggest that the histaminergic activity is also enhanced by cerebral ischemia (71).

Whereas H_1 and H_2 receptors are apparently not involved, inhibition mediated by the H_3-autoreceptor constitutes a major regulatory mechanism for histaminergic neuron activity under physiologic conditions. Administration of selective H_3 receptor agonists reduces histamine turnover (2) and release, as shown by microdialysis (72). In contrast, H_3-receptor antagonists enhance histamine turnover (2 ,49) and release *in vivo* (73 ,74), a finding suggesting that autoreceptors are tonically activated.

Agents inhibiting histamine release *in vitro* through stimulation of presynaptic α_2-adrenergic or muscarinic heteroreceptors reduce histamine release and turnover *in vivo*, but systemic administration of antagonists of these receptors does not enhance histamine turnover, a finding suggesting that these heteroreceptors are not tonically activated under basal conditions.

Activation of central nicotinic receptors inhibits histamine turnover (75). Several types of serotonergic receptors are likely to modulate histamine neuron activity. 5-Hydroxytryptamine $(5\text{-HT})_{1A}$-receptor agonists inhibit (76), whereas 5-HT_2-receptor antagonists enhance (77), histamine turnover in various brain regions. Stimulation of D2 (but not D3) dopamine receptors by endogenous dopamine released by amphetamine increases histamine neuron activity (77 ,78).

Histamine turnover in the brain is rapidly reduced after administration of various sedative drugs such as ethanol, Δ^9-tetrahydrocannabinol, barbiturates, and benzodiazepines (2),

presumably as a result of their interaction with GABA receptors present on nerve endings and on perikarya of histaminergic neurons (63 ,79).

In contrast, stimulation of μ-opioid receptors enhances histamine turnover in brain (2). NMDA receptors increase *in vivo* release of histamine from the anterior hypothalamus (80). Activation of NMDA and non-NMDA receptors in the diagonal band of Broca, the lateral preoptic area, and the anterior hypothalamic area led to inhibition or enhancement of firing rates of tuberomammillary neurons (15).

PHYSIOLOGIC ROLES OF HISTAMINERGIC NEURONS

Part of "14 - Histamine "

In spite of many different suggestions mainly derived from the observations of responses to locally applied histamine, only a few physiologic roles of histaminergic neurons appear relatively well documented.

Arousal

Our initial proposal in 1977 (81) that histaminergic neurons play a critical role in arousal has been confirmed by data from a variety of experiments mainly performed by Lin and Jouvet in cats (33) and Monti in rats (2). In agreement with this concept, ablation of these neurons and inhibition of histamine synthesis, release, or action by the H_1 receptor decrease wakefulness and increase deep slow-wave sleep; conversely, inhibition of histamine methylation or facilitation of histamine release by H_3 receptor blockade increase arousal (49).

The arousing effect of histamine may result from H_1 and H_2-receptor–mediated depolarization of thalamic relay neurons that induces a shift of their activity from burst firing (predominating in deep sleep during which they are poorly responsive to sensory inputs) to single spike activity (predominating in arousal during which sensory information is more faithfully relayed) (31). Arousal may also result from H_1-receptor–mediated excitation of neocortical pyramidal neurons by the same mechanism as in thalamus, that is, reduction of a background leakage potassium current (28). Finally, arousal may occur indirectly by H_1-receptor–mediated excitation of ascending cholinergic neurons within the nucleus basilis or mesopontine tegmentum (32 ,33), which also induces cortical activation.

All these cellular actions of histamine, together with observations that tuberomammillary neuron firing is maximal during wakefulness, suggest that histaminergic systems make an important contribution to the control of arousal. The circadian changes in histaminergic neuron activity seem to be directed by two major neuronal inputs arising from the anterior hypothalamus. The first ones are slow-wave sleep-activated inhibitory GABA- and galanin-containing neurons arising from the ventrolateral preoptic area (12 ,13); in contrast, neurons releasing the neuropeptide orexin that emanate from the lateral hypothalamus appear to exert opposite actions because disruption of the orexin gene is associated with narcolepsy in dogs and knockout mice (14 ,82). Other monoaminergic neurons participating in control of sleep and wakefulness states as well as GABA/galanin ventrolateral preoptic neurons also receive inputs from orexin neurons, which are, themselves, likely influenced by photic signals from the suprachiasmatic nucleus. In turn, neurons from the suprachiasmatic nucleus and the preoptic area seem to be influenced in a complex manner by histaminergic inputs (83 ,84). Hence a complex neuronal network in the hypothalamus with reciprocal influences involving histaminergic neurons seems to control wakefulness.

The major part played by the H_1 receptor in these processes, confirmed in mutant mice lacking this receptor (85), accounts for the sedating effects of the first generation of “antihistamines,” that is, antagonists that easily enter the brain and are still ingredients of over-the-counter sleeping pills (86). It may also account for the sedative side effects of many antipsychotic or antidepressant drugs that are potent H_1 antagonists.

Cognitive Functions

The idea that histaminergic neurons may improve cognitive performance is consistent with projections of these neurons to brain areas such as the prefrontal and cingulate cortices or hippocampus, their excitatory influences therein, and their positive role in wakefulness.

Ciproxifan, a potent and selective H_3-receptor antagonist (or inverse agonist), which strongly enhances histamine turnover in brain, improved attentional performances in the rat five-choice test under conditions similar to those of drugs enhancing cholinergic transmissions (49). Various H_3 antagonists facilitate various forms of learning. They improve short-term social memory in rats (87), reverse the scopolamine- or senescence-induced learning deficit in a passive avoidance test in mice (88), and facilitate retention in a footshock avoidance test in mice (89).

Generally, H_3 agonists exert opposite effects, and the effects of H_3-antagonists are reversed by H_1 antagonists, a finding that suggests that these effects are attributable to enhanced histamine release. In contrast to the large body of experiments indicating a “procognitive” role of tuberomammillary neurons, Huston and coworkers repeatedly showed that excitotoxic lesions aimed at these neurons ablation result in improvement of learning in a variety of tests (e.g., ref. 90). The discrepancy with data from pharmacologic approaches may result from the difficulty to achieve selective histamine neuron ablation.

Control of Pituitary Hormone Secretion

Histamine affects secretion of several pituitary hormones (2 ,91). Magnocellular neurons of the supraoptic and paraventrical

nuclei are typically excited, an essentially H_1-receptor–mediated response resulting in enhanced blood levels of vasopressin and oxytocin. Histaminergic neurons are activated during dehydration, parturition, and lactation, and histamine release onto magnocellular neurons participates in the control of these physiologic processes by the neurohypophysial hormones (92 ,93).

Histaminergic neurons may also participate in the hormonal responses to stress. In agreement with this concept, they are activated during various forms of stress and heavily project to hypothalamic or limbic brain areas (e.g., the amygdala or bed nucleus of the stria terminalis) involved in these responses. Various pharmacologic studies have shown the participation of endogenous histamine by H_1- and H_2-receptor stimulation in the adrenocorticotropic hormone–corticosterone, prolactin, or renin responses to stressful stimuli such as restraint, endotoxin, or dehydration (94). Many histaminergic neurons contain estrogen receptors, project to luteinizing hormone–releasing hormone neurons in preoptic and infundibular regions, and may constitute, by H_1-receptor stimulation, an important relay in the estradiol-induced preovulatory luteinizing hormone surge (95).

Satiation

Weight gain is often experienced by patients receiving H_1 antihistamines as well as by patients taking antipsychotics or antidepressants displaying potent H_1-receptor antagonist properties. These effects reflect the inhibitory role of endogenous histamine on food intake mediated by the H_1 receptor, namely, on the ventromedial nucleus (97). Histamine neurons projecting to the hypothalamus may be responsible for the food intake suppression induced by leptin (70).

Seizures

The anticonvulsant properties of endogenous histamine were initially suggested from the occurrence of seizures in patients with epilepsy, particularly children, after administration of high doses of H_1-receptor antagonists crossing the blood–brain barrier, even those agents devoid of anticholinergic activity (86). These drugs, by completely occupying the H_1 receptor, as assessed by positron emission tomography studies (30), could block the histamine-induced reduction of a background-leakage K^+ current.

Drug-induced changes in histamine synthesis, release or metabolism confirmed the role of the endogenous amine acting through the H_1 receptor in preventing seizure activity elicited by pentetrazol, transcranial electrical stimulation, or amygdaloid kindling. Acquired amygdaloid kindling susceptibility appears associated with reduced histamine synthesis in limbic brain areas (97). In addition, kainic acid–induced limbic seizures are accompanied by up-regulation of the H_1-receptor mRNA in striatum and dentate gyrus, a finding consistent with a regulatory role of this receptor in seizure activity (98).

Nociception

The antinociceptive effects of histidine loads, H_3-receptor antagonists, and histamine *N*-methyltransferase inhibitors, as well as opposite effects of histamine synthesis inhibitors or H_3 agonists, support the idea that brain histamine inhibits nociceptive responses such as the mouse hot plate jump (99). In contrast, peripherally acting H_3-receptor agonists prevent nociceptive responses such as mouse abdominal constriction by inhibiting sensory C-fiber activity (100).

ROLE OF HISTAMINERGIC NEURONS IN NEUROPSYCHIATRIC DISEASES

Part of "14 - Histamine "

Among the various approaches that tend to establish the implication of other neuronal systems in neuropsychiatric diseases, so far only a few have been applied to histamine.

Histamine, Schizophrenia, and Antipsychotic Drug Actions

Overdose of various classic H_1 antagonists was repeatedly reported to result in toxic psychoses with hallucinations resembling schizophrenia, and the hallucinogenic potential of these drugs has even led to abuse (86). Conversely, metamphetamine, a drug with hallucinogenic potential and to which patients with schizophrenia seem hyperresponsive, releases histamine in rodent brain areas, an indirect effect mediated by stimulation of D2 and not D3 dopamine receptors (77 ,78). Even more, endogenous dopamine appears to exert a tonic stimulation of histamine neurons because typical neuroleptics, such as haloperidol, decrease their activity. In contrast, atypical neuroleptics, such as clozapine, enhance histamine turnover, an effect related to 5-HT_2 receptor blockade and possibly underlying their procognitive properties (77). The locomotor activation elicited in rodents by amphetamine and other dopaminergic agonists is attenuated by H_3-receptor blockade (101). Repeated amphetamine administration to rodents that results in behavioral sensitization to dopamine agonists, a cardinal feature of schizophrenia, is accompanied by enhanced histamine release, a finding that presumably reflects an enhanced tonic dopaminergic influence on histaminergic neurons (77 ,78). In one comprehensive study, an enhanced level of *t*-methylhistamine, the major histamine metabolite, was detected in the cerebrospinal fluid of patients with schizophrenia, who were either treated or untreated with neuroleptic agents (102).

In several open studies, famotidine, an H_2 antagonist, was found to improve schizophrenia in patients, a finding

that remains to be confirmed in control studies. A previous claim of association between polymorphisms of the H_2-receptor gene and schizophrenia could not be confirmed (103).

These various observations, although not readily forming a coherent picture, suggest that histaminergic neuron activity is enhanced in patients with schizophrenia, and blockade of H_2 or H_3 receptors could be useful in the treatment of this disease.

Histamine and Alzheimer's Disease

Neuropathologic studies have documented a deficit in histaminergic neurotransmission in Alzheimer's disease. In some, but not all, cortical areas (e.g., the frontal or temporal cortex) of brains affected by Alzheimer's disease, there is a decrease of histamine and histidine decarboxylase levels that may reach up to approximately 50% (104), and the expression of the *hdc* gene in neurons of the tuberomammillary nucleus is also reduced (S. Trottier, personal communication). Decreased histaminergic input may affect cholinergic neuron activity in the nucleus basilis (32) and acetylcholine release in cortical areas.

If one takes into account an additional direct positive influence of histamine on attention and memory, this indicates that enhancing histaminergic neurotransmission may constitute a novel symptomatic therapeutic approach to Alzheimer's disease. The drug tacrine was even more potent in inhibiting histamine-*N*-methyltransferase, the main histamine-metabolizing enzyme, than acetylcholinesterase (105).

Histamine and Other Neuropsychiatric Disorders

Anxiety may be increased by endogenous histamine acting at the H_1 receptor. In agreement with this concept, H_1-receptor knockout mice display significantly less anxiety in the elevated maze test (106). However, the utility of H_1-receptor antagonists in anxiety disorders is not established.

Patients with *attention-deficit disorders* may benefit from enhanced histamine release, as suggested by the therapeutic effect of amphetamine in children and the attention-enhancing effects of an H_3-receptor antagonist in the rat (49).

Antidepressant-like effects in the mouse forced swim test result from enhanced histamine release and H_1-receptor activation (107).

CONCLUSION

Part of "14 - Histamine "

This chapter describes how our knowledge of the molecular neurobiology of cerebral histaminergic systems and their implications in physiologic functions, such as arousal or hormonal regulations, have progressed over the years. In contrast, little is known, so far, about their possible implications in neuropsychiatric diseases and the therapeutic utility of psychotropic drugs to affect their activity. H_3-receptor antagonists (or inverse agonists) that markedly enhance brain histamine release are currently undergoing clinical trials. It seems likely that the next edition of this book will see their place in therapeutics established.

REFERENCES

1. Hill SJ, Ganellin CR, Timmermann H, et al. International Union of Pharmacology. XIII. Classification of histamine receptors. *Pharmacol Rev* 1997;49:253–278.

2. Schwartz JC, Arrang JM, Garbarg M, et al. Histaminergic transmission in the mammalian brain. *Physiol Rev* 1991;71:1–51.

3. Garbarg M, Barbin G, Feger J, et al. Histaminergic pathway in rat brain evidenced by lesions of the MFB. *Science* 1974;186:833–835.

4. Panula P, Airaksinen MS. The histaminergic neuronal system as revealed with antisera against histamine. In: Watanabe T, Wada H, eds. *Histaminergic neurons: morphology and function*. Boca Raton, FL: CRC, 1991:127–144.

5. Staines WA, Nagy JI. Neurotransmitter coexistence in the tuberomammillary nucleus. In: Watanabe T, Wada H, eds. *Histaminergic neurons: morphology and function*. Boca Raton, FL: CRC, 1991:163–176.

6. Tohyama M, Tamiya R, Inagaki N, et al. Morphology of histaminergic neurons with histidine decarboxylase as a marker. In: Watanabe T, Wada H, eds. *Histaminergic neurons: morphology and function*. Boca Raton, FL: CRC, 1991:107–126.

7. Wouterlood FG, Steinbusch HWM. Afferent and efferent fiber connections of histaminergic neurons in the rat brain: comparison with dopaminergic, noradrenergic and serotonergic systems. In: Watanabe T, Wada H, eds. *Histaminergic neurons: morphology and function*. Boca Raton, FL: CRC, 1991:145–162.

8. Airaksinen MS, Paetau A, Paliärui L, et al. Histamine neurons in human hypothalamus: anatomy in normal and Alzheimer diseased brains. *Neuroscience* 1991;44:465–481.

9. Bayliss DA, Wang YM, Zahnow CA, et al. Localization of histidine decarboxylase mRNA in rat brain. *Mol Cell Neurosci* 1990;1:3–9.

10. Peter D, Liu Y, Sternini C, et al. Differential expression of two vesicular monoamine transporters. *J Neurosci* 1995;15:6179–6188.

11. Ericson H, Blomqvist A, Köhler C. Origin of neuronal inputs to the region of the tuberomammillary nucleus of the rat brain. *J Comp Neurol* 1991;311:45–64.

12. Sherin JE, Shiromani PJ, McCarley RW, et al. Activation of ventrolateral preoptic neurons during sleep. *Science* 1996;271:216–219.

13. Sherin JE, Elmquist JK, Torrealba F, et al. Innervation of histaminergic tuberomammillary neurons by GABAergic and galaninergic neurons in the ventrolateral preoptic nucleus of the rat. *J Neurosci* 1998,18:4705–4721.

14. Chemelli RM, Willie JT, Sinton CM, et al. Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell* 1999;98:437–451.

15. Yang QZ, Hatton GI. Electrophysiology of excitatory and inhibitory afferents to rat histaminergic tuberomammillary nucleus neurons from hypothalamic and forebrain sites. *Brain Res* 1997;773:162–172.

16. Ericson H, Blomqvist A, Köhler C. Brainstem afferents to the tuberomammillary nucleus in the rat brain with special reference to monoaminergic innervation. *J Comp Neurol* 1989;281:169–192.

17. Bekkers JM. Enhancement by histamine of NMDA-mediated synaptic transmission in the hippocampus. *Science* 1993;261:104–106.

18. Vorobjev VS, Sharonova IN, Walsh IB, et al. Histamine potentiates *N*-methyl-D-aspartate responses in acutely isolated hippocampal neurons. *Neuron* 1993;11:837–844.

19. Garbarg M, Traiffort E, Ruat M, et al. Reversible labelling of H_1, H_2 and H_3-receptors. In: Schwartz JC, Haas HL, eds. *The histamine receptor*. New York: Wiley Liss, 1992:73–95.

20. Pollard H, Bouthenet ML. Autoradiographic visualization of the three histamine receptor subtypes in the brain. In: Schwartz JC, Haas HL, eds. *The histamine receptor*. New York: Wiley Liss, 1992:179–192.

21. Yamashita M, Fukui H, Sugawa K, et al. Expression cloning of a cDNA encoding the bovine histamine H_1 receptor. *Proc Natl Acad Sci USA* 1991;88:11515–11519.

22. Traiffort E, Leurs R, Arrang JM, et al. Guinea pig histamine H_1 receptor. I. Gene cloning, characterization and tissue expression revealed by *in situ* hybridization. *J Neurochem* 1994;62:507–518.

23. De Backer MD, Loonen I, Verhasselt P, et al. Structure of the human histamine H_1 receptor gene. *Biochem J* 1998;335:663–670.

24. Leurs R, Traiffort E, Arrang JM, et al. Guinea pig histamine H_1 receptor. II. Stable expression in chinese hamster ovary cells reveals the interaction with three major signal transduction pathways. *J Neurochem* 1994;62:519–527.

25. Bakker RA, Wieland K, Timmerman H, et al. Constitutive activity of the histamine H_1 receptor reveals inverse agonism of histamine H_1 receptor antagonists. *Eur J Pharmacol* 2000;387:R5–R7.

26. Haas HL. Electrophysiology of histamine receptors. In: Schwartz JC, Haas HL, eds. *The histamine receptor*. New York: Wiley Liss, 1992:161–177.

27. McCormick DA, Williamson A. Modulation of neuronal firing mode in cat and guinea pig LGNd by histamine: possible cellular mechanisms of histaminergic control of arousal. *J Neurosci* 1991;11:3188–3199.

28. Reiner PB, Kamondi A. Mechanisms of antihistamine-induced sedation in the human brain: H_1 receptor activation reduces a background leakage potassium current. *Neuroscience* 1994;59:579–588.

29. Martinez Mir MI, Pollard H, Moreau J, et al. Three histamine receptors (H_1, H_2 and H_3) visualized in the brain of human and non-human primates. *Brain Res* 1990;526:322–327.

30. Yanai K, Watanabe T, Yokoyama H, et al. Mapping of histamine H_1 receptors in the human brain using (^{11}C)pyrilamine and positron emission tomography. *J Neurochem* 1992;59:128–136.

31. Steriade M, McCormick DA, Sejnowski TJ. Thalamocortical oscillations in the sleeping and aroused brain. *Science* 1993;262:679–685.

32. Khateb A, Fort P, Pegna A, et al. Cholinergic nucleus basalis neurons are excited by histamine *in vitro*. *Neuroscience* 1995;69:495–506.

33. Lin JS, Hou Y, Sakai K, et al. Histaminergic descending inputs to the mesopontine tegmentum and their role in the control of cortical activation and wakefulness in the cat. *J Neurosci* 1996;16:1523–1527.

34. Gantz I, Schaffer M, Delvalle J, et al. Molecular cloning of a gene encoding the histamine H_2 receptor. *Proc Natl Acad Sci USA* 1991;88:429–433.

35. Selbach O, Brown RE, Haas HL. Long-term increase of hippocampal excitability by histamine and cyclic AMP. *Neuropharmacology* 1997;36:1539–1548.

36. Ganellin CR. Pharmacochemistry of H_1 and H_2 receptors. In: Schwartz JC, Haas, HL, eds. *The histamine receptor*. New York: Wiley Liss, 1992:1–56.

37. Traiffort E, Ruat M, Schwartz JC. Interaction of mianserin, amitriptyline and haloperidol with guinea pig cerebral histamine H_2 receptors studied with [^{125}I]iodoaminopotentidine. *Eur J Pharmacol* 1991;207:143–148.

38. Vizuete ML, Traiffort E, Bouthenet ML, et al. Detailed mapping of the histamine H_2 receptor and its gene transcripts in guinea pig brain. *Neuroscience* 1997;80:321–343.

39. Arrang JM, Garbarg M, Schwartz JC. H_3-receptor and control of histamine release. In: Schwartz JC, Haas HL, eds. *The histamine receptor*. New York: Wiley Liss, 1992:145–159.

40. Schlicker E, Malinowska B, Kathmann M, et al. Modulation of neurotransmitter release via histamine H_3 heteroreceptors. *Fundam Clin Pharmacol* 1994;8:128–137.

41. Ligneau X, Garbarg M, Vizuete ML, et al. [^{125}I]Iodoproxyfan, a new antagonist to label and visualize cerebral histamine H_3 receptors. *J Pharmacol Exp Ther* 1994;271:452–459.

42. Cherifi Y, Pigeon C, Le Romancier M, et al. Purification of a histamine H_3 receptor negatively coupled to phosphoinositide turnover in the human gastric cell line HGT1. *J Biol Chem* 1992;267:25315–25320.

43. Takeshita Y, Watanabe T, Sakata T, et al. Histamine modulates high-voltage–activated calcium channels in neurons dissociated from the rat tuberomammillary nucleus. *Neuroscience* 1998;87:797–805.

44. Lovenberg TW, Roland BL, Wilson SJ, et al. Cloning and functional expression of the human histamine H_3 receptor. *Mol Pharmacol* 1999;55:1101–1107.

45. Tardivel-Lacombe J, Rouleau A, Héron A, et al. Cloning and cerebral expression of the guinea pig histamine H_3 receptor: evidence for two isoforms. *Neuroreport* 2000;11:755–759.

46. Arrang JM, Morisset S, Pillot C, et al. Subclassification of histamine receptors, H_3-receptor subtypes? Localization of H_3 receptors in the brain. In: Leurs R, Timmerman H, eds. *The histamine H_3 receptor: a target for new drugs*. Amsterdam: Elsevier, 1998:1–12.

47. West RE, Wu RL, Billah MM, et al. The profiles of human and primate [^{3}H]N$^\alpha$-methylhistamine binding differ from that of rodents. *Eur J Pharmacol* 1999;377:233–239.

48. Ligneau X, Morisset S, Tardivel-Lacombe J, et al. Distinct pharmacology of rat and human histamine H_3 receptors: role of two amino acids in the third transmembrane domain. *Br J Pharmacol* 2000;131:1247–1250.

48a. Morisset S, Rouleau A, Ligneau X, et al. High constitutive activity of native H_3 receptors regulates histamine neurons in brain. *Nature* 2000;408:860–864.

49. Ligneau X, Lin JS, Vanni-Mercier G, et al. Neurochemical and behavioral effects of ciproxifan, a potent histamine H_3-receptor antagonist. *J Pharmacol Exp Ther* 1998;287:658–666.

50. Cumming P, Shaw C, Vincent SR. High affinity histamine binding site is the H_3 receptor: characterization and autoradiographic localization in rat brain. *Synapse* 1991;8:144–151.

51. Pollard H, Moreau J, Arrang JM, et al. A detailed autoradiographic mapping of histamine H_3-receptors in rat brain areas. *Neuroscience* 1993;52:169–189.

52. Goodchild RE, Court JA, Hobson I, et al. Distribution of histamine H_3-receptor binding in the normal human basal ganglia: comparison with Huntington's and Parkinson's disease cases. *Eur J Neurosci* 1999;11:449–456.

53. Williams K. Subunit-specific potentiation of recombinant *N*-methyl-D-aspartate receptors by histamine. *Mol Pharmacol* 1994;46:531–541.

54. Payne GW, Neuman RS. Effects of hypomagnesia on histamine H_1 receptor–mediated facilitation of NMDA responses. *Br J Pharmacol* 1997;121:199–204.

55. Meyer JL, Hall AC, Harrington ME. Histamine phase shifts the hamster circadian pacemaker via an NMDA dependent mechanism. *J Biol Rhythms* 1998;13:288–295.

56. Yanovsky Y, Reymann K, Haas HL. pH-dependent facilitation of synaptic transmission by histamine in the CA1 region of mouse hippocampus. *Eur J Neurosci* 1995;7:2017–2020.

57. Carter McRee R, Terry-Ferguson M, Langlais PJ, et al. Increased histamine release and granulocytes within the thalamus of a rat model of Wernicke's encephalopathy. *Brain Res* 2000;858:227–236.

58. Uteshev V, Stevens DR, Haas HL. A persistent sodium current in acutely isolated histaminergic neurons from rat hypothalamus. *Neuroscience* 1995;66:143–149.

59. Stevens DR, Haas HL. Calcium-dependent prepotentials contribute to spontaneous activity in rat tuberomammillary neurons. *J Physiol (Lond)* 1996;493:747–754.

60. Kamondi A, Reiner PB. Hyperpolarization-activated inward current in histaminergic tuberomammillary neurons of the rat hypothalamus. *J Neurophysiol* 1991;66:1902–1911.

61. Arrang JM, Gulat-Marnay C, Defontaine N, et al. Regulation of histamine release in rat hypothalamus and hippocampus by presynaptic galanin receptors. *Peptides* 1991;12:1113–1117.

62. Schwartz JC, Arrang JM, Garbarg M, et al. Modulation of histamine synthesis and release in brain via presynaptic autoreceptors and heteroreceptors. *Ann NY Acad Sci* 1990;604:40–54.

63. Stevens DR, Kuramasu A, Haas HL. $GABA_B$-receptor–mediated control of GABAergic inhibition in rat histaminergic neurons *in vitro*. *Eur J Neurosci* 1999;11:1148–1154.

64. Ono J, Yamatodani A, Kishino J, et al. Cholinergic influence of K^+-evoked release of endogenous histamine from rat hypothalamic slices *in vitro*. *Methods Find Exp Clin Pharmacol* 1992;14:35–40.

65. Preuss CV, Wood TC, Szumlanski CL, et al. Human histamine *N*-methyltransferase pharmacogenetics: common genetic polymorphisms that alter activity. *Mol Pharmacol* 1998;53:708–717.

66. Morisset S, Traiffort E, Arrang JM, et al. Changes in histamine H_3 receptor responsiveness in mouse brain. *J Neurochem* 2000;74:339–346.

67. Mochizuki T, Yamatodani A, Okakura K, et al. Circadian rhythm of histamine release from the hypothalamus of freely moving rats. *Physiol Behav* 1992;51:391–394.

68. Itoh Y, Oishi R, Saeki K. Feeding-induced increase in the extracellular concentration of histamine in rat hypothalamus as measured by *in vivo* microdialysis. *Neurosci Lett* 1991;125:235–237.

69. Deleted in press.

70. Yoshimatsu H, Itateyama E, Kondou S, et al. Hypothalamic neuronal histamine as a target of leptin in feeding behavior. *Diabetes* 1999;48:2286–2291.

71. Adachi N, Itoh Y, Oishi R, et al. Direct evidence for increased continuous histamine release in the striatum of conscious freely moving rats produced by middle cerebral artery occlusion. *J Cereb Blood Flow Metab* 1992;12:477–483.

72. Itoh Y, Oishi R, Adachi N, et al. A highly sensitive assay for histamine using ion-pair HPLC coupled with postcolumn fluorescent derivatization: its application to biological specimens. *J Neurochem* 1992;58:884–889.

73. Itoh Y, Oishi R, Nishibori M, et al. Characterization of histamine release from the rat hypothalamus as measured by *in vivo* microdialysis. *J Neurochem* 1991;56:769–774.

74. Mochizuki T, Yamatodani A, Okakura K, et al. *In vivo* release of neuronal histamine in the hypothalamus of rats measured by microdialysis. *Naunyn Schmiedebergs Arch Pharmacol* 1991;343:190–195.

75. Oishi R, Adachi N, Okada K, et al. Regulation of histamine turnover via muscarinic and nicotinic receptors in the brain. *J Neurochem* 1990;55:1899–1904.

76. Oishi R, Itoh Y, Saeki K. Inhibition of histamine turnover by 8-OH-DPAT, buspirone and 5-hydroxytryptophan in the mouse and rat brain. *Naunyn Schmiedebergs Arch Pharmacol* 1992;345:495–499.

77. Morisset S, Sahm UG, Traiffort E, et al. Atypical neuroleptics enhance histamine turnover in brain via 5-hydroxytryptamine$_{2A}$ receptor blockade. *J Pharmacol Exp Ther* 1999;288: 590–596.

78. Ito C, Onodera K, Sakurai E, et al. Effects of dopamine antagonists on neuronal histamine release in the striatum of rats subjected to acute and chronic treatments with methamphetamine. *J Pharmacol Exp Ther* 1996;279:271–276.

79. Okakura-Mochizuki K, Mochizuki T, Yamamoto Y, et al. Endogenous GABA modulates histamine release from the anterior hypothalamus of the rat. *J Neurochem* 1996;67:171–176.

80. Okakura K, Yamatodani A, Mochizuki T, et al. Glutamatergic regulation of histamine release from rat hypothalamus. *Eur J Pharmacol* 1992;213:189–192.

81. Schwartz JC. Histaminergic mechanisms in brain. *Annu Rev Pharmacol Toxicol* 1977;17:325–339.

82. Lin L, Faraco J, Li R, et al. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* 1999;98:365–376.

83. Cote NK, Harrington ME. Histamine phase shifts the circadian clock in a manner similar to light. *Brain Res* 1993;613:149–151.

84. Lin JS, Sakai K, Vanni-Mercier G, et al. Hypothalamo-preoptic histaminergic projections in sleep-wake control in the cat. *Eur J Neurosci* 1994;66:618–625.

85. Inoue I, Yanai K, Kitamura D, et al. Impaired locomotor activity and exploratory behavior in mice lacking histamine H_1 receptors. *Proc Natl Acad Sci USA* 1996;93:13316–13320.

86. Sangalli BC. Role of the central histaminergic neuronal system in the CNS toxicity of the first generation H_1-antagonists. *Prog Neurobiol* 1997;52:145–157.

87. Prast H, Argyriou A, Philippu A. Histaminergic neurons facilitate social memory in rats. *Brain Res* 1996;734:316–318.

88. Onodera K, Miyazaki S, Imaizumi M, et al. Improvement by FUB 181, a novel histamine H_3-receptor antagonist, of learning and memory in the elevated plus-maze test in mice. *Naunyn Schmiedebergs Arch Pharmacol* 1998;357:508–513.

89. Flood JF, Uezu K, Morley JE. Effect of histamine H_2 and H_3 receptor modulation in the septum on post-training memory processing. *Psychopharmacology* 1998;140:279–284.

90. Frisch C, Hasenöhrl RU, Haas HL, et al. Facilitation of learning after lesions of the tuberomammillary nucleus region in adult and aged rats. *Exp Brain Res* 1998;118:447–456.

91. Knigge U, Warberg J. Minireview: the role of histamine in the neuroendocrine regulation of pituitary hormone secretion. *Acta Endocrinol* 1991;124:609–619.

92. Kjaer A, Larsen PJ, Knigge U, et al. Dehydration stimulates hypothalamic gene expression of histamine synthesis enzyme: importance for neuroendocrine regulation of vasopressin and oxytocin secretion. *Endocrinology* 1995;136:2189–2197.

93. Luckman SM, Larsen PJ. Evidence for the involvement of histaminergic neurones in the regulation of the rat oxytocinergic system during pregnancy and parturition. *J Physiol (Lond)* 1997;501:649—655.

94. Kjaer A, Knigge U, Jorgensen H, et al. Dehydration-induced renin secretion: involvement of histaminergic neurons. *Neuroendocrinology* 1998;67:325–329.

95. Fekete CS, Strutton PH, Cagampang FRA, et al. Estrogen receptor immunoreactivity is present in the majority of central histaminergic neurons: evidence for a new neuroendocrine pathway associated with luteinizing hormone—releasing hormone-synthesizing neurons in rats and humans. *Endocrinology* 1999;140:4335–4341.

96. Sakata T, Yoshimatsu H, Kurokawa M. Hypothalamic neuronal histamine: implications of its homeostatic control of energy metabolism. *Nutrition* 1997;13:403–411.

97. Toyota H, Ito C, Ohsawa M, et al. Decreased central histamine in the amygdaloid kindling rats. *Brain Res* 1998;802:241–246.

98. Lintunen M, Sallmen T, Karlstedt K, et al. Postnatal expression of H_1-receptor mRNA in the rat brain: correlation to L-histidine decarboxylase expression and local upregulation in limbic seizures. *Eur J Neurosci* 1998;10:2287–2301.

99. Malmberg-Aiello P, Lamberti C, Ipponi A, et al. Effects of two histamine-N-methyltransferase inhibitors, SKF 91488 and BW 301 U, in rodent antinociception. *Naunyn Schmiedebergs Arch Pharmacol* 1997;355:354–360.

100. Rouleau, A, Garbarg, M, Ligneau, X, et al. Bioavailability and antiinflammatory properties of BP 2-94, a histamine H_3 receptor agonist prodrug. *J Pharmacol Exp Ther* 1997;281:1085–1094.

101. Clapham J, Kilpatrick GJ. Thioperamide, the selective histamine H_3 receptor antagonist, attenuates stimulant-induced locomotor activity in the mouse. *Eur J Pharmacol* 1994;259:107–114.

102. Prell GD, Green JC, Kaufmann CA, et al. Histamine metabolites in cerebrospinal fluid of patients with chronic schizophrenia: their relationship to levels of other aminergic transmitters and ratings of symptoms. *Schizophr Res* 1995;14:93–104.

103. Ito C, Morisset S, Krebs MO, et al. Histamine H_2 receptor gene variants: lack of association with schizophrenia. *Mol Psychiatry* 2000;5:159–164.

104. Panula P, Rinne J, Kuokkanen K, et al. Neuronal histamine deficit in Alzheimer's disease. *Neuroscience* 1998;82:993–997.

105. Morisset S, Traiffort E, Schwartz JC. Inhibition of histamine versus acetylcholine metabolism as a mechanism of tacrine activity. *Eur J Pharmacol* 1996;315:R1–R2.

106. Yanai K, Son LZ, Endou M, et al. Behavioural characterization and amounts of brain monoamines and their metabolites in mice lacking histamine H_1 receptors. *Neuroscience* 1998;87:479–487.

107. Lamberti C, Ipponi A, Bartolini A, et al. Antidepressant-like effects of endogenous histamine and of two histamine H_1 receptor agonists in the mouse forced swim test. *Br J Pharmacol* 1998;123:1331–1336.

15
Purinergic Neurotransmission

Michael Williams

Michael Williams: Department of Molecular Pharmacology and Biological Chemistry, Northwestern University School of Medicine, Chicago, Illinois.

The purine nucleoside, adenosine, its nucleotides, adenosine triphosphate (ATP), adenosine diphosphate (ADP), and adenosine monophosphate (AMP), and the pyrimidine nucleotide, uridine triphosphate (UTP) (Fig. 15.1), play a critical role in central and peripheral nervous system homeostasis and function as extracellular messengers to regulate cell function. The effects of adenosine and the nucleotides are mediated by activation of distinct P1 (adenosine) and P2 (ATP) cell-surface receptors present on neurons, astrocytes, and microglia, as well as other cells that are present in the central nervous system (CNS) under different conditions (e.g., macrophage infiltration). These receptors are generically known as *purinergic receptors* (1).

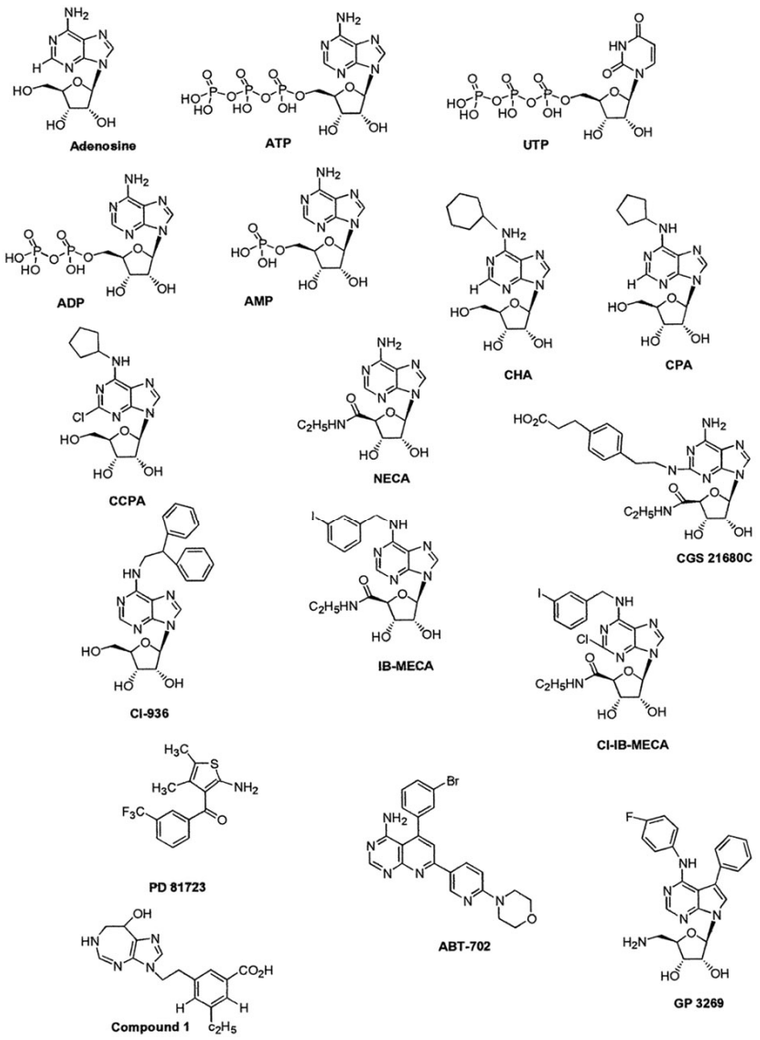


FIGURE 15.1. Structures of P1 and P2 agonists and modulators of adenosine availability.

Adenosine, ADP, and ATP, and, to a lesser extent, UTP are well-known intracellular constituents, intimately involved in all aspects of cell function acting as enzyme cofactors, sources of energy, and building blocks for DNA. Thus, the factors regulating their availability in the extracellular space as chemical messengers have been an area of active research and considerable debate since the late 1970s (2).

ATP can be released as a cotransmitter together with acetylcholine, norepinephrine, glutamate, γ-aminobutyric acid (GABA), calcitonin gene–related peptide, vasoactive intestinal peptide, and neuropeptide Y (3). ATP is available on demand, and the body can synthesize its own weight in ATP per day (4). Even though extracellular ATP levels can reach millimolar concentrations in the extracellular local environment after release or cellular perturbation (1), these concentrations are miniscule compared with the overall steady-state nucleotide content of the cell. Once released, in addition to interacting directly with P2 receptors, ATP can be hydrolyzed by a family of approximately 11 ectonucleotidases that metabolize ATP, ADP, diadenosine polyphosphates such as Ap4A, Ap5A (Fig. 15.2), and nicotinamide-adenine dinucleotide (5). Ecto-ATPases hydrolyze ATP to ADP, ectoapyrases convert both ATP and ADP to AMP, and ecto-5′-nucleotidase converts AMP to adenosine. The activities of ectoapyrase and ecto-5′-nucleotidase can change with cellular dynamics (6), and in guinea pig vas deferens, soluble nucleotidases are released together with ATP and norepinephrine (7), a finding representing a potential mechanism to limit the actions of extracellular ATP by enhancing its inactivation. The metabolic pathways linking ATP, ADP, AMP, and adenosine and the potential for each of these purines to elicit distinct receptor-mediated effects on cell function form the basis of a complex, physiologically relevant, purinergic cascade comparable to those involved blood clotting and complement activation (8) (Fig. 15.3).

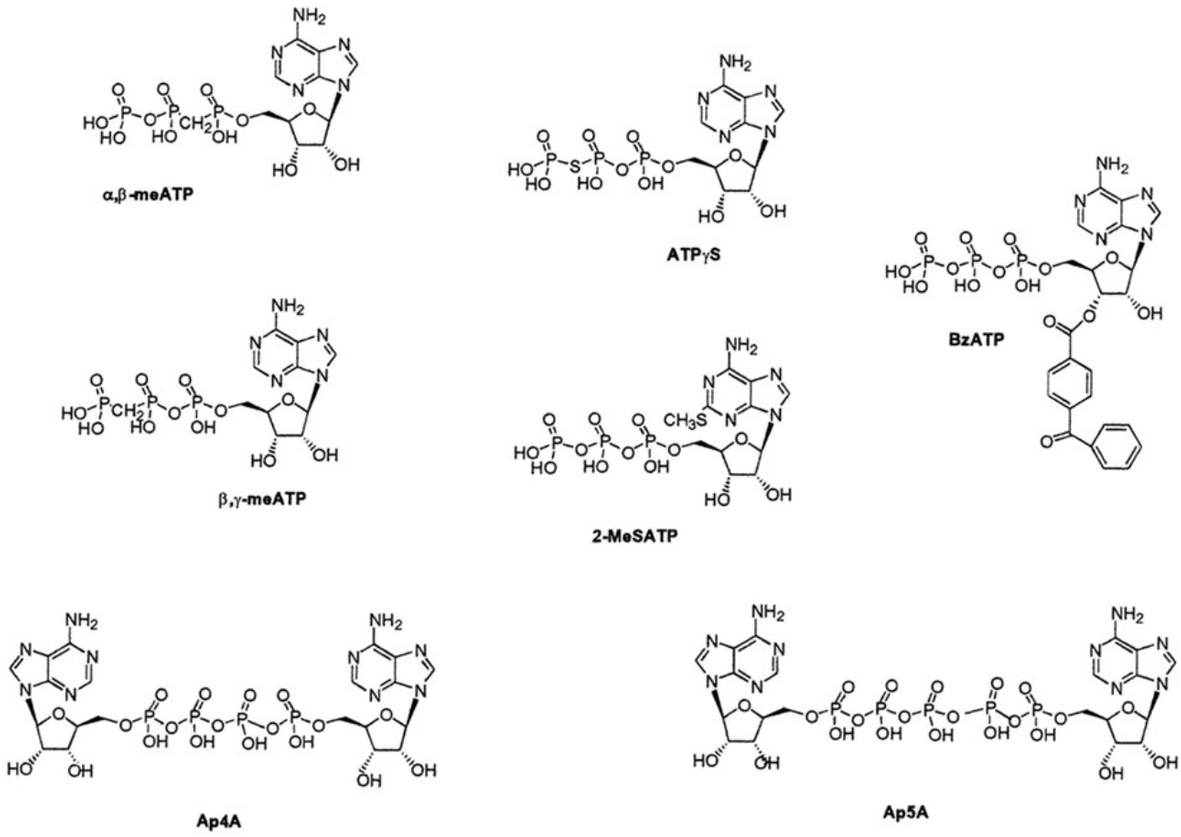


FIGURE 15.2. Structures of P2 receptor agonists.

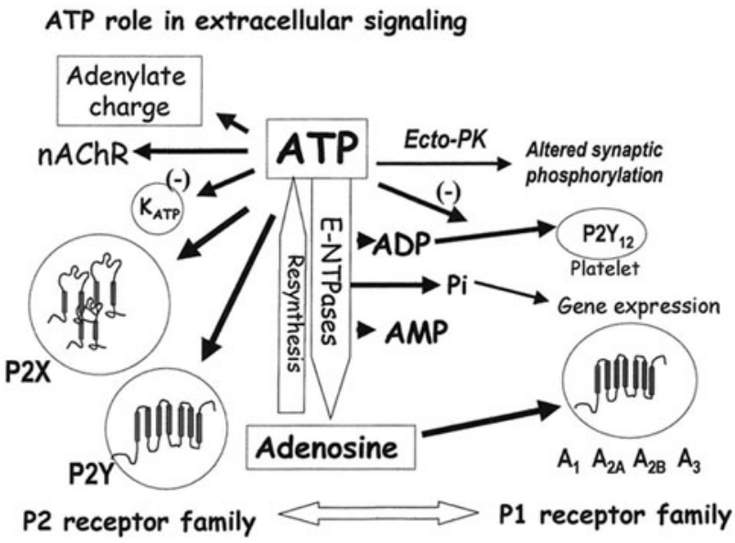


FIGURE 15.3. The purinergic cascade. ATP is released into the extracellular milieu from nerves or cells, where they can interact to form a purinergic cascade. ATP acts at a variety of P2 receptors (see text) and is sequentially degraded to ADP and AMP by ectonucleotidase activity. ADP interacts with P_{2T} receptors. AMP gives rise to adenosine, which can interact with the various P1 receptors (A_1, A_{2A}, A_{2B}, A_3). Adenosine can also be formed by intracellular 5′-nucleotidase activity. Adenylate charge indicates the transfer of energy in the form of adenine nucleosides or nucleotides from one cell to another (see ref. 13). KATP is an ATP-modulated potassium channel.

The extracellular effects of ATP on the various members of the P2 receptor family are terminated either by receptor desensitization or by dephosphorylation of the nucleotide, leading to the formation of ADP, AMP, and adenosine. These latter compounds have their own receptor-mediated functional activities, some of which are antagonistic to one another. For instance, ATP antagonizes ADP actions on platelet aggregation, whereas adenosine-elicited CNS sedation contrasts with the excitatory actions of ATP on nerve cells (9). In the broader framework of ATP-modulated proteins (or ATP-binding cassette proteins), ATP-sensitive potassium channels (K_{ATP}) undergo activation when intracellular ATP levels are reduced (10 ,11). Thus, as P2 receptor–mediated responses decrease because of ATP hydrolysis to adenosine, P1-mediated responses and K_{ATP}-mediated responses are enhanced. In addition to activating the as yet uncloned platelet $P2_T$ receptor, ADP also enhances its own availability. Activation of A_1 and A_{2A} receptors can inhibit ATP availability (1), and activation of hippocampal A_{2A} and A_3 receptors can desensitize A_1 receptor–mediated inhibition of excitatory neurotransmission (12). The transfer of purines transfer from one cell to another in the context of cellular adenylate charge (13) reflects another means by which purines can modulate cellular communication, in terms of both information transfer and alteration of the target environment. ATP also functions as a substrate for synaptic ectokinases, which modulate the phosphorylation state of the synaptic membrane (14) and, consequently, the intrinsic properties of the synapse. Once in the extracellular space, ATP thus has the ability to function as a pluripotent modulator of synaptic function.

The corresponding role of UTP in terms of functional synaptic signaling is less well understood (15), and although high concentrations of exogenous uracil have been shown to modulate dopaminergic systems in the CNS (16), data on the existence of a "uridine receptor" equivalent to the P1 receptor are limited (17).

Extracellular adenosine levels at rest are in the range of 30 to 300 nM (18), and they subserve a physiologic role in tissue homeostasis as reflected by the CNS stimulant actions of caffeine, a natural methylxanthine that acts as an antagonist to counteract the sedative actions of endogenous adenosine, and the role of the nucleoside as an endogenous hypnotic (19). Adenosine also acts as an autocrine homeostatic agent or, as conceptualized by Newby (20), a "retaliatory metabolite," to regulate the tissue energy supply-and-demand balance in response to changes in blood flow and energy availability and thus to conserve tissue function under adverse conditions. Reduced oxygen or glucose availability resulting from tissue trauma, such as during stroke, epileptogenic activity, and reduced cerebral blood flow, leads to ATP breakdown and the formation of ADP, AMP, and adenosine.

Under basal conditions, extracellular levels of adenosine are tightly regulated by ongoing metabolic activity. Bidirectional nucleoside transporters and the enzymes adenosine deaminase (ADA) and adenosine kinase (AK) regulate adenosine removal from the extracellular space (21). Numerous studies have shown that inhibition of AK is physiologically more relevant in increasing extracellular adenosine levels than inhibition of ADA or adenosine transport. AK inhibitors are also more effective in enhancing the neuroprotective actions of endogenous adenosine than inhibitors of ADA or adenosine transport (8 ,21). Selective AK inhibitors, such as GP 3269 and ABT-702 (Fig. 15.1), and ADA inhibitors, such as compound 1 (Fig. 15.1), are effective site- and event-specific agents that can locally enhance levels of adenosine in areas of tissue trauma and thus have the potential to avoid the potential cardiovascular side effects associated with a general elevation of extracellular levels of the purine (8). However, data have shown that *in vivo* administration of AK inhibitors (22), even at single doses close to where these agents show efficacy in animal models of epilepsy and pain, results in brain microhemorrhaging that can lead to minicerebral infarcts and cognitive impairment. Based on this finding, the AK approach to selective modulation of endogenous adenosine function does not appear to have a sufficient therapeutic window in CNS tissues to be a viable drug discovery target.

Adenosine has both presynaptic and postsynaptic effects on neurotransmission processes (12), whereas ATP has excitatory actions in a variety of neuronal systems including rat trigeminal nucleus, nucleus tractus solitarius, dorsal horn, and locus ceruleus. The nucleotide also functions as a fast transmitter in guinea pig celiac ganglion and rat medial habenula (9). Electrophysiologic studies on P2X and neuronal nicotinic receptor (nAChR)–mediated responses suggested that these two ligand-gated ion channels (LGICs) interacted with one another, with each receptor containing a inhibitory binding site for agonists active at the corresponding receptor and resulting in functional cross-talk between the systems (23). Recombination of nAChR α-subunits with P2X receptor subunits to form functional receptor constructs has also been reported (24), a finding further suggesting that heterooligimerization between these two different classes of LGICs may occurs and represents a molecular basis for the cross-talk hypothesis. This finding also adds a layer of further complexity to an already complex situation in understanding the precise subunit composition of functional P2X receptors. Dimerization of the A_1 subtype of the P1 (adenosine) G-protein–coupled receptor (GPCR) also occurs (25), a finding consistent with the emerging view that functional homooligimerization and heterooligimerization of a variety of GPCRs is the norm rather than the exception (26).

P1 AND P2 RECEPTORS

Part of "15 - Purinergic Neurotransmission "

Four distinct P1 receptors sensitive to adenosine and 12 P2 receptors sensitive to ADP, ATP, and UTP have been cloned and characterized (1), thus providing a diversity of discrete cellular targets through which adenosine, ADP, ATP, and UTP can modulate tissue function (Table 15.1). The four adenosine-sensitive P1 receptors are designated A_1, A_{2A}, A_{2B}, and A_3. Functional P2 receptors are divided into ionotropic P2X receptors, a family of eight LGICs ($P2X_1$ to $P2X_8$), and the metabotropic P2Y family, which consists of the $P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$, $P2Y_{11}$, $P2Y_{12}$, and $P2Y_{13}$ GPCRs. The missing numbers in the P2Y family sequence are proposed receptors that have been subsequently found to lack functional responses, are species variants, or have been inadvertently assigned to the P2 receptor family (1 ,8).

	Agonist Rank Order Potency	Antagonist Rank Order Potency
$P2X_1$	2-MeSATP ≥ ATP ≥ α,β-meATP > BzATP	Ip_5I >> suramin, PPADS > MRS 2220
$P2X_2$	2-MeSATP > ATP >> α,β-meATP	Suramin, PPADS ≥ TNP-ATP
$P2X_3$	2-MeSATP > ATP > BzATP	TNP-ATP >> suramin, PPADS
$P2X_4$	ATP > 2-MeSATP > α,β-meATP	TNP-ATP >> suramin
$P2X_5$	ATP > 2-MeSATP >> α,β-meATP	Suramin, PPADS
$P2X_6$	ATP > 2-MeSATP >> α,β-meATP	Suramin
$P2X_7$	BzATP > ATP > UTP >> α,β-meATP	KN-62 >> suramin, PPADS
$P2X_8$	ATP = 2meSATP> α,β-meATP > ATPγS	Suramin, PPADS

[a]Functional heteromers composed of $P2X_{1/5}$, $P2X_{2/3}$, $P2X_{2/6}$, and $P2X_{2/6}$ subunits have been described.

TABLE 15.1. P2X RECEPTORS

Adenosine (P1) Receptors

All four P1 GPCRs—A_1, A_{2A}, A_{2B}, and A_3 (Table 15.1)—are heterogeneously distributed in a variety of mammalian tissues including heart, smooth muscle, kidney, testis, platelets, leukocytes, and adipocytes. in addition to the CNS. The A_1 receptor is widely distributed in the CNS and is functionally coupled to inhibition of cyclic AMP (cAMP) formation, stimulation of potassium conductance, inhibition of N-channel–mediated calcium conductance, stimulation of phospholipase C production, and modulation of nitric oxide production (1 ,12). Selective agonists for the A_1 receptor are all adenosine analogues and include cyclohexyl (CHA; A_1 K_i = 1 to 5 nM), cyclopentyl (CPA; A_1 K_i = 0.6 nM), and 2-chlorocyclopentyl (CCPA; K_i = 0.6 nM) adenosine (27) (Fig. 15.1). Agonist effects at the

A_1 receptor are selectively blocked by the 8-substituted xanthines, cyclopentylxanthine (Fig. 15.3) (CPX; K_i = 0.46 nM) and CPT (K_i = 24 nM), and by nonxanthines such as N-0861 (K_i = 10 nM). The A_1 receptor shows distinct species pharmacology (8). Like other GPCRs, it can be allosterically modulated (28) by compounds such as PD 81,723 (Fig. 15.1) that, although not directly interacting with the agonist binding site of the receptor, stabilize an agonist-preferring conformation of the A_1 receptor independent of G-protein interactions (29).

The A_2 receptor exists in two distinct molecular and pharmacologically subtypes, both of which are linked to activation of adenylate cyclase (1). The A_{2A} receptor has high affinity for adenosine, may also use N- and P-type Ca^{2+} channels as signal transduction mechanisms, and is localized in the striatum, nucleus accumbens, and olfactory tubercle regions of mammalian brain. The lower-affinity A_{2B} receptor is more ubiquitously distributed throughout the CNS (1 ,30). The adenosine analogue, CGS 21680C (Fig. 15.1) (K_i A_{2A} = 15 nM), is the present agonist of choice for the A_{2A} receptor with the xanthine antagonists KF 17837 (Fig. 15.4) (K_i A_{2A} = 24 nM) and CSC 8-(3-chlorylstyryl) caffeine (K_i A_{2A} = 9 nM) and nonxanthines such as SCH 58261 (K_i A_{2A} = 2.3 nM) and ZM 241385 (K_i A_{2A} = 0.3 nM), being up to 6,800-fold selective for the A_{2A} receptor (27). Like the A_1 receptor, the A_{2A} receptor also shows species-dependent pharmacology (8). The A_{2B} receptor has been cloned and is widely distributed in brain and peripheral tissues (1 ,30). However, its functional characterization has proven difficult because of a paucity of selective ligands. Responses more potently elicited by the nonselective adenosine agonist, NECA (Fig. 15.1) and not by selective A_1-, A_{2A}-, or A_3-receptor agonists can be attributed to A_{2B}-receptor activation. Enprofylline (Fig. 15.3) is a selective, albeit weak, A_{2B}-receptor antagonist (30).

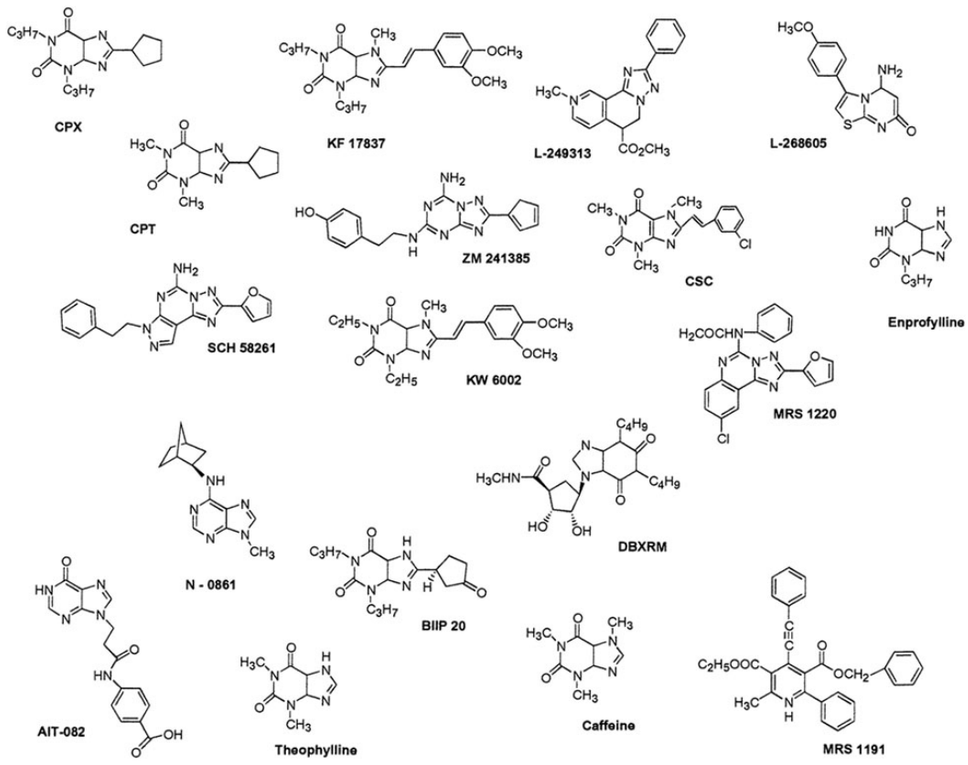


FIGURE 15.4. Structures of P2 receptor antagonists.

The A_3 receptor was the first P1 receptor identified by cloning rather than by pharmacologic characterization and is linked to inhibition of adenylate cyclase and elevation of cellular inositol 1,4,5-triphosphate (IP_3) levels and intracellular Ca^{2+}. It also shows distinct species-dependent pharmacology, especially in regard to xanthine antagonist blockade of the rat A_3 receptor (1 ,31 ,32), and it shows widespread distribution with low levels in brain. IB-MECA (Fig. 15.1) (K_i = 1nM) and its 2-chloro analogue (Cl^- IB-MECA; K_i = 0.3 - 0.7 nM) are potent and selective A_3-receptor agonists (27). The human, but not the rat, A_3 receptor is selectively blocked by xanthines (31), such as DBXRM (K_i A_3 = 229 nM), and by nonxanthines such as MRS 1191, MRS 1220, L-249313, L-268605, and MRE-1008-21M (Fig. 15.4). A_3 receptors are involved in mast cell function, eosinophil apoptosis, and the phenomenon known as preconditioning that occurs during ischemic reperfusion of the heart that protects against myocardial infarction (1 ,33).

P2 (ATP and UTP) Receptors

P2 receptors were originally classified on the basis of the rank-order potency of agonists structurally related to ATP. Most of these putative receptors (with the exception of the P_{2T} receptor) have been subsequently cloned and functionally characterized in various heterologous expression systems (1). However, their functional characterization in native tissues and in animals has been limited by a paucity of potent, selective, and bioavailable ligands, both agonists and antagonists. All known P2-agonist ligands are analogues of ATP, UTP, and ADP and, irrespective of their degree of chemical modification, show varying degrees of susceptibility to extracellular degradation and differences in intrinsic activity (27 ,34). The selectivity and potency of these agonists are thus very much dependent on the tissue preparation and species used and also on the experimental protocol. Indeed, few studies exist in which a systematic evaluation of the relative selectivity of P2-receptor agonists has been determined. BzATP (Fig. 15.2), which is widely used as a selective agonist for the $P2X_7$ receptor (EC_{50} = 18 μM), is, however, far more potent at transfected rat and human $P2X_1$ (EC_{50} = 1.9 nM) and $P2X_3$ (EC_{50} = 98 nM) receptors (35) and thus cannot be used, *a priori*, in defining a $P2X_7$-receptor–mediated response.

P2 receptors are present on excitable tissues, such as neurons,

glia, and smooth muscle cells (1), and can be grouped into three classes based on agonist effects (36). Group 1, comprising the $P2X_1$ and $P2X_3$ receptors, has high ATP affinity for ATP (EC_{50} = 1 μM) and is rapidly activated and desensitized. Group 2 includes the $P2X_2$, $P2X_4$, $P2X_5$, and $P2X_6$ receptors that have lower ATP affinity (EC_{50} = 10 μM), have a slow desensitization profile, and exhibit sustained depolarizing currents. The only receptor in Group 3 is the $P2X_7$ LGIC, which has low ATP affinity (EC_{50} = 300 - 400 μM) and shows little or no desensitization on agonist exposure.

P2X Receptors

P2X receptors are ATP-gated LGICs formed from various P2X subunits that share a common motif of two transmembrane-spanning regions (2TM). Like the amiloride-sensitive epithelial Na^+ channel, P2X receptor subunits have a large extracellular domain with both the N- and C-termini located intracellularly (1 ,37). A functional P2-receptor channel consists of multimeric combinations of the various P2X subunits to form a nonselective pore permeable to Ca^{2+}, K^+, and Na^+ that mediates rapid (approximately 10-millisecond) neurotransmission events. Available evidence indicates that functional P2X receptors are trimeric, in contrast to the typical pentameric structure of other LGICs (38). In addition to putative $P2X_1$ to $P2X_7$ homomers, $P2X_{1/5}$, $P2X_{2/3}$, and $P2X_{4/6}$ functional heteromers have been identified (1 ,39). $P2X_5$ and $P2X_6$ receptors do not appear to exist in homomeric form, but rather as heteromers with other P2X receptor subtypes. Unlike other LGICs, such as nAChRs, the 5-hydroxytryptamine ($5\text{-}HT_3$) receptor, little is known regarding the agonist (ATP) binding site on P2X receptor constructs or of ancillary sites that may modulate receptor function.

The utility of current P2-receptor antagonists, such as PPADS, DIDS, reactive blue-2, and suramin (Fig. 15.5) (27), is limited by their lack of selectivity for different P2

receptors and other proteins (34) or by their limited potency and bioavailability. These compounds can also inhibit the ectonucleotidases responsible for ATP breakdown, thus confounding receptor characterization (40). Radioligand-binding assays for P2 receptors are also far from robust; available ligands binding to cell lines lack any type of P2 receptor (41). The use of high throughput screening techniques to identify novel ligands thus depends on functional fluorescence assays such as FLIPR (fluorescence imaging plate redder) rather than binding.

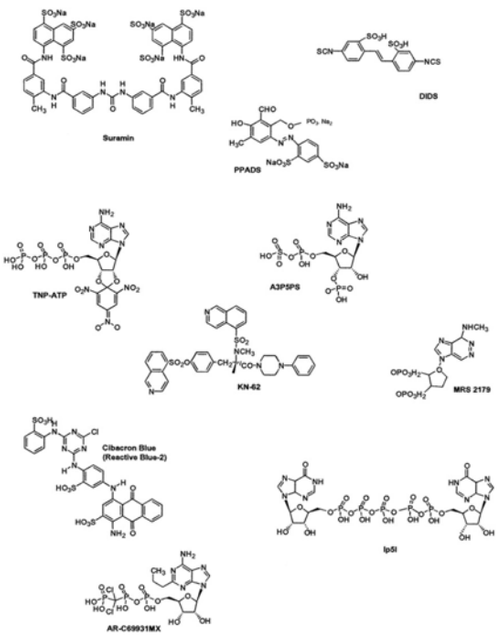


FIGURE 15.5. Structures of P1 receptor antagonists.

Among the newer P2X-receptor antagonists (Fig. 15.5) are the following: TNP-ATP, a noncompetitive, reversible allosteric antagonist at $P2X_1$ and $P2X_3$ receptors with nanomolar affinity (42) that also has weak activity at $P2X_4$ and $P2X_7$ receptors; Ip5I, a potent, selective $P2X_1$ antagonist (K_i < 100 nM) antagonist (43); KN-62, a potent (IC_{50} = 9 to 13 nM), noncompetitive antagonist of the human $P2X_7$ receptor that is inactive at the rat $P2X_7$ receptor (44). The ATP analogue, A3P5PS (Fig. 15.5), is a partial agonist–competitive antagonist at the turkey erythrocyte $P2Y_1$ receptor (27), with the derivative, MRS 2179 being a full $P2Y_1$-receptor antagonist (IC_{50} =330 nM). AR-C 69931-MX (Fig. 15.5) is a potent, selective antagonist at the ADP-sensitive $P_{2T}/P2Y_{AC}$ receptor involved in platelet aggregation that is currently in clinical trials as a novel antithrombotic agent (45).

$P2X_1$ receptors are activated by 2MeSATP, ATP, and αß-meATP (Fig. 15.2), and they exhibit rapid desensitization kinetics (Table 15.2) (1). $P2X_1$ subunits are present in the dorsal root, in trigeminal and celiac ganglia, and in spinal cord and brain. The $P2X_2$ receptor is activated by 2MeSATP and ATPγS but is insensitive to αß-meATP and ßγ-meATP (Fig. 15.2). It is present in brain, spinal cord, superior cervical ganglia, and adrenal medulla. $P2X_{2\text{-}1}$, $P2X_{2\text{-}2}$, $P2X_{2\text{-}3R}$, and $P2X_{2\text{-}3}$ receptors are splice variants of the $P2X_2$ receptor that have been localized, among other places, to the cochlear endothelium, an area in the ear associated with sound transduction (46). $P2X_1$ and $P2X_2$ receptors can be blocked by PPADS and suramin (1). The $P2X_3$ receptor has a rank order of activation in which 2MeSATP >> ATP > αß-meATP and is localized to a subset of sensory neurons that includes the dorsal root, trigeminal, and nodose ganglia (1). It has similar properties to the $P2X_1$ subtype including αß-meATP sensitivity and rapid desensitization kinetics. $P2X_2$ and $P2X_3$ subunits can form a functional heteromeric $P2X_{2/3}$ receptor *in vitro* (39) that combines the pharmacologic properties of $P2X_3$ (αß-meATP sensitivity) with the kinetic properties of $P2X_2$ (slow desensitization). $P2X_4$ receptors are activated by 2MeSATP and are only weakly activated by αß-meATP. The rat and human homologues of the $P2X_4$ receptor differ in their sensitivity to suramin and PPADS; the human $P2X_4$ receptor is weakly sensitive, and the rat $P2X_4$ receptor is insensitive to these P2X-receptor antagonists (1). The $P2X_4$ receptor is present in rat hippocampus, superior cervical ganglion, spinal cord, bronchial epithelium, adrenal gland, and testis, as well as human brain. The agonist profile for the $P2X_5$ receptor is ATP > 2MeSATP > ADP with αß-meATP being inactive. This receptor does not exhibit rapid desensitization kinetics but is blocked by suramin and PPADS. Message for the $P2X_5$ receptor is present in the central horn of the cervical spinal cord, in trigeminal and dorsal root ganglia neurons, and in the brain in the mesencephalic nucleus of the trigeminal nerve. The $P2X_6$ receptor is present in the superior cervical ganglion, cerebellar Purkinje cells, spinal motoneurons of lamina IX of the spinal cord, superficial dorsal horn neurons of lamina II, and trigeminal, dorsal root, and celiac ganglia (1). $P2X_4$ and $P2X_6$ subunits form functional heteromers *in vitro* (39).

TABLE 15.2. CLASSIFICATION OF P2Y RECEPTORS

	Agonist Rank Order Potency	Antagonist Rank Order Potency
$P2Y_1$	2-MeSADP > 2-MeSATP > HT-AMP > ADP > ADPβS > ATP > α,β-meATP > UTP inactive	MRS2179 > isoPPADS > A3′P5′P ≥ PPADS suramin
$P2Y_2$	ATP = UTP (100) > ATPγS = Ap_4A	
$P2Y_4$	UTP ≥ UTPγS > ATP	PPADS > reactive blue 2 > suramin > ATP (human)
$P2Y_6$	UDP >> UTP ≥ 2-MeSADP	Suramin > PPADS
$P2Y_{11}$	ATP > ADP >>> UTP	
$P2Y_{12}$	ADP	AR-C 69931MX = CT5054 >>> ATP
$P2Y_{13}$	ADP	2Me5ADP >> ATP

The $P2X_7$ receptor, also known as the $P_{2Z}/P2Z$ receptor before it was cloned (47), is present in the superior cervical ganglion and spinal cord, mast cells, and macrophages (48). Cerebral artery occlusion results in an increase in $P2X_7$ immunoreactivity in the stroke-associated penumbral region (49). The $P2X_7$ receptor has a long (240 amino acid) intracellular C-terminal region that allows the receptor to form

a large nonselective cytolytic pore on prolonged or repeated agonist stimulation (1 ,48). Exposure of the $P2X_7$ receptor to ATP for brief periods (1 to 2 seconds) results in transient pore opening that mediates cell-to-cell communication. Prolonged receptor activation triggers cytolytic pore formation with the initiation of an apoptotic cascade involving caspase-1 (interleukin 1ß convertase) and an associated release of IL-1ß from macrophages (48). The $P2X_7$ receptor is partially activated by saturating concentrations of ATP and is fully activated by the ATP analogue BzATP (Fig. 15.2). The ability to form a cytolytic pore was considered unique for the $P2X_7$ receptor, but other P2X receptors such as P2X and P2X show the same phenomenon on prolonged exposure to ATP, a finding indicating that the intracellular C-terminal tail is not a prerequisite for cytolytic pore formation (50 ,51).

P2Y Receptors

P2Y receptors are GPCRs activated by purine or pyrimidine nucleotides (1 ,52). The seven mammalian functional subtypes, $P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$, $P2Y_{11}$, $P2Y_{12}$, and $P2Y_{13}$, have been cloned and are coupled to Gq_{11}. Receptor activation results in stimulation of phospholipase C and IP_3 activation and subsequent release of calcium from intracellular stores. The P_{2T} receptor, present in platelets and preferentially sensitive to ADP, has been cloned, as the $P2Y_{12}$ receptor.

The $P2Y_1$ receptor is preferentially activated by adenine nucleotides, with 2MeSATP the most potent. UTP and UDP are inactive at this receptor. Suramin, PPADS, cibacron blue, A3P5PS, and MRS 2179 (Fig. 15.5) are competitive antagonists at this receptor (1 ,27). The $P2Y_2$ receptor is activated by both ATP and UTP; nucleotide diphosphates are inactive (1 ,52). Antagonists such as suramin are less efficacious at the $P2Y_2$ receptor. UTP is the preferred agonist for the $P2Y_4$ receptor, with ATP and the nucleotide diphosphates inactive. Diphosphates are more active at the $P2Y_6$ receptor than triphosphates, and this has led to the classification of the $P2Y_6$ receptor as a UDP-preferring receptor. The $P2Y_{11}$ receptor is unique among other P2Y receptors in that only ATP serves as an agonist for this receptor (53). The $P2Y_{12}$ and $P2Y_{13}$ receptors are ADP-selective receptors.

Diadenosine polyphosphates including Ap4A and Ap5A (Fig. 15.2) comprise another group of purine-signaling molecules that modulate cell function by activation of cell-surface P2 receptors (1 ,54). Whereas an Ap4A receptor that modulates neurotransmitter release and amphetamine-elicited Ap4A release has been pharmacologically characterized in nervous tissue, it is unclear whether diadenosine polyphosphate actions involve distinct receptor subtypes or reflect activation of known P2 receptors. Receptors for the diadenosine polyphosphates have not yet been cloned.

MITOCHONDRIAL PURINE RECEPTORS?

Part of "15 - Purinergic Neurotransmission "

In addition to functioning as the key source of ATP within the cell, mitochondria play a key role in the apoptotic cascade as the source of the cytochrome C that is released after changes in mitochondrial transition pore function elicited by members of the bcl-2 family of cell death proteins (55). The ability of the $P2X_7$ receptor to initiate an apoptotic cascade by activation of caspase-1 (48) and the key role of mitochondria in various degenerative diseases (56 ,57) raise the question whether intracellular P2 receptors are present on the outer mitochondrial membrane and may provide a direct mechanism for ATP to influence mitochondrial function.

THERAPEUTIC POTENTIAL OF PURINES IN NERVOUS TISSUE

Part of "15 - Purinergic Neurotransmission "

Adenosine potently inhibits the release of the neurotransmitters dopamine, GABA, glutamate, acetylcholine, serotonin, and norepinephrine and acts through presynaptic A_1 receptors (12). Adenosine acts preferentially on excitatory versus inhibitory neurotransmitter release, a finding suggesting a degree of physiologic specificity in modulating brain function. Adenosine also directly modulates postsynaptic neuronal excitability by activating A_1 and A_{2A} receptors resulting in hyperpolarization of the postsynaptic membrane.

Over the past 2 decades, many studies have provided evidence of involvement of purines in the actions of various CNS-active drugs including antipsychotics, antidepressants, anxiolytics, and cognition enhancers. These studies have come from experiments in which the effects of known CNS drugs representative of these therapeutic classes were examined for their ability to modulate adenosine-mediated responses in the CNS, or, alternatively, they studied the effects of various P1 ligands, both agonists or antagonists, on the effects of such prototypic CNS agents. In many instances, only single, somewhat high, concentrations of an isolated compound, or limited numbers of compounds, were used to generalize to a complete class of psychotherapeutic agents, often with no negative control data, thus limiting the value of the data (58).

For P2 receptors, the absence of ligands, agonists, and antagonists has limited the functional characterization of the various receptor subtypes. The delineation of a role for P2 receptors in CNS disorders has been postulated largely on the basis of *in situ* localization of the mRNAs encoding the different P2 receptor subtypes or of immunohistochemical studies.

Transgenic Models of P1 and P2 Receptor Function

For both P1 and P2 receptors, the use of mice either deficient in, or overexpressing, a targeted receptor can potentially

provide a unique means to assess the role of the given receptor, the phenotype of which will provide information on the role of the receptor. Although this approach is not always straightforward because some phenotypes are fatal and, in others, a knockout of one receptor leads to compensation in associated receptor systems such that the resultant mouse phenotype is atypical, knockouts can be helpful in the absence of selective antagonists or antisense probes.

P1 receptor knockouts show altered cardiovascular function (A_1) and reduced exploratory activity, aggressiveness, hypoalgesia, and high blood pressure (A_{2A}) (59). P2 knockouts are associated with decreased male fertility ($P2X_1$) (60), decreased nociception and bladder hyporeflexia ($P2X_3$) (61), decreased platelet aggregation and bleeding time ($P2Y_1$) (62), and reduced chloride secretion ($P2Y_2$) (63). A preliminary report on a $P2X_7$ knockout has appeared (64).

PURINERGIC THERAPEUTICS

Part of "15 - Purinergic Neurotransmission "

Three distinct classes of compound can modulate P1 and P2 receptor function: (a) conventional agonist, partial agonist and antagonist ligands; (b) allosteric modulators of receptor function; and (c) modulators of the endogenous systems that regulate the extracellular availability of ATP, adenosine, UTP, and their respective nucleotides. This last group includes the various ecto-ATPases that catalyze the degradation of nucleotides (5), ADA, AK, and the bidirectional member transporter systems that remove adenosine from the extracellular environment (21 ,65). From data on AK effects in brain tissue (22), it appears that modulation of endogenous adenosine levels by inhibition of AK is not a viable drug discovery approach.

Efforts over the last 25 years to develop directly acting P1-receptor agonists and antagonists as therapeutic agents (8) have proven less than successful because of a combination of the choice of disease states in which other therapeutic modalities are clearly superior (58) and side effects are associated with global receptor modulation. Partial agonists, allosteric modulators, and novel modulators of ATP metabolism may prove clinically useful agents with improved therapeutic indices (65).

Stroke and Ischemia

Extracellular adenosine levels are markedly increased after hypoxia and focal ischemia, a finding providing additional evidence that the purine acts as a homeostatic neuroprotective agent (8). Adenosine-receptor agonists such as CHA reduce stroke-related cell death and hippocampal neurodegeneration, whereas adenosine antagonists exacerbate ischemic brain damage by enhancing glutamate release. The neuroprotective effects of adenosine are mediated by several P1 receptors: A_1-receptor activation stabilizes neuronal membrane potential, inhibits neuronal excitability and glutamate release (8 ,12), and thus prevents initiation of the stroke cascade (66).

Adenosine also hyperpolarizes astrocyte membranes limiting extracellular glutamate and potassium accumulation and modulates local cerebral blood flow and local inflammatory responses, such as platelet aggregation, neutrophil recruitment, and adhesion acting through the A_{2A} receptor (67). A_3-receptor agonists have biphasic effects on cell survival. At nanomolar concentrations, they are neuroprotective and inhibit apoptosis, but at micromolar concentrations they are neurotoxic (31).

mRNA for the $P2X_7$ receptor is up-regulated on microglial cells in the ischemic penumbral region 24 hours after middle cerebral artery occlusion in the rat (49), a finding indicating that cytolytic pore formation and inflammatory cytokine release are associated with neural trauma and neurodegeneration. Antisense to the $P2X_7$ receptor or selective receptor antagonists may represent a novel approach to the treatment of stroke.

Epilepsy

Seizure activity is associated with rapid and marked increases in CNS adenosine concentrations in animals (68), as well as in patients with epilepsy with spontaneous-onset seizures (69). Seizure activity induced by a variety of chemical and electrical stimuli in animal models is reduced by adenosine and related agonists (68) acting through A_1receptors. In electrically kindled seizure models, adenosine agonists reduce seizure severity and duration without significantly altering seizure threshold. These anticonvulsant effects are blocked by doses of methylxanthines that, when given alone, have no observable effect on seizure activity (68), a finding leading to the hypothesis that adenosine functions as an endogenous anticonvulsant.

Neurodegeneration: Alzheimer's Disease and Parkinson's Disease

The nerve cell death that follows excessive glutamate release and changes in calcium homeostasis after ischemia and hypoxia may reflect an acute manifestation of more subtle, long-term changes associated with apoptotic and necrotic cell death in Alzheimer's disease (AD) and Parkinson's disease (PD). Adenosine antagonists including caffeine, theophylline, and BIIP 20 (Fig. 15.4) are potent CNS stimulants (8 ,18), and they can enhance cognition in animal models by blocking the actions of endogenous adenosine. Certain compounds acting by purinergic mechanisms, such as BIIP 20 and propentofylline, have been examined in the clinic for their efficacy in cognitive disorders. Although provocative clinical data have been generated, neither compound showed sufficiently robust efficacy in larger AD trials

to warrant continuation. However, aged patients with rheumatoid arthritis who consume large quantities of antiinflammatory agents such as indomethacin show an inverse correlation for the incidence of AD, a finding highlighting the pivotal role of inflammation in disease origin. Adenosine agonists and AK inhibitors have marked antiinflammatory activity (67), inhibiting free radical production, and thus they may be effective in maintaining cell function in AD, in addition to modulating cytotoxic events.

Trophic factors in nervous tissue act to ensure neuronal viability and regeneration. Withdrawal of nerve growth factor, which exerts a tonic cell death–suppressing signal, leads to neuronal death. Polypeptide growth factors linked to receptor tyrosine kinases, such as fibroblast growth factors, epidermal growth factor, and platelet-derived growth factor, are increased with neural injury (70). ATP can act in combination with various growth factors to stimulate astrocyte proliferation and to contribute to the process of reactive astrogliosis, a hypertrophic-hyperplastic response typically associated with brain trauma, stroke and ischemia, seizures, and various neurodegenerative disorders. In reactive astrogliosis, astrocytes undergo process elongation and express glial fibrillary acidic protein, an astrocyte-specific intermediate filament protein with an increase in astroglial cellular proliferation. ATP increases glial fibrillary acidic protein and activator protein-1 (AP-1) complex formation in astrocytes and mimics the effects of basic fibroblast growth factor (70). Both ATP and guanosine triphosphate induce trophic factor (nerve growth factor, neurotrophin-3, fibroblast growth factor) synthesis in astrocytes and neurons. The effects of guanosine triphosphate are, however, not consistent with any known P2-receptor profile. Nonetheless, these studies have focused research on the hypoxanthine analogue, neotrofin (AIT-082) (Fig. 15.1), which up-regulates neurotrophin production and enhances working memory and restores age-induced memory deficits in mice (71). This compound has shown positive effects in early phase II trials for AD.

In 1974, Fuxe showed that methylxanthines such as caffeine could stimulate rotational behavior and could potentiate the effects of dopamine agonists in rats with unilateral striatal lesions Conversely, adenosine agonists blocked the behavioral effects of dopamine (72). Anatomic links between central dopamine and adenosine systems are well established; adenosine A_{2A} receptors are highly localized in striatum, nucleus accumbens, and olfactory tubercle, brain regions that also have high densities of dopamine D1 and D2 receptors. mRNAs for adenosine A_{2A} receptors and dopamine D2 receptors are co-localized in GABAergic-enkephalin striatopallidal neurons in the basal ganglia (Fig. 15.6) that form an "indirect" pathway from the striatum to the globus pallidus that originates from striatal GABA-enkephalinergic neurons. Through GABAergic relays, this pathway interacts with a glutaminergic pathway from the subthalamic nucleus that can activate the internal segment of the pars reticulata, which, turn, through a pars reticulata–thalamic GABAergic pathway, inhibits the thalamic-cortical glutaminergic pathway. Dysfunction of this pathway may underlie the movement disorders seen in Huntington chorea and PD. A direct pathway originating in striatal GABAergic–substance P–dynorphinergic neurons inhibits the internal segment of the pars reticulata to disinhibit the ascending thalamic glutaminergic pathway and to activate the cortex (Fig. 15.6). The balance between the direct (cortical activating) and indirect (cortical inhibiting) striatal dopaminergic pathways provides a tonic regulation of normal motor activity. These studies indicate that striatal adenosine A_{2A} receptors may play a pivotal role in neurologic disorders involving basal ganglia dysfunction such as PD. The A_{2A} agonist, CGS 21680, given intrastriatally, attenuates the rotational behavior produced by dopamine agonists in unilaterally lesioned rats. Mechanistically, radioligand-binding studies have shown an increased efficacy of CGS 21680 in reducing the binding affinity of supersensitive D2 receptors, a finding supporting the increased sensitivity of animals with supersensitive dopamine receptors to CGS 21680 treatment. Repeated administration of the dopamine antagonist, haloperidol can up-regulate the density of both D2 and A_{2A} receptors in rat striatum.

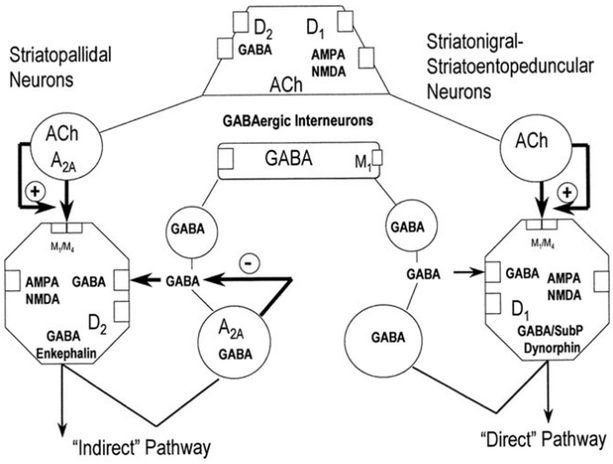


FIGURE 15.6. Dopamine–adenosine (ADO) interactions in the substantia nigra. An indirect pathway dopaminergic pathway arises from the striatal GABA-enkephalinergic dopaminergic neurons on which both dopamine D1 and adenosine A_{2A} receptors are co-localized. Through a GABAergic interneuron originating in the external globus pallidus, the indirect pathway connects to a glutaminergic pathway arising in the subthalamic nucleus. This, in turn, can activate the internal segment of the pars reticulata and, through another GABA pathway, inhibit ascending glutaminegic neurons arising from the thalamus that innervate the cortex. The direct pathway arises from striatal GABA–substance P–dynorphinergic neurons that, through a GABAergic relay, inhibit the internal segment of the pars reticulata to disinhibit the ascending thalamic-cortical glutaminergic pathway. The balance between the direct (activating) and indirect (inhibitory) striatal dopaminergic pathways can then tonically regulate normal motor activity. Dopaminergic inputs arising from the substantia nigra pars compacta can facilitate motor activity, inhibiting the indirect pathway by activation of D2 receptors and activating the direct pathway by D1 receptor activation. (Adapted from Svenningsson P, Le Moine C, Fisone G, et al. Distribution, biochemistry and function of striatal adenosine A_{2A} receptors. *Prog Neurobiol* 1999;59:355–396; and Richardson PJ, Kase H, Jenner PG. Adenosine A_{2A} receptor antagonists as new agents for the treatment of Parkinson's disease. *Trends Pharmacol Sci* 1997;18:338–344.)

Adenosine A_1 receptor activation can reduce the high-affinity state of striatal dopamine D1 receptors, the A_1 receptor agonist, and CPA blocking D1-receptor–mediated locomotor activation in reserpinized mice (72). The nonselective adenosine agonist, NECA, can attenuate the perioral dyskinesias induced by D1-receptor activation in rabbits. Acting through striatal A_{2A} and A_1 receptors, adenosine directly modulates dopamine-receptor–mediated effects on striatal GABA-enkephalinergic neurons and striatal GABA–substance P neurons (Fig. 15.6). These adenosine agonist–mediated effects are independent of G-protein coupling and may involve an intramembrane modulatory mechanism involving receptor heterooligimerization (26).

The dynamic interactions between dopaminergic and purinergic systems in striatum suggest that dopaminergic dysfunction may be indirectly ameliorated by adenosine receptor modulation. Selective adenosine A_{2A} receptor antagonists such as KF 17837 and KW 6002 (Fig. 15.4) have shown positive effects in 1-methyl-4-[henyl-1,2,3,6-tetrahydropyridine–lesioned marmosets and cynomolgus monkeys, well characterized animal models of PD, enhancing the effects of L-dopa (73 ,74). KW-6002 has successfully completed human phase I trials. More recently, a 30-year longitudinal study of 8,004 Japanese-American man enrolled in the Honolulu Heart Program showed an inverse association of the incidence of PD with caffeinated coffee consumption. In men who drank no coffee, the incidence of PD was 10.4 per 10,000 person-years, and it was 1.9 per 10,000 person-years in men drinking at least 28 oz of coffee per day (75).

Adenosine agonists can mimic the biochemical and behavioral

actions of dopamine antagonists in animal models by activation of A_{2A} receptors (9 ,72), a process that inhibits dopamine synthesis and attenuates dopamine transductional processes. CGS 21680, like typical and atypical neuroleptics, can reverse apomorphine-induced loss of prepulse inhibition (76). These actions involve a decrease in dopaminergic neurotransmission, with adenosine receptor agonists acting as functional dopamine antagonists. Adenosine agonists have a behavioral profile similar to that of dopamine antagonists in a conditioned avoidance response paradigm (77), in which they potently disrupting avoidance responding without significantly impairing escape behavior. They also produce catalepsy at the same dose levels effective in attenuating conditioned avoidance response, a property shared by typical neuroleptic agents such as haloperidol. CI-936, an A_{2A} agonist (Fig. 15.1), entered clinical trials in the mid 1970s as a novel antipsychotic agent, but its development was discontinued for unstated reasons.

Sleep

The hypnotic and sedative effects of adenosine are well known, as are the central stimulant activities of the various xanthine adenosine antagonists including caffeine (18). Direct adenosine administration into the brain elicits an EEG profile similar to that observed in deep sleep, an increase in rapid eye movement (REM) sleep with a reduction in REM sleep latency resulting in an increase in total sleep. In contrast, caffeine suppresses REM sleep and decreases total sleep time. Microdialysis studies have shown that extracellular adenosine concentrations are increased in basal forebrain in direct proportion to periods of sustained wakefulness and decline during sleep, a finding indicating that adenosine functions as a endogenous sleep regulator (19). Infusion of the A_{2A} agonist, CGS 21680, into the subarachnoid space associated with the ventral surface of the rostral basal forebrain, an area designated the prostaglandin D_2–sensitive sleep-promoting zone, increased slow-wave and paradoxical sleep, effects that were blocked by the A_{2A} antagonist, KF 17837 (78). The A_1-selective agonist, CHA, suppressed slow-wave and paradoxical sleep before eliciting an increase in low-wave sleep.

Pain

The role of purines in pain perception is well established (79 ,80 and 81), and both P1 agonists and P2X antagonists may represent novel approaches to nociception. ATP application to sensory afferents results in neuronal hyperexcitability and the perception of intense pain (79). These pronociceptive effects are mediated by $P2X_3$ and $P2X_{2/3}$ receptors present on sensory afferents and in the spinal cord. The nucleotide also induces nociceptive responses at local sites of administration and can facilitate nociceptive responses to other noxious stimuli, such as substance P. P2 receptor antagonists such as suramin and PPADS, even though they are limited in their *in vivo* effects, reduce nociceptive responses in animal models of acute and persistent pain (1 ,79). ATP is released from certain cell types (e.g., sympathetic nerves, endothelial cells, visceral smooth muscle) in response to trauma (1 ,8 ,79), and $P2X_3$-receptor expression is up-regulated in sensory afferents and spinal cord after damage to peripheral sensory fibers. $P2X_3$-receptor knockout mice have reduced nociceptive responses (61). The effects of adenosine are opposite effects to those of ATP (80), a finding suggesting that the nociceptive effects of ATP can be autoregulated by adenosine production from the nucleotide. Adenosine, adenosine-receptor agonists, and AK inhibitors inhibit nociceptive processes in the brain and spinal cord. When given intrathecally, these agents have analgesic activity in a broad spectrum of animal models (e.g., mouse hot plate, mouse tail flick, rat formalin, mouse abdominal constriction, rat neuropathic pain models), effects that are blocked by systemic or intrathecal administration of adenosine antagonists. Adenosine A_1-receptor agonists modulate acutely evoked and inflammation-evoked responses of spinal cord dorsal horn nociceptive neurons and can also inhibit pain behaviors elicited by spinal injection of substance P and the glutamate agonist, *N*-methyl-D-aspartate (NMDA). Glutamate is a key mediator of the abnormal hyperexcitability of spinal cord dorsal horn neurons (central sensitization) associated with clinical pain states. A_1 agonists can inhibit the spinal cord release of glutamate and can also reduce cerebrospinal fluid levels of substance P in rat, another key mediator of nociceptive responses. Adenosine has both presynaptic and postsynaptic effects on transmission from primary afferent fibers to neurons of the substantia gelatinosa of the spinal dorsal horn (12 ,80 ,81), and it involves both peripheral and supraspinal mechanisms. Adenosine agonists such as CHA and NECA, were 10- to 1,000-fold more potent in inhibiting acetylcholine-induced writhing in mice when these agents were administered intracerebroventricularly than orally, a finding indicating a supraspinal site of action. The ability of adenosine to inhibit peripheral neurotransmitter (12), and inflammatory processes (67), may block peripheral sensitization, a key feature of the pain resulting from tissue injury and inflammation.

Adenosine agonists are also active in human pain states (81). Spinal administration of the A_1 agonist, R-PIA, relieved allodynia in a patient with neuropathic pain without affecting normal sensory perception, whereas adenosine infusion at doses without effect on the cardiovascular system improved pain symptoms and reduced spontaneous pain and ongoing hyperalgesia and allodynia in patients with neuropathic pain. Low-dose infusion of adenosine during surgical procedures reduced the requirement for volatile anesthetic and also for postoperative opioid analgesia (82). AK inhibitors, such as CP 3269 and ABT-702 (Fig. 15.1), are effective analgesic agents in animal pain models by effects that can be blocked by xanthine adenosine antagonists.

CHALLENGES IN THE DEVELOPMENT OF CNS-SELECTIVE THERAPEUTIC AGENTS

Part of "15 - Purinergic Neurotransmission "

The field of purinergic molecular biology and pharmacology has exploded as more is learned about the cellular targets through which ATP, ADP, AMP, and adenosine (and UTP) produce their effects on mammalian tissues. A clear historical delineation between the P1 and P2 fields is that in the former, more than 20 years of pharmacology and medicinal chemistry resulted in the identification of receptor selective ligands before the receptors were cloned. In contrast, definitive evidence for the existence of the P2-receptor family resulted from both pharmacologic and cloning studies. The latter have resulted in the identification of a remarkable diversity of receptors responsive to ATP, unfortunately in the absence of selective, bioavailable ligands, especially antagonists, that will allow a clearer understanding of P2-receptor

function in normal and pathologic states. Evidence of the oligomerization of GPCRs and the emerging data on P2X heteromers both within the P2-receptor family and with other LGICs, such as nAChRs, suggest that the dynamics and the actual composition of systems targeted by purinergic receptors are potentially very complex (83).

Early efforts to develop therapeutics based on the modulation of P1-receptor–mediated processes met with limited success. Only adenosine has been approved for use as a cardiac imaging agent and for the treatment of supraventricular tachycardia, acute systemic uses that avoid some of the side effects seen with long-acting adenosine agonists. Similarly, the unexpected *in vivo* effects of AK inhibitors suggest that this is not a viable approach to the discovery of new drugs. The use of the adenosine antagonist theophylline for the treatment of asthma and the widespread use of caffeine as a CNS stimulant represent other P1-targeted therapeutics. The evaluation of A_{2A} antagonists as indirect dopamine agonists for use in PD (73 ,74 and 75) is an intriguing and novel approach to treating this neurodegenerative disorder, although the side effect liabilities are unknown at present.

In contrast, the highly discrete localization of $P2X_3$ receptors to sensory nociceptive neurons (79) has led to an intensive effort to identify $P2X_3$ antagonists as novel analgesic agents. Similarly, the discrete localization of other P2 receptors and evidence from mouse knockout studies suggest that selective agonists and antagonists for these receptor subtypes may represent very novel therapeutic agents as well as research tools to understand target function.

A caveat in the drug discovery process, as in all life's endeavors, is that the less that is known regarding the functional liabilities of a molecular target, the more attractive it is as drug target. In the area of purinergic medications, the identification of new ligands in combination with a broader-based evaluation of compound efficacy and side effect liability will greatly assist in the prioritization of therapeutic targets that are amenable to modulation by purinergic ligands (57). Finally, the renewed interest in mitochondria as cellular organelles that have function beyond energy production (56) represents an additional level of molecular targeting for P1- and P2-receptor ligands that may have benefit in treating human disease states, especially those involving apoptosis (55).

ACKNOWLEDGMENTS

Part of "15 - Purinergic Neurotransmission "

I would like to thank Mike Jarvis for his contributions to the previous CD-ROM version of this chapter. Because of space limitations, it is not possible to cite primary literature sources exhaustively. The reader is referred to reference 1 for a more comprehensive bibliography.

REFERENCES

1. Ralevic V, Burnstock G. Receptors for purine and pyrimidines. *Pharmacol Rev* 1998;60:413–492.

2. Barnard EA, Burnstock G. ATP as a neurotransmitter. P2 purinoceptors: localization, function and transduction mechanisms. *CIBA Found Symp* 1996;198:262–265.

3. Burnstock G. Purinergic cotransmission. *Brain Res Bull* 1999;50:355–357.

4. Noji H. The rotary enzyme of the cell: the rotation of F1-ATPase. *Science* 1998;282:1844–1845.

5. Zimmerman H, Braun N. Ecto-nucleotidases: molecular structures, catalytic properties, and functional roles in the nervous system. *Prog Brain Res* 1999;120:371–385.

6. Clifford EE, Martin KA, Dalal P, et al. Stage specific expression of P2Y receptors, ecto-apyrase and ecto-5′-nucleotidase in myeloid leukocytes. *Am J Physiol* 1997;273:C973–C987.

7. Tordorov LD, Mihaylova-Todorova S, Westfall TD, et al. Neuronal release of soluble nucleotidases and their role in neurotransmitter inactivation. *Nature* 1997;387:76–79.

8. Williams M, Jarvis MF. Purinergic and pyrimidinergic receptors as potential drug targets. *Biochem Pharmacol* 2000;59:1173–1185.

9. Ross FM, Brodie MJ, Stone TW. Nucleotide and dinucleotide effects on rates or paroxysmal depolarising bursts in rat hippocampus. *Prog Brain Res* 1999;120:251–262.

10. Cooper EC, Jan LY. Ion channel genes and human neurological disease: recent progress, prospects, and challenges. *Proc Natl Acad Sci USA* 1999;96:4759–4766.

11. Choe S. Structure and function of potassium channels. *Annu Rev Biophys Biomol Struct* 2000;29–46.

12. Masino SA, Dunwiddie TV. Role of purines and pyrimidines in the central nervous system. In: Abbracchio MP, Williams M, eds. *Purinergic and pyrimidinergic signaling. I. Molecular, nervous and urogenitary system function handbook of experimental pharmacology*, vol 151/I. Heidelberg: Springer-Verlag, 2001:251–287.

13. Atkinson DE. *Cellular energy metabolism and its regulation*. San Diego: Academic, 1977.

14. Ehrlich YH, Kornecki E. Ecto-protein kinases as mediators for the action of secreted ATP in the brain. *Prog Brain Res* 1999;120:411–426.

15. Connolly GP, Diuley JA. Uridine and its nucleotides: biological actions, therapeutic potential. *Trends Pharmacol Sci* 1999;20:218–226.

16. Agnati LR, Fuxe K, Ruggeri M, et al. Effects of chronic uridine on striatal dopamine release and dopamine related behaviours in the absence of presence of chronic treatment with haloperidol. *Neurochem Int* 1989;15:107–113.

17. Kardos J, Kovács I, Szárics E, et al. Uridine activates fast transmembrane Ca^{2+} ion fluxes in rat brain homogenates. *Neuroreport* 1999;10:1577–1582.

18. Fredholm BB, Bättig K, Holmén J, et al. Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacol Rev* 1999;51:83–133.

19. Porkka-Heiskanen T, Strecker RE, Thakkar M, et al. Adenosine: a mediator of the sleep-inducing effects of prolonged wakefulness. *Science* 1997;276:1265–1268.

20. Newby AC. Adenosine and the concept of a retaliatory metabolite. *Trends Biochem Sci* 1984;9:42–44.

21. Kowaluk EA, Bhagwat SS, Jarvis MF. Adenosine kinase inhibitors. *Curr Pharmaceut Design* 1998;4:403–416.

22. Erion MD, Wiesner JB, Rosengren S, et al. Therapeutic potential of adenosine kinase inhibitors as analgesic agents. *Drug Dev Res* 2000;50:22.

23. Searl TJ, Redman RS, Silinsky EM. Mutual occlusion of P2X ATP receptors and nicotinic receptors on sympathetic neurons of the guinea-pig. *J Physiol (Lond)* 1998;510:783–791.

24. Khakh BS, Zhou X, Sydes J, et al. State-dependent cross-inhibition between transmitter-gated cation channels. *Nature* 2000;406:405–410.

25. Cirulea F, Casado V, Mallol J, et al. Immunological identification of A_1 adenosine receptors in brain cortex. *J Neurosci Res* 1995;42:818–828.

26. Milligan G. Receptors as kissing cousins. *Science* 2000;288:65–67.

27. Jacobson KA, van Rhee AM. Development of selective purinoceptor agonists and antagonists. In: Jacobson KA, Jarvis MF, eds. *Purinergic approaches in experimental therapeutics*. New York: Wiley Liss, 1997:101–128.

28. Bruns RF, Fergus J. Allosteric enhancement of adenosine A_1 receptor binding and function by 2-amino-3-benzothophenes. *Mol Pharmacol* 1990;38:939–949.

29. Linden J. Allosteric enhancement of adenosine receptors. In: Jacobson KA, Jarvis MF, eds. *Purinergic approaches in experimental therapeutics*. New York: Wiley Liss, 1997:85–98.

30. Feoktistov I, Biaggioni I. Adenosine A_{2B} receptors. *Pharmacol Rev* 1998;49:381–402.

31. Jacobson KA. Adenosine A_3 receptors: novel ligands and paradoxical effects. *Trends Neurosci* 1998;19:184–191.

32. Linden J. Cloned adenosine A_3 receptors: pharmacological properties, species differences and receptor functions. *Trends Pharmacol Sci* 1994;15:298–306.

33. Liu G-S, Downey JM, Cohen MC. Adenosine, ischemia and preconditioning. In: Jacobson KA, Jarvis MF, eds. *Purinergic approaches in experimental therapeutics*. New York: Wiley Liss, 1997:153–172.

34. Bhagwat SS, Williams M. P2 Purine and pyrimidine receptors: emerging superfamilies of G-protein coupled and ligand gated ion channel receptors. *Eur J Med Chem* 1997;32:183–193.

35. Bianchi B, Lynch KJ, Touma E, et al. Pharmacological characterization of recombinant human and rat P2X receptor subtypes. *Eur J Pharmacol* 1999;376:127–138.

36. Dubyak GR, Clifford EE, Humphreys BD, et al. Expression of multiple ATP subtypes during the differentiation and inflammatory activation of myeloid leukocytes. *Drug Dev Res* 1996;39:269–278.

37. North RA, Barnard EA. Nucleotide receptors. *Curr Opin Neurobiol* 1997;7:346–357.

38. Torres GE, Haines WR, Egan TM, et al. Coexpression of $P2X_1$ and $P2X_5$ receptor subunits reveals a novel ATP-gated ion channel. *J Biol Chem* 1999;274:6653–6659.

39. North RA, Surprenant A. Pharmacology of cloned P2X receptors. *Annu Rev Pharmacol Toxicol* 2000;40:563–580.

40. Kennedy C, Leff P. How should P2X receptors be classified pharmacologically? *Trends Pharmacol Sci* 1995;16:168–174.

41. Yu H, Bianchi B, Metzger R, et al. Lack of specificity of [^{35}S]-ATPγS and [^{35}S]-ADPβS as radioligands for inotropic and metabotropic P2 receptor binding. *Drug Dev Res* 1999;48:84–93.

42. Virginio C, Robertson G, Surprenant A, et al. Trinitrophenyl-substituted nucleotides are potent antagonists selective for $P2X_1$, $P2X_3$, and heteromeric $P2X_{2/3}$ receptors. *Mol Pharmacol* 1998;53:969–973.

43. King BF, Liu L, Pintor J, et al. Diinosine pentaphosphate (IP_5I) is a potent antagonist at recombinant rat P2X receptors. *Br J Pharmacol* 1999;128:981–988.

44. Gargett CE, Wiley JS. The isoquinoline derivative KN-62 a potent antagonist of the P2Z-receptor of human lymphocytes. *Br J Pharmacol* 1997;120:1483–1490.

45. Ingall AH, Dixon J, Bailey A, et al. Antagonists of the platelet P_{2T} receptor: a novel approach to antithrombotic therapy. *J Med Chem* 1999;42:213–220.

46. Thorne PR, Housley GD. Purinergic signalling in sensory systems. *Semin Neurosci* 1996;8:233–246.

47. Surprenant A, Rassendren F, Kawashima E, et al. The cytolytic P2Z receptor for extracellular ATP identified as a P2X receptor. *Science* 1996;272:735–738.

48. Di Virgilio F, Sanz JM, Chiozzi P, et al. The $P2Z/P2X_7$ receptor of microglial cells: a novel immunomodulatory receptor. *Prog Brain Res* 1999;120:355–370.

49. Collo G, Neidhart S, Kawashima E, et al. Tissue distribution of the $P2X_7$ receptor *Neuropharmacology* 1997;36:1277–1283.

50. Virginio C, MacKenzie A, Rassendren FA, et al. Pore dilation of neuronal P2X receptor channels. *Nat Neurosci* 1999;2:315–321.

51. Khakh BS, Bao XR, Labarca C, et al. Neuronal P2X transmitter-gated cation channels change their ion selectivity in seconds. *Nat Neurosci* 1999;2:322–330.

52. Communi O, Boeynaems JM. Receptors responsive to extracellular pyrimidine nucleotides. *Trends Pharmacol Sci* 1997;18:83–86.

53. Communi D, Robaye B, Boeynaems JM. Pharmacological characterization of the human $P2Y_{11}$ receptor. *Br J Pharmacol* 1999;128:1199–1206.

54. Miras Portugal MT, Gualix J, Mateo J, et al. Diadenosine polyphosphates, extracellular function and catabolism. *Prog Brain Res* 1999;120:397–410.

55. Honig LS, Rosenberg RN. Apopotosis and neurologic disease. *Am J Med* 2000;108:317–330.

56. Dykens JA, Davis RE, Moos WH. Introduction to mitochondrial function and genomics. *Drug Dev Res* 1999;46:2–13.

57. Simon DK, Johns DR. Mitochondrial disorders: clinical and genetic features. *Annu Rev Med* 1999;50:111–127.

58. Williams M, Burnstock G. Purinergic neurotransmission and neuromodulation: a historical perspective. In: Jacobson KA, Jarvis MF, eds. *Purinergic approaches in experimental therapeutics*. New York: Wiley Liss, 1997:3–26.

59. Ledent C, Vaugeois JM, Schiffmann SN, et al. Aggressiveness, hypoalgesia and high blood pressure in mice lacking the adenosine A_{2a} receptor. *Nature* 1997;388:674–678.

60. Mulryan K, Gitterman DP, Lewis CH, et al. Reduced vas deferens contraction and male infertility in mice lacking $P2X_1$ receptors. *Nature* 2000;403:86–89.

61. Cockayne DA, Zhu Q-M, Hamilton S, et al. $P2X_3$-deficient mice display urinary bladder hyporeflexia and reduced nocifensive behavior. *Nature* 2000;407:1011–1015.

62. Fabre JE, Nguyen M, Latour A, et al. Decreased platelet aggregation, increased bleeding time and resistance to thromboembolism in $P2Y_1$-deficient mice. *Nat Med* 1999;5:1199–1202.

63. Cressman VL, Lazorowski E, Homolya L, et al. Effect of loss of $P2Y_2$ receptor gene expression on nucleotide regulation of murine epithelial Cl^- transport. *J Biol Chem* 1999;274:26461–26468.

64. Sikora A, Liu J, Brosnan C, et al. Cutting edge: purinergic signaling regulates radical-mediated bacterial killing mechanism in macrophages through a $P2X_7$-independent mechanism. *J Immunol* 1999;163:558–561.

65. Ijzermann AP, van der Wenden NM. Modulators of adenosine uptake, release, and inactivation. In: Jacobson KA, Jarvis MF, eds. *Purinergic approaches in experimental therapeutics*. New York: Wiley Liss, 1997:129–148.

66. Dirnagl U, Iadecola C, Moskowitz MA. Pathobiology of ischaemic stroke: an integrated view. *Trends Neurosci* 1999;22:391–397.

67. Firestein GS. Anti-inflammatory effects of adenosine kinase inhibitors in acute and chronic inflammation. *Drug Dev Res* 1996;39:371–376.

68. Knutsen LJS, Murray TF. Adenosine and ATP in epilepsy. In: Jacobson KA, Jarvis MF, eds. *Purinergic approaches in experimental therapeutics*. New York: Wiley Liss, 1997:423–447.

69. During MJ, Spencer DD. Adenosine: a potential mediator of seizure arrest and postictal refractoriness. *Ann Neurol* 1992;32:618–624.

70. Neary JT, Rathbone MP, Cattabeni F, et al. Trophic actions of extracellular nucleotides and nucleosides on glial and neuronal cells. *Trends Neurosci* 1996;19:13–18.

71. Rathbone MP, Middlemiss PJ, Gysbers JW, et al. Trophic effects of purines in neurons and glial cells. *Prog Neurobiol* 2000;59:663–690.

72. Svenningsson P, Le Moine C, Fisone G, et al. Distribution, biochemistry and function of striatal adenosine A_{2A} receptors. *Prog Neurobiol* 1999;59:355–396.

73. Richardson PJ, Kase H, Jenner PG. Adenosine A_{2A} receptor antagonists as new agents for the treatment of Parkinson's disease. *Trends Pharmacol Sci* 1997;18:338–344.

74. Kanda T, Jackson MJ, Smith LA, et al. Adenosine A_{2A} antagonist: a novel antiparkinsonian agent that does not provoke dyskinesia in parkinsonian monkeys. *Ann Neurol* 1998;43:507–513.

75. Webster RG, Abbott RD, Petrovitch H, et al. Association of coffee and caffeine intake with the risk of Parkinson disease. *JAMA* 2000;283:2674–2679.

76. Hauber W, Koch M. Adenosine A_{2A} receptors in the nucleus accumbens modulate prepulse inhibition of the startle response. *Neuroreport* 1997;8:1515–1518.

77. Martin GE, Rossi D, Jarvis MF. Adenosine agonists reduce conditioned avoidance responding in the rat. *Pharmacol Biochem Behav* 1993;45:951–958.

78. Satoh S, Matsumura H, Suzuki F, et al. Promotion of sleep mediated by the A_{2a}-adenosine receptor and possible involvement of this receptor in the sleep induced by prostaglandin D2 in rats. *Proc Natl Acad Sci USA* 1996;93:5980–5984.

79. Burnstock G. A unifying purinergic hypothesis for the initiation of pain. *Lancet* 1996;347:1604–1605.

80. Sawynok J. Purines in pain management. *Curr Opin CPNS Invest Drugs* 1999;1:27–38.

81. Salter MW, Sollevi A. Roles of purines in nociception and pain. In: Abbracchio MP, Williams M, eds. *Purinergic and pyrimidinergic signaling. I. Molecular, nervous and urogenitary system function handbook of experimental pharmacology*, vol 151/I. Heidelberg: Springer-Verlag, 2001:371–401.

82. Fukunaga AF. Purines in anesthesia. In: Jacobson KA, Jarvis MF, eds. *Purinergic approaches in experimental therapeutics*. New York: Wiley Liss, 1997:471–494.

83. Williams M. Purines: from promise to premise. *J Auton Nerv Syst* 2000;81:285–288.

16

Neurotrophic Factors and Intracellular Signal Transduction Pathways

David S. Russell

Ronald S. Duman

David S. Russell and Ronald S. Duman: Laboratory of Molecular Psychiatry, Departments of Psychiatry and Pharmacology, Yale University School of Medicine, Connecticut Mental Health Center, New Haven, Connecticut.

A clear limitation of treating diseases of the central nervous system arises from the loss of regenerative potential of the brain at a very early age. Improper development of neural circuits or injury of neurons appears to be permanently fixed in the adult, with seemingly little hope of restorative therapy. This is most clearly true of the spectrum of neurologic diseases, such as neurodegenerative diseases (e.g., Alzheimer's or Parkinson's disease and stroke) and inflammatory disease (e.g., multiple sclerosis). In addition, breakthrough disorders such as epilepsy or myoclonus reflect inappropriate “wiring” or lack of appropriate feedback control of neuronal activity. It has also become apparent that some psychiatric diseases represent similar fixed deficits or lack of appropriate functional adaptation. Neurochemical and neuroanatomic deficits have been documented in affective diseases, schizophrenia, and anxiety disorders. Methods to recreate the early restorative potential of the brain would have great potential for significantly, and possibly permanently, reversing the often debilitating features of these neurologic and psychiatric diseases.

Over the last several years, of body of literature has delineated the important roles of neurotrophic factors in guiding the development of the nervous system. The first identified neurotrophic factor, nerve growth factor (NGF), was identified and characterized in the pivotal studies of Levi-Montalcini (1). The general properties of NGF now essentially define what neuroscientists consider a *neurotrophic factor*. A neurotrophic factor is capable of supporting the survival of at least one population of neurons in culture. It is secreted by a target tissue (either neuronal or nonneuronal) and acts on the neurons that innervate that tissue to support their survival or differentiation. Finally, a neurotrophic factor is expressed in the appropriate region and at the appropriate time in development to support the survival of a particular neuronal population. Although several variations and extensions of these principles have been delineated, the basics of defining a neurotrophic factor remain the same.

An exciting development in the neurosciences has been the realization that neurotrophic factors play important roles in the adult brain. The time course of neurotrophic factor expression is intriguing and indicates important function in the adult nervous system, as well as during development (2 ,3 and 4). The expression of these factors usually is very high during early development, a time of substantial growth, differentiation, and modeling of the nervous system. Later the levels generally drop, but they do not subside completely. In fact, in most cases in which it has been explored, the continued presence of these factors is substantial and is critical throughout adulthood (e.g., see refs. 5 ,6 and 7). Neuronal populations continue to depend on these factors for survival and optimal functioning.

More intriguing, although perhaps not surprising, studies have clearly shown that after development, these factors participate in the ongoing remodeling of neuronal function that underlies the adaptability or plasticity of neurons. In some cases, specific neurotrophic factors have been found to be necessary and sufficient for these changes to occur, from hippocampal plasticity and long-term potentiation (8 ,9 ,10 and 11) to the acquisition of new songs by songbirds (12 ,13 and 14). Models of neuronal now often incorporate components of neurotrophic factor signaling to explain synaptic alterations or strengthening. Contrary to the original models, these signaling events have been found not only to be retrograde signals from target neurons or other tissues, but also to be anterograde or autocrine signals. Furthermore, numerous studies have demonstrated that the expression of at least some of these factors can be rapidly regulated in the adult, a finding supporting a dynamic role in mediating responses to the environment.

Ongoing work in the neurotrophic factor field has been

devoted to characterizing the pathways that underlie the intracellular signaling of these factors. This information may then be used to treat many different neurologic and psychiatric disorders, given the apparent critical roles of neurotrophic factors in the normal functioning and adaptability of the brain, as well as their potential to recapitulate early developmental processes to restore damaged or maladaptive neural systems. This review focuses on the neurobiology of these factors, their interactions with classical neurotransmitter systems, and their potential roles in the origin and possible treatment of psychiatric illnesses.

NEUROTROPHINS

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

The number of neurotrophic factors has now been expanded in to the dozens, each with unique a specificity in terms of biological activity, regional and temporal expression, and target specificity. Some factors previously known for their effects in other systems, such as insulin-like growth factors or tumor-derived factors, have now clearly been found also to be neurotrophic factors. These have been grouped into different families largely based on homology at the level of nucleotide sequence and therefore evolutionary relatedness. Perhaps the best understood and most widely expressed in the brain of these families are the *neurotrophins* (NTs) (2 ,15). NGF is the prototype of the NT family, which also now includes brain-derived neurotrophic factor (BDNF), NT-3, and NT-4 (or NT-4/5). NGF has the most restricted specificity among these. NGF in the brain acts specifically on cholinergic neurons. In the rest of the nervous system, it also acts on sympathetic and sensory neurons. BDNF and NT-3 are widely and highly expressed, particularly in cortical and neocortical structures. NT-4 is also widely expressed, although generally at lower levels in the adult than are the others. The NTs are small, secreted proteins of about 12 kd that contain characteristic intramolecular disulfide bonds. These then form active noncovalent homodimers. They are found stored in vesicles clustered near the membrane. Each has been cloned and expressed in active recombinant forms.

TRK RECEPTORS

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

Generally better conserved than their ligands, the neurotrophic factor receptors also form families of related proteins (16 ,17). These receptors can be found in many different forms, from single, active proteins to large heteromeric complexes. Common to these are an extracellular ligand-binding portion, a mechanism to transduce this signal across the membrane, and at least one intracellular signaling apparatus. These may be contained in single proteins or distributed among several interacting proteins. Most, if not all, of the neurotrophic factor receptor complexes include a protein tyrosine kinase. Although phosphorylation on tyrosine residues represents a relatively small proportion of all protein phosphorylation in the cell, it seems to be a critical part of neurotrophic factor signal transduction and function.

The NTs act through receptors known as *Trk receptors*. Given their name from a troponin/receptor kinase gene fusion identified from colon carcinoma, it has now been found that in their normal (protooncogene) forms, each Trk receptor contains a ligand-binding domain, a single transmembrane domain, and an intrinsic, intracellular tyrosine kinase domain (16 ,17). Each receptor, when transfected into a cell line, is capable of transducing the appropriate NT signals independently of other receptor proteins (18). There is specificity among the Trk receptors at physiologic NT concentrations. TrkA is a receptor for NGF, whereas TrkC is preferentially bound by NT-3. TrkB serves as a receptor for both BDNF and NT-4. The expression patterns of these receptors correlate with known sensitivity of those neurons to specific NTs. Several studies, particularly in mice with engineered deletions of NTs or their receptors, have shown significant complexity to these interactions. A review of this work falls beyond the scope of this review.

TRK DOCKING PROTEINS

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

On binding of an NT, the Trk receptor tyrosine kinase becomes activated. The most critical substrate of this activity appears to be the receptor itself. The receptor becomes rapidly autophosphorylated, and this is critical to receptor function. The receptor autophosphorylation sites form *docking sites* for the interaction of downstream signaling molecules (Fig. 16.1). Many signaling proteins contain domains that specifically bind to tyrosine residues when they are phosphorylated. Further binding specificity is mediated by the amino acids surrounding the autophosphorylated tyrosine. The domains of the signaling molecules that are used to bind the tyrosine residues seem to fall into a small number of conserved motifs (19).

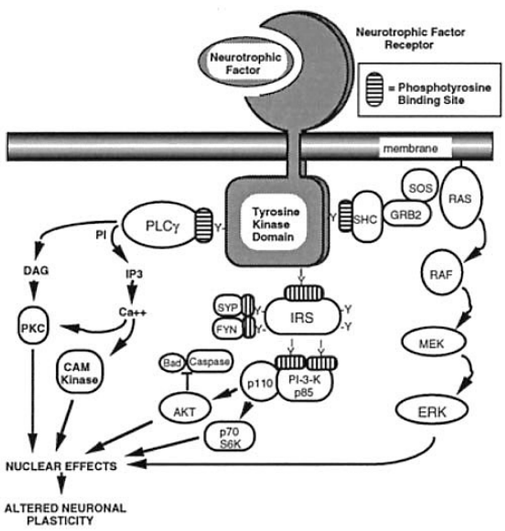


FIGURE 16.1. Neurotrophin intracellular signal transduction pathways. This schematic represents the three major signaling pathways emanating from a Trk-like tyrosine kinase receptor. The ligand, here a neurotrophin dimer, binds to its receptor and activates the tyrosine kinase activity. This results in autophosphorylation and phosphorylation of substrates. A complex forms in which docking proteins bind to the autophosphorylated receptor and are activated. These docking proteins bind to the receptor phosphorylation sites by SH2 domains (PLC-γ) or PTB domains (IRS and Shc). The three pathways shown here are as follows: (1) The PLC-γ pathway leads to the production of diacylglycerol and intracellular calcium. This results in activation of CAM kinases and PKC. (2) PI-3-Kinase (PI-3-K), comprising an 85-kd regulatory subunit and a 100-kd catalytic subunit, generally becomes activated by binding to an insulin receptor substrate (IRS)–like adaptor protein, which interacts with the receptor and then activates signaling proteins. IRS proteins also bind other signaling molecules such as Fyn and Syp. The PI-3-K then activates AKT and p70 S6-kinase (p70 S6K). AKT has an antiapoptotic effect through actions on Bad and caspaces. (3) Ras becomes activated through the stimulation of the GDP/GTP exchange activity of SOS, which is in a complex with Shc and Grb2. Activated Ras then turns on the cascade of kinases including Raf, MEK, and erks (also known as MAP kinases). Erks stimulate many known effectors. Each of these pathways then exerts a number of nuclear and nonnuclear actions with short- and long-term consequences for the cells.

Most proteins that bind to phosphorylated tyrosines fall into one of two groups. The most common phosphotyrosine binding motif is the *src-homology domain 2, or SH-2 domain*. SH-2 domains are typically identified based on their homology to other SH-2 domain–containing proteins. Some of these have been shown directly to possess specificity for phosphorylated tyrosines in the appropriate amino acid context. The SH-2 domain–containing proteins also often contain, or interact with proteins containing, an *src-homology domain 3, or SH-3* (20). The other, unrelated conserved motif that directs a separate type of specific protein–protein interaction has been termed, appropriately, a *phosphotyrosine binding domain, or PTB domain*. Proteins containing a PTB domain bind to a distinct set of phosphorylated tyrosine residues from those with SH-2 domains (21).

After these signaling proteins bind to the activated, autophosphorylated

receptor, they become activated. The mechanisms by which they are activated are not entirely clear. Often they, too, become phosphorylated on tyrosine residues. In addition, their recruitment to the membrane or into signaling complexes plays a role in the initiation of their activity (22 ,23). Activation of each of these NT-signaling proteins triggers distinct downstream cascades of target enzymes and other biological effects. Although there is great diversity of neurotrophic factor receptors, they seem to trigger only a few well-conserved types of downstream signaling pathways. Among the best characterized of the pathways include the Ras/extracellular signal regulated kinase (ERK) pathway, the phosphotidylinositol-3′-OH-kinase (PI-3-K) pathway, and the phospholipase C-γ (PLC-γ) pathway (24). In addition, specific tyrosine phosphatases are activated that modulate these responses and may contain pathway-activating properties of their own.

NEUROTROPHIC FACTOR INTRACELLULAR SIGNALING PATHWAYS: RAS/ERK (MAPK) CASCADE

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

The Ras/ERK pathway is regulated by the activity of the Ras proteins. Ras is a small, membrane-associated protein that serves as a transducer of signal from tyrosine kinase activity to ERK proteins, among other activities (25). The activity of Ras depends on the type of the guanine nucleotide it is bound to. Hence, Ras is a G protein, although distinct from the heterotrimeric G proteins coupled to many neurotransmitter receptors. Ras is active when binding guanosine triphosphate (GTP), but at rest it is inactive and bound to guanosine diphosphate (GDP). Activation of Trk receptor tyrosine kinases leads to the binding of an adapter protein in the Shc family. Shc becomes tyrosine phosphorylated and binds a Grb protein, such as Grb2. Shc contains a PTB domain and an SH-2 domain, whereas Grb2 contains two SH-2 domains and a SH-3 domain. This complex then activates a GDP-GTP exchange factor, such as SOS, which, in turn, activates Ras through GTP binding.

Once activated, Ras recruits a serine kinase of the Raf family to the membrane, where it is activated. This initiates a cascade in which Raf activates MEK (from *M*APK or *E*RK *k*inase), and then MEK phosphorylates and activates ERKs (26). ERKs, also known as mitogen-activated protein kinase (MAPKs), are abundant, multifunctional, intracellular kinases with many different cellular activities. ERKs have been shown to phosphorylate such diverse proteins as tyrosine hydroxylase, transcription factors, regulators of protein translation, microtubule proteins, and many others (27 ,28). In cells *in vitro*, ERKs have been shown to mediate neuron survival, neuritic process elongation (29), and levels of specific neuronal enzymes and ion channels, among other effects (30). ERKs have been shown to be important in hippocampal long-term potentiation in brain slices (31). Some effects of ERK activation are very rapid, whereas others are delayed and persistent.

PLC-γ CASCADE

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

A second NT-signaling pathway involves PLC-γ activation (32 ,33). Like the better-understood PLC-ß, PLC-γ cleaves phosphatidylinositol phosphates into diacylglycerol and inositol phosphates. Diacylglycerol can activate protein kinase C (PKC), whereas inositol-1,4,5-tris phosphate releases intracellular stores of calcium. Intracellular calcium can exert numerous effects from the activation of Ca^{2+}/calmodulin–dependent

protein kinases to the production of cyclic adenosine monophosphate through some adenylyl cyclases. All these are known to have powerful effects on neurons. Unlike PLC-ß, which is regulated by heterotrimeric G-protein–coupled receptors, PLC-γ is regulated by tyrosine phosphorylation (34). PLC-γ contains SH-2 and SH-3 domains. When bound to tyrosine phosphorylated receptors, it is recruited to the membrane and becomes phosphorylated, which activates its PLC activity. Virtually nothing is known about the role of PLC-γ in the intact brain, although it is likely to exert important effects on neuronal function.

PI-3-K CASCADE

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

Somewhat less well understood is the PI-3-K pathway (35 ,36 and 37). Several types of PI-3-Ks have been identified. Type 1 PI-3-Ks are regulated by tyrosine kinase activity. They are heterodimers of a catalytic α subunit and a regulatory ß subunit. The ß subunit contains SH-2 and SH-3 domains. When bound to phosphorylated tyrosines, the ß subunit activates the catalytic activity of the α subunit. This phosphorylates phosphatidylinositol on the 3′-hydroxl group (distinct from the 4′,5′ phosphorylated forms mentioned earlier). Furthermore, PI-3-K has been shown to possess protein kinase activity and can bind Ras (38). The PI-3-K lipid product, phosphotidylinositol-3′-phosphate, activates at least two protein kinases, AKT and S6-kinase. AKT is best known for its powerful ability to oppose programmed cell death (i.e., antiapoptotic effects), although it has other metabolic actions as well. S6-kinase is named for its ability to phosphorylate the ribosomal subunit S6 (although not to be confused with ribosomal S6-kinase RSK), and it has numerous other cellular effects as well. Currently, these effects are less well elucidated than those of the ERKs. Within neurons, PI-3-K has been shown to mediate cell survival, initiation of neuritic process outgrowth, and acquisition of sensitivity to glutamate excitotoxicity (39), among other actions. Again, little is known about its role in intact brain.

Although PI-3-K possesses some ability to bind phosphorylated receptors itself, it seems largely to be activated by receptors through docking proteins. Important PI-3-K docking proteins are the insulin receptor substrate (IRS) family of proteins (40). IRS1, IRS2, and IRS4 are expressed in brain (41). More distantly related PI-3-K docking proteins are the GAB family of proteins. All these bind to the receptors through PTB domains and become phosphorylated on numerous tyrosines. Most of these tyrosines are then bound by PI-3-K, leading to a substantial amplification of PI-3-K signaling. IRS proteins can be bound by other signaling molecules such as protein tyrosine phosphatases and also possess numerous serine and threonine phosphorylation sites. Therefore, it is likely that the IRS proteins are important sites of convergence of numerous types of signaling pathways.

INTERACTION OF NEUROTROPHIN SIGNALING CASCADES

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

Numerous levels of complexity have been found in the downstream signaling pathways of the NT receptors. PI-3-K has been shown to contribute to ERK activation by Ras-dependent and Ras-independent process pathways (36 ,38 ,42). Ras can contribute to activation of AKT, and both SHC and IRS can bind to the same phosphorylated tyrosine site, although with differing affinities (43). PLC-γ can activate ERK in neurons and can theoretically terminate PI-3-K signaling by cleaving phosphotidylinositol-3′-phosphate. In addition, most of these proteins exist in multiple isoforms arising either from different genes or differential splicing of the same gene. These isoforms are differentially expressed during development and in different brain regions, although there is also a great deal of overlap. The complement of signaling proteins and adaptors would be expected to determine the effects of the NT on particular populations of neurons. Regulation of these proteins may influence plasticity or other neuronal responses. Furthermore, this complexity of expression and cross-talk allows tremendous opportunity for potential sites of therapeutic intervention.

REGULATION OF NEUROTROPHIN SIGNALING BY ACTIVATION OF G-PROTEIN–COUPLED RECEPTORS

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

There is also evidence of cross-talk between G-protein–coupled receptor signaling pathways and the NT cascades. Many different types of interactions occur between the G-protein receptor–coupled second messenger-dependent pathways and the Trk signaling pathways. Only a few are discussed here, to demonstrate the complexity and possible types of interactions.

Activation of Trk Docking Proteins and the Ras/ERK Pathway by G-Protein–Coupled Receptors

G-protein–coupled receptors are reported to activate the Ras/ERK by a pathway that is independent of their respective second messenger systems (44). This alternate pathway is dependent on internalization of the receptor and recruitment of a soluble tryosine kinase that directly phosphorylates the adaptor proteins (Shc and Gab) that lead to activation of Ras and subsequently the Ras/ERK pathway. For example, internalization of the ß-adrenergic receptor (ßAR) leads to binding of ß-arrestin, which inhibits activation of the receptor. Studies demonstrated that ß-arrestin also functions as an adaptor protein that binds both ßAR and a soluble tyrosine kinase, Src. Studies demonstrated that 5-hydroxytryptamine (5-HT_{1A}) receptors activate the Ras/ERK pathway, possibly through this mechanism (45). Regulation

of the Ras/ERK pathway by internalization of G-protein–coupled receptors is not observed in all cases (e.g., 5-HT_{2A} receptor), a finding indicating receptor or cellular specificity in the control of this pathway.

Numerous other potential mechanisms of cross-talk between G protein and neurotrophic factor signaling pathways have been identified. For instance, several other forms of PI-3-Ks have been elucidated, some of which are activated by G-protein–coupled receptors, but presumably they lead to at least some similar downstream signaling effects through the production of similar phosphorylated inositol lipids (35 ,36 and 37). Similarly, PLC-ß, the G-protein–coupled form of PLC, would be expected to produce at least some of the same effects as PLC-γ by activation of PKCs and mobilization of calcium. Furthermore, both PKC and levels of cyclic adenosine monophosphate modulate the activity of Raf and hence the ERK pathway (46 ,47 ,48 and 49). Calcium mobilization can activate ERKs through the intracellular tyrosine kinase, Pyk2 (50). Obviously, the potential for cross-talk with G-protein–coupled systems is great, and sorting out the mechanisms relevant in various brain regions and mechanisms of plasticity is an important area for investigation.

SOME OTHER CLASSES OF NEUROTROPHIC FACTORS AND THEIR SIGNALING SYSTEMS

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

Although this review focuses on the NT family of neurotrophic factors, several others have generated significant interest (Fig. 16.2). For example, the glial cell line–derived neurotrophic factor (GDNF) family of proteins has been found to have profound effects on dopaminergic and motor neurons, as well as enteric neurons (51 ,52 and 53). Thus far, this family includes GDNF, artemin, persephin, and neurturin. This family, distantly related to the transforming growth factor-ß (TGF-ß) family, signals through a heteromeric signaling complex (54). A common component of all the GDNF receptor complexes identified is the Ret protein. Ret spans the membrane and contains an intracellular tyrosine kinase domain. This domain, which is less well characterized than for the Trk receptor, can nonetheless signal at least some of the same intracellular signaling pathways, including the Ras/ERK pathway and PI-3-K (55). Ret associates with various receptor-binding proteins (GFR-α subunits 1 through 4). These subunits do not span the membrane, but are linked to the extracellular surface through a glycosyl-phosphatidylinositol moiety. The GFR-α subunits provide specificity for ligand binding and participate in the activation of the Ret tyrosine kinase. Furthermore, evidence indicates that they can stimulate activation of src-like tyrosine kinases independent of Ret. The known activities of the GDNF family of proteins has spurred interest in its role in the pathogenesis and possible treatment of diseases such as Parkinson's disease, addiction, and amyotrophic lateral sclerosis (52 ,53).

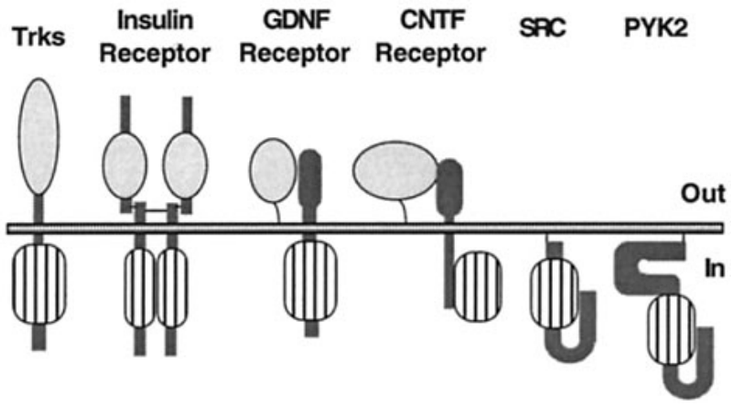


FIGURE 16.2. Other neurotrophic factor signaling cascades. This cartoon illustrates the general motifs used by neurotrophic factor receptors to transduce a signal across the membrane and to couple with the intracellular signaling pathways. The *light gray ovals* represent extracellular (Out) ligand-binding domains. The *vertical striped* intracellular (In) ovals are tyrosine kinase domains. The Trk receptors, like EGF receptors, are single polypeptide chains that span the membrane. These contain intrinsic extracellular ligand binding domains and intracellular tyrosine kinase activity. The insulin and IGF receptors are derived from the cleavage and subsequent disulfide linking of two identical chains. These receptors also contain intrinsic binding and kinase activities. The GDNF receptor is a heteromeric complex in which the ligand-binding specificity resides within a separate phosphatidylinositolglycan-linked subunit (GFR-α). The signal from the ligand–GFR-α complex is then transduced across the membrane by Ret, which contains an intracellular tyrosine kinase domain. The CNTF receptor is typical of the cytokine receptors, which contain separate binding and transmembrane subunits, and then couple to a soluble intracellular tyrosine kinase (a JAK). Other systems couple to the tyrosine kinase intracellular signaling pathway by interacting with one of several membrane-associated intracellular tyrosine kinases. Some G proteins, for instance, act through src-like proteins, whereas intracellular calcium can activate Pyk2, an intracellular tyrosine kinase in the focal adhesion kinase family.

Another large class of receptors couples to an intracellular tyrosine kinase known as the Janus kinase, or JAK (56 ,57). These receptor kinases include extracellular binding components, one of which spans the membrane but does not have intrinsic kinase activity. Instead, they bind to and activate specific members of the JAK family of kinases. The JAKs then activate certain intracellular effector molecules, including IRS proteins, SHC, and others. Moreover, they interact with a unique group of proteins called STATs. STAT proteins bind to JAK. After tyrosine phosphorylation, they are released and translocated to the nucleus, where they function directly as DNA-binding transcriptional activators. Many neurotrophic factors activate the JAK/STAT pathway, including ciliary neurotrophic factor, growth hormone, leptin, and many cytokines. Characterization of the role of STAT-mediated transcription in brain lags behind that of other earlier identified transcription factors, but it is an area of intensive study.

Even factors known for their hormone functions or peripheral

effects have been found to have substantial activity in the central nervous system. Insulin and insulin-like growth factor 1 (IGF-1), and their respective transmembrane receptor tyrosine kinases, are expressed widely in brain and play roles in development and behavior (58 ,59). Furthermore, epidermal growth factor (EGF) and EGF-like ligands in the TGF-α family are expressed in brain along with their receptor, the EGF receptor. Evidence is accumulating for roles for these factors in the adult central nervous system as well (60). Although not discussed in detail, these serve as further examples of the complexity of neurotrophic factor signaling at many levels in the brain. These receptors also share coupling to the same modules of signaling pathway proteins as the Trks and Ret. Apparently, there is a tremendous diversity of neurotrophic ligands and ligand-binding domains within their receptors that allows fine anatomic and temporal specificity of action, along with the potential for synergistic or counterregulatory mechanisms. However, the relatively smaller number of conserved signaling pathways to which they couple suggests that they share common mechanisms of action to shape neuronal responses.

ROLE OF NEUROTROPHIC FACTORS IN THE ACTIONS OF PSYCHOTROPIC DRUGS

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

As investigations of the psychotropic drugs have been extended, it has become clear that these agents also influence the expression of neurotrophic factors and their signal transduction systems. Regulation of neurotrophic factor signaling could thereby contribute to the desired actions of therapeutic of agents, as well as the negative effects of other drugs. These possibilities are illustrated briefly in this section.

Regulation of Neurotrophin Systems by Stress and Antidepressant Treatment

Preclinical and clinical studies have reported an atrophy or loss of neurons in limbic brain structures that could be related to dysfunction of neurotrophic factor systems. Chronic physical or social stress can result in atrophy or death of stress-vulnerable neurons in the hippocampus of rodents and nonhuman primates (61 ,62). More recently, investigators conducting brain imaging studies reported that the volume of the hippocampus is reduced in patients suffering from depression of posttraumatic stress disorder (63). Postmortem studies also demonstrate that the numbers of neurons and glia in prefrontal cortex are reduced in patients with depression (63). The expression of BDNF in hippocampus is decreased by exposure of animals to stress (64). This effect could contribute to the atrophy and death of hippocampal neurons, although it is also likely that other pathways are involved in this effect (61 ,62).

In contrast to the actions of stress, antidepressant treatment increases the expression of BDNF, as well as its receptor TrkB, in hippocampus (65 ,66). Up-regulation of BDNF is dependent on long-term antidepressant treatment, consistent with the time course for the therapeutic action of these agents. Both norepinephrine and serotonin-selective reuptake inhibitor antidepressants increase BDNF expression, a finding suggesting that this NT system may be a common postreceptor target of these monoamines and antidepressant treatment. In addition, nonantidepressant psychotropic drugs do not increase BDNF expression in hippocampus, a finding indicating that this effect is specific to antidepressants. The possibility that BDNF contributes to the therapeutic actions of antidepressants is supported by behavioral studies. Infusion of BDNF into midbrain or hippocampus produces antidepressant-like effects in behavioral models of depression, the forced swim test, and learned helplessness paradigms (67 ,68). Additional studies will be required to elucidate further the role of BDNF, as well as other neurotrophic factors, in the pathogenesis and treatment of depression. However, these findings have contributed to an exciting new hypothesis of depression.

Role of Neurotrophic Factors in the Actions of Drugs of Abuse

A picture is emerging that neurotrophic factors and their signaling pathways play important roles in mediating acute and chronic changes in synaptic connectivity, neuronal physiology, and gene expression. A powerful and important model of environmentally induced acute and persistent alterations in brain function is the effect of chronic exposure to drugs of abuse (69 ,70). Within laboratory animals, exposure to any of several diverse addicting drugs leads a set of alterations in neuronal biochemistry, electrophysiology, and morphology in specific brain regions implicated in addictive behaviors. Concomitantly, these animals display alterations in behavior including tolerance, dependence, sensitization, craving, and drug-seeking behaviors reminiscent of the behaviors seen in humans suffering from drug addiction. Specifically, alterations in the dopaminergic nucleus, the ventral tegmental area (VTA), are reminiscent of the changes seen with neurotrophic factor withdrawal in cell culture: the cells become smaller with less prominent neuritic processes, they have decreased neurofilament expression and axoplasmic transport, and they have decreased expression and accumulation of tyrosine hydroxylase, the rate-limiting enzyme in dopamine syntheses and a critical neuron type-specific protein. Infusing NTs such as BDNF or GDNF into the VTA can restore most, or all, of these features to normal (71 ,72). Additional studies have also implicated endogenous NT-3 in the drug-induced changes (73). Furthermore, the role of ERK signaling in this system has been established. ERK activity is increased in the VTA by chronic morphine exposure (74). Infusion of a specific antisense oligonucleotide against ERK1 into the VTA again blocks the morphine-induced biochemical changes (75).

Dissecting the mechanisms of signaling protein regulation within specific brain nuclei in the intact animal poses special challenges. However, using tools from *in vitro* studies, headway is now starting to be made. For instance, the mechanism of this ERK up-regulation is unclear, but it has been shown that PLC-γ, which is capable of activating ERK, is up-regulated in VTA after chronic morphine exposure (76). Although levels of the neurotrophic factors themselves have not been found to be significantly altered in VTA by chronic drug exposure, they may be regulated indirectly by modulation of their signaling systems.

CONCLUSIONS

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

The neurotrophic factors and their signal transduction cascades represent a complex array of pathways that influence many aspects of neuronal function and survival during development as well as in the adult central nervous system. The characterization of these pathways has provided many new target sites for the development of novel agents that could be used to treat a variety of neurologic and psychiatric illnesses. There is currently a tremendous amount of interest in this area, and agents are already available for selective blockade of certain components of the Ras/ERK pathway. Moreover, characterization of the roles these pathways play in the normal nervous system may lead to identification of abnormal conditions that underlie pathologic states. The opening of the field of growth factor action into the neurosciences opens opportunities that may be as rich or as powerful, if not more, as those that have been presented with the more traditional neurotransmitter systems.

ACKNOWLEDGMENTS

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

We would like to acknowledge the support of United States Public Health Service grants DA00302, MH45481, MH53199, and 2 PO1 MH25642, the Veterans Affairs National Center Grant for Posttraumatic Stress Disorder, the Veterans Affairs Medical Center, and the National Alliance for Research on Schizophrenia and Depression (NARSAD).

DISCLAIMER

Part of "16 - Neurotrophic Factors and Intracellular Signal Transduction Pathways "

Dr. Duman serves as a consultant to Pfizer, Psychogenics, Janssen, Lilly, and Pharmacia-Upjohn. In addition, he currently receives research support from Pfizer and serves as a member of the scientific advisory board for Psychogenics.

REFERENCES

1. Levi-Montalcini R. The nerve growth factor 35 years later. *Science* 1987;237:1154–1162.

2. Barde YA. The nerve growth factor family. *Prog Growth Fact Res* 1990;2:237–248.

3. Lu B, Figurov A. Role of neurotrophins in synapse development and plasticity. *Rev Neurosci* 1997;8:1–12.

4. Chao MV. Trophic factors: an evolutionary cul-de-sac or door into higher neuronal function? *J Neurosci Res* 2000;59:353-355.

5. Maisonpierre PC, Belluscio L, Friedman B, et al. NT-3, BDNF, and NGF in the developing rat nervous system: parallel as well as reciprocal patterns of expression. *Neuron* 1990;5:501–509.

6. Friedman WJ, Ernfors P, Persson H. Transient and persistent expression of NT-3/HDNF mRNA in the rat brain during postnatal development. *J Neurosci* 1991;11:1577–1584.

7. Timmusk T, Belluardo N, Metsis M, et al. Widespread and developmentally regulated expression of neurotrophin-4 mRNA x in rat brain and peripheral tissues. *Eur J Neurosci* 1993;5:605–613.

8. Kang H, Schuman EM. Long-lasting neurotrophin-induced enhancement of synaptic transmission in the adult hippocampus. *Science* 1995;267:1658–1662.

9. Korte M, Carroll P, Wolf E, et al. Hippocampal long-term potentiation is impaired in mice lacking brain-derived neurotrophic factor. *Proc Natl Acad Sci USA* 1995;92:8856–8860.

10. Patterson SL, Abel T, Deuel TA, et al. Recombinant BDNF rescues deficits in basal synaptic transmission and hippocampal LTP in BDNF knockout mice. *Neuron* 1996;16:1137–1145.

11. Messaoudi E, Bårdsen K, Srebro B, et al. Acute intrahippocampal infusion of BDNF induces lasting potentiation of synaptic transmission in the rat dentate gyrus. *J Neurophysiol* 1998;79:496–499.

12. Akutagawa E, Konish M. Transient expression and transport of brain-derived neurotrophic factor in the male zebra finch's song system during vocal development. *Proc Natl Acad Sci USA* 1998;95:11429–11434.

13. Rasika S, Alvarez-Buylla A, Nottebohm F. BDNF mediates the effects of testosterone on the survival of new neurons in an adult brain. *Neuron* 1999;22:53–62.

14. Dittrich F, Feng Y, Metzdorf R, et al. Estrogen-inducible, sex-specific expression of brain-derived neurotrophic factor mRNA in a forebrain song control nucleus of the juvenile zebra finch. *Proc Natl Acad Sci USA* 1999;96:8241–8246.

15. Thoenen H. The changing scene of neurotrophic factors. *Trends Neurosci* 1991;14:165–170.

16. Barbacid M. The Trk family of neurotrophin receptors. *J Neurobiol* 1994;25:1386–1403.

17. Ip NY, Yancopoulos GD. Neurotrophic factors and their receptors. *Ann Neurol* 1994;35:S13–S16.

18. Ip NY, Stitt TN, Tapley P, et al. Similarities and differences in the way neurotrophins interact with the Trk receptors in neuronal and nonneuronal cells. *Neuron* 1993;10:137–149.

19. Ullrich A, Schlessinger J. Growth factor signaling by receptor tyrosine kinases. *Neuron* 1992;9:383–391.

20. Schlessinger J. SH2/SH3 signaling proteins. *Curr Opin Genet Dev* 1994;4:25–30.

21. van der Geer P, Pawson T. The PTB domain: a new protein module implicated in signal transduction. *Trends Biochem Sci* 1995;20:277–280.

22. Stokoe D, Macdonald SG, Cadwallader K, et al. Activation of Raf as a result of recruitment to the plasma membrane. *Science* 1994;264:1463–1467.

23. Leevers SJ, Paterson HF, Marshall CJ. Requirement for Ras in Raf activation is overcome by targeting Raf to the plasma membrane. *Nature* 1994;369:411–414.

24. Russell DS. Neurotrophins: mechanisms of action. *Neuroscientist* 1995;1:3–6.

25. Denhardt DT. Signal-transducing protein phosphorylation cascades mediated by Ras/Rho proteins in the mammalian cell: the potential for multiplex signalling. *Biochem J* 1996;318:729–747.

26. Ahn NG, Seger R, Krebs EG. The mitogen-activated protein kinase activator. *Curr Opin Cell Biol* 1992;2:992–999.

27. Davis RJ. The mitogen-activated protein kinase signal transduction pathway. *J Biol Chem* 1993;268:14553–14556.

28. Cobb MH. MAP kinase pathways. *Prog Biophys Mol Biol* 1999;71:479–500.

29. Fukuda M, Gotoh Y, Tachibana T, et al. Induction of neurite outgrowth by MAP kinase in PC12 cells. *Oncogene* 1995;11:239–244.

30. Grewal SS, York RD, Stork PJ. Extracellular-signal-regulated kinase signalling in neurons. *Curr Opin Neurobiol* 1999;9:544–553.

31. English JD, Sweatt JD. A requirement for the mitogen-activated protein kinase cascade in hippocampal long term potentiation. *J Biol Chem* 1997;272:19103–19106.

32. Wahl MI, Jones GA, Nishibe S, et al. Growth factor stimulation of phospholipase C-gamma 1 activity: comparative properties of control and activated enzymes. *J Biol Chem* 1992;267:10447–10456.

33. Carpenter G, Ji QS. Phospholipase C-gamma as a signal-transducing element. *Exp Cell Res* 1999;253:15–24.

34. Nishibe S, Wahl MI, Hernandez-Sotomayor SM, et al. Increase of the catalytic activity of phospholipase C-gamma 1 by tyrosine phosphorylation. *Science* 1990;250:1253–1256.

35. Fruman DA, Meyers RE, Cantley LC. Phosphoinositide kinases. *Annu Rev Biochem* 1998;67:481–507.

36. Wymann MP, Pirola L. Structure and function of phosphoinositide 3-kinases. *Biochim Biophys Acta* 1998;1436:127–150.

37. Anderson RA, Boronenkov IV, Doughman SD, et al. Phosphatidylinositol phosphate kinases, a multifaceted family of signaling enzymes. *J Biol Chem* 1999;274:9907–9910.

38. Rodriguez-Viciana P, Warne PH, Dhand R, et al. Phosphatidylinositol-3-OH kinase as a direct target of Ras. *Nature* 1994;370:527–532.

39. Fryer HJ, Wolf DH, Knox RJ, et al. Brain-derived neurotrophic factor induces excitotoxic sensitivity in cultured embryonic rat spinal motor neurons through activation of the phosphatidylinositol 3-kinase pathway. *J Neurochem* 2000;74:582–595.

40. Myers MG Jr, White MF. Insulin signal transduction and the IRS proteins. *Annu Rev Pharmacol Toxicol* 1996;36:615–658.

41. Numan S, Russell DS. Discrete expression of insulin receptor substrate-4 mRNA in adult rat brain. *Brain Res Mol Brain Res* 1999;72:97–102.

42. Carpenter CL, Cantley LC. Phosphoinositide kinases. *Curr Opin Cell Biol* 1996;8:153–158.

43. Wolf G, Trub T, Ottinger E, et al. PTB domains of IRS-1 and Shc have distinct but overlapping binding specificities. *J Biol Chem* 1995;270:27407–27410.

44. Luttrell LM, Ferguson SSG, Daaka Y, et al. β-Arrestin-dependent formation of β_2 adrenergic receptor-Src protein kinase complexes. *Science* 1999;283:655–661.

45. Mendez J, Kadia TM, Somayazula RK, et al. Differential coupling of serotonin 5-HT1A and 5-HT1B receptors to activation of ERK2 and inhibition of adenylyl cyclase in transfected CHO cells. *J Neurochem* 1999;73:162–168.

46. Sozeri O, Vollmer K, Liyanage M, et al. Activation of the c-raf protein kinase by protein kinase C. *Oncogene* 1992;7:2259–2262.

47. Kolch W, Heldecker G, Kochs G, et al. Protein kinase C activates RAF-1 by direct phosphorylation. *Nature* 1993;364:249–252.

48. Wu J, Dent P, Jelinek T, et al. Inhibition of the EGF-activated MAP kinase signaling pathway by adenosine 3′,5′-monophosphate. *Science* 1993;262:1065–1069.

49. Cook SJ, McCormick F. Inhibition by cAMP of ras-dependent activation by raf. *Science* 1993;262:1069–1072

50. Lev S, Moreno H, Martinez R, et al. Protein tyrosine kinase PYK2 involved in Ca(2+)-induced regulation of ion channel and MAP kinase functions. *Nature* 1995;376:737–745.

51. Lin LF, Doherty DH, Lile JD, et al. GDNF: a glial cell line-derived neurotrophic factor for midbrain dopaminergic neurons. *Science* 1993;260:1130.

52. Shen L, Figurov A, Lu B. Recent progress in studies of neurotrophic factors and their clinical implications. *J Mol Med* 1997;75:637–644.

53. Gash DM, Zhang Z, Gerhardt G. Neuroprotective and neurorestorative properties of GDNF. *Ann Neurol* 1998;44:S121–S125.

54. Treanor JJ, Goodman L, de Sauvage F, et al. Characterization of a multicomponent receptor for GDNF. *Nature* 1996;382:80–83.

55. van Weering DH, Bos JL. Signal transduction by the receptor tyrosine kinase Ret. *Recent Results Cancer Res* 1998;154:271–281.

56. Stahl N, Yancopoulos GD. The tripartite CNTF receptor complex: activation and signaling involves shared with other cytokines. *J Neurobiol* 1994;25:1454–1466.

57. Cattaneo E, Conti L, De-Fraja C. Signalling through the JAK-STAT pathway in the developing brain. *Trends Neurosci* 1999;22:365–369.

58. Adamo M, Raizada MK, LeRoith D. Insulin and insulin-like growth factor receptors in the nervous system. *Mol Neurobiol* 1989;3:71–100.

59. Arsenijevic Y, Weiss S. Insulin-like growth factor-I is a differentiation factor for postmitotic CNS stem cell-derived neuronal precursors: distinct actions from those of brain-derived neurotrophic factor. *J Neurosci* 1998;18:2118–2128.

60. Yamada M, Ikeuchi T, Hatanaka H. The neurotrophic action and signalling of epidermal growth factor. *Prog Neurobiol* 1997;51:19–37.

61. Sapolsky RM. Stress, glucocorticoids, and damage to the nervous system: the current state of confusion. *Stress* 1996;1:1–19.

62. McEwen BS. Stress and hippocampal plasticity. *Annu Rev Neurosci* 1999;22:105–122.

63. Duman RS, Malberg J, Thome J. Neural plasticity to stress and antidepressant treatment. *Biol Psychiatry* 1999;46:1181–1191.

64. Smith MA, Makino S, Kvetnansky R, et al. Stress alters the expression of brain-derived neurotrophic factor and neurotrophin-3 mRNAs in the hippocampus. *J Neurosci* 1995;15:1768–1777.

65. Nibuya M, Morinobu S, Duman RS. Regulation of BDNF and trkB mRNA in rat brain by chronic electroconvulsive seizure and antidepressant drug treatments. *J Neurosci* 1995;15:7539–7547.

66. Nibuya M, Nestler EJ, Duman RS. Chronic antidepressant administration increases the expression of cAMP response element binding protein (CREB) in rat hippocampus. *J Neurosci* 1996;16:2365–2372.

67. Siucak JA, Lewis DR, Wiegaud SJ, et al. Antidepressant-like effect of brain derived neurotrophic factor (BDNF). *Pharmacol Biochem Beh*; 56:121–127.

68. Shirayama Y, Chen ACH, Duman RS. Antidepressant-like effects of BDNF and NT-3 in behavioral models of depression. *Soc Neurosci Abstr* 2000;387:2.

69. Nestler EJ. Molecular mechanisms of opiate and cocaine addiction. *Curr Opin Neurobiol* 1997;7:713–719.

70. Koob GF, Nestler EJ. The neurobiology of addiction. *J Neuropsychiatry Clin Neurosci* 1997;9:482–497.

71. Berhow MT, Russell DS, Terwilliger RZ, et al. Influence of neurotrophic factors on morphine- and cocaine-induced biochemical changes in the mesolimbic dopamine system. *Neuroscience* 1995;68:969–979.

72. Messer CJ, Eisch AJ, Carlezon WA Jr, et al. Role for GDNF in biochemical and behavioral adaptations to drugs of abuse. *Neuron* 2000;26:247–257.

73. Pierce RC, Pierce-Bancroft AF, Prasad BM. Neurotrophin-3 contributes to the initiation of behavioral sensitization to cocaine by activating the ras/mitogen-activated protein kinase signal transduction cascade. *J Neurosci* 1999;19: 8685–8695.

74. Ortiz J, Harris HW, Guitart X, et al. Extracellular signal-regulated protein kinases (ERKs) and ERK kinase (MEK) in brain: regional distribution and regulation by chronic morphine. *J Neurosci* 1995;15:1285–1297.

75. Berhow MT, Hiroi N, Nestler EJ. Regulation of ERK (extracellular signal regulated kinase), part of the neurotrophin signal transduction cascade, in the rat mesolimbic dopamine system by chronic exposure to morphine or cocaine. *J Neurosci* 1996;16:4707–4715.

76. Wolf DH, Numan S, Nestler EJ, et al. Regulation of phospholipase C-gamma in the mesolimbic dopamine system by chronic morphine administration. *J Neurochem* 1999;73:1520–1528.

17
Regulation of Gene Expression

Eric J. Nestler

Steven E. Hyman

Eric J. Nestler: Department of Psychiatry, The University of Texas Southwestern Medical Center, Dallas, Texas.

Steven E. Hyman: National Institute of Mental Health, Bethesda, Maryland.

For all living cells, regulation of gene expression by extracellular signals is a fundamental mechanism of development, homeostasis, and adaptation to the environment. Indeed, the ultimate step in many signal transduction pathways is the modification of transcription factors that can alter the expression of specific genes. Thus, neurotransmitters, growth factors, and drugs are all capable of altering the patterns of gene expression in a cell. Such transcriptional regulation plays many important roles in nervous system functioning, including the formation of long-term memories. For many drugs, which require prolonged administration for their clinical effects (e.g., antidepressants, antipsychotics), the altered pattern of gene expression represents therapeutic adaptations to the initial acute action of the drug.

Mechanisms that underlie the control of gene expression are becoming increasingly well understood. Every conceivable step in the process is subject to dynamic regulation in the cell. This includes structural changes in the chromatin to make a particular gene accessible for transcription, transcription of DNA into RNA, splicing of RNA into mRNA, editing and other covalent modifications of the mRNA, translation of mRNA into protein, and, finally, post-translational modification of the protein into its mature, functional form.

Molecular details of each of these regulatory steps are becoming increasingly available. In this chapter, we focus on the regulation of gene expression by transcription factors because their role in mediating the ability of extracellular signals to alter gene expression remains the best characterized.

OVERVIEW OF TRANSCRIPTIONAL CONTROL MECHANISMS

Part of "17 - Regulation of Gene Expression "

Regulation of Gene Expression by the Structure of Chromatin

In eukaryotic cells, DNA is contained within a discrete organelle called the nucleus, which is the site of DNA replication and transcription. Within the nucleus, chromosomes—which are extremely long molecules of DNA—are wrapped around histone proteins to form nucleosomes, the major subunits of chromatin (1 ,2 and 3). To fit within the nucleus, much of the DNA is tightly packed into a “coiled coil.” Compared with transcriptionally quiescent regions, actively transcribed regions of DNA may be more than 1,000-fold further extended. Chromatin does not just serve a structural role, however; in eukaryotes, chromatin plays a critical role in transcriptional regulation. Chromatin can inhibit access of transcription factors to the DNA and can thereby repress gene expression. In eukaryotic organisms, with their very large number of genes (approximately 40×10^3 in mammals), this means that the ground state of gene expression is for genes to be turned off. Activation of gene expression requires that cells alleviate nucleosome-mediated repression of an appropriate subset of genes. This is accomplished by means of activator proteins that modify chromatin structure. The activation process, which involves transcription factors, along with histones and cofactors, displaces or remodels chromatin, and opens up regions of the DNA, including the core promoters (see later) of genes, for the binding of regulatory proteins.

Transcription occurs when particular activator proteins displace nucleosomes. This permits a complex of proteins (described later) called *general transcription factors*, to bind DNA at a particular type of element, called a *core promoter*, and to recruit RNA polymerase. The construction of this protein complex at the transcription start site and the synthesis of the first phosphodiester bond between nucleotides are referred to as *transcription initiation* (3). The RNA polymerase must successfully transcribe an appropriate length of RNA without premature termination *(elongation)*. Premature

termination appears to be a regulated mechanism that controls expression of a small number of genes. Transcription of the RNA must also terminate appropriately *(termination)*.

Transcription Initiation: A Critical Biological Control Point

As described in the preceding section, transcription can be divided into three discrete steps: initiation, mRNA chain elongation, and chain termination. Although biologically significant regulation may occur at any step in the process, transcription initiation appears to be one of the most significant control points that gates the flow of information out of the genome. Certainly, as far as we know now, transcription initiation is the step in gene expression that is most highly regulated by extracellular signals (3).

Transcription initiation involves two critical processes: positioning of the appropriate RNA polymerase at the correct start sites of transcription and controlling the efficiency of initiations to produce the appropriate transcriptional rate for the circumstances of the cell. These control functions depend on regulatory elements that recruit appropriate transcription factors to the DNA (Fig. 17.1). Many transcription factors bind DNA directly; others interact indirectly through protein–protein interactions with factors that do bind DNA themselves. Those regulatory elements that set the transcription start sites of a gene are called the *basal or core promoters*. Other regulatory elements tether additional activator and repressor proteins to the DNA.

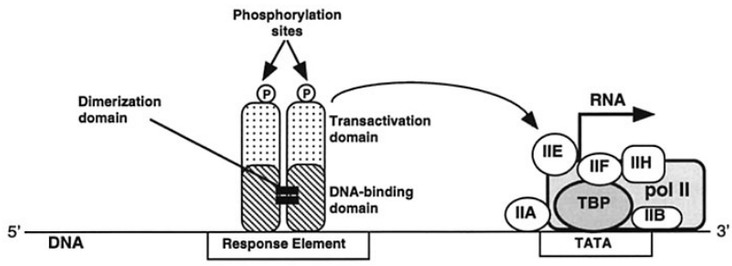


FIGURE 17.1. Scheme of a generalized polymerase II promoter. The figure shows two regulatory elements *(open rectangles)* along the stretch of DNA *(thin black line)*. These include the TATA element and a hypothetic activator or response element. The TATA element is shown binding the TATA-binding protein, TBP. Multiple general transcription factors and RNA polymerase II (pol II) associate with TBP. The general transcription factors are referred to with the nomenclature of TFII(x), for transcription factors of a pol II promoter. Shown are general factors, TFIIA, B, E, F, and H. Each of these transcription "factors" is actually composed of multiple individual proteins complexed together. TBP and its associated proteins are collectively called TFIID. This basal transcription apparatus recruits RNA polymerase II. It also forms the substrate for interactions with various activator proteins that bind to activator elements such as the one shown. Typical activator proteins contain DNA-binding domains, dimerization domains, and transcription activation domains. The latter interact with the basal transcription apparatus and may be modified by phosphorylation. Adapted from reference 14.

Core Promoters: Setting the Start Site and Direction of Transcription

In eukaryotes, transcription is carried out by three distinct RNA polymerases: RNA polymerases I, II, and III (4). These three polymerases interact with different classes of genes, each of which contains distinct promoter elements. Polymerase I (pol I) promoters are used by genes that encode large rRNAs (ribosomal RNAs). Polymerase II (pol II) promoters are used by genes that are transcribed to yield mRNAs and hence proteins. Pol II promoters are also used by a subset of the genes that encode snRNAs that are involved in RNA splicing. Polymerase III (pol III) promoters are used by genes that encode other small RNAs, including the remaining snRNAs, small rRNAs, and tRNAs (transfer RNAs).

None of the RNA polymerases bind DNA directly; rather, the polymerases are recruited to the DNA by other proteins. The core promoters for each of the three polymerases contain distinct elements on which different types of basal transcription complexes are assembled, each using different transcription factors. Because the main focus of this chapter is regulated expression of protein-encoding genes, only transcription by pol II is described.

The core promoters of genes transcribed by pol II are surprisingly diverse, but they share certain key features. By far the most common core promoter element for pol II promoters is the *TATA box* (Fig. 17.1), a sequence rich in the nucleotides A and T located between 25 and 30 bases upstream of the transcription start site. In TATA box–containing genes, mutation of this sequence can inhibit transcription

initiation or make it inaccurate. In addition to setting the start site of transcription, the TATA box sets the orientation of the basal transcription complex and therefore the 5′ to 3′ direction in which pol II synthesizes the RNA. Many pol II promoters (including those for many neurally expressed genes) lack a TATA box; in these cases, a poorly conserved core promoter element called an *initiator* is found.

The TATA-binding protein (TBP) initiates the formation of the basal transcription complex along with multiple TBP-associated proteins (TAFs) and multiple additional general transcription factors (Fig. 17.1). Each of the transcription factors represented in Fig. 17.1 was originally identified as a chromatographic fraction derived from cell nuclei, and it is a mixture of proteins. Thus, TBP together with its TAFs was originally identified as a fraction called TFIID, where TFII is a nomenclature identifying general transcription factors associated with pol II, and the final letter designates the fraction. TFIID, but not TBP by itself, is required to build a basal transcription complex from TATA-less promoters.

Transcription Factors: Key Regulators of Gene Expression

The basal transcription apparatus is not adequate to initiate more than low levels of transcription. To achieve significant levels of transcription, this multiprotein assembly requires help from additional transcriptional activators that recognize and bind to regulatory elements found elsewhere within the gene. Because they are tethered to DNA—by their binding to specific recognition sequences in the DNA—such proteins can be described as sequence-specific transcription factors (5 ,6 and 7).

Functional regulatory elements are generally found within several hundred bases of the start site of the gene to which they are linked, but they can occasionally be found many thousands of base pairs (bp) away, either upstream or downstream of the start site. Regulatory elements that exert control near the core promoter itself have been called *promoter elements*, and those that act at a distance have been called *enhancer elements*, but the distinction between promoter and enhancer elements is artificial from a mechanistic point of view. Both are generally composed of small, modular elements (generally 7 to 12 bp in length), each of which is a specific binding site for one or more transcription factors. The fundamental logic of transcriptional regulation in eukaryotes is that it is combinatorial: each gene has a particular combination of regulatory elements, the nature, number, and spatial arrangement of which determines the gene's unique pattern of expression. These promoter or enhancer elements control the cell types in which the gene is expressed, the times during development in which it is expressed, and the level at which it is expressed in adults both basally and in response to physiologic and environmental signals (7).

Sequence-specific transcription factors typically contain several physically distinct functional domains (these are shown in Fig. 17.1): (a) the DNA-binding domain recognizes and binds to a specific nucleotide sequence (i.e., response element); (b) the transcription activation domain interacts with coactivators or with general transcription factors (i.e., components of the pol II complex) to form a mature or fully active transcription complex; and (c) the multimerization domain permits the formation of homomultimers and heteromultimers with other transcription factors. The modularity of these proteins is emphasized by the finding that particular binding domains, activation domains, and interaction domains are used in different combinations in many naturally occurring proteins. Experimentally, domains can be swapped from different activators to produce novel hybrid proteins that are functionally active.

Many transcription factors are active only as dimers or higher-order complexes. Multimerization domains are diverse and include so-called leucine zippers (described later), Src homology (SH-2) domains, and certain helical motifs (8 ,9 and 10). Within transcription factor dimers, whether they are homodimers or heterodimers, both partners commonly contribute jointly to both the DNA binding domain and to the activation domain. In some cases, dimerization can be a mechanism of negative control of transcription. This is illustrated by the CREB (cyclic adenosine monophosphate [cAMP]–response element binding protein) family of transcription factors discussed later.

Regulation of transcription factors by the formation of heterodimers is not an “all or none” proposition, however. Within the Fos family of transcription factors (described later), some family members, such as c-Fos, are strong activators when dimerized with a partner from the Jun family, such as c-Jun. Other Fos-related proteins, such as Fra1 (Fos-related antigen-1), bind DNA as heterodimers with c-Jun, and they may still activate transcription, but at lower levels than c-Fos (11). Overall, the ability of transcription factors to form heterodimers and other multimers increases the diversity of transcription factor complexes that can form in cells and, as a result, increases the types of specific regulatory information that can be exerted on gene expression.

Sequence-specific transcriptional activator and repressor proteins may contact several proteins within the basal transcription complex directly. In other cases, they interact with the basal transcription apparatus through the mediation of coactivator or adapter proteins. In either of these situations, transcription factors that bind at a distance from the core promoter can still interact with the basal transcription apparatus, because the DNA forms loops that bring distant regions in contact with each other.

Many activator proteins become involved only in the assembly of the mature transcription apparatus after modification, most commonly phosphorylation, that occurs in response

to extracellular signals. An important effect of many phosphorylation events is to alter the ability of the phosphoprotein to interact with other proteins. This is illustrated by CREB, which can activate transcription only when phosphorylated on a particular serine residue (ser133) (12). As seen later, phosphorylation of ser133 permits CREB to interact with an adapter protein, CBP (CREB-binding protein), which, in turn, contacts and activates the basal transcription apparatus (13).

REGULATION OF GENE EXPRESSION BY EXTRACELLULAR SIGNALS

Part of "17 - Regulation of Gene Expression "

Transcription Factors: Targets of Signaling Pathways

Most genes probably contain response elements that confer responsiveness to physiologic signals. Response elements work by binding transcription factors that are activated (or inhibited) by specific physiologic signals, of which the most common is *phosphorylation*. Two general mechanisms of transcriptional regulation by extracellular signals are illustrated Fig. 17.2 (14). In one mechanism, transcription factors that are present at significant levels in cells under basal conditions are rapidly activated by signaling cascades to activate or repress transcription of responsive target genes. In the other major mechanism, transcription factors that are expressed at very low levels under basal conditions are themselves induced by a physiologic signal, after which they can regulate expression of a series of additional genes.

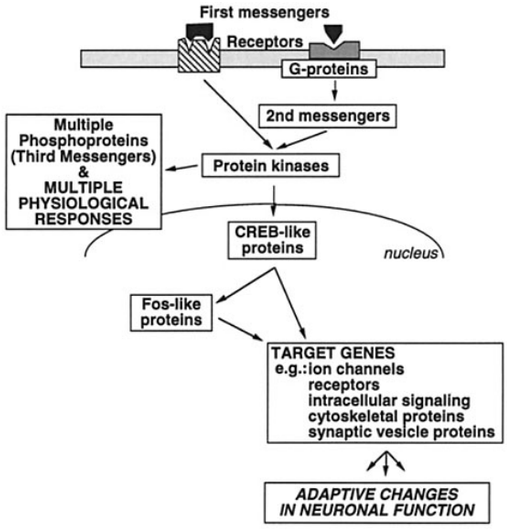


FIGURE 17.2. Scheme of intracellular pathways underlying regulation of gene expression. Activation of neurotransmitter, hormone, or neurotrophic factor receptors leads to the activation of specific second messenger and protein phosphorylation pathways, which produce multiple effects on neuronal function through the phosphorylation of numerous proteins. Among the effects of these intracellular pathways on neuronal function is the regulation of gene expression. This can be accomplished by two basic types of mechanisms. In one case, transcription factors, already in the nucleus, are phosphorylated by protein kinases; this alters their transcriptional activity and leads to alterations in the expression of specific target genes. CREB is an example of a transcription factor that functions in this manner. Among the target genes for CREB family transcription factors are those for other transcription factors, for example, Fos and Jun family proteins. Increased expression of Fos and Jun then leads to alterations in the expression of additional target genes.

A critical step in extracellular regulation of gene expression is the transduction of signals from the cell membrane to the nucleus; this can be accomplished by several different types of mechanisms. Some transcription factors themselves translocate to the nucleus on activation. One example is provided by the steroid hormone receptor transcription factors, discussed at length later. Another example is the transcription factor nuclear factor-κB (NF-κB) (15). This transcription factor is retained in the cytoplasm by its binding protein, IκB, which masks the nuclear localization signal within NF-κB. Signal-regulated phosphorylation of IκB by protein kinase C leads to dissociation of NF-κB, which permits it to enter the nucleus, where it can bind DNA; IκB is then proteolyzed within the cytoplasm.

Other transcription factors must be directly phosphorylated or dephosphorylated to bind DNA. For example, phosphorylation of STATs (signal transducers and activators of transcription) by protein tyrosine kinases in the cytoplasm permits their multimerization, which, in turn, permits nuclear translocation and construction of an effective DNA binding site within the multimer (16).

Still other transcription factors are already bound to their cognate *cis*-regulatory elements in the nucleus under basal conditions and are converted into transcriptional activators by phosphorylation. CREB, for example, is bound to DNA elements termed cAMP-response elements (CREs) (Fig. 17.3) before cell stimulation. The critical nuclear translocation step in CREB activation involves not the transcription factor itself, but activated protein kinases (cAMP-dependent protein kinase; also called protein kinase A) that, on entering the nucleus, phosphorylate CREB. Alternatively, CREB activation can involve the nuclear translocation of second messengers, such as Ca^{2+} bound to calmodulin, which, on entering the nucleus activate protein kinases that then phosphorylate CREB (Fig. 17.3). As stated earlier, phosphorylation converts CREB into a transcriptional activator by permitting it to recruit CBP into the transcription complex.

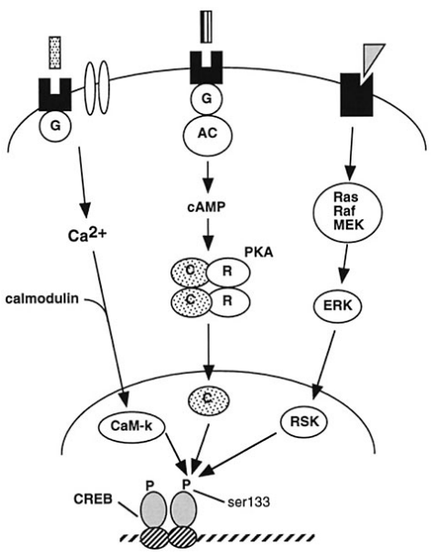


FIGURE 17.3. Scheme of the regulation of CREB phosphorylation by several signaling pathways. The figure illustrates how several signaling pathways converge on the phosphorylation of CREB at a single serine residue, ser133. Neurotransmitters that stimulate adenylyl cyclase would increase CREB phosphorylation by the activation of protein kinase A (PKA). On activation, free PKA catalytic subunits would translocate to the nucleus, where they would phosphorylate ser133 of CREB. Neurotransmitters that inhibit adenylyl cyclase would cause the opposite cascade and inhibit CREB phosphorylation. Any of several signals that increase cellular Ca^{2+} levels (e.g., certain inotropic or G-protein–coupled receptors, voltage-gated Ca^{2+} channels) would also increase CREB phosphorylation. Here, it appears that a wave of increased Ca^{2+} would permeate the nucleus, where it would activate certain Ca^{2+}/calmodulin-dependent protein kinases (CaM kinases), particularly CaM-K IV, which phosphorylates ser133 of CREB. In addition, growth factor regulated pathways lead to CREB phosphorylation, although the details are not as well established. One possibility, shown in the figure, is that activation of Ras-Raf-MEK pathways would lead to activation of ERK (a type of MAP kinase), which would translocate to the nucleus and phosphorylate and activate RSK (ribosomal S6 kinase). RSK would then phosphorylate ser133 of CREB. MEK, MAP kinase and ERK kinase; ERK, extracellular signal regulated kinase; RSK, ribosomal S6 kinase.

The remainder of this chapter provides a more in-depth discussion of several transcription factor families that have received a great deal of attention as mediators of neural and behavioral plasticity in the adult.

CREB Family of Transcription Factors

CREs were the first second messenger response element to be well characterized (12 ,17 ,18). As the name suggests, CREs confer activation by cAMP on genes to which it is linked. Subsequently, it was found that the same element confers response to Ca^{2+} and to the Ras pathway as well. CREs have been identified in many genes expressed in the nervous system, including those encoding neuropeptides (e.g., somatostatin, proenkephalin, vasoactive intestinal polypeptide), neurotransmitter synthetic enzymes (e.g., tyrosine hydroxylase), signaling proteins (e.g., adenylyl cyclase type VIII), and transcription factors (e.g., c-Fos and CREB itself) (12 ,17 ,18).

The idealized or “consensus” CRE sequence is TGACGTCA, although the actual CREs present in various genes differ slightly. The consensus CRE sequence illustrates an important principle, the palindromic nature of many transcription factor–binding sites. Examining the sequence TGACGTCA, it can be readily observed that the sequence on the two complementary DNA strands, which run in opposite directions, is identical. Many regulatory elements are perfect or approximate palindromes because many transcription factors bind DNA as dimers, with each member of the dimer recognizing one of the half-sites. The major protein that binds to CREs is CREB. CREB binds to a CRE as a homodimer, with a higher affinity for palindromic than for asymmetric CREs.

When bound to a CRE, CREB activates transcription only when it is phosphorylated on its critical ser133. It does so, as described earlier, because phosphorylated CREB, but not dephosphorylated CREB, can recruit the adapter protein, CBP, into the transcription complex. CBP, in turn, interacts with the basal transcription complex.

Regulation by cAMP, Ca^{2+}, and Growth Factors

As discussed in previous sections, the regulation of CREB activation by phosphorylation illustrates the requirement for nuclear translocation of protein kinases or second messengers when transcription factors are already found in the nucleus under basal conditions and the role of phosphorylation in regulating protein–protein interactions. An additional important principle illustrated by CREB is the convergence of signaling pathways. CREB is activated in response to activation of the cAMP or Ca^{2+} pathways. This occurs because the same serine residue (ser133) is phosphorylated both by protein kinase A and by Ca^{2+}/calmodulin–dependent protein kinases (CaM kinases) (Fig. 17.3). CaM kinase IV appears to be the most important form of the enzyme that mediates this phosphorylation (19 ,20). CREB also appears to be phosphorylated on ser133 by a growth factor–activated kinase, RSK—ribosomal S6 kinase—that is phosphorylated and activated by mitogenactivated protein (MAP) kinases (21).

Thus, diverse types of signaling pathways converge on the phosphorylation and activation of CREB. If each individual signal is relatively weak, convergence may be a critical mechanism for achieving significant gene regulation, with some genes activated only when multiple pathways are stimulated. Furthermore, some genes that contain CREs are known to be induced in a synergistic fashion by the interaction of cAMP and Ca^{2+} signaling pathways. In addition

to ser133, CREB contains other sites for phosphorylation by a variety of protein kinases, which may fine tune the regulation of CREB-mediated transcription. For example, CaM kinase II phosphorylates a distinct serine residue in CREB, which diminishes the ability of other kinases to phosphorylate ser133. Activation of CaM kinase II would therefore appear to mediate a dampening of the CREB signal (19 ,20).

Role in Neural Plasticity

The convergent activation of a single transcription factor by multiple signaling pathways is particularly important in the nervous system, because this is an important candidate mechanism for long-term neural adaptations, including those underlying long-term memory, drug addiction, and fear conditioning. As discussed elsewhere in this volume, it is reasonably well established that some forms of long-term memory require new gene expression. Furthermore, associative memory depends on the temporally coordinated arrival of two different signals, which must then be integrated within target neurons and their circuits. Activation of CREB is therefore a plausible candidate for playing an important role in long-term memory and related phenomena. Consistent with this prediction, *Drosophila* in which CREB was inactivated by a dominant negative transgene, and mice in which CREB was inactivated by homologous recombination (i.e., knockout), show deficits in long-term memory (22 ,23 ,24 and 25). Manipulation of CREB also influences long-term potentiation in the hippocampus (23 ,24 and 25) and aspects of drug addiction (26 ,27 ,28 and 29). Although much work remains to understand the precise role of CREB in these various phenomena, it does appear that CREB, which is regulated by several major neural signaling pathways, is a critical mediator of many types of plasticity.

CREB-like Proteins

CREB illustrates yet another important principle of transcriptional regulation: CREB is a member of a larger family of related proteins. Many transcription factors are members of families. CREB is closely related to proteins called the *activating transcription factors* (ATFs) and the *CRE modulators* (CREMs), each generated by distinct genes. In addition, several alternative splice forms are known for CREB, certain of the ATFs, and CREMs (30 ,31).

All these proteins bind CREs as dimers, and many can heterodimerize with CREB itself. ATF1 appears to be very similar to CREB; it can be activated by both the cAMP and Ca^{2+} signaling pathways (30 ,31). Many of the other ATF proteins and CREM isoforms also appear to activate transcription. However, certain CREMs (e.g., ICER—inducible cAMP element repressor) act to repress it (30). These CREM isoforms lack the glutamine-rich transcriptional activation domain found in CREB-ATF family proteins that are transcriptional activators. Thus, CREB-ICER heterodimers may occupy CREs, but fail to activate transcription. Like CREB, many of the ATF proteins are constitutively made in cells, but ATF3 and certain CREM isoforms (e.g., ICER) are inducible.

Leucine Zipper Dimerization Motif

The dimerization domain used by the CREB-ATF proteins and several other families of transcription factors is called a *leucine zipper* (8 ,9). This domain was first identified in the transcription factor C/EBP (CAATT-enhancing binding protein) (32), and it is also used by the AP-1 proteins, as described in detail later. The so-called leucine zipper actually forms a coiled coil. The dimerization motif is an α helix in which every seventh residue is a leucine; based on the periodicity of α helices, the leucines line up along one face of the α helix two turns apart. The aligned leucines of the two dimerization partners interact hydrophobically to stabilize the dimer. In CREB, C/EBP, and many AP-1 proteins, the leucine zipper is at the C-terminus of the protein. There is a region of highly basic amino acid residues just upstream of the leucine zipper that forms the DNA binding domain. Dimerization by the leucine zipper juxtaposes the adjacent basic regions of each of the partners; these juxtaposed basic regions undergo a conformational change when they bind DNA, which results in what has been described as a “scissors grip.” This combination of motifs has led this superfamily of proteins to be described as the basic-leucine zipper proteins or the bZIP proteins.

AP-1 Family of Transcription Factors

Another group of bZIP transcription factors that plays an important role in the regulation of neural gene expression by extracellular signals comprises the activator protein-1 (AP-1) proteins. The name AP-1 was originally applied to a transcriptional activity stimulated by protein kinase C activation (33). This activity was found to be composed of multiple proteins, which bind as heterodimers (and a few as homodimers) to the DNA sequence TGACTCA, the AP-1 sequence. The consensus AP-1 sequence is a heptamer that forms a palindrome flanking a central C or G, and differs from the CRE sequence by only a single base. Yet this one base difference strongly biases protein binding away from CREB (which requires an intact CGTCA motif) to the AP-1 family of proteins and means that, under most circumstances, this sequence will not confer cAMP responsiveness on a gene.

Many genes expressed in the nervous system contain AP-1 sites within their regulatory regions (34 ,35 and 36). Examples include genes encoding neuropeptides (neurotensin and substance P), neurotransmitter receptors (D1 dopamine, NR1 NMDA, and GluR2 AMPA glutamate receptor subunits), neurotransmitter synthetic enzymes (tyrosine hydroxylase),

and cytoskeletal proteins (neurofilament proteins), to name a few. In some cases, it has been possible to demonstrate a role for these sites in regulation of gene promoter activity *in vitro*, although it has been very difficult to identify with certainty those genes that are regulated by AP-1 transcription factors in the brain *in vivo* (35).

As alluded to earlier, the AP-1 sequence was described initially as a TPA-response element (TRE) because the phorbol ester, 12-O-tetradecanoyl-phorbol-13-acetate (TPA), which activates protein kinase C, can induce gene expression through AP-1 proteins (33). In addition, it is now thought that a major role of the AP-1 sequence is to confer responsiveness to growth factor–stimulated signaling pathways such as the Ras/MAP-kinase pathways. This occurs by phosphorylation of specific AP-1 proteins by certain MAP kinases.

AP-1 proteins generally bind DNA as heterodimers composed of one Fos family member and one Jun family member (34). Both families are bZIP proteins: they form dimers through their leucine zipper domains. The known members of the Fos family are c-Fos, Fra1, Fra2, and FosB and its alternatively spliced variant ΔFosB. The known members of the Jun family are c-Jun, JunB, and JunD. Unlike Fos proteins, c-Jun and JunD (but not JunB) can form homodimers that bind to AP-1 sites, albeit with far lower affinity than Fos-Jun heterodimers. The potential complexity of transcriptional regulation is greater still because some AP-1 proteins can heterodimerize through the leucine zipper with members of the CREB-ATF family, such as ATF2 with c-Jun. AP-1 proteins can also form higher-order complexes with apparently unrelated families of transcription factors. For example, AP-1 proteins can complex with and thereby apparently inhibit the transcriptional activity of steroid hormone receptors (see later).

Among Fos and Jun proteins, only JunD is expressed constitutively at high levels in many cell types. The other AP-1 proteins tend to be expressed at low or even undetectable levels under basal conditions, but, with stimulation, they may be induced to high levels of expression. Thus, unlike regulation by constitutively expressed transcription factors such as CREB, regulation by Fos-Jun heterodimers requires new transcription and translation of the transcription factors themselves (Fig. 17.2).

Cellular Immediate Early Genes

Genes that are transcriptionally activated by synaptic activity, drugs, or growth factors have often been classified roughly into two groups. *Immediate early genes* (IEGs), such as the c-*fos* gene itself, are activated rapidly (within minutes) and transiently and do not require new protein synthesis. *Late-response genes*, in contrast, are induced or repressed more slowly (over hours) and are dependent on new protein synthesis. The term IEG was applied initially to describe viral genes that are activated "immediately" on infection of eukaryotic cells, by commandeering host cell transcription factors for their expression. Viral IEGs generally encode transcription factors needed to activate viral "late" gene expression. This terminology has been extended to cellular (i.e., nonviral) genes, although this has created some confusion.

The terminology is problematic because there are many cellular genes induced independently of protein synthesis, but with a time course intermediate between "classic" IEGs and late-response genes. In fact, some genes may be regulated with different time courses or requirements for protein synthesis in response to different extracellular signals. Moreover, many cellular genes regulated as IEGs encode proteins that are not transcription factors: for example, any gene induced by CREB could potentially show temporal features of induction of an IEG. Despite these caveats, the concept of IEG-encoded transcription factors in the nervous system has proved useful in thinking about the complexities of gene regulation. In addition, because of their rapid induction from low basal levels in response to neuronal depolarization (the critical signal being Ca^{2+} entry) as well as various second messenger and growth factor pathways, several IEGs have been used as cellular markers of neural activation, and this has permitted novel approaches to functional neuroanatomy (37).

The protein products of those cellular IEGs that function as transcription factors bind to regulatory elements contained within a subset of late response genes to activate or repress them. IEGs such as c-fos have therefore been termed "third messengers" in signal transduction cascades, with neurotransmitters designated intercellular first messengers and small intracellular molecules, such as cAMP and Ca^{2+}, second messengers (34). There have, however, been misunderstandings among some neurobiologists that IEGs are a necessary step in the signal-induced expression of most neural genes involved in the differentiated function of neurons. In fact, as stated earlier, many genes involved in differentiated neural functions, including genes encoding certain neuropeptides and neurotrophic factors, to name a few, are activated in response to neuronal depolarization or cAMP through phosphorylation of CREB rather than through IEG third messengers.

Activation by Multiple Signaling Pathways

The most studied cellular immediate early gene is c-*fos*. The c-*fos* gene contains three binding sites for CREB (the strongest of which is shown in Fig. 17.4). As a result, it is not surprising that the gene can be activated rapidly by neurotransmitters or drugs that stimulate the cAMP or Ca^{2+} pathways (38).

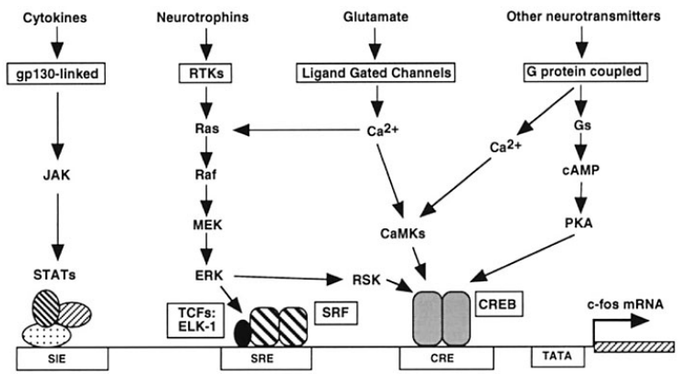


FIGURE 17.4. Scheme of the regulatory region of the c-*fos* gene. The figure shows only three of the many known transcription factor binding sites within this gene. These sites are as follows: a CRE (cAMP-response element), which binds CREB; a serum-response element (SRE), which binds serum response factor (SRF) and Elk-1 (also called the ternary complex factor or TCF); and an SIF-inducible element (SIE), which binds STAT proteins (signal transducers and activators of transcription). Proteins binding at each of these sites are constitutively present in cells and are activated by phosphorylation. CREB can be activated by protein kinase A, CaM kinases (CaM-Ks), or RSKs (ribosomal S6 kinases) (Fig. 17.3); Elk-1 can be activated by the MAP kinases ERK1 and ERK2 (extracellular signal regulated kinases 1 and 2); and the STAT proteins can be activated by the JAK protein tyrosine kinases. Thus, activation of the c-*fos* gene—by any of multiple signaling pathways—depends only on signal-induced phosphorylation rather than on new protein synthesis. This explains the rapidity of its induction by a wide array of stimuli. MEK, MAP kinase and ERK kinase; PKA, protein kinase A; RTKs, receptor tyrosine kinases. Adapted from reference 14.

The c-*fos* gene also can be induced by the Ras/MAP-kinase pathway (39 ,40). Neurotrophins, such as nerve growth factor, bind a family of *receptor tyrosine kinases* (called Trks) that activate Ras. Ras then triggers a cascade

of protein kinases, which results in the phosphorylation and activation of certain MAP kinases called *extracellular signal regulated kinases* (ERKs). These ERKs can phosphorylate and activate additional protein kinases, such as RSK, which, among its other substrates, can phosphorylate ser133 of CREB, as described earlier. However, an additional mechanism exists whereby ERK can induce the c-*fos* gene, and this mechanism appears to predominate in many cell types (41). Here, the ERKs translocate into the nucleus where they phosphorylate the transcription factor Elk-1 (also called the ternary complex factor or TCF). Elk-1 then complexes with the serum response factor (SRF) to bind to and activate the serum response element (SRE) within the c-*fos* gene (Fig. 17.4). SREs are present within many other growth factor-inducible genes as well. In comparison with cAMP- or Ca^{2+}-dependent phosphorylation of CREB, the Ras/MAP kinase pathway depends on a complex chain of phosphorylation events. Nonetheless, these events can occur very rapidly to induce gene expression.

Still another mechanism exists for c-*fos* induction: cytokine-activated signaling pathways that act through STATs (42). As stated earlier, STATs are activated on their phosphorylation by certain protein tyrosine kinases. This permits STATs to form multimeric complexes, translocate to the nucleus, and bind to their specific DNA response elements, generally now described at STAT sites. However, the STAT site in c-*fos* had already been named the SIE or SIF-inducible element (SIF itself is an acronym for sis-inducible factor, i.e., a factor induced by the oncogene v-*sis*, which activated c-*fos* though this site). STATs are activated by the class of cytokines that interact with gp130-linked receptors, which include ciliary neurotrophic factor, LIF (leukemia inhibitory factor), interleukin-6, leptin, and prolactin, to name a few (16 ,43). These receptors activate a cytoplasmic protein tyrosine kinase called JAK (Janus kinase), which then phosphorylates the STATs. As shown in Fig. 17.4 , the c-*fos* gene contains an SIE, which binds STAT proteins and mediates the induction of c-Fos by cytokines.

Most other Fos and Jun family proteins are also induced rapidly in response to diverse acute stimuli and, presumably, many of the same mechanisms operate for the genes encoding these proteins. However, the response elements within these genes are not as well characterized as are those for c-*fos*, and further research is needed to understand their regulation.

Regulation by Phosphorylation

Several AP-1 proteins are regulated at the post-translational level by phosphorylation. The best-established example is

the phosphorylation of c-Jun, which provides a critical mechanism of regulation of AP-1-mediated signaling (44 ,45). c-Jun phosphorylation occurs in response to activation of a MAP kinase–related signaling pathway that is activated by many forms of cellular stress. In this pathway, a Ras-like G protein is activated by any of several insults (e.g., ultraviolet radiation, osmotic stress, toxins, certain cytokines); this triggers a cascade of protein kinases analogous to that triggered by Ras and the neurotrophins outlined earlier. The culmination of this pathway is the phosphorylation and activation of certain MAP kinases called SAP kinases (stress-activated protein kinases) or alternatively JNKs (for Jun N-terminal kinases). JNKs phosphorylate c-Jun on serines 63 and 73 in its transcriptional activation domain and increase the ability of c-Jun to activate transcription. Phosphorylation and activation of c-Jun have been implicated in the regulation of apoptosis (programmed cell death) pathways (45).

Generation of Unique AP-1 Complexes by Repeated Stimulation

After acute stimulation of cells, different members of the Fos family are induced with varying time courses of expression, which leads to a progression of distinct AP-1 complexes over time (46). Thus, under resting conditions, c-*fos* mRNA and protein are barely detectable in most neurons, but c-*fos* expression can be induced dramatically in response to numerous stimuli. As just one example, experimental induction of a grand mal seizure causes marked increases in levels of c-*fos* mRNA in brain within 30 minutes and induces substantial levels of c-Fos protein within 2 hours. c-Fos is highly unstable and is degraded back to low, basal levels within 4 to 6 hours (46). Administration of cocaine or amphetamine causes a similar pattern of c-*fos* expression in striatum (47 ,48). In either of these stimulus paradigms, other Fos-like proteins are also induced, but with a longer temporal latency than c-Fos; their peak levels of expression lag behind c-Fos by approximately 1 to 2 hours. Moreover, expression of these proteins persists a bit longer than c-Fos, but it still returns to basal levels within 8 to 12 hours.

With repeated stimulation, however, the c-*fos* gene, and to an extent the genes for other Fos-like proteins, become refractory to further activation (i.e., their expression becomes desensitized) (49 ,50). Instead, other Fos-like proteins continue to be expressed. These proteins, originally termed chronic Fras (51 ,52 and 53), are now known to be biochemically modified isoforms of ΔFosB, which exhibit very long half-lives in brain (54 ,55 and 56). As a result, these proteins accumulate in specific neurons in response to repeated perturbations and persist long after cessation of these perturbations (Fig. 17.5). Although the precise physiologic significance of these stable ΔFosB isoforms remains unknown, there is now direct evidence that ΔFosB plays an important role in aspects of drug addiction (57) and in mediating various types of striatal-based movement disorders (e.g., see refs. 58 ,59 and 60). More generally, ΔFosB may function as a sustained molecular switch that gradually converts an acute response to a long-lived adaptation in the brain (61).

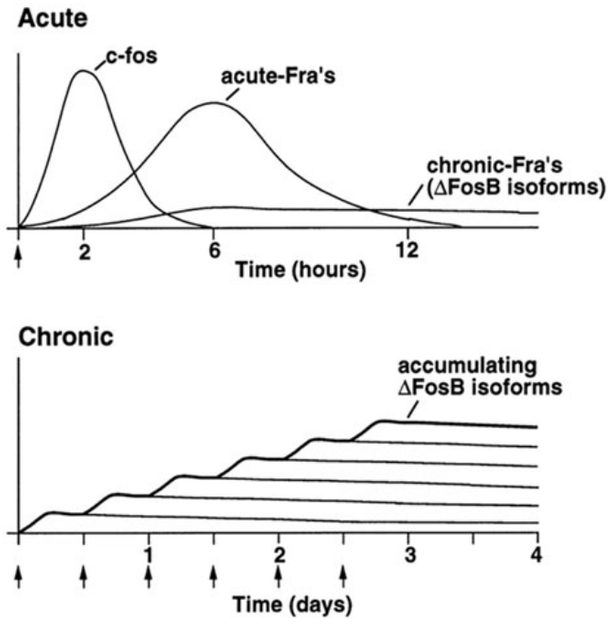


FIGURE 17.5. Scheme showing the composition of AP-1 complexes changing over time. **Top:** There are several waves of Fras (Fos-related antigens) induced by many acute stimuli in neurons. C-Fos is induced rapidly and degraded within several hours of the acute stimulus, whereas other "acute Fras" (e.g., FosB, ΔFosB, Fra1, Fra2) are induced somewhat later and persist somewhat longer than c-Fos. The "chronic Fras" are biochemically modified isoforms of ΔFosB; they, too, are induced (although at low levels) after a single acute stimulus but persist in brain because of their enhanced stability. In complexes with Jun family proteins, these waves of Fras form AP-1 binding complexes with shifting composition over time. **Bottom:** With repeated stimulation, each acute stimulus induces a low level of ΔFosB isoforms. This is indicated by the lower set of overlapping lines, which indicate ΔFosB-induced by each acute stimulus. The result is a gradual increase in the total levels of ΔFosB with repeated stimuli during a course of long-term treatment. This is indicated by the increasing *stepped line* in the graph. The increasing levels of ΔFosB with repeated stimulation would result in the gradual induction of significant levels of a long-lasting AP-1 complex, which could underlie persisting forms of neural plasticity in the brain. (Adapted from reference 53.)

Steroid Hormone Receptor (or Nuclear Receptor) Superfamily

The steroid hormones (e.g., glucocorticoids, gonadal steroids such as estrogen and testosterone, and mineralocorticoids), retinoids, thyroid hormones, and vitamin D_3 are small, lipid-soluble ligands that diffuse readily across cell

membranes. Unlike amino acid–derived neurotransmitters and neuropeptides, their receptors are cytoplasmic, rather than localized to the cell membrane. On ligand binding, these receptors translocate to the nucleus, whereupon they bind to specific hormone-response elements (HREs) located in the regulatory regions of specific genes and thereby regulate expression of those genes. These receptors are referred to as the *steroid hormone receptor, or nuclear receptor, superfamily* (62 ,63 and 64).

The glucocorticoid receptor (GR) exemplifies general mechanisms utilized by this superfamily (62 ,63 ,64 and 65). Under basal conditions, the GR is retained in the cytoplasm by a large multiprotein complex of chaperone proteins, including the heat shock protein Hsp90 and the immunophilin Hsp56. When bound by glucocorticoid, the GR dissociates from its chaperones and translocates to the nucleus. The first activity of the GR to be identified was its function as a ligand-regulated transcription factor, as stated earlier, by binding to glucocorticoid response elements (GREs). GREs are typically 15 bases in length; they consist of two palindromic half-sites of six bases each separated by a 3-bp spacer. As described earlier for other transcription factors, this palindromic organization of the GRE suggests that the GR binds as a dimer. Like many other transcription factors, the nuclear receptor superfamily has a modular structure. The GR has three domains: an N-terminal transcriptional activation domain, a C-terminal ligand binding domain, and an intervening DNA binding domain. The DNA-binding domain of the GR is characterized by a zinc finger motif, in which multiple cysteines are organized around a central zinc ion. This type of DNA binding domain is used by many other transcription factors, including the immediate early gene zif268/egr1 (see later). The DNA binding domain also contains a region that permits dimer formation after GR monomers bind GRE half-sites.

GREs can confer either positive or negative regulation on genes to which they are linked, depending on the particular GRE involved (62 ,63 ,64 and 65). One of the first positive GREs characterized is that within the metallothionein IIA gene, which encodes a protein that chelates heavy metals. A well-characterized negative GRE is found within the proopiomelanocortin *(POMC)* gene. This negative GRE permits glucocorticoids to repress the gene that encodes adrenocorticotropic hormone and is therefore an important mechanism of feedback inhibition on further glucocorticoid synthesis.

GRs have many important physiologic actions that do not appear to be mediated by DNA binding. GRs can interfere with transcriptional activity mediated by other transcription factors, particularly AP-1 and NF-κB. Although the mechanisms are not fully understood, GRs appear to interact directly with AP-1 and NF-κB proteins to block their ability to activate transcription (62 ,63 ,64 and 65). An alternative mechanism by which glucocorticoids may interfere with NF-κB activity is by inducing expression of IκB, the protein that holds NF-κB in the cytoplasm.

Other Transcription Factors in Neural Signaling

The CREB, AP-1, STATs, and steroid hormone receptor families are just a few of the literally hundreds, perhaps thousands, of transcription factors that are expressed in neurons and glia. Most of these other factors are best understood with respect to their roles in nonnervous tissues, although more recent work implicates them in neural signaling as well. Examples of these other transcription factors have already been mentioned. NF-κB is activated by protein kinase C and immunologic signals and likely plays an important role in the regulation of neural gene expression (15). C/EBP and its several family members are known to mediate some of the effects of the cAMP pathway on gene expression and have been implicated in neural plasticity (32 ,66). Specific protein-1 (SP-1), which binds to GC rich regions of promoters, is often thought of as a general transcription factor, that is, a regulator of the basal rate of transcription. However, more recent research has shown that certain SP-1 family members are subject to dynamic regulation and could mediate transcriptional changes induced by extracellular signals (5). AP-2 (activator protein-2) binding sites are present in many neural-expressed genes, although we still know very little about its physiologic role (5). Zif268 (also called Egr-1) and related Egr family members are zinc finger transcription factors that, like c-Fos, are induced rapidly and transiently in brain by many stimuli with temporal features of IEGs. Induction of Egr family proteins has been correlated with induction of hippocampal long-term potentiation; however, their specific target genes remain poorly characterized (67).

CONCLUSIONS

Part of "17 - Regulation of Gene Expression "

Our discussion of nuclear signaling mechanisms highlights several important points. The first is that the potential number of mechanisms by which the expression of a gene can be controlled is vast. This highlights the exquisite control over gene expression that is required both for the generation of a diversity of cell types during development and for adaptation of cells to the environment throughout life.

We devoted most attention to nuclear transcription factors, because these provide the best-understood mechanisms of how cells adapt to external cues with alterations in gene expression. However, even the complexity of mechanisms discussed represents the tip of the iceberg. Regulatory regions of genes are often far longer than the coding regions of genes. Regulatory information is contained not only within the 5′ promoter regions of genes, but throughout intronic (and sometimes exonic) sequences as well as 3′ untranscribed regions. Within the 5′ regions, we focused on relatively small response elements, such as CRE and AP-1 sites. It is extraordinary, indeed, that the difference of 1

nucleotide (e.g., from a CRE to an AP-1 site) in a sequence of literally thousands can confer specificity on a gene with respect to its regulation. Nonetheless, we know that any given gene likely contains many regulatory sites. Moreover, these sites do not function in isolation, but they influence one another. As a result, the rate of expression of a given gene represents the temporal and spatial synthesis of multiple signaling pathways. Unraveling these layers of complexity is a daunting task, particularly *in vivo*, but it could hold important clues for understanding neural and behavioral plasticity.

ACKNOWLEDGMENTS

Part of "17 - Regulation of Gene Expression "

This chapter is based on a more extensive, earlier chapter (see reference 14).

REFERENCES

1. Workman JL, Kingston RE. Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu Rev Biochem* 1998;67:545–579.

2. Struhl K. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* 1999;98:1–4.

3. Brown CE, Lechner T, Howe L, et al. The many HATs of transcription coactivators. *Trends Biochem Sci* 2000;25:15–19.

4. Asturias FJ, Kornberg RD. Protein crystallization on lipid layers and structure determination of the RNA polymerase II transcription initiation complex. *J Biol Chem* 1999;274:6813–6816.

5. Herdegen T, Leah JD. Inducible and constitutive transcription factors in the mammalian nervous system: control of gene expression by Jun, Fos and Krox, and CREB/ATF proteins. *Brain Res Rev* 1998;28:370–490.

6. Fickett JW, Wasserman WW. Discovery and modeling of transcriptional regulatory regions. *Curr Opin Biotechnol* 2000;11:19–24.

7. Collingwood TN, Urnov FD, Wolffe AP. Nuclear receptors: coactivators, corepressors and chromatin remodeling in the control of transcription. *J Mol Endocrinol* 1999;23:255–275.

8. Luscher B, Larsson LG. The basic region/helix-loop-helix/leucine zipper domain of Myc proto-oncoproteins: function and regulation. *Oncogene* 1999;18:2955–2966.

9. Hagerman PJ. Do basic region-leucine zipper proteins bend their DNA targets: does it matter? *Proc Natl Acad Sci USA* 1996;93:9993—9996.

10. Beattie J. SH2 domain protein interaction and possibilities for pharmacological intervention. *Cell Signal* 1996;8:75–86.

11. Cohen DR, Ferreira PC, Gentz R, et al. The product of a fos-related gene, fra-1, binds cooperatively to the AP-1 site with Jun: transcription factor AP-1 is comprised of multiple protein complexes. *Genes Dev* 1989;3:173–184.

12. Frank DA, Greenberg ME. CREB: a mediator of long-term memory from mollusks to mammals. *Cell* 1994;79:5–8.

13. Hardingham GE, Chawla S, Cruzalegui SH, et al. Control of recruitment and transcription-activating function of CBP determines gene regulation by NMDA receptors and L-type calcium channels. *Neuron* 1999;22:789–798.

14. Nestler EJ, Hyman SE, Malenta RC. *Molecular basis of neuropharmacology: a foundation for clinical neuroscience*. New York: McGraw-Hill, 2001.

15. Baldwin AS Jr. The NF-kappa B and I kappa B proteins: new discoveries and insights. *Annu Rev Immunol* 1996;14:649–683.

16. Chatterjee-Kishore M, van den Akker F, Stark GR. Association of STATs with relatives and friends. *Trends Cell Biol* 2000;10:106–111.

17. Montminy M. Transcriptional regulation by cyclic AMP. *Annu Rev Biochem* 1997;66:807–822.

18. De Cesare D, Sassone-Corsi P. Transcriptional regulation by cyclic AMP–responsive factors. *Prog Nucleic Acid Res Mol Biol* 2000;64:343–369.

19. Finkbeiner S, Greenberg ME. Ca^{2+} channel-regulated neuronal gene expression. *J Neurobiol* 1998;37:171–189.

20. Soderling TR. The Ca-calmodulin–dependent protein kinase cascade. *Trends Biochem Sci* 1999;24:232–236.

21. Bonni A, Greenberg ME. Neurotrophin regulation of gene expression. *Can J Neurol Sci* 1997;24:272–283.

22. Belvin MP, Yin JC. *Drosophila* learning and memory: recent progress and new approaches. *Bioessays* 1997;19:1083–1089.

23. Bito H, Deisseroth K, Tsien RW. CREB phosphorylation and dephosphorylation: a $Ca^{(2+)}$- and stimulus duration–dependent switch for hippocampal gene expression. *Cell* 1996;87:1203–1214.

24. Silva AJ, Kogan JH, Frankland PW, et al. CREB and memory. *Annu Rev Neurosci* 1998;21:127–148.

25. Mayford M, Kandel ER. Genetic approaches to memory storage. *Trends Genet* 1999;15:463–470.

26. Blendy JA, Maldonado R. Genetic analysis of drug addiction: the role of cAMP response element binding protein. *J Mol Med* 1998;76:104–110.

27. Cole RL, Konradi C, Douglass J, et al. Neuronal adaptation to amphetamine and dopamine: molecular mechanisms of prodynorphin gene regulation in rat striatum. *Neuron* 1995;14:813–823.

28. Lane-Ladd SB, Pineda J, Boundy V, et al. CREB in the locus coeruleus: biochemical, physiological, and behavioral evidence for a role in opiate dependence. *J Neurosci* 1997;17:7890–7901.

29. Carlezon WA Jr, Thome J, Olson VG, et al. Regulation of cocaine reward by CREB. *Science* 1998;282:2272–2275.

30. Sassone-Corsi P. Coupling gene expression to cAMP signalling: role of CREB and CREM. *Int J Biochem Cell Biol* 1998;30:27–38.

31. Hai T, Wolfgang CD, Marsee DK, et al. ATF3 and stress responses. *Gene Expr* 1999;7:321–335.

32. Lekstrom-Himes J, Xanthopoulos KG. Biological role of the CCAAT/enhancer-binding protein family of transcription factors. *J Biol Chem* 1998;273:28545–28548.

33. Boulikas T. Phosphorylation of transcription factors and control of the cell cycle. *Crit Rev Eukaryot Gene Expr* 1995;5:1–77.

34. Morgan JI, Curran T. Stimulus-transcription coupling in the nervous system: involvement of the inducible proto-oncogenes fos and jun. *Annu Rev Neurosci* 1991;14:421–451.

35. Hiroi N, Brown J, Ye H, et al. Essential role of the fosB gene in molecular, cellular, and behavioral actions of electroconvulsive seizures. *J Neurosci* 1998;18:6952–6962.

36. Wisdom R. AP-1: one switch for many signals. *Exp Cell Res* 1999;253:180–185.

37. Sagar SM, Sharp FR, Curran T. Expression of c-fos protein in brain: metabolic mapping at the cellular level. *Science* 1998;240:1328–1331.

38. Ahn S, Olive M, Aggarwal S, et al. A dominant-negative inhibitor of CREB reveals that it is a general mediator of stimulus-dependent transcription of c-fos. *Mol Cell Biol* 1998;18:967–977.

39. Rivera VM, Greenberg ME. Growth factor-induced gene expression: the ups and downs of c-fos regulation. *New Biol* 1990;2:751–758.

40. Karin M. The regulation of AP-1 activity by mitogen-activated protein kinases. *J Biol Chem* 1995;270:16483–16486.

41. de Belle I, Walker PR, Smith IC, et al. Identification of a multiprotein complex interacting with the c-fos serum response element. *Mol Cell Biol* 1991;11:2752–2759.

42. Leaman DW, Pisharody S, Flickinger TW, et al. Roles of JAKs in activation of STATs and stimulation of c-fos gene expression by epidermal growth factor. *Mol Cell Biol* 1996;16:369–375.

43. Ip NY. The neurotrophins and neuropoietic cytokines: two families of growth factors acting on neural and hematopoietic cells. *Ann NY Acad Sci* 1998;840:97–106.

44. Schwarzschild MA, Cole RL, Hyman SE. Glutamate, but not dopamine, stimulates stress-activated protein kinase and AP-1-mediated transcription in striatal neurons. *J Neurosci* 1997;17:3455–3466.

45. Chen YR, Tan TH. The c-Jun N-terminal kinase pathway and apoptotic signaling. *Int J Oncol* 2000;16:651–662.

46. Sonnenberg JL, Macgregor-Leon PF, Curran T, et al. Dynamic alterations occur in the levels and composition of transcription factor AP-1 complexes after seizure. *Neuron* 1989;3:359–365.

47. Graybiel AM, Moratalla R, Robertson HA. Amphetamine and cocaine induce drug-specific activation of the c-fos gene in striosome-matrix compartments and limbic subdivisions of the striatum. *Proc Natl Acad Sci USA* 1990;87:6912–6916.

48. Young ST, Porrino LJ, Iadarola MJ. Cocaine induces striatal c-fos-immunoreactive proteins via dopaminergic D1 receptors. *Proc Natl Acad Sci USA* 1991;88:1291–1295.

49. Hope B, Kosofsky B, Hyman SE, et al. Regulation of IEG expression and AP-1 binding by chronic cocaine in the rat nucleus accumbens. *Proc Natl Acad Sci USA* 1992;89:5764–5768.

50. Winston SM, Hayward MD, Nestler EJ, et al. Chronic electroconvulsive seizures down regulate expression of the c-fos proto-oncogene in rat cerebral cortex. *J Neurochem* 1990;54:1920–1925.

51. Chen J, Kelz MB, Hope BT, et al. Chronic FRAs: stable variants of ΔFosB induced in brain by chronic treatments. *J Neurosci* 1997;17:4933–4941.

52. Hope BT, Kelz MB, Duman RS, et al. Chronic electroconvulsive seizure (ECS) treatment results in expression of a long-lasting AP-1 complex in brain with altered composition and characteristics. *Neurosci* 1994;14:4318–4328.

53. Hope BT, Nye HE, Kelz MB, et al. Induction of a long-lasting AP-1 complex composed of altered Fos-like proteins in brain by chronic cocaine and other chronic treatments. *Neuron* 1994;13:1235–1244.

54. Hiroi N, Brown J, Haile C, et al. FosB mutant mice: loss of chronic cocaine induction of Fos-related proteins and heightened sensitivity to cocaine's psychomotor and rewarding effects. *Proc Natl Acad Sci USA* 1997;94:10397–10402.

55. Mandelzys A, Gruda MA, Bravo R, et al. Absence of a persistently elevated 37 kDa fos-related antigen and AP-1–like DNA-binding activity in the brains of kainic acid-treated fosB null mice. *J Neurosci* 1997;17:5407–5415.

56. Kelz MB, Chen S, Carlezon WA, et al. Expression of the transcription factor ΔFosB in the brain controls sensitivity to cocaine. *Nature* 1999;401:272–276.

57. Atkins J, Carlezon WA, Chlan J, et al. Region-specific induction of ΔFosB by repeated administration of typical versus atypical antipsychotic drugs. *Synapse* 1999;33:118–128.

58. Andersson M, Hilbertson A, Cenci MA. Striatal fosB expression is causally linked with LL-DOPA–induced abnormal involuntary movements and the associated upregulation of striatal prodynorphin mRNA in a rat model of Parkinson's disease. *Neurobiol Dis* 1999;6:461–474.

59. Crocker SJ, Morelli M, Wigle N, et al. D1-Receptor-related priming is attenuated by antisense-meditated “knockdown” of fosB expression. *Mol Brain Res* 1998;53:69–77.

60. Pennypacker KR, Hong JS, McMillian MK. Implications of prolonged expression of Fos-related antigens. *Trends Pharmacol Sci* 1995;16:317–321.

61. Kelz MB, Nestler EJ. ΔFosB: a molecular switch underlying long-term neural plasticity. *Curr Opin Neurol* 2000;13:715–720.

62. Lin RJ, Kao HY, Ordentlich P, et al. The transcriptional basis of steroid physiology. *Cold Spring Harb Symp Quant Biol* 1998;63:577–585.

63. Klein-Hitpass L, Schwerk C, Kahmann S, et al. Targets of activated steroid hormone receptors: basal transcription factors and receptor interacting proteins [see Comments]. *J Mol Med* 1998;76:490–496.

64. Karin M. New twists in gene regulation by glucocorticoid receptor: is DNA binding dispensable? *Cell* 1998;90:487–490.

65. Gottlicher M, Heck S, Herrlich P. Transcriptional cross-talk, the second mode of steroid hormone receptor action. *J Mol Med* 1998;76:480–489.

66. Alberini CM, Ghirardi M, Metz R, et al. C/EBP is an immediate-early gene required for the consolidation of long-term facilitation in Aplysia. *Cell* 1994;76:1099–1114.

67. O’Donovan KJ, Tourtellotte WG, Millbrandt J, et al. The EGR family of transcription-regulatory factors: progress at the interface of molecular and systems neuroscience. *Trends Neurosci* 1999;22:167–173.

Section II
Emerging Methods in Molecular Biology and Genetics

Samuel H. Barondes

Samuel H. Barondes: Center for Neurobiology and Psychiatry, Department of Psychiatry, University of California-San Francisco, San Francisco, California 94143-0984.

Emerging Methods in Molecular Biology and Genetics - Introduction

When the American College of Neuropsychopharmacology was founded in the mid-1950s, molecular biology and genetics were in their infancy and had little to offer neuropsychopharmacology. By 1967, when the first volume in this series was published, it still had not become apparent how greatly our field would be influenced by research on genes and on DNA. Of more than a hundred papers in that volume--*Psychopharmacology: A Review of Progress 1957-1967*--only a few used tools of molecular biology. None of those papers, including my own, envisioned the central role that such tools have now come to play.

This central role, which is documented in this section, is the result of the development of two types of new technologies. One of them, the automated sequencing of the nucleotides in DNA, facilitated the decoding of the structure of all genes, including those that make up the human genome. The other consists of ways to manipulate the structure of individual genes in isolated cells or in intact organisms, and to measure their levels of expression. This made it possible to directly study the biological actions of particular genes, and the effects of changes in their regulatory regions and coding regions.

These technologic advances are now being applied to a variety of problems in neuropsychopharmacology. For example, measurements of the levels of expression of large numbers of genes in various brain regions and nerve cells are providing information about the molecular basis of normal brain functions, and the effects of drugs on these functions. The same technologies are also being used to learn about the molecular pathogenesis of mental disorders. As work on the human genome continues it will lead to the identification of gene variants that influence susceptibility to mental disorders, as well as of gene variants that influence the alternative ways that individuals respond to particular psychiatric drugs. It will also provide new targets for the creation of better drugs, with greater efficacy and specificity.

The six chapters in this section provide a sampling of the molecular and genetic tools that are being used to advance neuropsychopharmacology. Because these tools are changing so rapidly, the authors provide overviews rather than extensive details. In this way they hope to make these tools comprehensible to the nonspecialist, and to invite their further application.

As you read these overviews you will become increasingly familiar with a series of new terms--from *gene chip*, *knockouts*, and *viral vectors*, to *genome scans* and *pharmacogenetics*. These terms are becoming commonplace, and are already scattered throughout this book. By the time the next volume in this series appears, it is likely that the methods that they refer to will be so widely used in our field that a separate section about them will no longer be needed.

18

Using Human Genomics to Advance Neuropsychopharmacology

L. Alison Mcinnes

Nelson B. Freimer

L. A. McInnes and Nelson B. Freimer: Department of Psychiatry and Neurogenetics Laboratory, University of California–San Francisco, San Francisco, California 94143.

Genomics, the study of genomes, includes gene mapping, sequencing, and investigation of gene functions. This field will advance neuropsychopharmacology in two complementary ways. First, it is hoped that application of genomics technologies to pedigree and population samples of patients with psychiatric disorders will allow the identification of genes contributing to the etiology and pathogenesis of these diseases and provide a rational basis for new drug development. Second, variations in the sequence of known genes whose products are the targets of current psychotropic drugs may influence the likelihood that an individual patient will have a therapeutic response to these drugs or experience a side effect. Identification and functional characterization of these sequence variants in large populations of patients of various ethnicities constitutes the discipline of pharmacogenomics. The scope of psychopharmacogenomics, however, is currently restricted by our limited knowledge of the genes that contribute to psychiatric disorders and the neural pathways altered by psychotropic agents. Fortunately, data and technologies provided by the U.S. Human Genome Project (HGP), including provision of the complete sequence of the human genome, will greatly facilitate identification of such genes (1 ,2). This chapter describes how technologic advances in genomics will shape the future of psychiatric genetics and psychopharmacogenomics, fields that may establish an objective basis for the restructuring of the nosology, diagnosis, and treatment of psychiatric disorders. The results of genetic studies of particular psychiatric disorders and of responses to specific drugs are considered in other parts of this book.

IDENTIFYING GENES FOR PSYCHIATRIC DISORDERS

Part of "18 - Using Human Genomics to Advance Neuropsychopharmacology "

Rational strategies for the advancement of psychopharmacology are dependent on furthering our currently sparse knowledge of the pathophysiologic basis of psychiatric disorders. To this end, human genetic approaches offer a promising alternative to traditional biochemical and neurophysiologic investigations as twin, family, and adoption studies all support the heritability of many psychiatric syndromes. Unfortunately, attempts to first map (i.e., localize a unique region of DNA shared by patients with a particular disorder) and then identify genes predisposing to psychiatric disorders have been frustrated by the complexity of the genetic mechanisms underlying behavioral phenotypes.

A phenotype is the observable physical manifestation of genetic variation at a particular site or locus in the DNA, whereas a genotype refers to the actual DNA sequence, at the responsible genetic locus. With single gene disorders (also referred to as mendelian disorders) there is a simple, direct relationship between variation in a single gene and the phenotype that results. Thus, all patients with a given mendelian disorder, such as cystic fibrosis, will carry abnormal genotypes at a single disease locus. In contrast, the relationship between phenotype and genotype is not straightforward for complex genetic traits. In this setting, multiple different susceptibility genes and environmental factors interact in varying combinations within individuals who appear to have clinically indistinguishable phenotypes. This means that in any given sample of patients diagnosed with a particular psychiatric disorder, the number of individuals actually sharing a disease gene or genes in common might be very small such that the “effective” sample size does not provide enough power to detect the responsible genes.

Fortunately, there are strategies for finding genes contributing to complex traits that have been successfully applied to the genetic dissection of other nonpsychiatric, genetically

heterogeneous disorders. One approach is to try to reduce genetic heterogeneity in the patient sample by studying genetically isolated populations or by narrowing the affected phenotype under study based on criteria of severity or the presence of a biological marker for the disease. Another approach is to greatly expand the sample size and number of DNA markers used in genetic association studies to increase the power to detect multiple possible genes contributing to disease in subsets of the sample. In either case, both pedigree and population-based genetic mapping studies are expected to yield more promising results in the future due in part to the extensive characterization of the human genome provided by the HGP. The HGP, begun in 1990, is a joint effort coordinated by the U.S. Department of Energy and the National Institutes of Health, with the cooperation in recent years of international entities such as the Wellcome Trust in Great Britain (3). One of the HGP's main goals is to finish the complete human genome sequence by the end of 2003 while concomitantly identifying all the estimated 100,000 genes in human DNA and creating the most dense and accurate genetic maps for genome screening studies. The next subsection describes how innovations in genetic maps and the structure of genetic mapping studies may eventually lead us to identify the as yet elusive genes for psychiatric disorders.

Genetic Maps

At the time of this writing, commonly used genetic maps consist mostly of microsatellite DNA markers (usually repeats of two, three, or four nucleotides that vary in length among individuals) that occur with fairly even spacing across the entire genome. These maps contain several thousand markers spaced at roughly 1 to 5 centimorgans (cM); 1 cM is a unit of genetic distance equivalent to a recombination frequency between two loci of 1%, i.e., one recombination would occur per hundred meioses. Much denser maps are under construction now, however, as part of the HGP. In fact, a major goal of the HGP is to characterize the extent of genetic variation that exists among humans in order to create a map of several hundred thousand markers to enable high-density genome screening studies of complex traits (4). Differences in single bases of DNA known as single nucleotide polymorphisms (SNPs) are thought to constitute roughly 90% of sequence variation in humans. Occurring at an average spacing of 1 SNP per 1,000 base pairs (1 kilobase, kb), they will thus constitute the majority of the markers in the planned high-density map. To facilitate identification of these SNPs, the National Institutes of Health (NIH) recently assembled cell lines and DNA from a collection of 450 anonymous, unrelated individuals representing the major ethnic groups of the United States; this collection is known as the DNA Polymorphism Discovery Resource (DPDR) (5). The DPDR is available to investigators to facilitate detection of population genetic variation in their loci of interest, with the expectation that they will share this information with the scientific community. For instance, SNP variants in the coding or regulatory regions of genes (cSNPs) may cause functional differences in gene expression. With such variants catalogued in advance, it will be relatively straightforward to test multiple candidate genes for association with a disease phenotype or a pharmacogenetic effect. The enormous task of identifying and scoring SNPs in large samples, or performing the projected high-density genome screening studies, has necessitated the development of high-throughput technologies such as DNA chips (6), which are discussed elsewhere in this volume.

Principles of Genetic Mapping: The Search for Identity by Descent

Genetic mapping methods are based on the expectation that a proportion of patients in a study population, whether members of an extended pedigree or more remotely related descendants of a common ancestor, will share segments of DNA identically by descent (IBD) in the vicinity of the disease gene under study. The principle behind this expectation is best illustrated by considering genetically isolated or founder populations (7 ,8). A founder population descends from a small number of ancestral individuals and grows in relative isolation with little admixture, so it is genetically homogeneous. The premise is that a disease mutation is introduced into a population on a particular “founder” chromosome, which is then transmitted to patients descended from a common ancestor carrying this mutation. Over subsequent generations, recombination will reduce the size of the segment of DNA that patients share around the disease gene. The shared segment includes the markers that flank the disease gene. These markers are said to be linked to the gene since, because of their close physical proximity, they seldom recombine with each other and hence are transmitted as a unit or segment. Detection of linkage is the goal of pedigree-based genetic studies. Alternatively, if alleles at markers co-occur more frequently than expected given the known allele frequencies and recombination fraction between the markers, they are said to be in linkage disequilibrium (LD). Evidence of LD between markers also indicates that they are probably close to each other. If LD is observed between the same markers over a region greater than occurs at random in a sample of patients, it may indicate that they share this segment of DNA IBD and that it harbors a disease gene. Detection of LD is a goal of population-based genetic studies.

The length of the IBD segment around a disease gene is inversely proportional to the number of generations by which patients are separated from a common ancestor (Fig. 18.1), and thus genomic screening strategies to detect such a segment will depend on the “age” of the study population (9). For example, a sample of affected individuals separated by roughly 15 generations from their common ancestors

might be expected to share segments around a disease gene detectable in genome screens using current microsatellite maps with markers spaced at 3 to 5 cM. In contrast, the Finnish population is a founder population in which present-day individuals are separated from their common ancestors by up to 100 generations. In this case, shared DNA segments harboring a particular disease gene may be so small that one would have to screen the genome with a much denser marker map (e.g., markers every 0.1 cM) to find them. Such screening studies in a nonhomogeneous population such as that of the United States, wherein common ancestors must be located in the very remote past, will require use of the planned SNP map of around several hundred thousand markers in order to detect regions of LD, which, it has been hypothesized, may be as small as 3 kb (10).

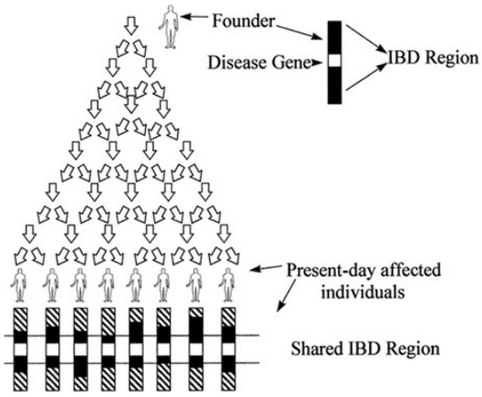


FIGURE 18.1. Genetic mapping studies in isolated populations take advantage of the fact that many recombination events separate affected individuals in the present day from a common disease ancestor. As a result, the majority of patients should share a segment of DNA around the disease gene that is longer than any other DNA segment they might share by chance, but still small enough for positional cloning purposes.

Alternatively, identification of much larger chromosomal regions that are IBD among a sample of patients may also be carried out in extended pedigrees wherein the small number of meioses separating affected individuals leads to a greater length of IBD sharing around the disease gene. Such regions may be easier to detect in pedigrees with genome screens using markers spaced at broader genetic intervals (on the order of 5 to 10 cM). It can be difficult, however, to find recombinant individuals that will allow refinement of the candidate interval to a sufficiently small region to facilitate positional cloning. We review the relative strengths and weaknesses of pedigree- and population-based genetic studies below.

Pedigree-Based Mapping Studies: Problems and Solutions

Linkage analysis of pedigree data has been a very successful method of mapping genes for rare single gene disorders with distinct phenotypes such as cystic fibrosis (11). There are many limitations on the use of linkage analysis for complex traits, however. Linkage analysis is a statistical means of quantifying the likelihood that the observed segregation of marker alleles within a family supports the hypothesis of linkage versus nonlinkage to a disease gene. Traditional linkage analysis is model-based or parametric, meaning that it requires specification of disease inheritance parameters that are not easily estimated for psychiatric disorders, such as the frequency of the disease allele, the genotype specific penetrance, or even the number of genes likely involved. The lack of knowledge of these parameters means that data must be analyzed under a number of different models. This process of multiple testing diminishes the strength of statistical conclusions such that it may be nearly impossible to distinguish between background “noise” and true but weak signals coming from multiple genes of small effect (12). However, several strategies have been successfully implemented to increase the power of linkage analysis for complex nonpsychiatric illnesses by attempting to reduce genetic heterogeneity in the patient sample or by using methods of statistical analysis that do not require specification of genetic model parameters.

Solution 1: Reduce Genetic Heterogeneity in the Sample

Strategies for reducing genetic heterogeneity include studying a small number of large, multiply affected pedigrees ascertained, if possible, from a founder population and narrowly defining the phenotype understudy. The premise in the former case is that the number of genes contributing to a particular disease phenotype within one or a few large families may be less than in many small families or the population at large. This premise is also more likely to hold if the number of disease loci in the population as a whole is limited, as may be the case for founder populations (7 ,8). Furthermore, a multiply affected family may indicate that the gene or genes involved are highly penetrant (penetrance refers to the likelihood that a person who has a disease gene will manifest the disease phenotype) and may be easier to find.

For nonpsychiatric complex traits, refinement of the affected phenotype can be accomplished in several ways. For instance, limiting the affected phenotype to include only the most extreme or distinct form of the illness under study has also been critical to the success of mapping studies for complex traits, as such phenotypes are expected to reflect a more homogeneous genetic etiology than more broadly defined phenotypes (13 ,14). Illustrating these points, a gene

for a severe form of Alzheimer's disease (AD) characterized by early age of onset was detected in a few large multiply affected pedigrees ascertained from a genetically isolated population of German descent (15 ,16). Although such a gene may not contribute significantly to AD in the general population, it may still provide clues as to relevant biological pathways that might suggest candidate genes for other mapping studies.

Another way to refine an affected phenotype is to require the presence of an objective measure associated with the disorder such as elevated immunoglobulin E (IgE) levels in patients with asthma (17). However, we still have no comparable biological markers for psychiatric disorders at least when these disorders are defined by current nosology. Hence, investigators are attempting to find endophenotypes or subcomponents within psychiatric syndromes that may be objectively measured and inherited in a more straightforward fashion than the constellation of symptoms that constitute the full psychiatric syndrome (18 ,19 and 20). What differentiates this strategy from other attempts to refine traditional psychiatric phenotypes is that family members who have not received a psychiatric diagnosis may still be segregating the trait of interest and may display an endophenotype that allows them to be considered “affected” for genetic study.

Many of the efforts to investigate familial transmission of endophenotypes have focused on schizophrenia. For example, investigators have hypothesized that abnormalities of sensory gating are biological markers for attentional dysfunction, which seems to be a core phenotypic characteristic of schizophrenia or psychosis. Abnormal ocular movements and failure to suppress evoked responses to auditory stimuli after a cue (the P50 response) are both thought to be transmitted within families of schizophrenic probands whether or not family members have a psychiatric diagnosis (21 ,22 ,23 ,24 and 25). Two studies have examined the same set of families for linkage to either an abnormal P50 response alone or in combination with abnormal oculomotor movements (26 ,27). Evidence for linkage to each phenotype implicated different loci (the α_7-nicotinic acetylcholine receptor subunit gene and a region in chromosome 22q11-12, respectively) and was stronger than evidence for linkage of the schizophrenia phenotype alone.

Few endophenotypes have been characterized so far for mood disorders; however, a possible endophenotype is that of suicide (28). Roy et al. (29) studied the monozygotic and dizygotic co-twins of twin suicide victims and found that significantly more monozygotic co-twins than dizygotic co-twins also attempted suicide, possibly arguing for a genetic component to this behavior. Investigators are continuing to develop brain banks of suicide victims for postmortem studies including genetic screens and searching for relevant biochemical markers (30).

Solution 2: Alternative Analytic Models

As an alternative to model-based methods, nonparametric or “model-free” methods can be utilized to detect linkage and may be more robust when the true mode of inheritance is unknown. These methods were originally developed for samples of affected sibling pairs but have now been modified for analysis of other types of relative pairs or whole pedigrees. Simply stated, nonparametric methods are designed to calculate the amount of IBD sharing of marker alleles among affected relatives where the null hypothesis is that transmission of alleles is independent of transmission of disease. For any pair of affected relatives, the probability that the pair will share zero, one, or two alleles IBD can be calculated based on their degree of relationship. Linkage is detected if the sharing of marker alleles among affected relatives is increased over the sharing expected given their relationship.

Demonstration that patients share a series of alleles over multiple markers on the same chromosome, also known as haplotypes, can definitively establish that a segment of DNA has been inherited IBD and thus likely harbors a disease gene. Methods that can quantitate the significance of haplotype sharing among affected individuals are thus particularly useful tools for determining candidate gene intervals in both pedigree and population samples, although the inheritance patterns within extended pedigrees need to be fully characterized to avoid misinterpretation of the allele sharing data.

In addition to the problems with statistical detection of linkage, another major shortcoming of pedigree and affected relative pair studies is that investigators may detect a disease gene but be able to localize it only to a very broad genetic interval. The extent to which a genetic interval containing the disease locus can be narrowed to a small-enough interval for positional cloning purposes depends on the number of opportunities for recombination of the disease haplotype in affected persons; in pedigrees where individuals are separated by only two or three generations, opportunities for recombination of the disease haplotype are limited. Although affected relative pairs are usually much easier to collect than multiply affected pedigrees, very large numbers are required to detect linkage, and the accuracy of gene localization is usually much less than that provided by pedigrees.

Finally, when studying complex traits it is very likely that some individuals will be phenocopies, which means that they exhibit the phenotype under study but due to different genetic or environmental factors. In this case, one may be misled by apparent recombination events even in a single individual, and may therefore incorrectly delineate the candidate interval for a disease gene in a region.

All of these factors can seriously impede localization and positional cloning efforts for disease genes in pedigree samples.

Population-Based Mapping Studies

Given the limitations on pedigree studies described above, analysis of a population-based sample is frequently a preferred strategy for high-resolution mapping of disease loci (31). One reason for this is that many meioses (and therefore opportunities for recombination) have occurred since the patients in such a sample were separated from their common ancestor, so there is a better chance to narrowly define a candidate gene interval (Fig. 18.1). Risch and Merikangas (32) also proposed that population mapping strategies might be a more efficient means of initially localizing disease genes (given a sufficient sample size and an appropriately dense marker map), particularly for loci of relatively small effect, as the sample sizes needed for affected relative pair strategies may be huge and thus not feasible.

Association studies are designed to be case-control or family based (see ref. 33 for review). In case-control association studies, allele frequencies at a particular marker are compared between a sample of patients and a sample of controls matched as closely as possible to cases in terms of ethnicity, age, gender, and other relevant socioeconomic variables. Unfortunately, perfect matching can never be guaranteed, and unknown population stratification can occur if many of the cases or controls share an uninvestigated variable. In this setting, the alleles of cases might appear to differ markedly from controls at a particular genetic locus because of such an unknown variable and not because of the presence of the disease phenotype; this could lead to a false-positive result. Such stratification can occur even within distinct ethnic groups. For example, an association study of type 2 diabetes mellitus in a Native-American tribe seemed to indicate that a particular allele of the immunoglobulin complex was protective against diabetes (34). However, after extensive genealogic examination of Native Americans with this allele, it appeared that they all had distant Caucasian relatives. As the allegedly associated allele was common in Caucasians, and diabetes was less common in Caucasians than in the tribe studied, overrepresentation of this allele in nondiabetic Native Americans reflected only the presence of Caucasian admixture and not a true protective effect from diabetes.

Alternatively, affected individuals and their parents are ascertained for family-based studies that utilize the alleles on the nontransmitted chromosomes of parents as controls for the patients' alleles to prevent ethnic mismatching. One commonly used approach to analyzing such family data is the transmission disequilibrium test (TDT) (35). In this test each allele of a heterozygous parent is measured to see if it is transmitted to an affected offspring significantly more often than the expected 50% by chance. In this case, the implicated allele would be both associated and linked to the disorder, obviating the possibility that the allele is falsely associated through population stratification. Other approaches for analyzing family-based association data use nontransmitted parental alleles as controls, but do not evaluate actual transmission of these alleles to the patient and do not exclude data from homozygous parents. One such approach is the likelihood-based method of Terwilleger (36). A disadvantage of family-based LD methods is that it can often be difficult to sample parents of affected individuals, especially for adult-onset disease.

The LD tests described above are often used to examine the association at single markers individually, which can also be problematic because a very large number of markers must be used for LD genome screening studies, even in isolated populations, and statistical correction for multiple testing is necessary. Interpreting the significance of single-point association tests in this setting becomes extremely difficult (37). Fortunately, the development of multipoint statistical methods for quantifying the significance of haplotypes shared over multiple markers could help to increase the power to detect even weak LD signals coming from a subset of the sample. Such approaches are inherently more powerful than single-point tests of association and will be essential for the evaluation of data generated from SNP maps.

One promising LD method, termed ancestral haplotype reconstruction (AHR), assesses the likelihood that a sample of patient haplotypes have descended from a common mutation-bearing founder haplotype (38). This method is currently being modified so that it will be useful both for genome screening and subsequent fine-mapping studies. At the genome screening stage, markers are generally spaced at sufficient distances such that they can be considered to be in linkage equilibrium with each other in distantly related affected persons. Detecting LD between two or more markers in this setting should thus point to the candidate gene interval as long as the underlying assumption is met that the markers tested are not in LD with each other independent of the disease phenotype (so-called background or random LD); it is still not certain how such background LD is distributed within the genome and between different populations. Once a candidate region has been identified by LD analysis, the next step is to type as many markers as closely spaced as possible within the area to determine the minimal interval of maximal IBD sharing that should contain the disease gene. Although this step can be accomplished in some cases just by observation (39), a statistical method that can assign some level of significance to the observed data would be very useful, especially if the disease haplotype is relatively common and a large sample is required to detect its contribution to the phenotype. However, multipoint analysis of markers typed at high density for fine mapping (or for genome screens with dense SNP maps) is complicated because, since the markers are so closely spaced, one cannot assume that these markers are not in LD with each other independent of the disease phenotype. Multipoint LD methods such as AHR will need to take into account the possibility that significant background LD could occur between

closely spaced markers in order to distinguish this background LD from what may be a very subtle increment of LD surrounding the true disease locus.

In summary, the ability to localize disease genes using LD methods in a given population sample depends on the amount and extent of LD present, the number of disease predisposing alleles at a given locus (allelic complexity), the degree to which the disease locus increases the likelihood of manifesting the affected phenotype, and the power of current statistical methods to measure existing LD. For an excellent review of the strengths and weaknesses of current statistical approaches for analyzing LD, see ref. 40 .

Identification of a Disease Gene

Should psychiatric geneticists overcome the many obstacles facing them and succeed in mapping a disease locus to a specific interval, the next step would be to identify the disease gene within it, a process termed *positional cloning* or, given the completion of genome sequencing, the *positional candidate approach* (41). Positional cloning in its purest sense is the process of identifying a disease gene based only on knowledge of its chromosomal location as determined by linkage analysis or LD mapping, without any knowledge of the gene's function. This process involves laborious efforts to build a physical map and sequence the region. Physical maps are made by isolating and linking together yeast and/or bacterial artificial chromosomes (YACs, BACs) containing fragments of human DNA from the region. These fragments are then sequenced and ordered so that the genomic DNA sequence across the candidate gene region is known. Then comes the arduous process of identifying all the genes in the interval and performing mutation detection. Advances in physical mapping and gene-finding technologies arising from the HGP have greatly speeded up this process, however. For instance, near-complete genomic sequence data for the region of interest may already have been deposited in the public databases, obviating the need for extensive genomic sequencing. Next, an investigator may explore large databases of partial complementary DNA (cDNA) sequences also known as “expressed sequence tags” (ESTs) to find most or all of the genes mapping within the area. In addition, there are already a multitude of Web-based sequence analysis programs that provide a host of information such as exon and promotor prediction, open reading frames, and protein homologies for translated sequence. This computer exercise is part of the positional candidate approach. Once all the ESTs have been identified, the complete sequence of the corresponding genes (including definition of intron/exon boundaries and the promotor, if possible) are elucidated and mutation detection begins. If the causative mutation is not detected in coding sequence, however, it can be very difficult to detect disease-causing mutations in the surrounding noncoding DNA or introns, as these regions are large and likely to display more natural variation than highly conserved coding sequence. One possible strategy in this situation is to search for an orthologous (similar) gene in a model organism such as the mouse. Comparison of these sequences may highlight strongly conserved regions of DNA outside of coding sequence that may be functionally relevant and important to examine closely for disease predisposing variants (42). Provision of the complete sequences of the genomes of model organisms amenable to genetic manipulations such as flies, mice, roundworms, and yeast by the HGP could speed understanding of comparable gene structure and function in humans and serve as a molecular confirmation and/or supplement to current gene prediction programs (43 ,44).

PHARMACOGENOMICS

Part of "18 - Using Human Genomics to Advance Neuropsychopharmacology "

As genes contributing to the development of psychiatric disorders are discovered, they will be added to the known array of neurotransmitters, receptors, and transporters that are already considered candidate genes for pharmacogenetic analysis. Pharmacogenetics is the study of the genetic mechanisms determining an individual's responsiveness to drugs (45). In this approach, genes involved in drug metabolic pathways and sites of action, or disease processes if known, are examined for naturally occurring variants or polymorphisms, which may then be shown to affect the expression of that gene. This effect on gene function may then be linked to the efficacy of the drug and/or a predisposition to particular side effects in individuals with that genotype. For example, the apolipoprotein (apo) E4 allele at the apo E locus has been shown to be associated with late-onset AD, as well as to a poor response to cholinesterase inhibitor treatment of AD (46). The downside of this strategy is that it is limited by the paucity of candidate genes with proven association to psychiatric phenotypes. Fortunately, investigators in the emergent field of pharmacogenomics are seeking to identify previously unsuspected genetic markers of drug responsiveness by using a variety of high-throughput genomics technologies to examine drug-induced changes of gene expression in different tissues throughout the body. We provide here a brief overview of the principles of pharmacogenetics and pharmacogenomics (see ref 47 . for review), focusing on how these approaches are being applied in an effort to provide a more rational basis for pharmacotherapy of psychiatric disorders than the trial-and-error approach we currently employ.

The impact of genetic variation on drug response is characterized broadly by changes in pharmacokinetic and pharmacodynamic parameters. Pharmacokinetic studies assess the processes of absorption, distribution, first-pass and general metabolism, and elimination of drugs. Transport processes in renal, intestinal, and hepatic epithelia and drug metabolizing enzymes exhibit genetic variability, which will in many cases be likely to influence the pharmokinetics of

relevant drugs. In the latter situation, the cytochrome P-450 system has been best studied (48) beginning with the classic example of the poor metabolism of the antihypertensive drug debrisoquine due to several mutant alleles of the polymorphic CYP2D6 gene also known as debrisoquine/sparteine hydroxylase. This enzyme is responsible for the metabolism of roughly a quarter of all drugs including most antipsychotics and antidepressants (49). This enzyme is also inhibited potently by fluoxetine. About 7% of Caucasians and an even greater percentage of Asians are poor metabolizers of such drugs due to polymorphisms in this enzyme. On the other hand, some persons carry different alleles and/or multiple copies of this gene, which predispose to more rapid metabolism; up to 13 copies have been documented in a single individual. A study of nortriptyline metabolism in these individuals clearly demonstrated that clearance of nortriptyline was proportional to the number of copies of the *CYP2D6* gene, especially the CYP2D6*2 allelic form (50 ,51). Knowledge that a person has a genotype predisposing to unusually slow or rapid metabolism could guide appropriate drug choice and dosing regimen. Unfortunately, despite intensive study, no definitive relationship between polymorphisms in cytochrome P-450 enzymes and drug efficacy or predisposition to side effects of antidepressant drugs has yet been discovered (52).

Pharmacodynamics concerns the relationship between the concentration of a drug and response at its site of action, for example at receptors and transporters for neurotransmitters. Pharmacodynamic effects may also vary temporally, and so both the acute and chronic nature of response to the drug must be considered. In the acute phase of responsiveness, receptor polymorphisms could alter any of the myriad steps in a pathway from receptor-drug binding through the cascade of signals that result; such variants may determine who is more prone to immediate drug reactions, for example the malignant hyperthermia that may occur in response to antipsychotics. Genetic variation could also play a role in the drug-induced neural plasticity that occurs as a result of the chronic treatment required for alleviation of psychiatric symptomatology as well as chronic use of addictive substances. Adaptive responses to drugs will vary among individuals, and genetic factors may predict such phenomena as waning of drug response over time, and proneness to side effects such as tardive dyskinesia induced by antipsychotics (53).

From a pharmacogenetic perspective, one of the most obvious candidates for studying psychiatric drug responsiveness identified to date is the serotonin (5-hydroxytryptamine, 5-HT) transporter (5-HTT) (52). This transporter plays a critical role in the termination of serotoninergic transmission and is the target of the most widely prescribed family of antidepressants, known collectively as the selective serotonin reuptake inhibitors (SSRIs). Heils et al. (54) identified in 1996 a variant of 5-HTT with a 44–base pair (bp) insertion within the promoter region, which is commonly referred to now as the long (versus the short) allele. Lesch et al. (55) then demonstrated that baseline levels of transcription of the long variant were more than double that of the short variant in transfected cells and that this difference was reflected in altered 5-HTT expression and 5-HT reuptake). Since then, investigators have studied depressed patients to see if they can correlate SSRI response with alleles at the 5-HTT promoter. Smeraldi et al. (56) presented data suggesting that patients with delusional depression responded better to fluvoxamine if they were homozygous for the long allele of the 5-HTT promoter. Other investigators have obtained similar findings in a samples of depressed patients treated with paroxetine (57), although in this study rapidity of response was improved in persons homozygous for the long allele, while overall outcome at 12 weeks was the same for all genotypes.

Besides the serotonin transporter, there are several thousand other genes for neuromodulatory molecules that have been cloned and expressed in some type of cellular system and that are viable candidates for drug targets (58). Pharmacogenomic technologies may aid in prioritizing these candidates for examination, or identifying yet more candidates, by determining those genes that are activated or deactivated in tissues during an acute psychiatric episode and in response to treatment. One approach to evaluating gene expression involves hybridization of fluorescent or radioactively labeled messenger RNA (mRNA) samples taken from the relevant cell population to cDNA arrays. Changes in gene expression (up, down, or none) can then be compared between different samples at a single time point or within a sample overtime. This technique is known as serial analysis of gene expression (SAGE) (59). However, if the changes in gene activation induced by disease or by drugs are localized to a specific population of cells in inaccessible tissues such as that of the brain, rather than, say, in fibroblasts from skin biopsies, the SAGE technique will not be helpful. Alternatively, large-scale analysis of proteins within clinical samples is also predicted by some to become a useful means of identifying biological markers indicative of a response to drugs (60). Again, any changes in protein expression would need to be detectable in easily obtainable fluids such as blood or urine to be of use in the evaluation of psychiatric disorders, and the process of informed consent for such experimentation would need to be reviewed thoroughly.

SUMMARY

Part of "18 - Using Human Genomics to Advance Neuropsychopharmacology "

Genomics has great potential to advance the field of psychiatry in general and neuropsychopharmacology in particular. This will occur through the identification of genes involved in the etiology of psychiatric disorders and the identification and characterization of genes involved in the response to psychotropic agents. Although we have not made much headway as of yet in either of these fields, impressive advances

in our knowledge of the genomes of human and model organisms as well as access to the technologies developed to provide this information are likely to stimulate swift progress. Success in these endeavors will lead to more efficient identification and development of drug targets, the provision of an objective basis for choice of drug, “custom drugs” based on an individuals genotype, improved diagnosis of disease, and earlier detection of genetic predispositions to disease.

ACKNOWLEDGMENTS

Part of "18 - Using Human Genomics to Advance Neuropsychopharmacology "

L. Alison McInnes is supported by National Institute of Mental Health (NIMH) grant K01 MH01748-01. Nelson B. Freimer is supported by NIMH grants R01-MH49499 and K02-MH01375. We would also like to thank Susan Service for her original drawing.

REFERENCES

1. Collins FS, Patrinos A, Jordan E, et al. New goals for the U.S. Human Genome Project: 1998–2003. *Science* 1998;282(5389):682–689.

2. Collins FS. Shattuck lecture—medical and societal consequences of the Human Genome Project. *N Engl J Med* 1999;341(1):28–37.

3. http://www.ornl.gov/TechResources/HumanGenome/publicat/publications.html

4. Roberts L. SNP mappers confront reality and find it daunting. *Science* 2000;287(5460):1898–1899.

5. Collins FS, Brooks LD, Charavarti A. A DNA Polymorphism Discovery Resource for research on human genetic variation. *Genome Res* 1998;8:1229–1231.

6. Wang DG, Fan J-B, Siao C-J, et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 1998;280:1077–1082.

7. de la Chapelle A. Disease gene mapping in isolated human populations: the example of Finland. *J Med Genet* 1993;30(10):857–65.

8. Wright AF, Carothers AD, Pirastu M. Population choice in mapping genes for complex diseases. *Nature Genet* 1999;23:397–404.

9. Houwen RH, Baharloo S, Blankenship K, et al. Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. *Nature Genet* 1994;8:380–386.

10. Kruglyak L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genet* 1999;22:139–144.

11. Kerem B, Rommens JM, Buchanan JA, et al. Identification of the cystic fibrosis gene: genetic analysis. *Science* 1989;245(4922):1073–1080.

12. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genet* 1995;11:241–247.

13. Cannon-Albright LA, Goldgar DE, Meyer LJ, et al. Assignment of a locus for familial melanoma, MLM, to chromosome 9p13–p22. *Science* 1992;258(5085):1148–1152.

14. Keating M, Dunn C, Atkinson D, et al. Consistent linkage of the long-QT syndrome to the Harvey ras-1 locus on chromosome 11. *Am J Hum Genet* 1991;49(6):1335–1339.

15. Levy-Lahad E, Wijsman EM, Nemens E, et al. A familial Alzheimer's disease locus on chromosome 1. *Science* 1995;269(5226):970–973.

16. Sherrington R, Rogaev EI, Liang Y, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 1995;375(6534):754–760.

17. Laitinen T, Kauppi P, Ignatius J, et al. Genetic control of serum IgE levels and asthma: linkage and linkage disequilibrium studies in an isolated population. *Hum Mol Genet* 1997;6(12):2069–76.

18. Freedman R, Adler LE, Leonard S. Alternative phenotypes for the complex genetics of schizophrenia. *Biol Psychiatry* 1999;45:551–558.

19. Adler LE, Freedman R, Ross RG, et al. Elementary phenotypes in the neurobiological and genetic study of schizophrenia. *Biol Psychiatry* 1999;46:8–18.

20. Weinberger DR. Schizophrenia: new phenes and new genes. *Biol Psychiatry* 1999;46:3–7.

21. Clementz BA, Geyer MA, Braff DL. Poor P50 suppression among schizophrenia patients and their first-degree biological relatives. *Am J Psychiatry* 1998;155:1691–1694.

22. Crawford TJ, Sharma T, Puri BK, et al. Saccadic eye movements in families multiply affected with schizophrenia: the Maudsley family study. *Am J Psychiatry* 1998;155:1703–1710.

23. Crawford TJ, Haeger B, Kennard C, et al. Saccadic abnormalities in psychotic patients. I. Neuroleptic-free psychotic patients. *Psychol Med* 1995;25:461–471.

24. Crawford TJ, Haeger B, Kennard C, et al. Saccadic abnormalities in psychotic patients. II. The role of neuroleptic treatment. *Psychol Med* 1995;25.473–483.

25. Ross RG, Harris JG, Olincy A, et al. Anticipatory saccades during smooth pursuit eye movements and familial transmission of schizophrenia. *Biol Psychiatry* 1998;44:690–697.

26. Freedman R, Coon H, Myles-Worsley M, et al. Linkage of a neurophysiological deficit in schizophrenia to a chromosome 15 locus. *Proc Natl Acad Sci USA* 1997;94:587–592.

27. Myles-Worsley M, Coon H, McDowell J, et al. Linkage of a composite inhibitory phenotype to a chromosome 22q locus in eight Utah families. *Am J Med Genet* 1999;88:544–550.

28. Mann JJ. Neurobiology of suicide. *Nature Med* 1998;4(1):25–30.

29. Roy A, Segal NL, Sarchiapone M. Attempted suicide among living co-twins of twin suicide victims. *Am J Psychiatry* 1995;152(7):1075–1076.

30. Huang YY, Grailhe R, Arango V, et al. Relationship of psychopathology to the human serotonin 1B genotype and receptor binding kinetics in postmortem brain tissue. *Neuropsychopharmacology* 1999;21(2):238–246.

31. Lander ES, Botstein D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 1989;121(1):185–199.

32. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;273:1516–1517.

33. Burmeister M. Basic concepts in the study of diseases with complex genetics. *Biol Psychiatry* 1999;45:522–532.

34. Knowler WC, Willliams RC, Pettitt DJ, et al. GM3;5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. *Am J Hum Genet* 1988;43:520–526.

35. Spielman RS, Ewens WJ. The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet* 1996;59:983–989.

36. Terwilliger JD. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am J Hum Genet* 1995;56:777–787.

37. Terwilliger JD, Weiss KM. Linkage disequilibrium mapping of complex disease: fantasy or reality? *Curr Opin Biotechnol* 1998;9:578–594.

38. Service S, Temple Lang D, Freimer N, et al. Linkage-disequilibrium mapping of disease genes by reconstruction of ancestral haplotypes in founder populations. *Am J Hum Genet* 1999;64:1728–1738.

39. Bull LN, Juijn JA, Liao M, et al. Fine-resolution mapping by haplotype evaluation: the examples of PFIC1 and BRIC. *Hum Genet* 1999;104(3):241–248.

40. Clayton D. Linkage disequilibrium mapping of disease susceptibility genes in human populations. *Int Statistical Rev* 2000;68 (*in press*).

41. Collins FS. Positional cloning moves from perditional to traditional. *Nature Genet* 1995;9:347–350.

42. Spengler SJ. Bioinformatics in the information age. *Science* 2000;287:1221–1223.

43. Brenner S. The end of the beginning. *Science* 2000;287:2173–2174.

44. Rubin GM, Yandell MD, Wortman JR, et al. Comparative genomics of the eukaryotes. *Science* 2000;287:2204–2215.

45. Weber WW. *Pharmacogenetics*. Oxford: Oxford University Press, 1997.

46. Poirier J, Davignon J, Bouthillier D, et al. Apolipoprotein E polymorphism and Alzheimer's disease. *Lancet* 1993;342(8873):697–699.

47. Bailey D, Bondar A, Furness LM. Pharmacogenomics—it's not just pharmacogenetics. *Curr Opin Biotechnol* 1998;9:595–601.

48. Meyer U, Zanger UM. Molecular mechanisms of genetic polymorphisms of drug metabolism. *Annu Rev Pharmacol Toxicol* 1997;37;269–296.

49. Bertilsson L, Dahl ML, Tybring G. Pharmacogenetics of antidepressants: clinical aspects. *Acta Psychiatr Scand Suppl* 1997;391:14–21.

50. Dalén P, Dahl ML, Ruiz ML, et al. 10-Hydroxylation of nortriptyline in white persons with 0, 1, 2, 3, and 13 functional CYP2D6 genes. *Clin Pharmacol Ther* 1998;63(4):444–452.

51. Yue QY, Zhong ZH, Tybring G, et al. Pharmacokinetics of nortriptyline and its 10-hydroxy metabolite in Chinese subjects of different CYP2D6 genotypes. *Clin Pharmacol Ther* 1998;64(4):384–390.

52. Catalano M. The challenges of psychopharmacogenetics. *Am J Hum Genet* 1999;65:606–610.

53. Kapitany T, Meszaros K, Lenzinger E, et al. Genetic polymorphisms for drug metabolism (CYP2D6) and tardive dyskinesia in schizophrenia. *Schizophr Res* 1998;32:101–106.

54. Heils A, Teufel A, Petri S, et al. Allelic variation of human serotonin transporter gene expression. *J Neurochem* 1996;66:2621–2624.

55. Lesch KP, Bengel D, Heils A, et al. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 1996;274:1527–1531.

56. Smeraldi E, Zanardi R, Benedetti F, et al. Polymorphism within the promotor of the serotonin transporter gene and antidepressant efficacy of fluvoxamine. *Mol Psychiatry* 1998;3:285–290.

57. Zanardi R, Benedetti F, Di Bella D, et al. Efficacy of paroxetine in depression is influenced by a functional polymorphism within the promoter of the serotonin transporter gene. *J Clin Pharmacol* 2000;20:105–107.

58. Tallman JF. Neuropsychopharmacology at the new millennium: new industry directions. *Neuropsychopharmacology* 1999;20:99–105.

59. http://www.sagenet.org/

60. Kleyn PW, Vesell ES. Genetic variation as a guide to drug development. *Science* 1998;281(5384):1820–1821.

19

Gene Targeting and Transgenic Technologies

Laurence H. Tecott

David S. Johnson

Laurence H. Tecott and David S. Johnson: Department of Psychiatry, Langley Porter Psychiatric Institute, University of California–San Francisco, San Francisco, California 94143-0984.

Recent progress in the development of molecular genetic methods enables the manipulation of genes in intact mammalian organisms. The power of such techniques to elucidate complex biological systems was initially recognized and exploited by developmental biologists and immunologists. More recently, the utility of these approaches for examining neural gene function in the context of the intact organism has led to their use in neuropsychopharmacology. Since the publication of the previous edition of this book, there has been an explosion in the application of molecular genetic technologies to study the regulation of complex behavior and its modulation by psychoactive drugs.

For several decades, the ability to manipulate genes in organisms such as yeast, fruit flies, and nematodes has produced important insights into the regulation of a wide variety of complex biological processes. Limitations in the use of such organisms for research in neuropsychopharmacology arise from the marked organizational differences between the mammalian brain and the systems that govern behavior in these organisms. By contrast, a substantial degree of homology exists in the organization of the central nervous system (CNS) and in the complement of genes expressed across mammalian species. Currently, the mouse genome is by far the most accessible mammalian genome to manipulation. Procedures exist in the mouse for introducing new genes, expressing elevated levels of endogenous genes, and eliminating or altering the function of identified target genes.

Mutant mouse models may be used for a number of purposes relevant to neuropsychopharmacologic research. For example, the impact of genetic mutations on the behavior of mutant mice may be examined, providing insights into the functional significance of particular gene products. In some cases, the manifestations (phenotypes) of these mutations may resemble features of human neuropsychiatric diseases, providing animal models for studying neural processes relevant to such disorders. Furthermore, as genes that confer susceptibility to human diseases are identified, it will be possible to introduce corresponding mutations into the mouse genome, generating useful models for studying disease pathophysiology and treatment. Finally, genetic models will be useful for investigating mechanisms through which nonselective drugs influence neural function and behavior. For example, the contribution of a particular receptor subtype to the actions of a nonselective drug may be examined by studying its actions in animals with targeted loss-of-function mutations of that receptor gene.

This chapter provides an overview of the transgenic and gene targeting approaches used to manipulate the mammalian genome. We have divided these techniques into three categories: (a) transgenic technologies, in which exogenous gene sequences are inserted into the mouse genome; (b) gene targeting technologies, in which mutations are targeted to inactivate or otherwise modify an endogenous gene of interest; and (c) conditional genetic manipulations, in which mutations are restricted to particular stages of development or to particular regions of the CNS. In addition to a brief description of these technologies, examples of their application to neuropsychopharmacology are provided, as well as discussions of the benefits and limitations of each approach.

TRANSGENIC PROCEDURES

Part of "19 - Gene Targeting and Transgenic Technologies "

The ability to insert an exogenous (or foreign) gene into the mouse genome by direct injection into the pronuclei of zygotes was achieved just two decades ago (1) . The term *transgenic* was applied to mice expressing exogenous DNA that had been produced using this technique (2) . With this method, the gene of interest is inserted into a random locus in the mouse genome, and is expressed “in trans,” i.e., not in its usual genetic locus. The techniques required for introducing

transgenes into the mouse genome have been highly refined, permitting their widespread use. Since the development of this technique, many thousands of lines of transgenic mice have been generated, and it has been the most widely utilized technique of genetic manipulation in mice.

Methods of Production of Transgenic Mice

Techniques for producing transgenic mice involve the microinjection of DNA constructs into fertilized mouse eggs (Fig. 19.1) . DNA constructs used for the generation of transgenic mice typically consist of a gene of interest located 3′ to promoter sequences selected to produce a desired distribution of gene expression. The maximum length of the DNA sequence that may be successfully incorporated into the mouse genome is not known, and up to 70 kilobase (kb) DNA fragments have been successfully integrated. The transgene is linearized and purified from prokaryotic vector sequences. For optimal integration efficiency, about 1 to 2 picoliter (pL) of DNA at a concentration of 1 to 2 ng/μL (corresponding to a few hundred molecules of a 5-kb DNA fragment) is microinjected into the male pronucleus of a fertilized mouse egg. Although labor intensive, direct injection of DNA into the pronucleus results in much higher rates of integration of transgenes than other known methods of transformation. After microinjection, the embryos are surgically transferred into the oviduct of pseudopregnant mice. Pseudopregnant females are generated by matings with vasectomized males. The act of copulation initiates the endocrine changes of pregnancy, providing a suitable uterine environment for the survival and implantation of the transferred embyros. The foster mothers give birth 19 to 21 days after oviduct transfer. For genotyping, DNA is typically isolated from mouse tail biopsies and screened for the presence of the transgene by Southern blotting or polymerase chain reaction (PCR). Typically, about 20% to 40% of the mice that develop to term possess the transgene. In the majority of cases, integration of the transgene occurs during the one-cell stage, so that the transgene is present in every cell of the transgenic mouse. Integration usually occurs at a single random chromosomal location, and, for reasons that are not fully understood, there are usually multiple copies of the transgene inserted as head-to-tail concatamers. Mice identified to possess the integrated transgene are referred to as *founders*. The founders are typically used in a breeding strategy to produce animals that are homozygous for the transgene insertion.

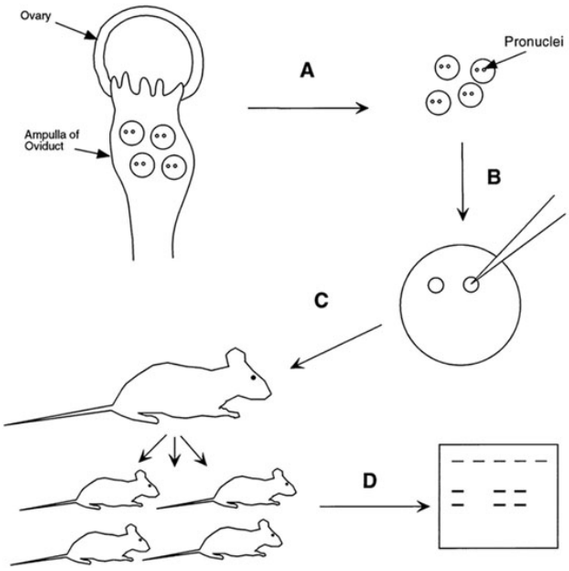


FIGURE 19.1. Procedure for production of transgenic mice. **A:** One-celled fertilized zygotes located in the oviduct ampullae of pregnant donor mice are surgically harvested. **B:** DNA encoding the gene of interest is microinjected directly into the pronucleus of the zygotes. **C:** Injected zygotes are surgically transferred into the oviducts of pseudopregnant female mice. **D:** DNA from the progeny can be analyzed by Southern blot or polymerase chain reaction (PCR) for the presence of the transgene.

Uses of Transgenic Mice

Because transgenic mice often possess multiple copies of the transgene, this method can be used to produce animals with increased levels of expression of particular genes, i.e., mice that “overexpress” genes of interest. In addition, it can be used to express altered forms of a gene product in the distribution of the endogenous gene. One example is a transgenic line bearing a transgene composed of the Ca^{2+}/calmodulin-dependent protein kinase α subunit (CaMKIIα) promoter driving expression of a mutant form of CaMKIIα that conferred Ca^{2+}-independent activation. These mice exhibited an increased stimulation threshold for the induction of synaptic plasticity in the hippocampus, as well as deficits in spatial memory (3, 4) . Studies of these animals led to an enhanced understanding of the role of CaMKIIα in synaptic plasticity and spatial memory acquisition.

In many cases, it is desirable to express a gene with an anatomic distribution that does not mirror its native expression pattern in the mouse. Such ectopic expression of a gene may be achieved using a transgenic construct in which the gene of interest is preceded by promoter elements that direct expression in an anatomic distribution characteristic of another gene. An example of this approach is a transgenic line in which the D1 dopamine receptor promoter was used to drive expression of a cholera toxin subunit (which constitutively activates G_s) in cells that express D1 dopamine receptors (5) . Studies of these animals revealed that chronic overstimulation (by constitutively activated G_s) of forebrain neurons expressing D1 receptors results in an abnormal behavioral phenotype that was likened to human compulsive

behaviors. For most genes, the promoter elements necessary to reproduce the native patterns of expression are not well defined. A useful approach for identifying important promoter elements for genes of interest involves the generation of transgenic mice in which putative promoter sequences are used to direct expression of *reporter genes*, whose expression is readily determined in brain tissue. Comparisons may then be made between the pattern of reporter gene expression and that of the gene of interest (6, 7 and 8) .

It is also possible use transgenic approaches to reduce the expression of a particular gene product. This may be achieved using “dominant-negative” mutations, mutations that induce loss of function of a gene product when expressed in the heterozygous state. For example, transgenic constructs may be designed to express antisense RNAs that hybridize to native messenger RNA (mRNA) sequences, thus decreasing production of the gene product of interest (9, 10 and 11) . Alternatively, the function of gene products that aggregate into multimeric complexes may be disrupted by dominant-negative mutations that produce dysfunctional subunits of the complex. The most prevalent approach used to generate loss of function mutations, gene targeting, is described in the next section.

Considerations in the Interpretation of Transgenic Mouse Phenotypes

An important factor that frequently complicates the interpretation of studies with transgenic mice is the difficulty that may be encountered in achieving a desired anatomic distribution of transgene expression. Promoter elements are often quite large, and additional regulatory elements may at times be located great distances from the gene of interest. In addition, the site of integration often affects the pattern and level of transgene expression, so that founder mice generated with a common targeting construct may display different expression patterns. It may therefore be difficult to accurately duplicate a promoter's endogenous pattern of gene expression in the setting of a transgenic mouse. Commonly, expression patterns are assessed in multiple founders, and those with the most appropriate transgene expression would then be selected for a particular experiment.

Several additional considerations in the interpretation of phenotypes in transgenic mice warrant mention. For example, the number of copies of the transgene incorporated into the genome varies between founder mice. In some cases, concatamers can be unstable and susceptible to deletion of one or more copies of the transgene. In addition, the integration of the transgene may occasionally disrupt an endogenous gene. This could lead to the development of a phenotypic abnormality unrelated to the function of the transgene—this is estimated to occur in 5% to 10% of transgenic mice (12) . This possibility may be assessed by determining whether similar phenotypes are present in animals derived independently from different founders, because the likelihood of two founders possessing the same transgene integration site is minimal.

GENE TARGETING PROCEDURES

Part of "19 - Gene Targeting and Transgenic Technologies "

A mutational approach has proved to be invaluable to investigators examining the roles of gene products in complex biological processes within prokaryotic and cultured eukaryotic cells. Recently it has become possible to apply this approach to a mammalian system. Gene targeting procedures enable the precise (site-specific) introduction of a mutation to one of the estimated 100,000 murine genes. Typically, mutations have been designed to eliminate gene function, resulting in the generation of “knockout” or “null mutant” mice. The introduction of mutations that produce more subtle alterations in gene function has also been achieved. Two major developments have made gene targeting experiments feasible: (a) the generation of totipotent embryonic stem (ES) cells, and (b) the elucidation of techniques to achieve homologous recombination in mammalian cells.

Gene Targeting Methods

ES cells are derived from 3.5-day-old mouse embryos, at the blastocyst stage of development (Fig. 19.2) . Blastocysts are cultured individually under conditions that permit the proliferation of the inner-cell mass cells, which are those cells that would normally become the fetus. These cells are then disaggregated, and individual ES cells clones are grown. Under optimal conditions, ES cells retain the ability to contribute to all of the tissues of the developing fetus. The derivation of ES cells was pioneered using embryos derived from the 129/Sv strain of mice, a strain that has been most commonly used in studies of early embryonic development. Although this mouse strain is not ideal for the study of behavior (see below), most ES cell lines in current use are 129/Sv-derived.

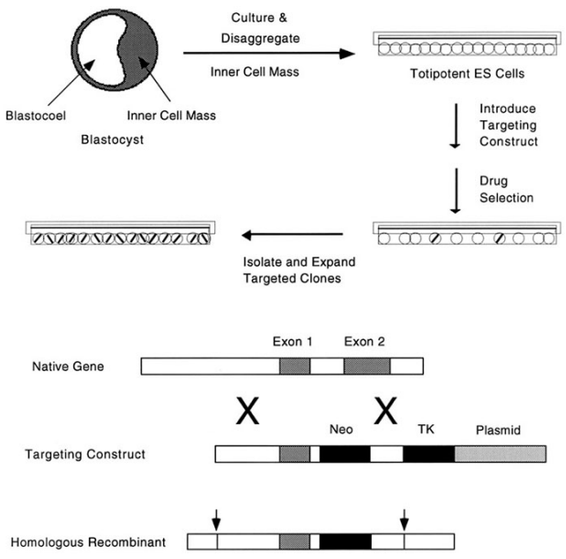


FIGURE 19.2. Gene targeting in embryonic stem (ES) cells. **A:** ES cells are derived from the blastocyst inner cell mass. A targeting construct is introduced into ES cells by electroporation. Cells are subjected to drug selection to enrich for homologous recombinant clones (*striped cells*). Homologous recombinant clones are isolated for blastocyst injection. **B:** A targeting construct is produced in which the second exon of the gene of interest is replaced by a neomycin resistance cassette (Neo). A thymidine kinase (TK) cassette is included for negative selection. Homologous recombination results in the incorporation of engineered mutation into endogenous gene locus. *Arrows* indicate the junction of construct sequences and native locus.

Homologous recombination is the process by which a mutation is targeted to a precise location in the genome. A targeting construct is generated that typically consists of a target gene sequence into which a loss-of-function mutation has been engineered (Fig. 19.2) . Most targeting constructs are designed to achieve homologous recombination events in which recombination at the target locus results in replacement of native target sequences with construct sequences. In mammalian cells, fragments of DNA preferentially integrate into the genome at random locations, at rates that greatly exceed homologous recombination. Therefore, targeting constructs are designed for use in selection strategies that enrich for ES clones in which homologous recombination has occurred. In the commonly used positive-negative selection strategy (13), a portion of a protein-coding exon is replaced by sequences that confer resistance to the drug neomycin. This mutation serves two purposes: (a) to inactivate

the gene product, and (b) to provide a marker that enables the selection of cells that have integrated the construct. This exogenous DNA fragment is flanked by regions of DNA that are homologous to the native gene. Adjacent to one of these homologous regions is a gene encoding thymidine kinase. Treatment with the drug ganciclovir will kill cells that express this gene.

The targeting construct is typically introduced into ES cells by electroporation. In this step, cells are subjected to an electric current that facilitates the internalization of the DNA construct. Those cells that failed to incorporate the targeting construct are killed by the addition of neomycin to the culture medium (positive selection). The majority of the remaining cells have incorporated the entire DNA construct (including the thymidine kinase gene) at random sites throughout the genome. By contrast, during homologous recombination, nonhomologous regions of the construct that are not flanked by homologous sequences are excluded from the integration event. Therefore, homologous recombinant clones will not contain the thymidine kinase gene. Thus, the addition of a second drug, ganciclovir, will selectively kill cells that have randomly incorporated the construct (negative selection), thereby enriching for targeted clones. ES cell clones that survive this double drug selection are then screened for homologous recombination by PCR or Southern blot analysis. The homologous recombinant clones, which are heterozygous for the introduced mutation, are then used to generate chimeric mice.

Following the isolation of homologous recombinant ES cell clones, these cells are microinjected into the fluid-filled blastocoele cavity of 3.5-day-old embryos at the blastocyst stage (Fig. 19.3) . The injected embryos are then surgically transferred into the uterus of pseudopregnant females. These animals will then give birth to chimeric mice, which are derived partly from the injected ES cells and partly from the host embryo. For example, ES cells derived from a brown strain of mice are often injected into embryos derived from black C57BL/6 mice, resulting in chimeras with coats

containing black and brown patches. The extent to which the ES cells have colonized the animal may be roughly approximated by the extent of the brown contribution to the coat. It is most important that ES cell derivatives colonize the germ cells of the chimera, so that the targeted mutation can be propagated to subsequent generations. If chimeras are mated with C57BL/6 mice, then the germ line transmission of ES cell–derived genetic material is indicated by the generation of brown offspring. Half of these brown mice will be heterozygous for the targeted mutation. These heterozygous mice are then bred to produce homozygous mutant mice that completely lack the normal gene product.

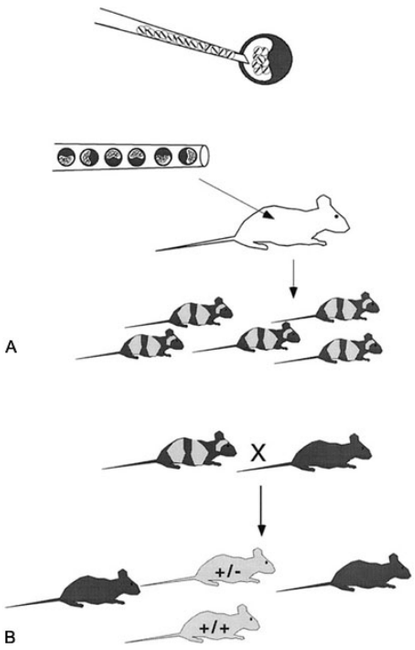


FIGURE 19.3. Generation of gene targeted mice. **A:** Homologous recombinant ES cells are injected into the blastocoele cavity. **B:** Injected blastocysts are surgically transferred into the uterus of pseudopregnant female mice for the production of chimeric mice. **C:** In this example, chimeric mice are bred with C57BL/6 animals. Germ-line transmission is indicated by coat color. One-half of animals with the coat color indicative of the ES cell line will be heterozygous for the targeted mutation and the other half wild type. Heterozygous animals may be bred for the production of homozygous mutant mice.

Uses of Gene Targeted Mice

Studies of null mutant mice provide novel insights into the functional roles of neural genes and, in some cases, animal models relevant to human neuropsychiatric disorders. An illustrative example is a recent study of mice lacking the hypothalamic neuropeptide orexin (14) . Observations of homozygous mutant mice revealed an unanticipated phenotypic abnormality. The mutants displayed frequent episodes of inactivity characterized by the sudden collapse of the head and buckling of the extremities. Subsequent electroencephalogram (EEG) analysis revealed these episodes to be similar to narcoleptic attacks observed in humans and in a strain of narcoleptic dogs. Moreover, a mutation of an orexin receptor gene was found to underlie the canine syndrome (15) . Thus, the null mutant phenotype revealed a novel role for orexin in sleep regulation. In addition, this line of mice represents an important animal model for examining the pathophysiology and treatment of narcolepsy.

Another example illustrates the potential utility of null mutant mice to uncover mechanisms underlying the behavioral effects of psychoactive drugs. The nonselective serotonin (5-hydroxytryptamine, 5-HT) receptor agonist m-chlorophenylpiperazine (mCPP) interacts with several subtypes of 5-HT receptors. Although this compound typically reduces locomotor activity in rodents, it produced a paradoxical hyperlocomotor response in a line of 5-HT_{2C} receptor null mutant mice (16) . This response to mCPP was blocked by pretreatment with a 5-HT_{1B} receptor antagonist, indicating that the absence of 5-HT_{2C} receptors unmasked a hyperlocomotor effect of mCPP on 5-HT_{1B} receptors in mutant mice. These results provide a model whereby genetic endowment may contribute to the development of a paradoxical drug response. When a compound alters the function of multiple gene products with opposing influences on behavior, then mutations or allelic variation of these genes may lead to paradoxical effects.

Although gene targeting techniques are most commonly used to generate animals with null mutations, subtle mutations may also be introduced that alter, but do not eliminate, function. The benefits of such an approach are highlighted by a mutation of a gene encoding the α_1 subunit of the γ-aminobutyric acid A ($GABA_A$) receptor (17) . The mutation produced a single amino acid change, rendering the α_1 subunit–containing subpopulation of $GABA_A$ receptors insensitive to benzodiazepines, without affecting their responsiveness to GABA. The resulting animals exhibited reduced sensitivity to the sedative and amnestic effects of diazepam, but no change in sensitivity to the anxiolytic-like effects of this drug. These results indicate that benzodiazepine site ligands devoid of activity at α_1 subunit–containing $GABA_A$ receptors may act as anxiolytics devoid of some of the side effects typically associated with benzodiazepines, a prediction borne out by a recent report of the behavioral effects

of such a compound (18) . These insights would not have been obtained using a conventional gene targeting approach, because a null mutation of the α_1 subunit gene would profoundly perturb brain GABA signaling.

Considerations in the Interpretation of Targeted Mutant Phenotypes

In interpreting behavioral phenotypes, attention must be paid to the effects of genetic background. The phenotypic consequences of many targeted mutations may be influenced by modifying genes that differ among various inbred strains (19) . In some cases, phenotypic abnormalities have been lost when mutations were bred to a new genetic background (20) . It may therefore be useful to examine the persistence of mutant phenotypes in the context of several genetic backgrounds. In one example, three groups independently generated lines of mice with null mutations of the 5-HT_{1A} receptor subtype (21, 22 and 23) . Interestingly, although each group placed this mutation on a different genetic background, all observed enhanced anxiogenic-like behaviors in the mutant lines. Thus, particularly strong evidence is provided for a contribution of the 5-HT_{1A} receptor to the regulation of anxiety.

Another potential problem arises from the common use of ES cells derived from 129/Sv mice. Mice of this strain are susceptible to structural abnormalities of the CNS, such as agenesis of the corpus callosum, and are impaired in several behavioral assays (19, 24) . This potential problem may be addressed through breeding programs to place targeted mutations on different inbred backgrounds, and by the generation of ES cell lines derived from other inbred strains. It has been recommended that the C57BL/6 and DBA strains be used as standards, due to the extensive body of data relating to the behavioral characterization of these strains (19) .

In addition to the above strain considerations, the standard application of gene targeting technology has several inherent limitations. The null mutations engineered into knockout mice are typically constitutive, i.e., they are present throughout embryonic and postnatal development. Therefore, the potential for developmental perturbations is a major caveat to the interpretation of mutant phenotypes in adult animals. It may be difficult to determine whether a mutant phenotype reflects a normal adult role for the targeted gene or an indirect effect of the mutation attributable to perturbed development. Such an effect may lead to an overestimation of the functional significance of the gene product in the adult animal. Conversely, if significant compensation for the loss of a gene product occurs during development, then the severity of the mutant phenotype may underestimate the functional significance of the gene product. The nature of such compensatory mechanisms and the extent to which they exist may be difficult to determine. The above considerations also pertain to the analysis of transgenic mice carrying constitutive mutations.

Another limitation of the standard gene targeting technology is that the mutations are ubiquitous, present in all of the cells of the animal. Thus, if a neural gene of interest is also expressed in peripheral tissues, then the absence of the gene product peripherally could lead to a lethal or altered phenotype, independent of its neural role. Moreover, for genes that are widely expressed in the CNS, it may also be difficult to anatomically localize the brain region(s) that underlie the mutant phenotypes. New techniques to overcome these problems by achieving region-specificity and inducibility of targeted mutations are under development, and are described in the next section.

PROCEDURES FOR ENGINEERING CONDITIONAL MUTATIONS

Part of "19 - Gene Targeting and Transgenic Technologies "

New technologies are under development for circumventing the limitations of standard gene targeting approaches by creating mutations that may be induced in adult animals and/or restricted to particular brain regions. Although these strategies are not yet in widespread use, it is likely that rapid advances in this area will lead to an exponential increase in the generation of such “conditional mutations” over the next several years.

Cell Type–Specific Mutation Strategies

When a null mutation of a gene results in a mutant phenotype, limitations in the interpretation of that phenotype can arise because the gene is inactivated in every cell in which it was expressed in the mouse in the wild-type (WT) state. Therefore, the observed abnormal phenotype may arise from the absence of the functioning gene product in a peripheral organ system, the peripheral nervous system, or the CNS—i.e., in any of those regions in which the gene is normally expressed. It is possible that the absence of a gene product in the periphery may lead to embryonic lethality, precluding use of the mutant for the studies of neural function. For genes that are widely expressed within the CNS, it may be difficult to identify neural circuits through which mutations produce behavioral perturbations. The ability to inactivate genes in restricted subpopulations of the cells that normally express them will be a valuable asset in studies to uncover the neural mechanisms underlying neural phenotypes.

Recent efforts have focused on a mutational strategy developed to exert spatial control over the pattern of expression of genetic changes introduced into mice. This approach utilizes somatic cell recombination rather than germ cell (or embryonic stem cell) recombination to inactivate a gene in restricted populations of cells or tissues. In this approach, a tissue-specific promoter is used to direct expression of one

of the site-specific recombinases (25) to limit gene inactivation to only those cells expressing the recombinase. The two recombinase systems that have been utilized for genetic manipulation in mice have been the Flp-frt system from yeast (26), and the Cre-lox system from bacteriophage P1 (27, 28 and 29), with the large majority of reports using this technique utilizing the Cre-lox system.

Cre (cyclization of recombination) recombinase is a 38-kd protein from bacteriophage P1, which recognizes and catalyzes reciprocal DNA recombination between two loxP (locus of crossing over of P1) sites. The loxP site is the 34–base pair (bp) recognition sequence for Cre composed of a palindromic 13-bp sequence separated by a unique 8-bp core sequence (Fig. 19.4A) . A gene or gene segment with flanking loxP ("floxed") sites will be excised by homologous recombination in the presence of Cre, leaving a single loxP site marking the point of excision and re-ligation of upstream and downstream DNA (Fig. 19.4B) . This approach, then, involves generating two independent lines of mice—a line bearing loxP sites, and a transgenic line in which Cre expression is driven by a tissue-specific promoter. Animals with a gene or gene region of interest flanked by loxP sites (floxed) are generated by gene targeting. Because the loxP sites are relatively small and placed in intronic regions, they do not typically interfere with normal gene transcription. Of course, WT patterns and levels of expression need to be documented in these floxed mouse lines, because inadvertent placement of lox sites into promoter elements or RNA splice sites could disrupt gene expression. The Cre mice are most commonly generated by creating a transgenic line of mice in which Cre expression is driven by a tissue-specific promoter. As discussed above in the section on transgenic mice, variability in transgenic expression patterns requires several lines of Cre mice that need to be generated and assayed for patterns of Cre expression. An alternative strategy is to use gene targeting procedures to place Cre under the control of an endogenous promoter (30) . The advantage of this approach is that Cre expression should closely approximate the WT expression pattern of the gene it is replacing because the original gene's promoter remains in its endogenous location. A potential disadvantage is that Cre may disrupt expression of the gene into which it has integrated. Once a line exhibiting the desired pattern of Cre expression is identified, it is crossed with an appropriate floxed line to commence a breeding strategy resulting in the generation of animals with a restricted pattern of gene inactivation (Fig. 19.4C) .

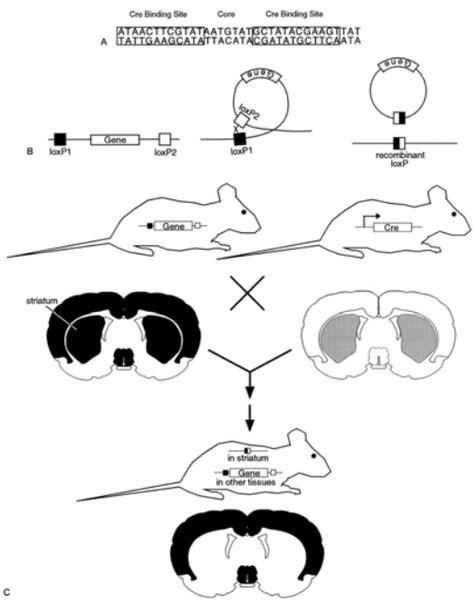


FIGURE 19.4. Strategy for cell-type–selective mutations using Cre-mediated recombination. **A:** The loxP DNA sequence indicating the core region and the inverted repeats that constitute Cre binding sites. **B:** In the presence of Cre, a gene flanked by tandomly oriented loxP sites (floxed gene) will be excised by homologous recombination. The recombination occurs in the core region of the loxP site, leaving a single recombinant loxP site in the genome after Cre excision. **C:** Use of the Cre-lox system to generate cell-type–specific gene inactivation in mice. A "floxed mouse" is generated by gene targeting to introduce loxP sites flanking a gene of interest. The wild-type expression pattern of this gene is shown (*black*) in the coronal section beneath the floxed mouse. In this example, there is expression of the gene in the cortex (ctx), striatum (str), and hypothalamus (hypothal). This floxed mouse can be crossed to a transgenic mouse expressing Cre in a distribution dictated by the promoter used in the transgene construct. In this example, Cre expression is shown (*gray*) and is limited to the striatum. A breeding program is pursued to produce animals in which expression of the floxed gene is normal except in the striatum, where the expression of Cre results in the excision of the floxed gene.

Several lines of Cre mice have been reported in which expression is restricted to subpopulations of cells within the CNS (31, 32, 33, 34 and 35) . The first example of this approach was the inactivation of the glutamate receptor subunit NMDAR1 in CA1 pyramidal neurons of the hippocampus, with expression in other brain areas mostly intact (31) . NMDAR1 is the predominant *N*-methyl-D-aspartate (NMDA) receptor subunit and is widely expressed in most CNS neurons. It had been previously demonstrated that widespread gene inactivation in NMDAR1 null mutants resulted in perinatal lethality (36, 37) . When the mutation was restricted to hippocampal CA1 neurons, animals were viable and exhibited impaired spatial learning and impaired plasticity at CA1 synapses (31) . Thus, spatial restriction of neural mutations can be used to uncover particular brain regions or cell type through which gene inactivation alters behavior.

The utility of this approach for producing cell-type–specific inactivation of genes is enhanced by the fact that the components of the system produced in one laboratory can be "mixed and matched" with components from another laboratory. That is, Cre lines generated for use with a particular floxed gene may also be used with other floxed genes when a similar pattern of gene inactivation is desired. Collaborative efforts to generate databases of Cre and floxed lines will speed and simplify the production of animals with restricted patterns of gene inactivation.

Inducible Mutation Strategies

As described above, the absence of a gene product throughout development complicates the interpretation of mutant phenotypes. Efforts are currently under way to overcome this limitation through the use of gene expression systems that may be induced in the adult animal. Strategies are under exploration for achieving this goal using a variety of compounds, such as tetracycline, steroid receptor antagonists, and ecdysone to induce gene expression. Although these approaches have yet to be optimized for general use, this development is likely to be close at hand. The tetracycline system has been the most utilized and best developed approach to inducible gene regulation.

Since the introduction of the Tet system by the Bujard laboratory in 1992, many laboratories have validated the utility of this approach to inducible gene regulation, and many refinements/improvements in the system have been introduced (38) . This system is based on the regulatory elements of a tetracycline resistance operon of *Escherichia coli*, in which the transcription of tetracycline resistance genes is negatively regulated by the tetracycline repressor (tetR) (Fig. 19.5) . When tetracycline is present, tetR binds the tetracycline and loses its capacity to bind to the operator sequences (tetO) located within the promoter of the tetracycline resistance operon, and transcription is permitted. By creating a fusion protein composed of the tetR and a potent transcriptional activator, VP16, a tetracycline-dependent transactivating factor (tTA) was produced that retained the DNA binding and activation specificity of the tetR. The desired regulatable gene of interest is placed under tetO plus a minimal promoter (P_{min}), that contains the basic promoter elements required for transcription in all cell types. Activation of this system requires the binding of the tTA to the tetO operator sequence (39) . The presence of tetracycline,

or other suitable ligand such as doxycycline, prevents tTA from binding to tetO and activating transcription of the gene of interest. This is referred to as the *tet-off* system—that is, when tetracycline is present, transcription is off. A *tet-on* system has also been developed, in which tetracycline induces transcription of the gene of interest. It utilizes a reverse tetracycline transcriptional activator (rtTA), designed so that it would bind to tetO and activate transcription only in the presence of tetracycline-related compounds (38) . Doxycycline is most frequently used because it is a potent regulator in both the tet-off and tet-on systems (38), and can be easily supplied to mice through their water or food supply (40, 41) .

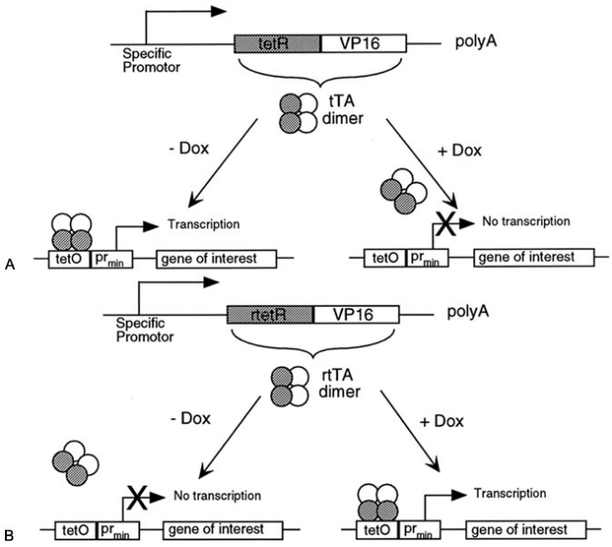


FIGURE 19.5. Tetracycline-regulated expression systems. **A:** Tet-off system. The tetracycline-controlled transactivator, tTA, is a fusion protein consisting of the tetracycline repressor (tetR) domain and a transcriptional activation domain (VP16). tTA homodimerizes, and in the absence of tetracycline (or the tetracycline analogue doxycycline) activates transcription of the gene of interest that has been placed downstream from tetO and a minimal promoter. Binding of doxycycline (Dox) to the tTA dimer prevents the binding of tTA to tetO, and transcription of the gene is prevented. Therefore, the tTA system has been called the Tet-off system, because in the presence of doxycycline, transcription is prevented. **B:** Tet-on system. In the Tet-on system, the tetracycline-controlled activator has been mutated to reverse the action of doxycycline on the transactivator. By contrast with tTA, doxycycline binding to rtTA enables the complex to bind to tetO and activate gene transcription. In the absence of doxycycline, rtTA is unable to bind to tetO and cannot activate transcription. Therefore, the rtTA system is also called the Tet-on system, because doxycycline activates transcription of the regulated gene.

The tet-off and tet-on systems are binary systems—i.e., they require two genetic elements to be introduced into mice. First, a tissue-specific promoter can be used to express tTA or rtTA in a region or cell-type specific manner; then the gene of interest is inserted behind tetO and a minimal promoter. This can be achieved by creating two separate transgenic lines of mice and then cross-breeding to produce bigenic lines. In these lines, expression of the gene of interest may be induced by doxycycline (tet-on) or by the discontinuation of doxycycline treatment (tet-off). For example, the tet-off system has been used to investigate the effects of the transcription factor ΔFosB on psychostimulant responses. A line of mice was generated in which expression of a ΔFosB transgene was suppressed by continuous doxycycline treatment throughout development. Discontinuation of treatment in adult animals led to overexpression of the transgene in the nucleus accumbens and to augmentation of the rewarding and locomotor stimulant properties of cocaine (42) . The utility of the tet-on system has also been demonstrated. For example, a line of mice was developed to examine the role of the Ca^{2+}-activated protein phosphatase calcineurin in synaptic plasticity. Treatment of these animals with doxycycline induced calcineurin overexpression in restricted forebrain regions, associated with deficits of neuronal plasticity and spatial learning (41) .

Rather than generating regulatable gain of function mutants with the Tet system, regulatable loss of function mutants can also be generated by combining the Tet system with the Cre-lox system (43, 44) . In this arrangement, a cell-type–specific promoter drives rtTA expression and Cre is linked to tetO and a minimal promoter. In the presence of doxycycline, Cre is expressed in the cell type specified by the promoter used to drive rtTA expression, and somatic

cell recombination excises floxed DNA fragments in those cells—achieving an inducible cell-type–specific knockout. This inducible knockout approach may be utilized to circumvent concerns discussed above in the interpretation of gene knockout phenotypes.

In these inducible knockout mice it must be remembered that although the excision of the floxed gene can be induced relatively quickly, the appearance of any phenotype resulting from the absence of the gene product will occur gradually, depending on the degradation rate of the relevant mRNA and the half-life of the protein. Another important limitation of strategies utilizing the Tet system relates to the inherent “leakiness” of the tetO operator; i.e., low levels of unwanted gene expression may occur during periods in which gene expression is expected to be turned off. This may be problematic when the inducible transgene is toxic or has significant effects even when expressed at very low levels. Recent findings with tetracycline controlled transcriptional silencers indicate that it may be possible to modify the tet system to substantially reduce unwanted gene expression (45) . In addition, work has begun on alternative inducible gene expression systems with low levels of basal expression. One such system utilizes the insect hormone ecdysone as an induction signal.

SUMMARY

Part of "19 - Gene Targeting and Transgenic Technologies "

The development of transgenic and gene targeting technologies is significantly enhancing understanding of cellular and molecular functions of genes and their contributions to neural processes relevant to clinical disorders. In some instances, novel roles for receptor subtypes have been revealed by the observation of unexpected phenotypic abnormalities in mutant animals. Mutant strains are also providing models for studying the pathophysiology and treatment of particular neuropsychiatric diseases. In addition, some mutant mouse models are useful for investigating the mechanisms of action of psychoactive drugs. Although mutant mouse models represent powerful tools in neuropsychopharmacology, they do not replace older methods of investigation. These models may be best used as components of integrated multidisciplinary research efforts spanning multiple levels of analysis.

This chapter briefly surveyed the most common strategies used for manipulating the mouse genome, and cited the advantages and limitations of each approach. Progress in the development of conditional mutagenesis strategies will address many of the current limitations, and facilitate uncovering the neural mechanisms through which mutations alter neural systems to impact behavior. Although this field is still in its infancy, an exponential increase in the application of molecular genetic technologies is anticipated to contribute substantially to our understanding of brain function in health and disease.

REFERENCES

1. Gordon JW, Sangos GA, Plotkin DJ, et al. Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc Natl Acad Sci USA* 1980;77:7380–7384.

2. Gordon JW, Ruddle FH. Integration and stable germ line transmission of genes injected into mouse pronuclei. *Science* 1981;214:1244–1246

3. Bach ME, Hawkins RD, Osman M, et al. Impairment of spatial but not contextual memory in CaMKII mutant mice with a selective loss of hippocampal LTP in the range of the theta frequency. *Cell* 1995;81:905–915.

4. Mayford M, Qang J, Kandel E, et al. CaMKII regulates the frequency-response function of hippocampal synapses for the production of both LTD and LTP. *Cell* 1995;81:891–894.

5. Cambell KM, Lecea L, Severynse DM, et al. OCD-Like behaviors caused by a neuropotentiating transgene targeted to cortical and limbic D1+ neurons. *J Neurosci* 1999;19:5044–5053.

6. Toth M, Ding D, Shenk T. The 5’ flanking region of the serotonin 2 receptor gene directs brain specific expression in transgenic animals. *Mol Brain Res* 1994;27:315–319.

7. Bahn S, Jones A, Wisden W. Directing gene expression to cerebellar granule cells using gamma-aminobutyric acid type A receptor alpha6 subunit transgenes. *Proc Natl Acad Sci USA* 1997;94:9417–9421.

8. Yang X, Yang F, Fyodorov D, et al. Elements between the protein-coding regions of the adjacent beta 4 and alpha 3 acetylcholine receptor genes direct neuron-specific expression in the central nervous system. *J Neurobiol* 1997;32:311–324.

9. Katsuki M, Sato M, Kimura M, et al. Conversion of normal behavior to shiverer by myelin basic protein antisense cDNA in transgenic mice. *Science* 1988;241:593–595.

10. Pepin M, Pothier F, Barden N. Impaired type II glucocorticoid-receptor function in mice bearing antisense RNA transgene. *Nature* 1992:355:725–728.

11. Jouvenceau A, Potier BBR, Ferrari S, et al. Glutamatergic synaptic responses and long-term potentiation are impaired in the CA1 hippocampal area of calbindin D(28k)-deficient mice. *Synapse* 1999;33:172–180.

12. Meisler MH. Insertional mutation of ‘classical’ and novel genes in transgenic mice. *Trends Genet* 1992;8:341–344.

13. Mansour SL, Thomas KR, Capecchi MR. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* 1988;336:348–352.

14. Chemelli RM, Willie JT, Sinton CM, et al. Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell* 1999;98(4):437–451.

15. Lin L, Faraco J, Li R, et al. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* 1999;98(3):365–376.

16. Heisler LK, Tecott LH. A paradoxical locomotor response in serotonin 5-HT2C receptor mutant mice. *J Neurosci* 2000;20(8):RC711–RC715.

17. Rudolph U, Crestani F, Benke D, et al. Benzodiazepine actions mediated by specific gamma-aminobutyric acid(A) receptor subtypes. *Nature* 1999;401(6755):796–800.

18. McKernan RM, Rosahl TW, Reynolds DS, et al. Sedative but not anxiolytic properties of benzodiazepines are mediated by the GABA(A) receptor alpha 1 subtype. *Nature Neurosci* 2000;3:587–592.

19. Crawley JN, Belknap JK, Collins A, et al. Behavioral phenotypes of inbred mouse strains: implications and recommendations for molecular studies. *Psychopharmacology* 1997;132(2):107–124.

20. Phillips TJ, Hen R, Crabbe JC. Complications associated with genetic background effects in research using knockout mice. *Psychopharmacology* 1999;147(1):5–7.

21. Parks CL, Robinson PS, Sibille E, et al. Increased anxiety of mice lacking the serotonin1A receptor. *Proc Natl Acad Sci USA* 1998;95(18):10734–10739.

22. Ramboz S, Oosting R, Amara DA, et al. Serotonin receptor 1A knockout: an animal model of anxiety-related disorder. *Proc Natl Acad Sci USA* 1998;95(24):14476–14481.

23. Heisler L, Chu HM, Brennan T, et al. Elevated anxiety and antidepressant-like responses in serotonin 5-HT1A receptor mutant mice. *Proc Natl Acad Sci USA* 1998;95:15049–15054.

24. Ozaki HS, Wahlsten D. Cortical axon trajectories and growth cone morphologies in fetuses of acallosal mouse strains. *J Comp Neurol* 1993;336:595–604.

25. Kilby NJ, Snaith MR, Murray JAH. Site-specific recombinases: tools for genome engineering. *Trends Genet* 1993;9:413–421.

26. Dymecki SM. Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. *Proc Natl Acad Sci USA* 1996;93:6191–6196.

27. Sauer B, Henderson N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci USA* 1988;85:5166–5170.

28. Sauer B. Manipulation of transgenes by site-specific recombination: use of Cre recombinase. *Methods Enzymol* 1993;225:890–900.

29. Sauer B. Inducible Gene Targeting In Mice Using The Cre/Lox System. *Methods*. 1998;14:381–392.

30. Voiculescu O, Charnay P, Schneider-Maunoury S. Expression pattern of a Krox-20/Cre knock-in allele in the developing hindbrain, bones, and peripheral nervous system. *Genesis* 2000;26:123–126.

31. Tsien JZ, Chen DF, Gerber D, et al. Subregion- and cell type-restricted gene knockout in mouse brain. *Cell* 1996;87(7):1317–1326.

32. Drago J, Padungchaichot P, Wong JYF, et al. Targeted expression of a toxin gene to D1 dopamine receptor neurons by cre-mediated site-specific recombination. *J Neurosci* 1998;18:9845–9857.

33. Meyers EN, Lewandoski M, Martin GR. An Fgf8 mutant allelic series generated by Cre- and Flp-mediated recombination. *Nature Genet* 1998;18:136–141.

34. Tsujita M, Mori H, Watanabe M, et al. Cerebellar granule cell-specific and inducible expression of Cre recombinase in the mouse. *J Neurosci* 1999;19:10318–10323.

35. Guo H, Mao C, Jin X, et al. Cre-mediated cerebellum- and hippocampus-restricted gene mutation in mouse brain. *Biochem Biophys Res Commun* 2000;269:149–154.

36. Forrest D, Yuzaki M, Soares HD, et al. Targeted disruption of NMDA receptor 1 gene abolishes NMDA response and results in neonatal death. *Neuron* 1994;13:325–338.

37. Li Y, Erzurumlu RS, Chen CSJ, et al. Whisker-related neuronal patterns fail to develop in the trigeminal brainstem nuclei of NMDAR1 knockout mice. *Cell* 1994;76:427–437.

38. Gossen M, Freundlieb S, Bender B, et al. Transcriptional activation by tetracyclines in mammalian cells. *Science* 1995;268:1766–1769.

39. Gossen M, Bujard H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci USA* 1992;89:5547–5551.

40. Chen J, Kelz MB, Zeng G, et al. Transgenic animals with inducible, targeted gene expression in brain. *Mol Pharmacol* 1998;54:495–503.

41. Mansuy IM, Winder DG, Moallem TM, et al. Inducible and reversible gene expression with the rtTA system for the study of memory. *Neuron* 1998;21:257–265.

42. Kelz MB, Chen J, Carlezon WA, et al. Expression of the transcription factor deltaFosB in the brain controls sensitivity to cocaine. *Nature* 1999;401:272–275.

43. St-Onge L, Furth PA, Gruss P. Temporal control of the Cre recombinase in transgenic mice by a tetracycline responsive promoter. *Nucleic Acids Res* 1996;24:3875–3877.

44. Utomo ARH, Nikitin AY, Lee WH. Temporal, spatial, and cell type-specific control of Cre-mediated DNA recombination in transgenic mice. *Nat Biotechnol* 1999;17:1091–1096.

45. Freundlieb S, Schirra-Muller C, Bujard H. A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. *J Gene Med* 1998;1:4–12.

20

Gene Delivery into the Brain Using Viral Vectors

Rachael L. Neve

William A. Carlezon Jr.

Rachel L. Neve and William A. Carlezon, Jr.: Department of Psychiatry, Harvard Medical School and McLean Hospital, Belmont, Massachusetts 02478.

The delivery of recombinant genes into the brain is becoming an increasingly important strategy for answering questions about the molecular mechanisms of brain function. Answers to these questions may be applied to many of the disorders that affect the brain. For example, an understanding of the mechanisms by which repeated exposure to drugs of abuse increases their stimulant and rewarding properties in animal models will almost certainly lead to new ways of treating addiction in humans. If we are able to decipher the molecular events underlying long-term changes in neurotransmitter release, we will find new approaches to diseases such as the epilepsies, in which neurotransmitter release is altered. Knowledge of the molecular means by which neurotransmitters shape neuronal development and plasticity, or how trophic factors regulate neuronal health, will lead to insights into how defects in these pathways cause specific psychiatric and neurodegenerative diseases.

Unfortunately, the brain does not yield easily to genetic intervention. The terminally differentiated state of most neurons in the brain precludes the use of vectors, such as conventional retroviruses, that are dependent on cell replication for stable maintenance in the cell. In addition, the molecular mechanisms of specific brain disorders may be restricted to subsets of neurons at specific times during development and maturity. Therefore, strategies for manipulating gene expression in the brain must utilize vectors that persist stably in postmitotic cells and that can be targeted both spatially and temporally in the nervous system. A number of such gene delivery systems have been developed over the last decade. Rather than giving a superficial overview of the field, this chapter highlights the use of herpes simplex virus (HSV) as a vector for gene transfer into neurons, with emphasis not only on its potential but also on its flaws; compares and contrasts HSV-mediated genetic intervention in neurons with that mediated by other viral vectors; and gives detailed examples of the practical uses of this technology by describing the use of HSV vectors to study the molecular basis for drug addiction.

HERPES SIMPLEX VIRUS: THE PROTOTYPIC VECTOR

Part of "20 - Gene Delivery into the Brain Using Viral Vectors "

HSV possesses multiple features that make it an ideal vector for delivery of genes into the nervous system. In particular, it accepts large molecules of exogenous DNA; it infects nondividing cells from a wide range of hosts with high efficiency; it enables strong expression of foreign genes; it is episomal, and thereby does not cause integration effects; its infection of postmitotic cells is persistent; and HSV-1 particles can be concentrated to relatively high titers. Because of these characteristics of HSV-1, and because it is neurotropic, it is currently one of the best viral vectors available for functional analysis of genes in the nervous system.

Amplicon vs. Genomic HSV-1 Vectors

There are two types of replication-deficient HSV vectors: those in which the foreign DNA of interest is cloned into the viral genome itself (genomic vectors), and those that are composed of a plasmid carrying minimal HSV sequences that allow it to be packaged into virus particles with the aid of a helper virus (amplicon vectors). A number of genes within the wild-type HSV genome are dispensable for its growth in cells *in vitro*. This knowledge was used to create “crippled” recombinant HSV-1 viruses that could be used as vectors for gene transfer into cells (1) . This type of genetically engineered genomic vector has been used by a number of investigators and is described in detail by Fink et al. (2) .

The idea of the amplicon vector originated with the discovery of defective HSV-1 particles (3, 4) that appeared in and interfered with HSV-1 stocks that were passaged at high multiplicities of infection (MOIs). Examination of the

genomes of these defective HSV-1 particles revealed that they carried only a minimal subset of DNA sequences from the wild type genome (3, 4 and 5) . These sequences included an origin of DNA replication and a cleavage/packaging site (the “a” site). It was found that incorporation of these two sequences into a plasmid (the “amplicon”) gave the plasmid the ability to be replicated and packaged into virus particles when it was introduced into a cell that was superinfected with wild-type virus (which supplied HSV replication and virion assembly functions in trans). The plasmid sequences that were packaged into virus particles consisted primarily of 150-kilobase (kb) concatamers of the original plasmid (3, 4, 5 and 6) .

The chief advantage of this amplicon type of vector, which is now packaged with replication-defective helper viruses, is that cloning manipulations are relatively easy due to the small size (5 to 10 kb) of the plasmid. One disadvantage is that production of amplicon vectors requires a co-propagated HSV helper virus, resulting in viral stocks that are a mixture of helper and amplicon viruses. In the past, cytotoxic effects of these stocks limited the amount of vector that could be used to infect cells. Even though the replication-incompetent helper viruses could not cause lytic infections in normal cells, cytopathic effects resulted both from proteins present in the HSV-1 particles and from expression of HSV-1 immediate-early (IE) genes (7) . Occasionally, wild-type HSV-1 revertants appeared during the amplicon packaging process, exacerbating the cytotoxicity of the virus preparations (7, 8) . However, recent improvements in the amplicon packaging procedure, which are discussed in the following section, have largely overcome these problems. The development of a helper virus–free packaging system for the HSV vector (9) has virtually eliminated any lingering cytotoxicity in the preparations, although the helper-free system yields relatively low titers of virus and still needs improvement in that area.

Present-Day Amplicon Vectors: Advantages and Disadvantages

Genomic and plasmid defective HSV-1 vectors have been used to manipulate neuronal physiology both *in vitro* and *in vivo*. These studies have been promising, but they have also revealed limitations of the current HSV vector systems. Because HSV is too fragile to purify on cesium chloride gradients, it could not be concentrated in the same way that encapsulated viruses such as adenovirus were concentrated. Moreover, as noted above, nonspecific cytopathic effects of the defective vectors restricted the number of viral particles that could be used to infect neurons. Finally, lack of persistence of high expression levels from the viral recombinants has hampered long-term *in vivo* studies and has limited the usefulness of the vectors for both experimentation and gene therapy.

Recent improvements in the amplicon packaging procedure (Table 20.1) have corrected some of the limitations listed above. The most widely used second-generation helper virus was HSV-1 tsK, with a temperature-sensitive single-base mutation in the *ICP4* (*IE3*) gene. Because revertants of this mutant arose at a finite frequency during the packaging procedure, lytic virus was present in some preparations (7, 8) . The frequency of revertants was decreased with the development of an efficient packaging system using a deletion mutant of *IE3* (8) as helper virus. However, occasional lytic virus particles continued to appear, albeit at a greatly reduced frequency, presumably as a result of recombination between the helper virus and the sequences flanking both sides of the *IE3*-containing fragment present in the permissive host.

A key breakthrough was made by Lim et al. (10) when they compared three replication-defective HSV-1 mutants [KOS strain 5dl1.2, deleted in the *ICP27* (*IE2*) gene; and strain 17 D30EBA and KOS strain dl120, both deleted in the *ICP4* (*IE3*) gene] for their usefulness as helper virus for packaging an amplicon vector. Historically, *IE3* mutants have been preferred because they express fewer HSV-1 genes under nonpermissive conditions than do ICP27 mutants. However, Lim et al. found that use of the *IE2* mutant 5dl1.2 yielded higher vector titers than did use of the *IE3* mutants, with no increase in cytotoxicity. In addition, wild-type lytic virus was virtually absent in stocks made using 5dl1.2 in conjunction with the permissive host 2-2 cells, likely due to the fact that 5dl1.2 is a more complete deletion than D30EBA, sharing little sequence with the *IE2*-containing fragment present in the permissive host.

To achieve a favorable ratio of recombinant vector to helper virus, the stocks derived from transfection of the packaging cells followed by superinfection with helper virus are passaged three times on the permissive host. The recombinant vector is packaged as long concatamers, which contain multiple origins of replication, conferring a selective replicative advantage on the vector-containing virus relative

to the helper virus. Therefore, the efficiency of the initial transfection of vector DNA into the packaging line is critical to the success of the packaging. Lim et al. (10) showed that the transfection of the vector DNA at the start of the packaging procedure was significantly more efficient using Lipofectamine (Life Technologies) than using calcium phosphate, and thereby achieved a favorable ratio (≥1) of vector to helper. Since then, we have achieved vector/helper ratios greater than 100.

For a viral vector to have utility for gene therapy, cytopathic effects of the virus must not outweigh the beneficial effects of the transgene. Unfortunately, in the past, investigators have had difficulty generating nontoxic HSV-based replication-defective vectors. They often were toxic to neurons *in vitro* (7) . Significant necrosis, often accompanied by inflammation and gliosis, was identified at injection sites with some genetically engineered HSV-1 vectors used for *in vivo* studies (see, e.g., ref. 11) . However, the achievement of a more favorable ratio of vector to helper, and the virtual elimination of wild-type virus in the vector preparations (10) has greatly reduced the cytotoxicity of present-day defective HSV-1 amplicon vectors (see, e.g., refs. 12, 13 and 14) . An additional improvement to the packaging procedure, the banding of the virus on a sucrose step gradient, followed by a high-speed centrifugation to pellet the virus, has reduced further the cytotoxicity of the virus preparations. It simultaneously removes toxic factors present in the crude cell lysates, and enables concentration of the vector to titers exceeding 10^8/mL.

A troubling problem that has not yet been resolved for any viral vector used in the brain, except perhaps for the lentiviral vector (see below), is that of persistence of expression. Numerous investigators have had the experiences reported by During et al. (15) and Lim et al. (10), in which an initial peak in expression of an HSV transgene *in vivo* or *in vitro*, respectively, has been followed by loss of the bulk of the expression by 1 to 2 weeks postinfection. Interestingly, superinfection with helper virus 5dl1.2 1 week postinfection rescued expression of a transgene expressed under the control of the IE 4/5 promoter (10) . Apparently, transactivating factors provided by the helper virus reactivated transcription of the transgene.

Two recent developments suggest that the problem of persistence of expression is not insoluble. Use of a 9-kb fragment of the tyrosine hydroxylase promoter to drive reporter gene expression in an HSV-1 amplicon vector resulted in prolonged gene expression *in vivo* (16), suggesting that neuronal, unlike viral, promoters in HSV-1 vectors have the potential to produce stable gene expression. Additionally, the development of hybrid amplicons that incorporate elements that allow autonomous replication of the episome (17) or that incorporate adeno-associated virus (AAV) elements for genomic integration of the amplicon (18, 19 and 20) have resulted in vectors that support long-term gene expression both *in vitro* and *in vivo*.

An improvement in gene transfer methods in general has been the incorporation of the gene for the green fluorescent protein (GFP) into many vectors. Coexpression of GFP allows the investigator to detect cells infected by the vector with a fluorescence microscope, whether they are fixed or alive. There are at least three ways to “mark” vector-infected cells with GFP (Fig. 20.1) . In the example shown, the objective is to coexpress GFP with the AMPA receptor GluR1. They can be expressed on a single transcript (HSV-IRES/GFP) by putting an internal ribosome entry site (IRES) between the two genes, which are then transcribed from the same promoter. The IRES enables independent translation of the two coding sequences even though they are present on the same messenger RNA (mRNA). In theory, the two coding sequences should be expressed at similar levels, but in practice the translation of one of the two coding sequences on the mRNA often occurs at the expense of the other. Alternatively, they can be expressed from two independent transcriptional units, in a bicistronic vector. The addition of an extra transcriptional cassette to the vector makes it larger and more unwieldy to use; and the level of transcription from one promoter is independent of the level of transcription from the other, so that the transgenes may be expressed at very different levels. Third, the GFP can be fused to the transgene product, so that they are expressed

as a single protein. This is a trickier construction, since the two coding sequences must be placed in frame with each other. However, the added benefit of being able to track the subcellular location of the transgene product makes this the option of choice in many instances.

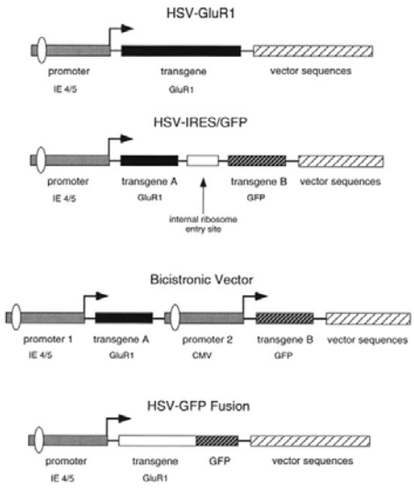


FIGURE 20.1. Different vector constructions for coexpressing two genes from a viral vector. Glutamate receptor GluR1 and green fluorescent protein (GFP) are used as examples.

COMPARISON OF HSV-1 WITH ALTERNATIVE VECTORS FOR GENE TRANSFER INTO NEURONS

Part of "20 - Gene Delivery into the Brain Using Viral Vectors "

Despite the problems that remain with the HSV-1 vector, it is a gene delivery system that has come of age. In addition, numerous alternative and increasingly user-friendly means of gene transfer into the brain are now available (Table 20.2) . Adenovirus vectors, like HSV-1 vectors, infect postmitotic cells and can enter a broad range of mammalian cell types. They have the additional advantages that they can be concentrated to very high titers (≥10^{10}/mL). However, the use of adenovirus vectors continues to be restricted by the robust host immune response that they elicit (21, 22) .

TABLE 20.1. CHRONOLOGY OF IMPROVEMENTS IN AMPLICON VECTOR PACKAGING SYSTEM

1985—Use of wild-type HSV-1 as helper virus and first amplicon plasmid (5)
1988—Use of HSV-1 *tsK* IE3 mutant as helper virus (64)
1990—Use of D30EBA IE3 deletion mutant and M64A complementing cell line (8)
1996—Use of 5*dl*1.2 IE2 deletion mutant and 2-2 complementing cell line (10)
1996—Improvements in transfection procedure and development of methods to purify and concentrate the virus result in amplicon/helper ratios of up to 100:1 (previously 1:100) and titers of up to 2×10^8 infectious units/mL
1997—Helper-free packaging by transient transfection of HSV cosmids (19)

HSV, herpes simplex virus.

TABLE 20.2. COMPARISON OF VIRUSES USED TO MANIPULATE GENE EXPRESSION IN THE BRAIN

Advantages	Disadvantages
Herpesvirus vectors	
Broad host and cell type range	Occasional cytotoxicity
Episomal (no possibility of insertional activation of host genes)	Lack of persistence of expression
Can accommodate up to 15 kb of foreign DNA	
High level of expression of foreign genes within hours	
Can be concentrated to high titers	
Helper virus–free stocks possible	
Adenovirus vectors	
Broad mammalian host and cell type range	Elicits host immune response
Episomal (no possibility of insertional activation of host genes)	
Growth to high titers (~10^{10}/mL)	Accommodates <10 kb of foreign DNA
High level of expression of foreign genes	
Expression is relatively persistent	Genetic manipulation is unwieldy
Adeno-associated vectors	
Broad mammalian host and cell type range	Can accommodate only 4.7 kb of foreign DNA
Nonpathogenic	
Helper virus–free stocks possible	
Expression is relatively persistent	
Lentivirus vectors	
Integrates into host chromosome	Can accommodate only 6–8 kb of foreign DNA
Expression is persistent	Low titers
	Potent human pathogen

kb, kilobase.

At present, HSV amplicon vectors can accommodate larger pieces of foreign DNA, on the order of 15 kb, than can the adenovirus vectors, which can only contain a maximum of 6 to 8 kb (although the new "gutless" adenovirus vectors can take up to 37 kb of foreign DNA). Foreign genes are cloned into easy-to-manipulate amplicon vectors that can be packaged directly into viral particles as head-to-tail repeats in the presence of the helper virus, with no intermediate recombination step required. This enables rapid construction of a large number of recombinant vectors simultaneously, and is particularly useful for those who are doing mutation analysis, and who wish to work with multiple genes. Such ease of cloning is not possible with the genomic HSV and adenovirus vectors.

Direct *in vivo* transfer of genes into the brain has been achieved using not only herpes virus and adenovirus vectors, but also adeno-associated virus vectors (see refs. 23, 24 for review) and lentivirus vectors (see ref. 25 for review). In contrast to other viral vectors, adeno-associated virus vectors do not cause an immune response or toxicity. Interestingly, the ability of adeno-associated virus vectors to transduce and express transgenes is not equivalent in all regions of the brain (26, 27) . Only in some regions will neurons bind and internalize the virus, with resultant long-term expression. Stability of expression of lentivirus vectors in the brain is their greatest advantage. Long-term expression of B-galactosidase and GFP was observed in rat neurons for at least 9 months following intracerebral injection of the vectors, with no sign of tissue pathology or immune response (28, 29) . Progress has been made in achieving biosafety with these vectors, by eliminating viral sequences nonessential for transduction.

GENE DELIVERY INTO THE BRAIN AS A TOOL FOR NEUROPSYCHIATRIC RESEARCH: DRUG ADDICTION

Part of "20 - Gene Delivery into the Brain Using Viral Vectors "

Overview

The bulk of the published work on gene transfer technologies has focused on their use for gene therapy (see ref. 30 for an excellent review). To date, the delivery of recombinant genes into the brain as a strategy for answering questions about the molecular mechanisms of brain function has utilized primarily HSV vectors. One of the most successful uses of this strategy has been in the field of addiction research. Exposure to drugs of abuse causes many changes in gene expression within the brain. A major challenge in addiction research is to determine which of these changes have a direct influence on behavior. Viral vectors offer the ability to study individual changes in gene expression in

discrete brain regions. In the case of addiction, it is thus possible to mimic certain aspects of the drug-exposed state without ever administering the drugs themselves. The ultimate goal of such studies is to understand the “biobehavioral” mechanisms of addiction, that is, to establish direct, causal relationships between drug-induced changes in biology and drug-induced changes in behavior. Examples of behavioral changes in addiction that may result from drug-induced alterations in biology (gene expression) are compulsive drug use (drug-taking) and craving (drug-seeking).

Biobehavioral Studies of Addiction

Addiction Circuitry

Much research on the neuronal circuitry involved in drug addiction has focused on the mesolimbic dopamine (DA) system. The dopaminergic projection from the ventral tegmental area (VTA) of the midbrain to the nucleus accumbens (NAc) of the forebrain has been implicated in the habit-forming (rewarding) effects of many types of abused drugs, including stimulants (cocaine, amphetamine) and opiates (heroin, morphine) (31, 32) . However, the neural events that mediate the acute rewarding effects of abused drugs are not understood, nor are the neuroadaptations that presumably underlie the transition from occasional drug use to compulsive drug use. In rats, repeated exposure to drugs of abuse appears to cause increases in sensitivity (“sensitization”) to the rewarding effects of drugs (33, 34 and 35), a phenomenon that may contribute to the addiction process (36) . This altered sensitivity is presumably a consequence of altered gene expression, and the VTA-NAc circuitry is a logical starting point for biobehavioral studies. Because several robust and reliable drug-induced neuroadaptations have been discovered within this circuitry (37), it has been the focus of gene transfer studies in which the behavioral significance of altered gene expression has been assessed. To date, the biobehavioral significance of three specific, drug-induced changes in gene expression have been studied using viral-mediated gene transfer—the ability of drugs (cocaine, morphine) (a) to increase expression of the AMPA (glutamate) receptor subunit GluR1 in VTA, (b) to alter expression of GluRs in the NAc, and (c) to increase the activity of the transcription factor CREB (cAMP response element binding protein) in the NAc.

GluRs in the VTA

AMPA receptors are made up of various combinations of the subunits GluR1, GluR2, GluR3, and GluR4 (collectively called GluRs) (38, 39) . Repeated intermittent exposure to morphine selectively elevates expression of GluR1 in the VTA (40) . Relative levels of GluR1 and GluR2 expression in the dopamine neuron-rich VTA are important because the subunit composition of AMPA receptors controls their function. High expression of GluR1 favors the formation of Ca^{2+}-permeable (GluR1-homomeric) AMPA receptors (13, 38, 39), which presumably increases sensitivity to the excitatory (depolarizing) effects of glutamate (12) . Conversely, high expression of GluR2 favors the formation of Ca^{2+}-impermeable (heteromeric) AMPA receptors, since this subunit contains a motif that blocks Ca^{2+} conductance (38) . Because repeated drug exposure is known to selectively increase the electrophysiologic responsiveness of VTA dopamine neurons to AMPA receptor agonists (41), it is possible that drug-induced alterations in GluR1 expression in this region contribute to drug-induced behavioral changes, such as sensitization. However, this question is difficult to address using traditional methods. For example, AMPA agonists and antagonists cannot be used to study the relationship between GluR1 expression and sensitized drug responses because they affect AMPA receptor function generally and thus do not mimic morphine's selective effects. This biochemistry-behavior relationship could be studied by performing intra-VTA microinjections of HSV-GluR1, which caused ˜75% increases in GluR1 expression within this region (13) .

It was first necessary to determine if viral-mediated elevations in GluR1 expression within the VTA would increase sensitivity to the locomotor-stimulating effects of morphine, a hallmark of behavioral sensitization (36, 42) . Low doses of morphine induced significantly more activity in rats with viral-mediated elevations in GluR1 expression within the VTA. Furthermore, increased activity in rats given HSV-GluR1 was seen only in response to morphine, and was not evident after saline. The transient nature of the HSV-mediated elevations in transgene expression allowed the behavioral adaptations to be correlated with the time course of GluR1 expression: when morphine was given only on days 7 and 8 after microinjection—when elevations in GluR1 expression had dissipated—the rats that were given HSV-GluR1 were no longer more sensitive to the stimulant effects of morphine. When some rats given HSV-GluR1 were tested with morphine on days 3 to 4 and on days 7 to 8, significant increases in sensitivity to the locomotor-stimulating effects of the persisted in rats given HSV-GluR1, despite the fact that GluR1 labeling in the VTA had dissipated. Thus when morphine is given while GluR1 expression in the VTA is elevated, increased sensitivity to morphine outlasts viral-induced increases in GluR1 expression. These data suggest that altered expression of GluR1 in the VTA underlies, at least in part, the development and expression of sensitized behavioral responses to morphine.

The effect of elevated GluR1 expression in the VTA on the rewarding effects of morphine was examined using place conditioning, a classic conditioning procedure in which rats learn to associate the rewarding effects of a drug with a distinctive environment. Rats given HSV-GluR1 into the rostral (anterior) portion VTA spent more time in morphine-associated environments than did control rats, indicating

an increase in morphine reward (13, 43) . Since prior treatment with morphine intensifies its rewarding actions in the place-conditioning paradigm (33), these data suggest that the behavioral consequences of morphine preexposure are mimicked by HSV-mediated expression of GluR1 in the VTA.

Together, these data demonstrate that specific changes in motivational states can result from altered expression of a single, localized gene product. Drug-related increases in GluR1 expression in the VTA, a region known to be involved in the induction of sensitization (42, 44), may themselves be sufficient to explain sensitization (13, 41), or they may lead to Ca^{2+}-dependent adaptations (45) that also contribute to changes in drug sensitivity (Fig. 20.2) . Thus these studies have added strength to the hypothesized association between the VTA and sensitization, and identified biobehavioral relevance for the drug-induced regulation of the GluR1 protein in the VTA.

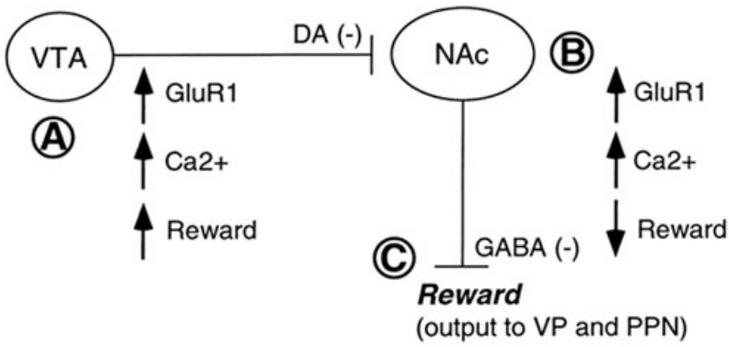


FIGURE 20.2. Simplified schematic of putative reward circuitry; see ref. 39 for detailed discussion. Treatments that increase excitation in the ventral tegmental area (VTA) (**A**) are rewarding, presumably because they promote the inhibitory actions of dopamine (a D2-like receptors) in the nucleus accumbens (NAc) (**B**). Inhibition of NAc GABAergic output neurons, in turn, decrease inhibitory influences on reward processes in other areas of brain reward circuitry (**C**), including the ventral pallidum (VP) and peduncular pontine nucleus (PPN) (63). Elevations in GluR1 expression in the VTA (**A**) increase drug reward, presumably because the accompanying changes in Ca^{2+} flux increase the excitability and/or neuronal function of VTA dopaminergic neurons (as in ref. 49). Conversely, elevations in GluR1 in the NAc decrease drug reward (**B**), presumably because the accompanying changes in Ca^{2+} flux increase the excitability of NAc GABAergic neurons that normally inhibit reward processes in distal regions (C). (Based on refs. 13, 43, and 55.)

GluRs in the NAc

There are two important reasons for interest in how altered expression of GluRs in the NAc affects the actions of drugs of abuse. First, GluR2 is a target gene of ΔFosB, a stable and long-lasting variant of the fos family of transcription factors that is regulated in the Nac by drugs of abuse (46) . Second, GluR1 expression is elevated the NAc in cocaine-sensitized rats during long-term drug withdrawal (47) .

Exposure to cocaine and other drugs of abuse causes the rapid and transient expression of the immediate-early gene c-*fos* in the striatum (including the NAc) (48, 49) . Repeated drug exposure decreases expression of the transient forms of c-*fos*, but increases expression of a more stable and long-lasting form of the transcription factor, ΔFosB (50, 51) . The accumulation and sustained transcriptional activity of ΔFosB in the NAc could mediate long-lasting neural and behavioral adaptations that accompany repeated drug exposure (37) . Consistent with this notion, inducible transgenic mice that express ΔFosB spontaneously (i.e., without prior drug treatment) in the NAc during adulthood have increased sensitivity to the locomotor-stimulating and rewarding effects of cocaine (46) . The proximal cause of the increase in drug sensitivity is presumably not increased expression of ΔFosB per se, but rather increased expression of a target gene (or genes) whose transcription is regulated by this factor. The ΔFosB-overexpressing mice also had large increases in GluR2 expression in the NAc, implicating this AMPA receptor subunit in the increased drug sensitivity. To examine whether elevated GluR2 expression in the NAc was sufficient to cause increases in sensitivity to the rewarding effects of cocaine, rats were tested in the place-conditioning paradigm after microinjections of HSV-GluR2 into this region (46) . This treatment dramatically increased sensitivity to the rewarding effects of cocaine, mimicking the effects of increased expression of ΔFosB. Together, these findings provide strong evidence that the increase in cocaine sensitivity seen in ΔFosB transgenic mice is attributable, at least in part, to elevated expression of GluR2 in the NAc.

For comparison, rats were given microinjections of vectors expressing other GluRs into the NAc, and tested with cocaine in the place-conditioning paradigm. Rats given microinjections of HSV-GluR1 into the NAc spent dramatically less time than control rats in the cocaine-associated environments, suggesting that elevated expression of this AMPA receptor subunit in this region increases sensitivity to the aversive effects of the drug. Additionally, some rats were tested after intra-NAc microinjections of HSV-GluR2Q, which expresses unedited GluR2. This form of GluR2 lacks the final transcriptional edit (Q → R) that produces the motif that blocks Ca^{2+} flux (38, 39) . Use of this construct showed that the ability of GluR2 to increase cocaine reward appears to be directly related to diminished Ca^{2+} permeability, because overexpression of GluR2Q does not increase cocaine reward, but rather causes effects that more closely resemble those of increased GluR1 expression (as would be expected).

There are many possible explanations for how altered Ca^{2+} flux in the NAc might influence drug reward, considering the role of Ca^{2+} in cellular functions including membrane depolarization, neurotransmitter release, signal transduction, and plasticity (52, 53) . Certainly, cocaine-induced changes in the excitability of NAc neurons have been reported:

repeated exposure to cocaine makes neurons in this region significantly less excitable than normal at short (3-day) withdrawal periods (54) . Studies with ΔFosB (46) suggest that these electrophysiologic adaptations are associated with increases in the rewarding efficacy of cocaine, because elevations in GluR2 expression (which would be expected to minimize Ca^{2+} flux and/or neuronal excitability) increase cocaine reward, whereas elevations in GluR1 (which would be expected to increase Ca^{2+} flux and/or neuronal excitability) decrease (or oppose) cocaine reward.

Together, these data support the working hypothesis (55) that altered Ca^{2+} flux and/or neuronal excitability in the NAc has important consequences on motivated behaviors (Fig. 20.2) . Moreover, they suggest that altered GluR1 expression in this region seen during long-term (3-week) cocaine withdrawal (47) might also be associated with important changes in the rewarding efficacy of the drug. Regardless, the use of HSV vectors has identified biobehavioral relevance for the drug-induced regulation of GluRs in the Nac.

CREB in the NAc

Chronic cocaine exposure increases 3′,5′-cyclic adenosine monophosphate (cAMP) formation and protein kinase A (PKA) activity in the NAc (37) . Direct stimulation of PKA in the NAc counteracts the rewarding properties of cocaine (56), suggesting that drug-induced up-regulation of the cAMP system is a neural mechanism of drug tolerance. Increased PKA activity leads to increased CREB phosphorylation, which activates CREB-mediated gene transcription and could be an important step in producing long-lasting neuroadaptations. To determine the functional role of CREB and its transcriptional consequences in the NAc, its expression in this region was increased directly by microinjecting HSV-CREB (57) . In other rats, a dominant negative mutant CREB (mCREB) was overexpressed, which is transcriptionally inactive and competes with endogenous CREB for cAMP response element binding sites (CREs) (58) .

The effects of elevated CREB expression in the NAc on cocaine reward was studied with place conditioning. Although the effects of a low (threshold) dose of cocaine were not altered by control treatments, this dose established dramatic conditioned place preferences in rats given bilateral microinjections of HSV-mCREB (which acts as a CREB antagonist) into this region. This rewarding effect was “inversed” to aversion in rats with elevated expression of CREB in the NAc; rats given HSV-CREB avoided drug-associated environments, suggesting that this dose of cocaine was made aversive by gene transfer. When cocaine was administered a week (rather than 3 days) after microinjections of the HSV vectors into the NAc, cocaine was devoid of rewarding or aversive effects. This finding confirms that the behavioral consequences of HSV viral vectors are transient and reversible, and have a time course of efficacy that parallels that of transgene expression.

Dose-response analyses revealed that microinjections of HSV-mCREB and HSV-CREB in the NAc were producing, respectively, approximately parallel leftward (more rewarding) and rightward (less rewarding) shifts in the effects of cocaine. At a high dose of cocaine, there were no differences in the preferences for the drug-associated environment between rats given HSV-mCREB and those given vehicle, consistent with observations that there is an upper limit to the magnitude of place preferences that can be established (59) . Treatment with high doses of cocaine established place preferences in some rats given HSV-CREB, suggesting that the aversive consequences of increased levels of CREB in the NAc can be counteracted by more drug.

One explanation for these findings is that elevated CREB expression in the NAc increases local dynorphin function. Dynorphin is the endogenous ligand for κ opioid receptors (60), and κ opioid agonists have aversive actions in the nucleus accumbens shell (NASh) (61) . To determine if dynorphin is involved in the cocaine aversion caused by HSV-CREB, brain receptors for dynorphin were blocked with the long-lasting κ receptor antagonist norBNI. Treatment with norBNI [intracerebroventricular (ICV)] before cocaine place conditioning blocked the aversive effects associated with cocaine in animals given HSV-CREB into the NAc, but not in rats given microinjections of vehicle or HSV-mCREB. The fact that only the aversive properties of cocaine are altered significantly by *nor*-Binaltorphimine (norBNI) suggests that microinjections of HSV-CREB into the NAc enhance the aversive aspects of cocaine via increased stimulation of κ opioid receptors by dynorphin.

These results suggest that drug-induced increases in CREB activity (62) is a homeostatic change that opposes drug reward. Mimicking increases in CREB activity by increasing levels with HSV-CREB or by stimulating PKA-induced phosphorylation (56) decreases the rewarding effects of cocaine. Moreover, these data implicate κ opioid receptors in the valence (reward versus aversion) of cocaine action, and suggest that CREB-mediated transcription in the NAc is a “drug reward rheostat” (Fig. 20.3), in part via effects on dynorphin expression. These data also suggest a sequence of D1 receptor–mediated intracellular events, culminating with altered gene transcription, through which exposure to cocaine influences subsequent responsiveness to the drug. Augmented release of dynorphin could inhibit local DA release through actions at κ opioid receptors on terminals of mesolimbic DA neurons that innervate the NAc (61) . Diminished release of dopamine in the NAc may itself be aversive, or it may unmask other actions of cocaine that are aversive or that oppose drug reward. Regardless, these viral vector studies have identified biobehavioral relevance for alterations in CREB function in the NAc.

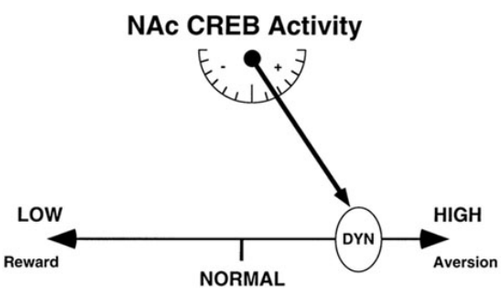


FIGURE 20.3. Schematic depiction of CREB activity in the NAc functioning as a cocaine “reward rheostat.” Elevated expression of CREB increases CREB-mediated transcription of dynorphin (a measure of CREB activity). Elevated dynorphin, in turn, decreases cocaine reward at high doses of drug, and makes cocaine aversive at low doses of drug. Conversely, disruption of CREB activity by overexpression of dominant-negative CREB (mCREB) decreases dynorphin transcription, which increases cocaine reward. (Based on ref. 57.)

Conclusions

The use of viral-mediated gene transfer in addiction research is leading to an understanding of where certain changes in gene expression occur within the cascade of molecular events that lead to the addicted phenotype. This approach complements and extends the predominantly pharmacologic approaches that previously have been used in addiction research, and in fact can be thought of as “genetic pharmacology.” Understanding the molecular basis of addiction will facilitate the development of effective pharmacologic treatments that specifically target proteins, enzymes, or transcription factors within the addiction cascade. These therapeutics may be the prototypes for a new generation of “smart” pharmacotherapies that are designed to negate or even reverse changes in the molecular structure of the brain that characterize specific brain disorders.

GENE DELIVERY INTO THE BRAIN AS A MEANS FOR GENE THERAPY

Part of "20 - Gene Delivery into the Brain Using Viral Vectors "

The recent rapid advancements in gene transfer technologies have raised hopes that central nervous system (CNS) gene therapy, the introduction of genes into the brain to ameliorate neuropsychiatric diseases, is closer to reality. However, a number of major methodologic advances must be made before it can become a reality. First, and most important, stability of transgene expression must be achieved. Second, not only stability but also inducibility and regulatibility of transgene expression are a priority, since the level of transgene product is often critical. Third, the transgene capacity of most vectors must be increased, so that not only the gene(s) of interest but also appropriate regulators or inducible promoters can be delivered to the brain. Fourth, because of the small volume of material that can be delivered stereotactically, it will be necessary to increase both the viral titers and the transduction efficiencies for all the known vectors. Fifth, a high degree of cell specificity of gene transfer must be achieved, by the use of targeted vectors that selectively infect particular cell types, cell-specific promoters, and routing via normal neuronal projections in the brain. Finally, nontoxic vectors that do not induce an immune response must be developed.

The development of gene therapy for neuropsychiatric diseases suffers, in addition, from many of the same problems that drug therapeutics research on these disease does. Neurodegenerative disorders such as Alzheimer's disease pose particular problems for gene therapy because neurons in the CNS cannot undergo regeneration. Therefore, gene therapeutic approaches must target the remaining brain cells, provide suitable replacements for the dying cells, or enable regeneration of CNS neurons. Another set of hurdles arises from the complex etiology of most neuropsychiatric disease. It is not clear that a single gene product will cure any of these diseases. In addition, the molecular mechanisms of different neuropsychiatric diseases may be restricted to subsets of neurons at specific times during development and maturity. Consequently, as noted above, optimal strategies for gene therapy must utilize vectors that persist stably in postmitotic cells and that can be targeted both spatially and temporally in the nervous system. Present-day viral vectors have come of age for use in basic research, but vectors useful for gene therapy in the brain are still a work in progress.

ACKNOWLEDGMENTS

Part of "20 - Gene Delivery into the Brain Using Viral Vectors "

We thank Dr. Frederick Boyce for helpful discussions. This work is supported by National Institutes of Health (NIH) grants AG12954 and HD34563 to R.L.N.

REFERENCES

1. Roizman B, Jenkins FJ. Genetic engineering of AOR1 genomes of large DNA virus. *Science* 1985;229:1208–1214.

2. Fink DJ, DeLuca NA, Goins WF, et al. Gene transfer to neurons using herpes simplex virus-based vectors. *Annu Rev Neurosci* 1996;19:265–287.

3. Spaete R, Frenkel N. The herpes simplex virus amplicon: a new eucaryotic defective-virus cloning-amplifying vector. *Cell* 1982;30:305–310.

4. Stowe N, McMonagle E. Propagation of foreign DNA sequences linked to a herpes simplex virus origin of replication. In: Gluzman Y, ed. *Eucaryotic viral vectors*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1982:199–204.

5. Spaete R, Frenkel N. The herpes simplex virus amplicon: analyses of cis-acting replication functions. *Proc Natl Acad Sci USA* 1985;82:694–698.

6. Vlazny D, Kwong A, Frenkel N. Site-specific cleavage/packaging of herpes simplex virus DNA and the selective maturation of nucleocapsids containing full-length viral DNA. *Proc Natl Acad Sci USA* 1982;79:1423–1427.

7. Johnson PA, Miyanohara A, Levine F, et al. Cytotoxicity of a replication-defective mutant of herpes simplex virus type 1. *J Virol* 1992;66:2952–2965.

8. Geller AI, Keyomarsi K, Bryan J, et al. An efficient deletion mutant packaging system for defective herpes simplex virus vectors: potential applications to human gene therapy and neuronal physiology. *Proc Natl Acad Sci USA* 1990;90:7603–7607.

9. Fraefel C, Song S, Lim F, et al. Helper virus-free transfer of herpes simplex virus type 1 plasmid vectors into neural cells. *J Virol* 1996;70:7190–7197.

10. Lim F, Hartley D, Starr P, et al. Generation of high-titer defective HSV-1 vectors using an IE 2 deletion mutant and quantitative study of expression in cultured cortical cells. *BioTechniques* 1996;20:460–470.

11. Isacson O. Behavioral effects and gene delivery in a rat model of Parkinson's disease. *Science* 1995;269:856–857.

12. Neve RL, Howe JR, Hong S, et al. Introduction of glutamate receptor subunit 1 into motor neurons in vivo using a recombinant herpes simplex virus alters the functional properties of AMPA receptors. *Neuroscience* 1997;79:435–447.

13. Carlezon WA Jr, Boundy VA, Haile CN, et al. Sensitization to morphine induced by viral-mediated gene transfer. *Science* 1997;277:812–814.

14. Bursztajn S, DeSouza R, McPhie DL, et al. Overexpression in neurons of human presenilin-1 or a presenilin-1 familial Alzheimer disease mutant does not enhance apoptosis. *J Neurosci* 1998;18:9790–9799.

15. During MJ, Naegele JR, O'Malley KL, et al. Long-term behavioral recovery in Parkinsonian rats by an HSV vector expressing tyrosine hydroxylase. *Science* 1994;266:1399–1403.

16. Jin BK, Belloni M, Conti B, et al. Prolonged in vivo gene expression by a tyrosine hydroxylase promoter in a defective herpes simplex virus amplicon vector. *Hum Gene Ther* 1996;7:2015–2024.

17. Wang S, Vos J. A hybrid herpesvirus infectious vector based on Epstein-Barr virus and herpes simplex virus type 1 for gene transfer into human cells in vitro and in vivo. *J Virol* 1996;70:8422–8430.

18. Johnston KM, Jacoby D, Pechan P, et al. HSC/AAV hybrid amplicon vectors extend transgene expression in human glioma cells. *Hum Gene Ther* 1997;8:359–370.

19. Fraefel C, Jacoby DR, Lage C, et al. Gene transfer into hepatocytes mediated by helper virus free HSV/AAV hybrid vectors. *Mol Med* 1997;3:813–825.

20. Costantini LC, Jacoby DR, Want S, et al. Gene transfer to the nigrostriatal system by hybrid herpes simplex virus/adeno-associated virus amplicon vectors. *Hum Gene Ther* 1999;10:2481–2494.

21. Byrnes AP, MacLaren RD, Charlton HM. Immunological instability of persistent adenovirus vectors in the brain: peripheral exposure to vector leads to renewed inflammation, reduced gene expression, and demyelination. *J Neurosci* 1996;16:3045–3055.

22. Tripathy SK, Black HB, Goldwasser E, et al. Immune responses to transgene-encoded proteins limit the stability of gene expression after injection of replication-defective adenovirus vectors. *Nature Med* 1996;2:545–550.

23. Xiao X, Li J, McCown TJ, et al. Gene transfer by adeno-associated virus vectors into the central nervous system. *Exp Neurol* 1997;144:113–124.

24. Rabinowitz JE, Samulski J. Adeno-associated virus expression systems for gene transfer. *Curr Opin Biotechnol* 1998;9:470–475.

25. Naldini L. Lentiviruses as gene transfer agents for delivery to non-dividing cells. *Curr Opin Biotechnol* 1998;9:457–463.

26. McCown TJ, Xiao X, Li J, et al. Differential and persistent expression patterns of CNS gene transfer by an adeno-associated virus (AAV) vector. *Brain Res* 1996;713:99–107.

27. Alexander IE, Russell DW, Spence AM, et al. Effects of gamma irradiation on the transduction of dividing and nondividing cells in brain and muscle of rats by adeno-associated virus vectors. *Hum Gene Ther* 1997;7:841–850.

28. Naldini L, Blömer U, Gage FH, et al. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc Natl Acad Sci USA* 1996;93:11382–11388.

29. Blömer U, Naldini L, Kafri T, et al. Highly efficient and sustained gene transfer in adult neurons with a lentivirus vector. *J Virol* 1997;71:6641–6649.

30. Costantini LC, Bakowska JC, Breakefield XO, et al. Gene therapy in the CNS. *Gene Ther* 2000;7:93–109.

31. Wise RA. Drug-activation of brain reward pathways. *Drug Alcohol Depend* 1998;51:13–22.

32. Koob GF, Sanna PP, Bloom FE. Neuroscience of addiction. *Neuron* 1998;21:467–476.

33. Lett BT. Repeated exposures intensify rather than diminish the rewarding effects of amphetamine, morphine, and cocaine. *Psychopharmacology* 1989;98:357–362.

34. Piazza PV, Deminiere JM, le Moal M, et al. Stress- and pharmacologically-induced behavioral sensitization increases vulnerability to acquisition of amphetamine self-administration. *Brain Res* 1990;514:22–26.

35. Horger BA, Shelton K, Schenk S. Preexposure sensitizes rats to the rewarding effects of cocaine. *Pharmacol Biochem Behav* 1990;37:707–711.

36. Robinson TE, Berridge KC. The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Res Rev* 1993;18:247–292.

37. Nestler EJ, Aghajanian GK. Molecular and cellular basis of addiction. *Science* 1997;278:58–63.

38. Hollmann M, Heinemann S. Cloned glutamate receptors. *Annu Rev Neurosci* 1994;17:31–108.

39. Seeberg PH, Higuchi M, Sprengel R.RNA editing of brain glutamate receptor channels: mechanism and physiology. *Brain Res Rev* 1998;26:217–229.

40. Fitzgerald LW, Ortiz J, Hamedani AG, et al. Drugs of abuse and stress increase the expression of GluR1 and NMDAR1 glutamate receptor subunits in the rat ventral tegmental area: common adaptations among cross-sensitizing agents. *J Neurosci* 1996;16:274–282.

41. Zhang Z-F, Hu X-T, White FJ. Increased responsiveness of ventral tegmental area dopamine neurons to glutamate after repeated administration of cocaine or amphetamine is transient and selectively involves AMPA receptors. *J Pharmacol Exp Ther* 1997;281:699–706.

42. Kalivas PW, Stewart J. Dopamine transmission in the initiation and expression of drug- and stress-induced sensitization of motor activity. *Brain Res Rev* 1991;16:223–244.

43. Carlezon WA Jr, Haile CN, Coopersmith R, et al. Distinct sites of opiate reward and aversion within the midbrain identified using a herpes simplex virus vector expressing GluR1. *J Neurosci* 2000;20:RC62(1,2,3,4 and 5).

44. Wolf ME. The role of excitatory amino acids in behavioral sensitization to psychomotor stimulants. *Prog Neurobiol* 1998;54:679–720.

45. Self DW, Nestler EJ. Molecular mechanisms of drug reinforcement and addiction. *Annu Rev Neurosci* 1995;18:463–495.

46. Kelz MB, Chen JS, Carlezon WAJ, et al. Expression of the transcription factor dFosB in the brain controls sensitivity to cocaine. *Nature* 1999;401:272–276.

47. Churchill L, Swanson CJ, Urbina M, et al. Repeated cocaine alters glutamate receptor subunit levels in the nucleus accumbens and ventral tegmental area of rats that develop behavioral sensitization. *J Neurochem* 1999;72:2397–2403.

48. Graybiel AM, Moratalla R, Robertson HA, et al. Amphetamine and cocaine induce drug-specific activation of the c-fos gene in striosome-matrix compartments and limbic subdivisions of the striatum. *Proc Natl Acad Sci USA* 1990;87:6912–6916.

49. Curran EJ, Akil H, Watson SJ. Psychomotor stimulant- and opiate-induced c-fos mRNA expression patterns in the rat forebrain: comparisons between acute drug treatment and a drug challenge in sensitized animals. *Neurochem Res* 1996;21:1425–1435.

50. Hope BT, Nye HE, Kelz MB, et al. Induction of a long-lasting AP-1 complex composed of altered Fos-like proteins in brain by chronic cocaine and other chronic treatments. *Neuron* 1994;13:1235–1244.

51. Nestler EJ, Kelz MB, Chen J. DeltaFosB: a molecular mediator of long-term neural and behavioral plasticity. *Brain Res* 1999;835:10–17.

52. Malenka RC, Nicoll RA. Long-term potentiation—a decade of progress? *Science* 1999;285:1870–1874.

53. Zucker RS. Calcium- and activity-dependent synaptic plasticity. *Curr Opin Neurobiol* 1999;9:305–313.

54. White FJ, Hu X-T, Zhang X-F. Neuroadaptations in nucleus accumbens neurons resulting from repeated cocaine administration. *Adv Pharmacol* 1998;42:1006–1009.

55. Carlezon WA Jr, Wise RA. Rewarding actions of phencyclidine and related drugs in nucleus accumbens shell and frontal cortex. *J Neurosci* 1996;16:3112–3122.

56. Self DW, Genova LM, Hope BT, et al. Involvement of cAMP-dependent protein kinase in the nucleus accumbens in cocaine self-administration and relapse of cocaine-seeking behavior. *J Neurosci* 1998;18:1848–1859.

57. Carlezon WA Jr, Thome J, Olson V, et al. Regulation of cocaine reward by CREB. *Science* 1998;282:2272–2275.

58. Gonzalez GA, Montminy MR. Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell* 1989;59:675–680.

59. Carr GD, Fibiger HC, Phillips AG. Conditioned place preference as a measure of drug reward. In: Liebman JM, Cooper SJ, eds. *The neuropharmacological basis of reward*. Oxford: Clarendon Press, 1989:364–319.

60. Chavkin C, James IF, Goldstein A. Dynorphin is a specific endogenous ligand of the kappa opioid receptor. *Science* 1982;215:413–415.

61. Shippenberg TS, Rea W. Sensitization to the behavioral effects of cocaine: modulation by dynorphin and kappa-opioid receptor agonists. *Pharmacol Biochem Behav* 1997;57:449–455.

62. Hyman SE. Addiction to cocaine and amphetamine. *Neuron* 1996;16:901–904.

63. Mogenson GJ, Jones DL, Yim CY. From motivation to action: functional interface between the limbic system and the motor system. *Prog Neurobiol* 1980;14:69–97.

64. Geller AI, Breakefield XO. Defective HSV-1 expresses escherichia coli beta-galactosidase in cultured peripheral neurons. *Science* 1988;241:1667–1669.

21

Neuropsychopharmacology of Worms and Flies

William R. Schafer

William R. Schafer: Division of Biology, University of California–San Diego, La Jolla, California 92093-0349.

Pharmacologic agents often biochemically interact with multiple receptor or channel proteins, and induce multiple changes in cellular physiology and signal transduction. Thus, identifying the biologically relevant targets and effectors of a given neuroactive substance can be a challenging problem. This chapter describes how genetic analysis in simple model organisms, primarily worms or flies, has been used to identify molecules that mediate drug responses in the nervous system.

Essentially all the studies described here rely on the same general strategy. The drug of interest is tested for its ability to affect worm or fly behavior. Once a behavioral response is defined for wild-type animals, it is then used as a behavioral assay to identify mutant worms/flies that exhibit abnormal drug responses and thus define genes whose products are involved in the drug's mechanism of action. Once these genes are identified and cloned, human homologues can be identified based on sequence similarity, and tested for involvement in human drug responses. This sort of approach has a number of potential advantages. For one, phenotype-driven genetic screens essentially make no prior assumptions about the types of molecules involved in the process being studied; any gene that is not essential for life and affects the behavioral response to a drug is in principle equally likely to be identified in a mutant hunt. Thus, this approach is well suited for identifying previously unknown receptors or signal transduction molecules that participate in drug responses. Furthermore, modern molecular genetics provides the ability to manipulate specific gene products in an intact animal, often in a cell-type–specific manner. By making it possible to assess a particular protein function within the context of an intact nervous system, this approach can provide a most compelling demonstration of *in vivo* function.

Among organisms with nervous systems, two are particularly amenable to genetic analysis: the nematode *Caenorhabditis elegans*, and the fruit fly *Drosophila melanogaster*. These organisms share a number of advantages that make them especially well suited for classic and molecular genetics. For example, both have short generation times (2 weeks for *Drosophila*, 3 days for *C. elegans*), can be maintained easily and in large numbers in the laboratory, and are amenable to germline transformation. In addition, detailed genetic maps of both organisms are available, and the genome sequences of both organisms are now virtually complete. Although both organisms contain relatively simple nervous systems, they differ significantly in scale and level of characterization. The *C. elegans* nervous system consists of exactly 302 neurons, whose precise position, cell lineage, and anatomic connectivity are known (1, 2 and 3) . Consequently, it is possible to identify the roles of specific neurons and muscle cells in behavior using techniques such as single-cell laser ablation, and to thereby understand in a precise manner how the action of a particular gene product in a defined set of neurons influences the whole animal's behavior (4) . *C. elegans* is particularly suitable for genetic analysis of basic intracellular processes in neurons because the worm's nervous system is nearly dispensable for growth in the laboratory. Thus, even mutants with defects in basic neuronal functions such as neurotransmitter release are often viable and fertile (5) . The *Drosophila* nervous system is somewhat more complex, and contains approximately 10^5 neurons. Consequently, it is somewhat less well characterized at the cellular level than the *C. elegans* nervous system; however, the increased behavioral complexity afforded by this bigger nervous system also makes it perhaps better suited for investigating more complex forms of behavior and learning (6) .

STUDIES OF DRUG MECHANISMS IN MODEL ORGANISMS

Part of "21 - Neuropsychopharmacology of Worms and Flies "

Genetic pharmacology has historically been a powerful approach for neurobiological studies in *C. elegans* and *Drosophila*. Many studies of drug-resistant flies or worms have made use of pesticides or antihelminthic drugs that target

the insect or nematode nervous system. For example, screens for *C. elegans* mutants resistant to the pesticide (and cholinesterase inhibitor) aldicarb have been used with notable success to identify genes involved in synaptic function; molecules first studied in this way include the vesicular acetylcholine transporter and the synaptic proteins UNC-13 and UNC-18 (7, 8) . Likewise, studies of *C. elegans* mutants resistant to the anthelminthic ivermectin have provided insight into the functions of the invertebrate-specific family of glutamate-gated chloride channels (9) . More recently, attention has turned to the possibility of using genetic pharmacology to study the mechanisms of action for psychotropic drugs, including therapeutic agents and drugs of abuse. The following sections describe some examples of drugs whose mechanism of action has been studied in worms and/or flies, and the information that these studies have provided so far.

Therapeutic Agents

Lithium

Lithium salts are widely used for the treatment of bipolar affective disorder (manic-depressive illness). Lithium remains among the most effective treatments for acute mania, and it is also an effective mood-stabilizing agent for the prevention of both manic and depressive episodes. However, lithium has a number of side effects; for example, it is a known teratogen in vertebrate embryos, and can mimic the action of insulin in inducing synthesis of glycogen (9a) . However, although lithium has been shown to affect a number of molecular and cellular processes in neurons and other cells, the mechanisms through which it exerts its therapeutic effects on mood are not well understood.

One of the most prevalent theories for lithium's mechanism of action, first proposed by Berridge et al. (10) , is the inositol depletion hypothesis. According to this model, the critical functional consequence of lithium treatment is to reduce the intracellular concentrations of inositol, a key component of the phosphoinositide signal transduction cycle that mediates the effects of many neuromodulators, including serotonin (11) . Lithium ions are uncompetitive inhibitors of both inositol monophosphatase (IMPase), the enzyme that catalyzes the conversion of inositol monophosphates (IMPs) to inositol, and inositol polyphosphatase (IPP), the enzyme that converts inositol 1,4-bisphosphate to inositol 4-monophosphate (Fig. 21.1) . Inositol is required for the generation of phosphatidyl 4,5-inositol bisphosphate (PIP_2), whose cleavage by phospholipase C yields the calcium mobilizing agent inositol 1,4,5-trisphosphate (IP_3) and the protein kinase C activator diacylglycerol (DAG). Since both of these phosphoinositide-derived second messengers are critical signal transduction molecules that mediate the effects of diverse neurotransmitters and neuromodulators, a severe depletion of intracellular inositol would be expected to dramatically alter neuronal function. Thus, it has been proposed that lithium exerts its psychoactive effects by depleting intracellular inositol pools and thereby attenuating phosphoinositide signaling in neurons and other cells. However, experiments in rats suggest that while clinically effective concentrations of lithium are sufficient to inhibit IMP activity in the brain, they result in only a modest decrease in inositol levels (12) . Thus, it is not clear that inositol depletion can account for the psychotropic effects of lithium.

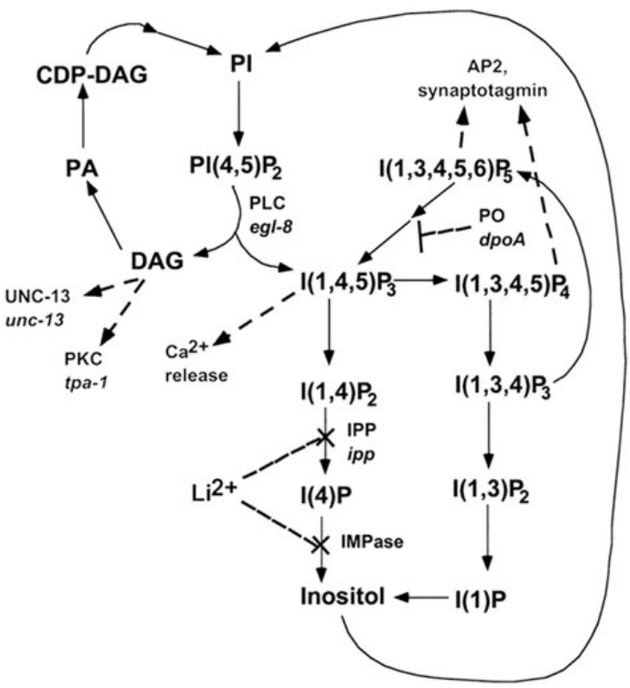


FIGURE 21.1. The phosphoinositide signaling pathway.

Recent genetic studies using simple eukaryotes has provided two plausible alternative hypotheses for lithium's mechanism of action. Interestingly, many of the key genetic findings on lithium response mechanisms have come from studies of a unicellular eukaryote that lacks a nervous system altogether, the slime mold *Dictyostelium discoideum*. Despite its considerable evolutionary divergence from the metazoa, many of the signal transduction mechanisms in *Dictyostelium* show remarkable conservation with those in human neurons. *Dictyostelium* usually exists as a free-living amoeba; however, during times of nutrient deprivation, these amoebae aggregate into a multicellular mass, or slug, which then develops into a fruiting body consisting of differentiated stalk and spore cells. Lithium has two effects on *Dictyostelium* development (13) . At high concentrations, lithium blocks the aggregation of amoebae. In contrast, low concentrations of lithium permit aggregation, but block spore cell differentiation, causing cells that normally would form the spore head to instead form stalk cells. This latter effect of

lithium on spore differentiation is mimicked by a mutation in the gene *gskA* (14) , which encodes a homologue of the signaling molecule glucogen synthase kinase 3 (GSK-3). GSK-3 molecules are conserved signaling molecules originally identified as negative regulators of glycogen synthesis, and subsequently implicated in the regulation of gene expression and cell movement. Since lithium's effects on both *Dictyostelium* development and glycogen synthesis were identical to those caused by inhibition of GSK-3, Klein and Melton (15) investigated whether lithium might affect GSK-3 signaling. They subsequently demonstrated that vertebrate GSK-3 is directly inhibited by lithium in *Xenopus* oocytes, and that GSK-3, but not IMPase, is responsible for the teratogenic effects of lithium on the embryo. Thus, at least some of the side effects of lithium, such as its teratogenic and insulin-mimetic effects, are almost certainly phosphoinositide-independent and instead mediated through the GSK-3 pathway. Because GSK-3 molecules are abundant in the brain, it is also possible that this pathway might also mediate some of lithium's therapeutic effects on mood.

However, other studies in both *Dictyostelium* and *Drosophila* support a link between the phosphoinositide pathway and the mood-altering effects of lithium. One such study concerned the mechanism of lithium's aggregation-inhibiting action in *Dictyostelium*, an effect that is independent of the *gskA* gene. A number of genes required for this high-concentration response to lithium were identified in genetic screens. One of these genes, *dpoA*, was shown to encode a proline oligopeptidase (PO), an enzyme involved in the degradation of bioactive peptides (16) . Interestingly, *dpoA* appeared to act via the phosphoinositide signaling pathway, since both mutations in *dpoA* and treatment with PO inhibitors elevated the levels of intracellular IP_3, but had no effect on GSK-3 activity. The elevation of IP_3 in *dpoA* mutants was a consequence of increased dephosphorylation of IP_5 (inositol 1, 3, 4, 5, 6-pentaphosphate), an alternate source of IP_3 utilized by both *Dictyostelium* and animal cells. Thus, inhibition of PO compensated for the decrease in PIP_2 levels induced by lithium by activating an alternative pathway for the production of IP_3 (and by extension inositol). Interestingly, abnormalities in PO activity have been observed in patients with both bipolar and unipolar depression (17, 18) . Thus, these results in *Dictyostelium* raise the possibility that PO may be linked to depression and mania through its effect on inositol signaling, and that lithium's efficacy in the treatment of depression may result from its ability to exert compensatory effects elsewhere in the inositol pathway.

However, the mechanism by which lithium-induced changes in the inositol pathway affect neuronal function may not involve inositiol depletion per se. This conclusion rests in part on a study of mutant flies defective in the enzyme IPP, a lithium-sensitive enzyme in the inositol pathway involved in the conversion of IP_3 to inositol (19) . ipp mutants were shown to be completely defective in the IPP activity, since they were unable to degrade I(1, 4) P_2, the IPP substrate. However, contrary to the prediction of the inositol depletion model, the phosphoinositol signaling pathway (which is necessary for *Drosophila* phototransduction) remained fully functional in photoreceptor neurons of the ipp mutant. Similar effects were seen when photoreceptor neurons were treated with lithium; IPP activity was inhibited, yet the inositol-dependent phototransduction cascade was still functional. Thus, neither genetic nor pharmacologic inhibition of IPP resulted in a depletion of inositol pools sufficient to interfere with the phosphoinositide signaling cascade. The ability to maintain high levels of inositol in the absence of IPP was apparently due to an alternate pathway involving synthesis and dephosphorylation of inositol 1,3,4,5-tetrakisphosphate (Fig. 21.1) . However, although ipp mutations and lithium treatment did not affect phosphoinositide signal transduction, they had unexpected and dramatic effects on synaptic function. Specifically, in ipp mutant and lithium-treated wild-type photoreceptor neurons, the probability of vesicular release was greatly increased, and affected neurons were unable to maintain a synaptic response to a prolonged tetanic stimulus. A variety of molecules involved in synaptic fusion and vesicular traffic, including synaptotagmin and adaptor protein 2 (AP2), are regulated through specific physical interactions with inositol polyphosphates (e.g., IP_4, IP_5, and IP_6) (20) . Thus, the effects of lithium on neuronal function in *Drosophila* as well as in humans may stem not from defects in inositol signaling per se, but from defects in synaptic function and plasticity due to alterations in inositol polyphosphate pools.

In summary, lithium provides a good example of the power of genetic neuropsychopharmacology in simple model systems. Studies in *Dictyostelium* were instrumental in identifying the GSK-3 pathway as a possible mediator of lithium's deleterious side effects, and have also provided insight into a possible link between neuroactive peptides and depression. Work in *Drosophila* has provided an important lead into discovering how lithium's effects on phosphoinositide signaling affect neuronal function. Future studies in both organisms have the potential to provide further insight into lithium's mechanism of action, in particular to address more precisely how lithium-induced changes in inositol lipid content alter synaptic transmission and plasticity in neurons.

Fluoxetine and Other Antidepressants

Another group of drugs that have been the subject of research in simple eukaryotes are those used in the treatment of unipolar depression. Such drugs include the monoamine oxidase (MAO) inhibitors, the tricyclic antidepressants (e.g., imipramine and clomipramine), and the selective serotonin reuptake inhibitors (SSRIs; e.g., fluoxetine). A common property of many of these molecules is their ability

to potentiate serotoninergic neurotransmission, either by interfering with reuptake of serotonin from the synapse (tricyclics and SSRIs) or by blocking enzymatic degradation of serotonin (MAO inhibitors). Thus, the therapeutic actions of all of these molecules are usually explained in terms of a model for depression known as the serotonin hypothesis. According to this model in its simplest form, levels of serotoninergic neurotransmission in the forebrain are a key determinant of mood, with high activity leading to euphoria and low activity to dysphoria. Thus the chronic dysphoria experienced by depressed patients could be a consequence of chronically low serotoninergic transmission, which could be compensated for by interfering with serotonin degradation. This serotonin hypothesis, or variations thereof, represents the most widely accepted explanation for antidepressant action (21, 22) .

However, the serotonin hypothesis, at least in its simplest form, fails to account for a number of observations about antidepressants. For one, a direct correlation between the level of serotoninergic transmission and mood has not been demonstrated; normal individuals treated with serotonin reuptake blockers do not typically experience euphoria, nor does dietary serotonin depletion induce depression in individuals not already prone to depression (23) . Moreover, the mood-altering effects of serotonin reuptake blockers in depressed patients occur on a different time scale from their effects on serotoninergic transmission; whereas SSRIs and most tricyclics elevate synaptic serotonin levels within hours, their effects on mood are not apparent for 2 to 6 weeks. Finally, a number of effective antidepressants appear to function independently from serotonin, including selective norepinephrine reuptake inhibitors (SNRIs) such as desipramine, MK869, which antagonizes substance P receptors, and bupropion, whose target is unknown (24, 25) . Because of these observations, many current models hypothesize that SSRIs are effective against depression not because of their acute effects on serotoninergic transmission, but because of long-term adaptive changes in monoamine neurotransmission that arise from chronic inhibition of serotonin reuptake (21) . An appealing feature of this type of model is that long-term activation of different direct targets by different classes of antidepressants (the serotonin transporter by SSRIs, other targets by atypical antidepressants) could in principle lead to a common set of adaptive responses in the brain. Alternatively, it is possible that antidepressants might act, at least in part, at serotonin-independent direct targets.

Studies in *C. elegans* have provided insight into potential serotonin-dependent and -independent activities of antidepressants. Nearly all antidepressants have at least two clear effects on *C. elegans* behavior: stimulation of egg laying and hypercontraction of muscles in the nose. Whereas the stimulation of egg laying by antidepressants is primarily due to potentiation of serotoninergic transmission (see below), the effect of antidepressants on the nose muscles appears to be independent of serotonin, since serotonin itself does not cause nose contraction, whereas antidepressants still contract the noses of serotonin-deficient mutants. Mutations conferring resistance to the induction of nose contraction by fluoxetine have been identified in seven genes, designated *Nrf* genes, for *n*ose *r*esistant to *f*luoxetine (26) . All the *Nrf* mutations are recessive and confer resistance to several chemically disparate antidepressants in addition to fluoxetine; thus, the products of the *Nrf* genes might potentially represent common, serotonin-independent antidepressant targets. So far, two *Nrf* genes have been cloned, *nrf-6* and *ndg-4*. These two genes define the first members of a novel gene family, and encode predicted multipass integral membrane proteins that are expressed in the nasal epidermis and the intestine. *nrf-6* and *ndg-4* have been shown to be defective in the transport of yolk proteins across the intestinal membrane, suggesting that NRF-6 and NDG-4 may be components of a complex that transports molecules across epithelial membranes. Based on this result, it is reasonable to suppose that the fluoxetine resistance of *nrf-6* and *ndg-4* mutants might reflect a defect in drug uptake rather than the absence of a functional drug target in the neuromuscular system. However, while NRF-6 and NDG-4 (and by extension their yet unidentified vertebrate homologues) may not represent antidepressant targets per se, they might represent molecules that function in transport of antidepressants across the blood–brain barrier.

Another *C. elegans* molecule that clearly represents a serotonin-independent antidepressant target is encoded by the gene *egl-2*. *egl-2* was originally defined by the dominant gain of function mutations that impaired the activity of the vulval muscles (which mediate egg laying) and enteric muscles (which mediate defecation) (27, 28) . Both of these defects in muscle activation could be relieved by treatment with the tricyclic antidepressant imipramine, though not by serotonin or fluoxetine. Thus, imipramine appeared to act through a serotonin-independent target to suppress the *egl-2* muscle activation phenotype (29) . The nature of this target was revealed when *egl-2* was cloned and shown to encode a potassium channel homologous to the *Drosophila* ether-a-go-go (eag) channel (30) . Studies on EGL-2 channels expressed in *Xenopus* oocytes demonstrated that the imipramine-suppressible dominant alleles of *egl-2* encoded mutant channels that opened inappropriately at low voltages. Remarkably, imipramine was shown to function as a specific antagonist of both the *EGL-2* channel and its mammalian homologue MEAG. Thus, this class of calcium channels appears to represent a conserved target of tricyclic antidepressants in both worms and humans. Interestingly, an important side effect of tricyclic antidepressants is a type of cardiac arrhythmia called long QT syndrome, a disorder which has also been linked to mutations in potassium channel genes (29a, 29b) . Thus, the blockade of eag-related potassium channels by tricyclics provides a likely explanation for this clinically important side effect of tricyclics.

Studies in *C. elegans* may also provide insight into the serotonin-dependent mechanisms of antidepressant action.

The ability of antidepressants (other than tricyclics) to stimulate egg laying in *C. elegans* depends on their ability to potentiate serotoninergic neurotransmission (29) , and can be mimicked by exogenous serotonin itself (31) . Serotonin is released from egg-laying motor neurons called HSNs (27) , and appears to function as a neuromodulator that modifies the functional state of the egg-laying muscles to potentiate contraction (32) . Serotonin also inhibits locomotion, apparently by inhibiting neurotransmitter release from excitatory motor neurons (32, 33) . The signal transduction mechanisms that mediate both of these actions of serotonin have been analyzed genetically, and in both cases the phospholipase C (PLC) homologue *egl-8* is required for serotonin response. In the egg-laying muscles, the effects of PLC appear to be mediated through the protein kinase C homologue tpa-1, whereas in the motor neurons the most important mediator appears to be the diacylglycerol-binding synaptic protein UNC-13. The involvement of the phosphoinositide signaling pathway in serotonin signal transduction in both the egg-laying muscles and the motor neurons of *C. elegans* has an interesting parallel in mammals, since a number of mammalian serotonin receptor subtypes also signal through activation of PLC.

The apparent conservation between the signaling pathways mediating serotonin response in *C. elegans* and humans raises the possibility that the long-term effects of elevated serotoninergic transmission might also be accessible to genetic analysis in *C. elegans*. As noted previously, the alleviation of depression by serotonin-potentiating antidepressants is thought to involve adaptive signaling pathways that are activated by prolonged elevation of serotoninergic neurotransmission. In *C. elegans*, prolonged exposure to serotonin has been shown to lead to adaptive down-regulation of egg-laying behavior and recovery from serotonin-induced paralysis (34) . Genes encoding possible components of serotonin adaptation pathways have been identified on the basis of serotonin hypersensitive or adaptation-defective phenotypes (35) ; however, at present little is known about how these or other genes affect long-term responses to serotonin. Future analysis of serotonin adaptation genes may provide insight into the molecular mechanisms underlying long-term responses to elevated serotonin transmission that may be important for the therapeutic action of antidepressants.

Volatile Anesthetics

A variety of volatile molecules, including diethyl ether, halothane, and isoflurane, are capable of inducing general anesthesia, a behavioral state involving loss of consciousness, analgesia, amnesia, and loss of motor activity. Although these agents have been widely used in surgery for over a century, their mechanism of action remains poorly understood. General anesthesia appears to result from defects in synaptic transmission rather than axonal firing; however, it is not clear whether anesthesia results from potentiation of inhibitory synapses, inhibition of excitatory synapses, or both. The potency of a given volatile anesthetic shows a very strong correlation to its lipid solubility; this observation, known as the Meyer–Overton rule, has led to the hypothesis that volatile anesthetics act by disrupting hydrophobic interactions between proteins and/or lipids in neurons. However, the biologically relevant targets for volatile anesthetics have not been conclusively identified. In principle, this problem appears ideally suited to attack by a phenotype-driven genetic approach; by identifying mutants that are resistant or hypersensitive to anesthetics and cloning and sequencing the mutant genes, it should be possible to identify anesthetic targets that are essential for anesthesia *in vivo*. In fact, such screens have been conducted in both *Drosophila* and *C. elegans*, and a variety of genes affecting sensitivity have been identified (36) . At present, none of the *Drosophila* anesthetic response genes have been cloned; thus, molecular information about their gene products is not available. However, the recent cloning of several *C. elegans* genes with quantitatively large effects on anesthetic sensitivity raises the possibility that they might define conserved molecular targets important for anesthetic action.

C. elegans has two distinct responses to volatile anesthetics. At lower concentrations (similar to the alveolar concentrations used in human anesthesia), volatile anesthetics rapidly induce abnormalities in the pattern of locomotion (37) . Although this effect is behaviorally quite dissimilar from anesthesia, it is similar to the effect of many mutations that affect synaptic transmission in *C. elegans*. In fact, treatment with volatile anesthetics confers resistance to the behavioral effects of cholinesterase inhibitors (38) , a hallmark of defective neurotransmitter release (7) . Thus, at these concentrations, volatile anesthetics appear to act presynaptically to interfere with synaptic transmission in *C. elegans*. A number of mutants with altered sensitivity to these low-concentration effects of volatile anesthetics have been identified. Potentially the most informative with respect to anesthetic mechanisms contain mutations in genes encoding components of the SNARE complex, the presynaptic machinery that mediates synaptic vesicle fusion. Recessive mutations in at least three SNARE genes, *unc-64* [encoding *C. elegans* syntaxin (39)], *snb-1* [encoding VAMP/synaptobrevin (40)], and *ric-4* (encoding SNAP-25), confer significant hypersensitivity on the effects of both halothane and isoflurane on coordinated movement. Furthermore, a novel mutation in *unc-64*, which affects a spice receptor site and consequently leads to the production of truncated syntaxin peptides, confers strong resistance to the effects of volatile anesthetics on both coordinated movement and cholinesterase sensitivity (38) . These results suggest that volatile anesthetics interfere with synaptic transmission through direct interaction with one or more members of the SNARE complex.

At approximately 10-fold higher concentrations, volatile anesthetics induce reversible paralysis in *C. elegans*, a behavioral

effect qualitatively reminiscent of anesthesia. Interestingly, none of the synaptic mutations affecting the low-concentration effects on coordinated movements affect this high-concentration paralytic response. However, a different, nonoverlapping group of genes has been identified that confers resistance or hypersensitivity to paralysis by anesthetics in *C. elegans*. Several of these genes have been cloned, including *unc-1*, which encodes a homologue of stomatin (41) , and *unc-8*, which encodes a subunit of the degenerin/ENaC family of passive sodium channels (42, 43) . Both *unc-1* and *unc-8* are expressed in neurons, and both genes can be mutated to confer either resistance or hypersensitivity to halothane (44) . Allele-specific genetic interactions between *unc-1*, *unc-8*, and the yet uncloned *unc-79* and *unc-80* genes suggest that their products may physically interact in a multimeric channel complex specifically involved in anesthetic responses. Since stomatin has been shown to function as a negative regulator of cation channels in erythrocytes, a reasonable hypothesis is that UNC-1/stomatin may modulate influx through UNC-8 degenerin channels in neurons that respond to anesthetics. Homologues of both stomatins and ENaC channels have been identified in mammals, and are known to be expressed in the central nervous system; thus, in principle stomatin-regulated ENaC channels could also affect anesthetic responses in humans.

In summary, there are two distinct sets of genes that affect responses to volatile anesthetics in *C. elegans*, which affect different behavioral responses to different concentrations of anesthetics. At present, it is not clear which of the two (or whether both) might encode homologues of biologically relevant human anesthetic targets. Although the genes involved in synaptic function alter anesthetic responses at clinically relevant concentrations, the behavioral responses they affect are qualitatively quite different from general anesthesia. Conversely, although the stomatin/degenerin genes affect a paralytic response that closely resembles anesthesia, the response also has a relatively long time delay and occurs at concentrations well above those clinically relevant in humans. Given the effective drug concentrations for these two behavioral responses, it is possible that the synaptic genes might encode targets relevant to anesthesia, while the stomatin/degenerin genes might encode targets relevant for side effects of anesthetics. Alternatively, it is possible that genes affecting high-concentration anesthetic responses do define molecules involved in anesthesia, especially since the nematode cuticle is relatively impermeant and presents a significant barrier for the entry of many drugs. Since well-defined mammalian homologues exist for both classes of anesthetic response genes, it should be possible in the future to examine these issues directly in mammalian systems.

Drugs of Abuse

Ethanol

Unlike many neuroactive substances, ethanol is not believed to have a single molecular target in neurons; rather, a number of receptors and channels, including the *N*-methyl-D-aspartate (NMDA), serotonin, and γ-aminobutyric acid (GABA) receptors and various voltage-gated ion channels, appear to be modulated by the presence of ethanol (45) . Very little information exists concerning the relative importance of each of these putative direct targets for the psychoactive effects of ethanol; however, a variety of experiments in cultured cells suggest that a critical short-term effect of ethanol is to enhance receptor-mediated synthesis of the second messenger 3′,5′-cyclic adenosine monophosphate (cAMP). Conversely, long-term ethanol exposure appears to decrease intracellular cAMP levels. Both the acute and chronic effects of ethanol have also been linked to changes in dopaminergic neurotransmission (46) . In particular, ethanol has been shown to promote release of dopamine in the mesolimbic pathways of the brain, in particular the so-called reward pathway synapses between the ventral tegmental area (VTA) and the nucleus accumbens (NAc). At present, the *in vivo* significance of these findings with respect to the psychoactive effects of ethanol in mammals remains to be determined. Moreover, although sensitivity to both the acute and chronic effects of ethanol are clearly affected by genetic factors, the nature of the genes affecting human ethanol sensitivity are not known.

Recent work in *Drosophila* has provided support for both the dopamine and cAMP hypotheses of ethanol action. Ethanol vapor has a number of effects on *Drosophila* behavior, including hyperactivity, disorientation, uncoordination, and ultimately immobilization. Using an instrument called an inebriometer (47) , lines of mutant flies have been identified that exhibit abnormal sensitivity to volatilized ethanol. Among the mutants showing significant hypersensitivity to ethanol were those containing a mutation in the learning gene amnesiac, which encodes a homologue of the mammalian pituitary adenylyl cyclase activating peptide (PACAP) (48, 49) . Consistent with the implications of this homology, the effects of amnesiac on ethanol response appeared to involve the adenylyl cyclase pathway, since the adenylyl cyclase activator foskolin blocks the ethanol sensitivity associated with amnesiac loss-of-function mutations. Moreover, several other loss-of-function mutations affecting cAMP pathway components, including the adenylyl cyclase gene rutabaga and the cAMP-dependent protein kinase gene *DCO*, also conferred ethanol sensitivity. Although one might suppose based on these results that the response to ethanol is simply a function of the level of cAMP signaling in the relevant neuronal targets (with increased ethanol response corresponding to low cAMP signaling), a variety of data are inconsistent with this simple model. For example, genetic or pharmacologic activation of the cAMP pathway does not lead to ethanol resistance. Nonetheless, these genetic data provide the first conclusive link between the activity of the cAMP pathway and the behavioral effects of ethanol in an intact organism; the precise nature of that link remains to be determined, but should be accessible to further genetic analysis.

Some of the behavioral effects of ethanol on *Drosophila* have also been shown to be dependent on dopamine (50) . Ethanol has varying effects on fly locomotion depending on the duration of exposure. During the first 7 to 10 minutes of ethanol treatment, animals become hyperactive and move at a greatly increased rate; subsequently, they become increasingly uncoordinated and eventually become completely immobile. When flies are depleted of dopamine through ingestion of a tyrosine hydroxylase inhibitor, they become significantly less susceptible to this stimulation of motor activity by ethanol. However, these dopamine-depleted flies exhibited no abnormalities in their sensitivities to ethanol-induced uncoordination or immobilization. Thus, the stimulation of motor activity by ethanol may involve ethanol-induced enhancement of dopaminergic transmission in brain areas controlling locomotion, whereas the other behavioral effects of ethanol are likely to involve other neurotransmitter systems.

The genetic analysis of ethanol response mechanisms in *Drosophila* is still in its early stages. However, it is already clear that mutants with altered responses to ethanol can be identified in straightforward genetic screens, and at least in some cases analyzed in the context of well-defined neuronal signaling cascades. Perhaps the greatest promise for future studies is the possibility that novel ethanol response genes, possibly including the direct molecular targets of ethanol, can be identified in ethanol-resistant or ethanol-hypersensitive screens.

Nicotine

Tobacco has been implicated in more deaths than any other addictive substance (51) , yet the biochemical basis for compulsive tobacco use remains poorly understood. The substance most responsible for the addictive properties of tobacco is nicotine, a potent stimulant and cholinergic agonist. Long-term exposure to nicotine is known to cause adaptive changes in the activity and number of nicotinic receptors in the brain, which are thought to be important for nicotine addiction (52) . For example, nicotinic receptors exist in multiple functional states, some of which are relatively refractory to channel opening though they retain affinity for agonists. Chronic exposure to nicotine or other agonists results in an increased fraction of receptors adopting the lower activity states, leading to an attenuation of the overall nicotine response (53) . Long-term nicotine treatment also causes a long-lasting functional inactivation of some nicotinic receptors (54) , which has a slower time course and is much longer lasting than the rapid, receptor-intrinsic desensitization induced by acute agonist exposure. Depending on the receptor and cell type, long-term nicotine treatment can also either increase or decrease the number of nicotinic receptors on the cell surface, effects that appear to be mediated at the level of protein turnover (55, 56) . The cellular pathways that promote these changes are not well understood; for example, little is known about the cellular pathways that regulate receptor turnover, or the molecular mechanisms that regulate the switching between different nicotinic acetylcholine receptor (nAChR) states.

Genetic analysis in *C. elegans* may provide insight into the mechanisms underlying long-term responses to nicotine. Both acute and chronic nicotine treatment have striking effects on the behavior of *C. elegans*, including hypercontraction of body wall muscles, stimulation of egg laying, and increased pharyngeal pumping. The effects of nicotine on the body and egg-laying muscles are mediated through a nicotinic receptor known as the levamisole receptor (57, 58) . The antihelminthic drug [and ganglionic nAChR agonist (59)] levamisole is a potent agonist of this receptor; like nicotine, levamisole causes body muscle hypercontraction and (at high doses) spastic paralysis. Although the levamisole receptor is found on nematode muscle, its pharmacologic profile generally resembles that of ganglionic nicotinic receptors of vertebrates. By screening for levamisole-resistant mutants, it has been possible to identify genes affecting the function of the levamisole receptor (60) . Mutations conferring strong resistance to levamisole have been identified in six genes. Three of these genes, *unc-38*, *unc-29*, and *lev-1*, encode nicotinic receptor subunits (61, 62) . The UNC-38 protein is most similar to the insect α-like subunits ALS and SAD (49% amino acid identity); among vertebrate receptor subunits, the closest similarity is to neuronal α subunits (61) . UNC-29 and LEV-1 are closely related proteins whose closest homologues in vertebrates are neuronal non-α subunits (approximately 55% sequence similarity). Three additional genes conferring strong levamisole resistance, *unc-50*, *unc-74*, and *unc-63*, have not been cloned, but have been shown to be required for assembly of a functional levamisole receptor as assayed *in vitro* (63) . In addition to conferring resistance to levamisole (and other nicotinic agonists), mutations in these genes cause defects in the coordination of body movement (60) . Mutations in three additional genes (*lev-8*, *lev-9*, and *lev-10*) confer weaker resistance to levamisole, do not cause defects in locomotion, and have no detectable effect on the biochemical properties of the receptor as assayed *in vitro* (60, 63) . Thus, the proteins encoded by these genes have been hypothesized to regulate the activity of the receptor indirectly.

Long treatments with nicotine and other nicotinic receptor agonists lead to adaptation (57) . Animals treated with exogenous nicotine initially hypercontract to the point of spastic paralysis; however, after several hours in the presence of nicotine, they recover their ability to move and regain much of their body length. In some *C. elegans* strains (for example, strains with weakly crippled nAChRs), long-term nicotine treatment eventually leads to almost complete inactivation of the response to nicotine. Moreover, when nicotine-adapted animals are removed from nicotine, their locomotive behavior becomes uncoordinated and resembles that of mutants with strong defects in the levamisole receptor (i.e. an *unc-29* or *unc-38* null mutant). Thus, long treatments with nicotine cause nicotine dependence in addition

to nicotine tolerance in the *C. elegans* body muscle. Long-term nicotine treatment also down-regulates levamisole receptors in the egg-laying muscles. Overnight treatment with nicotine leads to an almost complete attenuation of levamisole sensitivity with respect to egg laying, and this attenuation of levamisole response persists for up to 24 hours after removal from nicotine. This loss of levamisole responsiveness is accompanied by a corresponding decrease in the abundance of UNC-29–containing receptors in the vulval muscles, an effect that may be mediated at the level of protein turnover (64) . Interestingly, the nicotine-dependent decrease in UNC-29 receptor abundance requires the activity of TPA-1, a vulval muscle-expressed PKC isoform. Since UNC-29 and other nicotinic receptor subunits contain consensus sequences for PKC phosphorylation, this raises the possibility that direct phosphorylation of nicotinic receptors might represent a signal for increased turnover. In the future, it should be possible to test this hypothesis, as well as identify other genes required for long-term responses to nicotine in *C. elegans*.

Another set of genes, the weak levamisole-resistance genes *lev-8* and *lev-9*, appear to represent positive regulators of nicotinic receptor activity. Mutations in these genes confer partial resistance to levamisole and nicotine with respect to body muscle contraction and strong resistance with respect to egg laying (65) . However, *lev-8* and *lev-9* mutations do not affect the assembly of levamisole-binding nicotinic receptors as assayed *in vitro* (58) , and the abundance of UNC-29 receptors in the vulval muscles is not significantly reduced by mutations in these genes (65) . *lev-8* and *lev-9* may therefore encode regulatory proteins that stimulate the activity of nicotinic receptors *in vivo*, but are not subunits or essential accessory proteins. In principle, the inhibition of the *lev-8* or *lev-9* gene products might represent a plausible mechanism for functional inactivation of nicotinic receptors. Once *lev-8* and *lev-9* are cloned, it will be interesting to determine whether mammalian homologues exist for these molecules, and if so, whether they are involved in regulating the functional activity of nicotinic receptors in human neurons.

Cocaine

Cocaine is a potent psychostimulant, and among the most widespread addictive drugs of abuse. The psychoactive effects of cocaine are thought to result largely from its ability to potentiate aminergic neurotransmission in the limbic pathways of the brain. Cocaine inhibits the reuptake transporters for dopamine, serotonin, and norepinephrine, which leads to accumulation of monoamine transmitters at the synapse. The dopaminergic synapses of the nucleus accumbens are thought to be particularly important for cocaine addiction, since pharmacologic inhibition or surgical lesioning of these areas confers significant resistance to both the short-term and long-term effects of cocaine in rodents (46) . However, dopamine is probably not the only neurotransmitter involved in cocaine addiction, since mice lacking the vesicular dopamine transporter will still self-administer cocaine after repeated administration of the drug (66, 67) . Although dopaminergic transmission in the limbic reward pathways has been implicated in the reinforcing properties of a wide range of addictive substances in addition to cocaine, the molecular and cellular mechanisms that lead to addiction in these neurons are not well understood.

Recent work in *Drosophila* suggests that the mechanisms of cocaine action may be accessible to genetic analysis. When flies are exposed to volatized free-base cocaine, they exhibit dose-dependent stereotypical behaviors that are surprisingly reminiscent of cocaine's psychostimulant effects in mammals (68) . For example, at low doses treated flies become hyperactive and exhibit compulsive, continuous grooming behavior. At intermediate doses animals move more slowly and display stereotyped locomotive behaviors such as circling. Finally, at high doses animals undergo tremors, spastic paralysis, and finally death. Repeated treatment of flies with low doses of cocaine results in an increased behavioral response, a phenomenon known as sensitization; cocaine sensitization also occurs in mammals and is thought to underlie some aspects of addiction in humans. Interestingly, male flies are more sensitive to cocaine than females, a sexual dimorphism that also holds true in mammals (69) . Thus, cocaine has both short-term and long-term effects on fly behavior that are remarkably analogous to its effects on mammals.

These behavioral similarities between cocaine's action on flies and mammals raise the possibility that they might share a common functional basis as well. In fact, recent evidence indicates that cocaine's actions on fly behavior also involve effects on aminergic neurotransmission. Insects contain cocaine sensitive reuptake transporters for dopamine, serotonin, and octopamine (an invertebrate neurotransmitter chemically similar to norepinephrine); thus, cocaine at least in principle could increase synaptic levels of multiple monoamine neurotransmitters in the fly brain (70, 71 and 72) . The monoamine most convincingly implicated in cocaine's acute effects on flies is dopamine. Dopamine receptor antagonists have effects on grooming and locomotive behaviors that are the converse of the effects of cocaine, and these antagonists can also block the effects of cocaine and cocaethylene on these behaviors in decapitated *Drosophila* preparations (Fig. 21.2) (69, 73) . Moreover, when flies are depleted of endogenous dopamine using tyrosine hydroxylase inhibitors, they acquire resistance to the acute effects of cocaine treatment (50) . Paradoxically, however, transgenic animals in which dopamine and serotonin release is blocked by ectopic tetanus toxin expression are actually hypersensitive to cocaine (74) . Thus, although dopaminergic neurotransmission is clearly involved in behavioral responses to cocaine in *Drosophila*, the specific role that it plays in these responses is not completely clear.

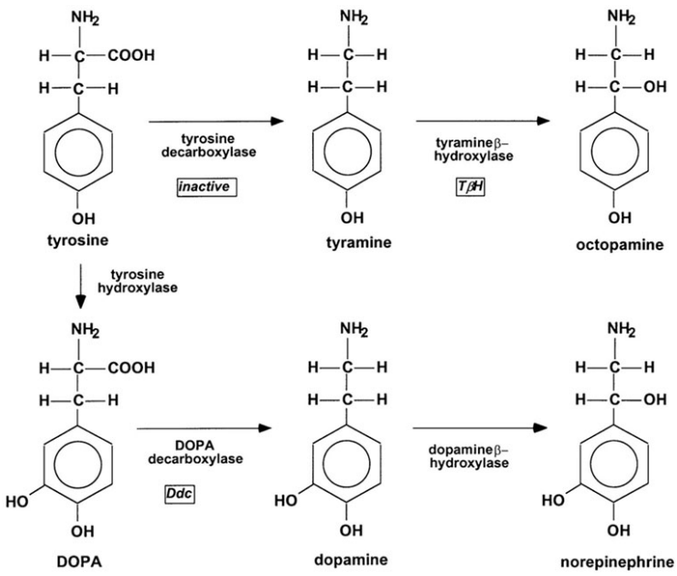


FIGURE 21.2. Biogenic amines and their biosynthesis.

Surprisingly, cocaine sensitization in *Drosophila* has been linked to a different biogenic amine—tyramine. Tyramine is present only in trace quantities in mammalian nervous systems; however, in insects it is a somewhat more abundant molecule and also serves as a precursor for the important neuromodulator octopamine (Fig. 21.3) . Mutants with defects in this biosynthetic pathway have been identified in *Drosophila* behavioral screens. For example, inactive mutants have low levels of the enzyme tyrosine decarboxylase, and consequently fail to efficiently synthesize both tyramine and octopamine; in contrast, TβH mutants are defective in the tyramine ß-hydroxylase enzyme, and thus synthesize tyramine but not octopamine. Interestingly, while inactive mutants display an essentially normal acute response to cocaine, they are strongly defective in sensitization (75) . This sensitization defect can be rescued by feeding the mutant flies tyramine but not octopamine; moreover, TβH mutants (which lack octopamine but not tyramine) and Ddc mutants (which fail to synthesize dopamine) show normal cocaine sensitization. Furthermore, cocaine actually increases the levels of tyrosine decarboxylase activity in treated flies, suggesting that cocaine sensitization may actually occur at least in part through induction of tyramine synthesis. Remarkably, both the induction of tyrosine hydroxylase activity by cocaine and cocaine sensitization itself require the activities of the period, clock, and double-time genes, three members of the conserved signal transduction pathway that controls circadian rhythms in animals and fungi (76) .

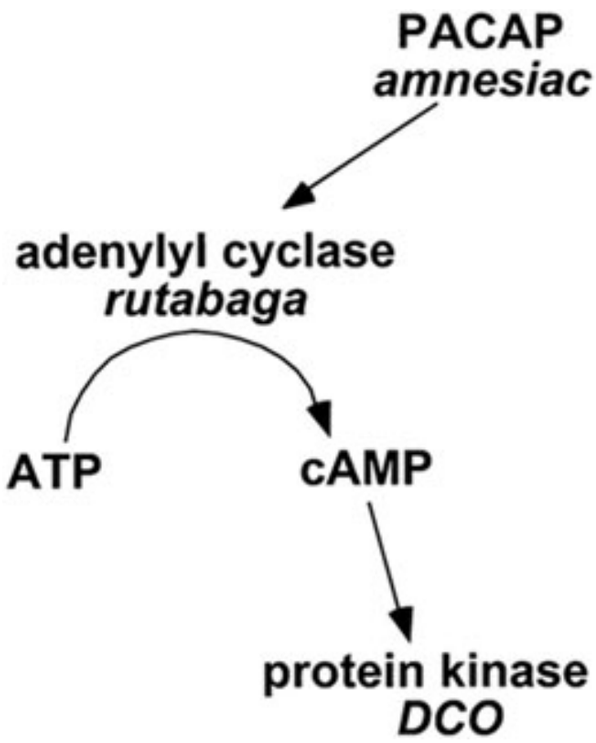


FIGURE 21.3. The adenylyl cyclase signaling pathway.

How might tyramine mediate cocaine sensitization in flies, and does it play a similar role in mammals? At present, these questions are difficult to answer. Although the function of tyramine in insect nervous systems has not been clearly established, putative tyramine receptors have recently been identified in both *Drosophila* and the honeybee (77) . Possibly cocaine might act in a period-dependent manner to facilitate tyramine release from nerve terminals, which could then induce plasticity in other monoamine pathways in the brain. Future studies will be needed to identify the specific tyramine receptors that might mediate such responses and to understand the neural basis for their effects on behavior. In vertebrates, tyramine receptors have not been identified; thus, it remains an open question whether tyramine plays a role in human sensitization to cocaine that parallels its role in *Drosophila*. However, the involvement in *Drosophila* of the circadian clock pathway, which is highly conserved between insects and humans, suggests that at least some components of the molecular mechanisms underlying this process may be shared between these widely divergent organisms.

QUESTIONS AND FUTURE PROSPECTS

Part of "21 - Neuropsychopharmacology of Worms and Flies "

Perhaps the major potential pitfall of using worm or fly genetics to investigate drug mechanisms is that there is no guarantee that those mechanisms will be conserved across the evolutionary gulf separating these disparate animals. Certainly at the anatomic level, the brains of humans, flies, and worms are vastly different organs. Nonetheless, for most pharmacologic studies, the critical issue is conservation at the molecular level, and with the worm and fly genomes essentially complete, it is clear that at the molecular level the *C. elegans* and *Drosophila* nervous systems are quite similar to their human counterpart. For example, the *C. elegans* and *Drosophila* genomes contain homologues of each of the basic types of potassium channels, calcium channels, and G proteins, as well as putative receptors for most human neurotransmitters (78, 79) . To be sure, there are a small number of nervous system molecules found in vertebrates and flies but not nematodes (e.g., voltage-gated Na channels), as well as molecules found in nematodes and flies but not vertebrates (e.g., the ivermectin-sensitive glutamate-gated Cl channel). However, on the whole the nematode,

fly, and vertebrate nervous systems appear to be remarkably similar at the molecular level given their vast differences in scale and functionality.

What are the prospects for model organism neuropsychopharmacology in the postgenomic future? The availability of substantial portions of the worm and fly genomes has already made the rate-limiting step of classic forward genetics—cloning a mutant gene—significantly easier and more straightforward. This cloning process will become easier still as high-resolution, single-nucleotide polymorphism maps of the worm and fly genomes become available. The imminent completion of the human genome will also provide great benefits to model organism studies, since it will allow rapid identification of human homologues for worm or fly genes and more reliable distinction of genuine mammalian orthologues from other members of a gene family. The great advantage of worm and fly studies for the elucidation of drug mechanisms is the ability to conduct unbiased, phenotype-driven mutant screens to identify unknown gene products involved in drug response. Since ethical considerations will always preclude such approaches in humans, and since time, space, and cost considerations make them inefficient even in simpler vertebrates, *C. elegans* and *Drosophila* are likely to serve as workhorses for basic neuroscience research for many years to come.

REFERENCES

1. White J, Southgate E, Thomson N, et al. The structure of the *Caenorhabditis elegans* nervous system. *Philos Trans R Soc Lond (Biol)* 1986;314:1–340.

2. Sulston JE, Horvitz HR. Post-embryonic cell lineages of the nematode *Caenorhabditis elegans*. *Dev Biol* 1977;56:110–156.

3. Sulston JE, Schierenberg E, White JG, et al. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol* 1983;100:64–119.

4. Bargmann CI, Avery L. Laser killing of cells in *Caenorhabditis elegans*. *Methods Cell Biol* 1995;48:225–250.

5. Nonet ML, Grundahl K, Meyer BJ, et al. Synaptic function is impaired but not eliminated in *C. elegans* mutants lacking synaptotagmin. *Cell* 1995;73:1291–1305.

6. Dubnau J, Tully T. Gene discovery in *Drosophila*: new insights for learning and memory. *Annu Rev Neurosci* 1998;21:407–444.

7. Rand JB, Nonet ML. Synaptic transmission. In: Riddle DL, et al., eds. *C. elegans II*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1997:611–643.

8. Miller KG, Alfonso A, Nguyen M, et al. A genetic selection for *Caenorhabditis elegans* synaptic transmission mutants. *Proc Natl Acad Sci USA* 1996;93:12593–12598.

9. Dent JA, Davis MW, Avery L. avr-15 encodes a chloride channel subunit that mediates inhibitory glutamatergic neurotransmission and ivermectin sensitivity in *Caenorhabditis elegans*. *EMBO J* 1997;16:5867–5879.9a. Bosch F, Gomez-Foix AM, Arino J, et al. Effects of lithium ions on glycogen synthase and phosphorylase in rat hepatocytes. *J Biol Chem* 1986;261:16927–16931.

10. Berridge MJ, Downes CP, Hanley MR. Neural and developmental actions of lithium: a unifying hypothesis. *Cell* 1989;59:411–419.

11. Divecha N, Irvine RF. Phospholipid signaling. *Cell* 1995;80:269–278.

12. Lubrich B, Patishi Y, Kofman O, et al. Lithium-induced inositol depletion in rat brain after chronic treatment is restricted to the hypothalamus. *Mol Psychiatry* 1997;2:407–412.

13. Maeda Y. Influence of ionic conditions on cell differentiation and morphogenesis of the cellular slime molds. *Dev Growth Differ* 1970;12:217–227.

14. Harwood AJ, Plyte SE, Woodgett J, et al. Glycogen synthase kinase 3 regulates cell fate in *Dictyostelium*. *Cell* 1995;80:139–148.

15. Klein PS, Melton DA. A molecular mechanism for the effect of lithium on development. *Proc Natl Acad Sci USA* 1996;93:8455–8459.

16. Williams RSB, Eames M, Ryves WJ, et al. Loss of a prolyl oligopeptidase confers resistance to lithium by elevation of inositol (1,4,5) triphosphate. *EMBO J* 1999;18:2734–2745.

17. Maes M, Goosens F, Scharpe S, et al. Lower serum prolyl endopeptidase enzyme activity in major depression: further evidence that peptidases play a role in the pathophysiology of depression. *Biol Psychiatry* 1994;35:545–552.

18. Maes M, Goosens F, Scharpe S, et al. Alterations in plasma prolyl oligopeptidase activity in depression, mania and schizophrenia: effects of antidepressants, mood stabilizers and antipsychotic drugs. *Psychiatry Res* 1995;58:217–225.

19. Acharya J, Labarca P, Delgado R, et al. Synaptic defects and compensatory regulation of inositol metabolism in inositol polyphosphate 1-phosphatase mutants. *Neuron* 1998;20:1219–1229.

20. De Camilli P, Emr SD, McPherson PS, et al. Phosphoinositides as regulators in membrane traffic. *Science* 1996;271:1533–1539.

21. Leonard BE. Second generation antidepressants: chemical diversity but unity of action? In: Montgomery SA, Corn TH, eds. *Psychopharmacology of depression*. Oxford: Oxford University Press, 1994:19–30.

22. Potter WZ. Adrenoreceptors and serotonin receptor function: relevance to antidepressant mechanisms of action. *J Clin Psychiatry* 1996;57(suppl):44–48.

23. McAllister-Williams RH, Young AH. The pathophysiology of depression: a synthesis of the role of serotonin and corticosteroids. In: Ebert D, Ebmeier KP, eds. *New models for depression*. Basel: Karger, 1998.

24. Baldessarini RJ. Drugs and the treatment of psychiatric disorders: depression and mania. In: Hardman JG, et al., eds. *The pharmacological basis of therapeutics*. New York: McGraw-Hill, 1996.

25. Kramer MS, Cutler N, Feighner J, et al. Distinct mechanism for antidepressant activity by blockade of central substance P receptors. *Science* 1998;281:1640–1645.

26. Choy RKM, Thomas JH. Fluoxetine-resistant mutants in *C. elegans* define a novel family of transmembrane proteins. *Mol Cell* 1999;4:143–152.

27. Trent C, Tsung N, Horvitz HR. Egg-laying defective mutants of the nematode *Caenorhabditis elegans*. *Genetics* 1983;104:619–647.

28. Reiner DJ, Weinshenker D, Thomas JH. Analysis of dominant mutations affecting muscle excitation in *Caenorhabditis elegans*. *Genetics* 1995;141:961–976.

29. Weinshenker D, Garriga G, Thomas JH. Genetic and pharmacological analysis of neurotransmitters controlling egg-laying in *C. elegans*. *J Neurosci* 1995;15:6975–6985.

29a. Trudeau MC, Warmke JW, Ganetzky B, et al. HERG, a human inward rectifier in the voltage-gated potassium channel family. *Science* 1995;269:92–95.

29b. Sanguinetti MC, Jiang C, Curran ME, et al. A mechanistic link between an inherited and an acquired cardiac arrhythmia: HERG encodes the IKr potassium channel. *Cell* 1995;81:299–307.

30. Weinshenker D, Wei A, Salkoff L, et al. Block of an ether-a-go-go-like K(+) channel by imipramine rescues egl-2 excitation defects in *Caenorhabditis elegans*. *J Neurosci* 1999;19:9831–9840.

31. Horvitz HR, Chalfie M, Trent C, et al. Serotonin and octopamine in the nematode *Caenorhabditis elegans*. *Science* 1982;216:1012–1014.

32. Waggoner L, Zhou GT, Schafer RW, et al. Control of behavioral states by serotonin in *Caenorhabditis elegans*. *Neuron* 1998;21:203–214.

33. Lackner MR, Nurrish SJ, Kaplan JM. Facilitation of synaptic transmission by EGL-30 Gqalpha and EGL-8 PLCbeta: DAG binding to UNC-13 is required to stimulate acetylcholine release. *Neuron* 1999;24:335–346.

34. Schafer WR, Kenyon CJ. A calcium channel homologue required for adaptation to dopamine and serotonin in *Caenorhabditis elegans*. *Nature* 1995;375:73–78.

35. Schafer WR, Sanchez BM, Kenyon CK. Genes affecting sensitivity to serotonin in *Caenorhabditis elegans*. *Genetics* 1996;143:1219–1230.

36. Krishnan KS, Nash HA. A genetic study of the anaesthetic response: mutants of *Drosophila melanogaster* altered in sensitivity to halothane. *Proc Natl Acad Sci USA* 1990;87:8632–8636.

37. Crowder CM, Shebester LD, Schedl T. Behavioral effects of volatile anesthetics in *Caenorhabditis elegans*. *Anesthesiology* 1996;85:901–912.

38. van Swinderen B, Saifee O, Shebester L, et al. A neomorphic syntaxin mutation blocks volatile anesthetic action in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 1999;96:2479–2484.

39. Saifee O, Wei L, Nonet ML. The *Caenorhabditis elegans* unc-64 locus encodes a syntaxin that interacts genetically with synaptobrevin. *Mol Biol Cell* 1998;9:1235–1252.

40. Nonet ML, Saifee O, Zhao H, et al. Synaptic transmission deficits in *Caenorhabditis elegans* synaptobrevin mutants. *J Neurosci* 1998;18:70–80.

41. Rajaram S, Sedensky M, Morgan PG. A stomatin homologue controls sensitivity to volatile anaesthetics in *C. elegans*. *Proc Natl Acad Sci USA* 1998;95:8761–8766.

42. Shreffler W, Magardino T, Shekdar K, et al. The unc-8 and sup-40 genes regulate ion channel function in *Caenorhabditis elegans* motor neurons. *Genetics* 1995;139:1261–1272.

43. Tavernarakis N, Shreffler W, Wang S, et al. unc-8, a DEG/ENaC family member, encodes a subunit of a candidate mechanically-gated channel that modulates *C. elegans* locomotion. *Neuron* 1997;18:107–119.

44. Rajaram S, Spangler TL, Sedensky MM, et al. A stomatin and a degenerin interact to control anesthetic sensitivity in *Caenorhabditis elegans*. *Genetics* 1999;153:1673–1682.

45. Diamond I, Gordon AS. Cellular and molecular neuroscience of alcoholism. *Physiol Rev* 1997;77:1–20.

46. Koob GF, Sanna PP, Bloom FE. Neuroscience of addiction. *Neuron* 1998;21:467–476.

47. Cohan FM, Hoffman AA. Genetic divergence under uniform selection. II. Different responses to selection for knockdown resistance to ethanol among *Drosophila melanogaster* populations and their replicate lines. *Genetics* 1986;114:145–163.

48. Moore MS, DeZazzo J, Luk AY, et al. Ethanol intoxication in *Drosophila*: genetic and pharmacological evidence for regulation by the cAMP signaling pathway. *Cell* 1998;93:997–1007.

49. Feany MB, Quinn WG. A neuropeptide gene defined by the *Drosophila* memory mutant amnesiac. *Science* 1995;268:869–873.

50. Bainton RJ, Tsai LT-Y, Singh CM, et al. Dopamine modulates acute responses to cocaine, nicotine and ethanol in *Drosophila*. *Curr Biol* 2000;10:187–194.

51. Peto R, Lopez AD, Boreham J, et al. Mortality from smoking worldwide. *Lancet* 1992;339:1268–1278.

52. Dani JA, Heinemann S. Molecular and cellular aspects of nicotine abuse. *Neuron* 1996;16:905–908.

53. Changeux J-P, Devillers-Thiery A, Chemouilli P. Acetylcholine receptor: an allosteric protein. *Science* 1984;225:1335–1345.

54. Simasko SM, Soares JR, Weiland GA. Two components of carmamylcholine-induced loss of nicotinic acetylcholine receptor function in the neuronal cell line PC12. *Mol Pharmacol* 1986;30:6–12.

55. Peng X, Gerzanich V, Anand R, et al. Chronic nicotine treatment up-regulates alpha3 and alpha7 acetylcholine receptor subtypes expressed by the human neuroblastoma cell line SH-SY5Y. *Mol Pharmacol* 1994;46:523–530.

56. Marks MJ, Pauly JR, Gross SD, et al. Nicotine binding and nicotinic receptor subunit RNA after chronic nicotine treatment. *J Neurosci* 1992;12(7):2765–2784.

57. Lewis JA, Wu C-H, Levine JH, et al. Levamisole-resistant mutants of the nematode *Caenorhabditis elegans* appear to lack pharmacological acetylcholine receptors. *Neuroscience* 1980;59:67–89.

58. Lewis JA, Fleming JT, McLafferty S, et al. The levamisole receptor, a cholinergic receptor of the nematode *Caenorhabditis elegans*. *Mol Pharmacol* 1987;31:185–193.

59. Eyre P. Some pharmacodynamic effects of the nematocides: methyridine, tetramisole, and pyrantel. *J Pharm Pharmacol* 1970;22:26–36.

60. Lewis JA, Wu C-H, Berg H, et al. The genetics of levamisole resistance in *Caenorhabditis elegans*. *Genetics* 1980;95:905–928.

61. Fleming JT, Squire MD, Barnes TM, et al. *Caenorhabditis elegans* levamisole resistance genes *lev-1*, *unc-29*, and *unc-38* encode functional nicotinic acetylcholine receptor subunits. *J Neurosci* 1997;17:5843–5857.

62. Fleming JT, Tornoe C, Riina HA, et al. Acetylcholine receptor molecules of the nematode *Caenorhabditis elegans*. *EXS* 1993;63:65–80.

63. Lewis JA, Elmer JS, Skimming J, et al. Cholinergic receptor mutants of the nematode *Caenorhabditis elegans*. *J Neurosci* 1987;7:3059–3071.

64. Waggoner LE, Dickinson KA, Poole DS, et al. Long-term nicotine adaptation in *Caenorhabditis elegans* involves PKC-dependent changes in nicotinic receptor abundance. *J Neurosci* 2000;20:8802–8811.

65. Kim J, Poole DS, Waggoner LE, et al. Genes affecting the activity of nicotinic receptors involved in *C. elegans* egg-laying behavior. *Genetics* 2001;157:1599–1610.

66. Rocha BA, Fumagalli F, Gainetdinov RR, et al. Cocaine self-administration in dopamine transporter knockout mice. *Nature Neurosci* 1998;1:132–137.

67. Sora I, Wichems C, Takahashi N, et al. Cocaine reward models: conditioned place preference can be established in dopamine- and serotonin-transporter knockout mice. *Proc Natl Acad Sci USA* 1998;95:7699–7704.

68. McClung C, Hirsch J. Stereotypic behavioral responses to free-base cocaine and the development of behavioral sensitization in *Drosophila*. *Curr Biol* 1998;8:109–112.

69. Yellman C, Tao H, He B, et al. Conserved and sexually dimorphic behavioral responses to biogenic amines in decapitated *Drosophila*. *Proc Natl Acad Sci USA* 1998;94:4131–4136.

70. Corey JL, Quick MW, Davidson N, et al. A cocaine-sensitive *Drosophila* serotonin transporter: cloning, expression, and electrophysical characterization. *Proc Natl Acad Sci USA* 1994;91:1188–1192.

71. Demchyshyn LL, Pristupa ZB, Sugamori KS, et al. Cloning, expression, and localization of a chloride-facilitated, cocaine-sensitive serotonin transporter from *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 1994;91:5158–5162.

72. Scavone C, McKee M, Nathanson JA. Monoamine uptake in insect synaptosomal preparations. *Insect Biochem Mol Biol* 1994;24:589–597.

73. Torres G, Horowitz JM. Activating properties of cocaine and cocaethylene in a behavioral preparation of *Drosophila melanogaster*. *Synapse* 1998;29:148–161.

74. Li H, Chaney S, Forte M, et al. Ectopic G-protein expression in dopamine and serotonin neurons blocks cocaine sensitization in *Drosophila melanogaster*. *Curr Biol* 2000;10:211–214.

75. McClung C, Hirsh J. The trace amine tyramine is essential for sensitization to cocaine in *Drosophila*. *Curr Biol* 1999;9:853–860.

76. Andretic R, Chaney S, Hirsh J. Requirement of circadian genes for cocaine sensitization in *Drosophila*. *Science* 1999;285:1066–1068.

77. Blenau W, Balfanz S, Baumann A. Amtyr1: characterization of a gene from honeybee (*Apis mellifera*) brain encoding a functional tyramine receptor. *J Neurochem* 2000;74:900–908.

78. Bargmann CI. Neurobiology of the *Caenorhabditis elegans* genome. *Science* 1998;282:2028–2033.

79. Rubin GM, Yandell MD, Wortman JR, et al. Comparative genomics of the eukaryotes. *Science* 2000;287:2204–2215.

22

Beyond Binding: Molecular and Cell Biological Approaches to Studying G-Protein–Coupled Receptors

Gabriel A. Vargas

Mark Von Zastrow

Gabriel A. Vargas: Department of Psychiatry, University of California–San Francisco, San Francisco, California 94143.

Mark Von Zastrow: Departments of Psychiatry, Cellular and Molecular Pharmacology, and Program in Cell Biology, University of California–San Francisco, San Francisco, California 94143.

The origins of the modern concept of receptors can be traced to the beginnings of the 20th century (1). Almost two decades passed until the first neurotransmitter, acetylcholine, was identified from classic physiologic studies on the vagus nerve performed by Otto Loewi in 1921. Since these seminal discoveries the pace of advance has increased enormously. A revolution in the field began in the 1950s, with the discovery that neurotransmitter receptors are targets of clinically relevant psychotropic drugs and the development of radioligand binding techniques (2 ,3). Radioligand binding methodologies remain a mainstay of modern neuropsychopharmacology, and have facilitated the identification of receptor subtypes as well as the discovery of novel receptors that mediate the actions of important drugs.

The application of recombinant DNA methodologies sparked a second revolution in neuropsychopharmacology. These methods facilitated the cloning of complementary DNAs (cDNAs) encoding distinct receptors, the identification of large families of homologous receptors, and unprecedented insight into subtype diversity within individual receptor families (4 ,5).

Important families of receptors include steroid hormone receptors, receptor tyrosine kinases, ligand-gated ion channels, and G-protein–coupled receptors (GPCRs). GPCRs comprise the largest class of signal-transducing receptors, with well over 1,000 members identified in humans. In some organisms, genes encoding GPCRs comprise 1% of the genome (6). GPCRs mediate the actions of the majority of neurotransmitters and neuromodulators, as well as other important biological ligands. These receptors are also critically important drug targets. Indeed, the majority of psychopharmaceuticals presently in use either bind directly to specific GPCRs (e.g., antipsychotics) or indirectly influence GPCR function by modulating the availability of endogenous ligands (e.g., selective serotonin reuptake inhibitors, SSRIs). Therefore elucidating mechanisms of GPCR function and regulation is of central importance to understanding the actions of clinically relevant drugs.

During the past several years there has been a great deal of progress in elucidating specific mechanisms of GPCR function and regulation. Much of this progress can be attributed to the application of newer molecular and cell biological techniques, which have complemented previously developed pharmacologic approaches for probing receptor function. This chapter discusses some of these molecular and cell biological approaches for isolating and studying cloned receptors, focusing specifically on GPCRs expressed in a variety of systems. Although we restrict our scope in this chapter to representative approaches applied to GPCRs, these methods have broad potential application and have been used to study other important receptor families.

ISOLATION AND IDENTIFICATION OF RECEPTORS

Part of "22 - Beyond Binding: Molecular and Cell Biological Approaches to Studying G-Protein–Coupled Receptors "

The identification of GPCRs by biochemical purification is a challenging task because of the generally low abundance of these proteins in cells and tissues, and because GPCRs are highly hydrophobic molecules that are easily denatured when solubilized in detergent solutions. Molecular cloning techniques have greatly facilitated the identification of GPCRs. Molecular cloning takes advantage of the ability to generate and screen “libraries” containing cDNAs corresponding to the messenger RNAs (mRNAs) that encode cellular proteins.

A detailed discussion of molecular cloning techniques is beyond the scope of the present review and has been described elsewhere (7). In general, a cDNA library is generated from a specific tissue and animal source (such as rat brain) by purifying mRNA from the tissue, using the enzyme reverse transcriptase to generate a strand of DNA complementary to each mRNA present in this mixture, and then using a DNA polymerase to generate double-stranded DNA from this sequence that is suitable for insertion into an appropriate plasmid or phage vector that facilitates faithful replication of the sequences and allows selection of individual clones corresponding to a single cDNA. The main challenge in receptor cloning is to isolate or "screen" for the appropriate receptor-encoding cDNA from a library. Several different approaches to library screening have been used successfully for cloning cDNAs encoding GPCRs.

Receptor Cloning from Protein Sequence

Early isolation and cloning of receptors relied on purifying sufficient quantities of receptor and then microsequencing peptide fragments. The hamster $ß_2$-adrenergic receptor was cloned using partial sequence information derived from protein purified from hamster lung (8). The sequence of amino acids present in a GPCR fragment allows one to predict the sequence of the corresponding region of the mRNA sequence by "back-translation" from the genetic code. This nucleic acid sequence can be searched for in the cDNA library by virtue of the ability of complementary strands of DNA to anneal together with extremely high specificity. Because of the degeneracy of the genetic code, many amino acids can be encoded by more than one nucleic acid codon. Therefore, nucleic acid "probes" used to screen cDNA libraries often contain mixtures of sequences representing multiple potential "spellings" of the known peptide fragment. cDNA clones encoding the sequence of interest can be isolated by hybridization of a labeled probe or by polymerase chain reaction (PCR)-mediated amplification from a cDNA mixture.

Receptor Cloning by Sequence Homology

Sequence homology within the GPCR superfamily of receptors enables the use of DNA sequence from a previously cloned receptor to probe libraries for other related receptors. Cloning of the dopamine D2 receptor is an example of where sequence homology was used to clone other members of the GPCR superfamily of receptors. Olivier Civelli's group (9) used the hamster $ß_2$-adrenergic receptor gene as a hybridization probe to isolate cDNA encoding the rat D2 dopamine receptor. The authors used two major criteria to determine that the cDNA isolated encoded a functionally relevant dopamine receptor. First, they demonstrated that mRNA corresponding to their cDNA sequence was expressed in tissues that had been previously shown by radioligand binding to express functional D2 receptors. They accomplished this by Northern blotting, a procedure by which RNAs isolated from cells or tissues is resolved by gel electrophoresis and the specific RNAs homologous to a particular sequence is detected by hybridization of a specifically labeled probe. Second, the authors demonstrated that the cDNA isolated from their library encoded a functional D2-class dopamine receptor. This was accomplished by applying conventional radioligand binding and receptor signaling assays to detect functional D2 receptor activity in fibroblast cells that do not normally express dopamine receptors and that were transfected with the cDNA isolated from library screening.

Receptor Cloning by Functional Expression

GPCRs can also be cloned based on their functional properties. The cloning of the serotonin (5-hydroxytryptamine) 5-HT_{1C} receptor used this approach (10). Taking advantage of the high expression level of the 5-HT_{1C} receptor in the choroid plexus, the authors isolated mRNA from this source and injected this preparation into *Xenopus* oocytes, which allows both injected mRNA and cDNA to be translated and expressed. The activation of the 5-HT_{1C} receptor by its ligand leads to an increase in intracellular calcium through the inositol phosphate signaling pathway. This rise in intracellular calcium leads to the opening of Ca^{2+}-dependent chloride channels that was detected by means of highly sensitive electrophysiologic techniques. Similar results could be obtained by injecting cDNAs produced from the mRNAs. Pools containing progressively smaller numbers of cDNAs were then analyzed until a single cDNA encoding a functional serotonin receptor was isolated.

The cloning of the delta opioid receptor (11 ,12) was also achieved by a functional assay. A cDNA library prepared from cells that endogenously express the delta opioid receptor was expressed in cultured fibroblast-like cells that do not normally express opioid receptors. Cells expressing opioid receptors were identified by binding of a radiolabeled opioid peptide. cDNA isolated from the selected cell was isolated and retransfected, to allow multiple rounds of purification until a single cDNA encoding the delta opioid receptor was isolated. The isolated cDNA was confirmed to encode a delta opioid receptor by analysis of radioligand binding properties and ligand-induced signal transduction in transfected cells.

Identification of Receptor Subtypes (Fig. 22.1)

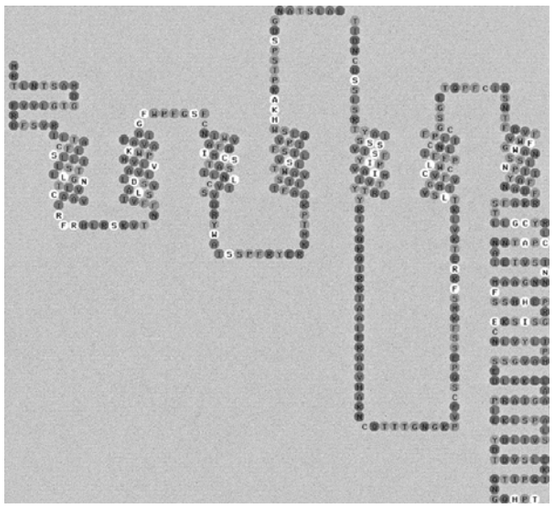


FIGURE 22.1. Dopamine D1 receptor two-dimensional (2D) snake diagram downloaded from the G-protein–coupled receptor database (GCRDb), http://gcrdb.uthscsa.edu/index.html, which is a very useful site for researchers working on GPCRs (13,14). This diagram shows the seven-transmembrane structure of GPCRs and the long carboxy terminus tail characteristic of D1-like dopamine receptors. In the GCRDb, amino acid mutations found in the receptor are listed in *white* in the 2D snake diagram and hyperlinked to the GRAP mutant database. [http://tinygrap.uit.no/]. (Used by permission of the GCRDb.) See color version of figure.

A general principle in molecular pharmacology is that multiple receptors exist that can recognize the same ligand. This "one ligand, multiple receptors" principle has led to great interest in the identification of multiple receptor subtypes.

For example, radioligand binding techniques and assays of receptor-mediated signal transduction originally defined two classes of dopamine receptor: D1 receptors that stimulate adenylyl cyclase, and D2 receptors that inhibit this enzyme. Molecular biological techniques confirmed that these receptors represent distinct gene products and led to the discovery of additional, structurally homologous subtypes of receptor protein. These subtypes of dopamine receptors would have been impossible to identify definitively using classic pharmacologic approaches, as their pharmacologic properties are quite similar and some subtypes are expressed at very low levels in native tissues. For example, the D4 dopamine receptor is a member of the “D2-class” of dopamine receptors that was cloned by sequence homology to the cloned D2 receptor (15). D4 receptors are of great interest because they bind the atypical antipsychotic clozapine with approximately 10-fold higher affinity than cloned D2 receptors (16).

Molecular biological methods also led to the discovery of additional diversity among closely related GPCRs encoded by genetic variants or by modification of the receptor gene after transcription. For example, multiple variants of D4 receptor, which differ only in the structure of the third cytoplasmic loop of the receptor protein, are encoded by closely related genes that are inherited in a mendelian manner ((17)). Two variants of the D2 dopamine receptor are

generated from the same gene by alternative splicing, which occurs during posttranscriptional processing of the RNA. Another interesting example of such receptor diversity is the 5-HT_{2C} receptor, which exists in variant forms determined by a posttranscriptional process called RNA editing (18). In many cases the functional significance of such variation among GPCRs is not known. However in some cases, such as RNA editing of the 5-HT_{2C} receptor, individual receptor variants have significantly different functional properties.

As a further extension of the “one ligand, multiple receptors” concept, molecular biological methods led to the discovery that a specific neurotransmitter can often bind to more than one class of receptor protein. The first example of this can be found in the acetylcholine receptors (AChRs). The idea that distinct muscarinic and nicotinic AChRs exist was first indicated by pharmacologic studies. Molecular cloning of the corresponding receptor cDNAs confirmed this idea and revealed very precisely the differences between these classes of AChR: muscarinic-type AChRs are GPCRs, whereas nicotinic-type AChRs are members of the structurally distinct family of ligand-gated ion channels (LGICs). There are now several examples of this type of diversity, including the existence of distinct GPCRs and LGICs for serotonin, glutamate, and γ-aminobutyric acid (GABA).

Discovery of Orphan Receptors

Orphanin FQ and Its Receptor

In addition to identification of receptor subtypes, the cloning approaches discussed above have also enabled the isolation of GPCRs before the identification of any known ligands. Putative receptors for which a ligand is not known, called orphan receptors, hinge on sequence similarity with other GPCRs. The cloning of an opioid-like GPCR, known as ORL1 or the orphanin receptor, was accomplished by several groups looking for additional members of the opioid receptor family by cloning structurally homologous cDNAs (19 ,20). Expression of the ORL1 cDNA revealed that this putative receptor did not bind any of the typical mu, delta, or kappa ligands with high affinity, despite having particularly high sequence homology to the kappa receptor. In addition, there are large differences in anatomic distribution of ORL1 mRNA compared to the known distribution of opioid receptors. Later, elegant studies led to the isolation of an endogenously expressed peptide ligand for this receptor, nociceptin or orphanin FQ, allowing the ORL1 gene product to be clearly established as a *bona fide* GPCR (21).

Orexins and Their Receptors

Many other cDNAs encoding candidate orphan receptors have been identified by DNA sequence analysis. Current estimates suggest that there are ˜100–200 such orphan GPCRs encoded in the human genome (not including predicted olfactory receptors). The isolation of so many candidate orphan receptors has led some groups to attempt systematic approaches to identifying their endogenous ligands. Identification of orexin and the orexin/hypocretin receptor was accomplished by using a cell-based detection system using cells transfected with a large number of candidate orphan receptor cDNAs (22). A cDNA (HFGAN72) was identified that was capable of causing an elevation in cytoplasmic free calcium in response to a crude peptide-containing extract prepared from brain. A specific peptide ligand was isolated from this mixture according to its ability to activate this orphan GPCR. This peptide is expressed highly in the lateral hypothalamus and influenced feeding behavior when introduced into rat brain, and is called orexin (from the Greek *orexis* meaning appetite). The receptor encoded by HFGAN72 was named the orexin receptor. Recently, a canine narcolepsy gene was identified by positional cloning as belonging to a subtype of orexin receptor (23). Thus the identification of orphan GPCRs can lead to powerful new insights relevant to diverse areas of neuropsychopharmacology.

EXPRESSION AND PURIFICATION OF CLONED GPCRS

Part of "22 - Beyond Binding: Molecular and Cell Biological Approaches to Studying G-Protein–Coupled Receptors "

Expression

The ability to express cloned cDNAs in various cell types has provided powerful tools for studying the functional properties of defined GPCRs. In many cases, receptors expressed in heterologous cell systems have remarkably similar functional properties to those in their native tissue of origin, although this is not always the case. For example, D2 dopamine receptors differ in their properties in pituitary GH4C1 cells and Ltk-fibroblasts. In the pituitary cells, D2 receptors fail to elicit phosphoinositide hydrolysis and induce a decrease of intracellular calcium. In contrast, in the fibroblast cells, the D2 receptor induced a rapid stimulation of inositol 1,4,5-trisphosphate and an increase of intracellular Ca^{2+} (24). Therefore it is important to compare results obtained from studies of cloned receptors in heterologous systems to the properties of receptors in native tissues.

In addition to facilitating functional studies, heterologous expression can also be used to produce large amounts of receptor protein, which is necessary for certain biophysical and structural studies. Mammalian cells are typically used for functional studies of expressed GPCRs; however, it is sometimes preferable to use nonmammalian cells (such as insect cells or yeast cells) for large-scale expression of GPCRs because these cells can be grown economically in very large amount. Prokaryotic expression systems (such as *Escherichia coli*) have potential advantages for large-scale production but have not been used widely in GPCR research because, in general, it has been difficult to obtain functional activity of GPCRs using these systems.

Purification

The ability to express cloned receptors makes it possible to modify the receptor protein to facilitate purification and detection. There are many strategies that have developed from this capability. For example, one useful approach is to insert an antigenic epitope into the GPCR, which can facilitate detection and purification by standard immunochemical procedures using well-characterized antibodies (see ref. 25 for review). This approach obviates the sometimes laborious path of generating antibodies that recognize the native receptor. Epitope tags are typically short sequences (˜10 residues) that bind tightly to a highly specific antibody. In many cases, epitopes chosen for this purpose are either synthetic (not derived from any known biological sequence) or derived from nonmammalian sources, in an attempt to reduce the probability of cross-reaction with endogenous antigens. Popular epitopes include a region from the hemagglutinin molecule of influenza virus (“HA” epitope) or the “Flag” epitope, a synthetic sequence recognized by a series of well-characterized antibodies. The use of such epitopes allows the purification of tagged proteins using commercially available reagents. Tagged receptors can be isolated from extracts prepared from cultured cells by immunoaffinity chromatography. A column can be made by immobilizing an antibody onto a solid support that will bind to tagged receptors as the extracts are passed over the column. Elution of the specifically bound receptor can be accomplished by changing either the salt concentration or the pH to disrupt the antigen-antibody interaction. An example of this approach is found in the high-level production of the human ß-adrenergic receptor (26).

MECHANISMS OF LIGAND BINDING AND ACTIVATION

Part of "22 - Beyond Binding: Molecular and Cell Biological Approaches to Studying G-Protein–Coupled Receptors "

Use of Site-Directed Mutagenesis

The availability of cDNAs encoding specific GPCRs and the development of various cellular expression systems provide powerful tools for examining the specificity of ligand binding to receptors. In particular, site-directed mutagenesis can be used to alter residues in the GPCR structure, and then the functional consequences of these modifications on ligand binding can be examined. Many techniques have been developed for introducing specific mutations into the receptor cDNA, as discussed in detail elsewhere (27). Selected examples of the use of these techniques for understanding the structural basis of ligand binding and receptor activation are outlined below.

Point Mutagenesis

Point mutations refer to modifications of the cDNA that result in the substitution of a single amino acid residue in the expressed receptor protein with another amino acid. Point mutagenesis can be used to dramatically change the physicochemical properties of a specific residue (e.g., substitution of a basic residue with an acidic one); or cause a subtle change in residue structure (such as substitution of a serine residue with cysteine, which results in a change of a single atom in the protein structure). Point mutations are extremely useful because they often do not cause global perturbations of receptor structure and therefore allow highly specific analysis of the function of defined receptor residues. For example, point mutations have been used to define important determinants of receptor-ligand interaction with considerable precision (28).

Point mutations can identify essential features in the receptor structure. Mutation of a single aspartic acid residue present in the predicted third transmembrane domain of catecholamine receptors abrogates high-affinity binding of catecholamines to the $ß_2$-adrenergic receptor. Since an aspartic acid residue is present at the corresponding position in other catecholamine receptors as well, this residue is said to be “conserved” in the structure of multiple receptors. Point mutagenesis applied to these receptors indicates that this aspartic acid is required for ligand binding to a number of receptors, apparently by serving a conserved function as a counter-ion for the positively charged amine moiety of catecholamine ligands (28 ,29). Like the aspartic acid residue discussed above, other amino acids that serve a specific function in the receptor generally are found to be conserved between receptors with similar properties.

Point mutations can identify nonconserved residues that determine the specificity with which drugs bind to structurally homologous receptors. Nonconserved residues present in the receptor sequence may play a pharmacologically important role in determining unique properties of individual receptor subtypes. Nonconserved residues can be essential for determining the specificity with which a drug binds to closely related subtypes of G-protein–coupled receptors. Point mutation analysis combined with appropriate pharmacologic assays can be used to identify such divergent receptor residues that are critical for drug binding, thus providing insight into the structural basis of ligand binding specificity that is useful for drug design.

Point mutations can provide insight into species- and population-specific differences in receptor pharmacology. For the very reason that nonconserved residues are often not essential for basic receptor function, these residues are often not conserved across species. Thus the pharmacology of many subtype-specific drugs can be highly dependent on the species of animal studied. For example, three homologous genes encoding distinct subtypes of α_2-adrenergic receptor are expressed in various mammalian species. However, the pharmacology with which subtype-selective drugs bind to receptor subtypes encoded by mouse or rat receptor genes can differ substantially from the pharmacology characteristic of the corresponding human receptor (30). This

has led to considerable confusion in the correspondence between pharmacologic and molecular biological definitions of specific receptor subtypes across species, and has important implications for the use of animal models for the development of subtype-specific drugs for humans. Furthermore, nonconserved residues involved in subtype-specific drug binding can also differ within the human population, as a result of random mutation and genetic drift. This concept has not yet been extensively explored but may be an important direction for the use of pharmacogenomics in clinical medicine.

Deletion Mutagenesis

Another mutational approach useful for probing receptor structure and function is removal of certain residues from the receptor structure entirely. Deletions of multiple residues in certain parts of the receptor protein (e.g., transmembrane helices) can be difficult to interpret because they often lead to massive disruption of receptor structure. However, deletion of limited regions in extracellular or cytoplasmic domains are often well tolerated and have been quite informative. For example, deletion of residues located in the amino-terminal extracellular domain of polypeptide receptors [such as the follicle-stimulating hormone (FSH) receptor] and the calcium receptor implicate this domain in ligand interaction. Deletion of residues located in the third cytoplasmic loop of various receptors, such as the muscarinic acetylcholine receptors, implicated this domain in functional coupling to heterotrimeric G proteins (27).

Substitution or Chimeric Mutagenesis

A very powerful approach to site-directed mutagenesis is to substitute entire series of residues from one receptor with the corresponding residues of another. This approach is based on the idea that receptors are composed of modular structural domains, and takes advantage of the fact that receptor domains that mediate similar functions often have conserved amino acid sequence. Chimeric substitutions are often less disruptive than deletions to the overall structure of the receptor protein. For example, chimeric mutagenesis has been useful for defining transmembrane residues that mediate subtype-specific and species-specific differences in ligand binding to adrenergic receptors. Receptor chimeras between α_2- and $ß_2$-adrenergic receptors defined multiple cytoplasmic domains that contribute to the specificity of receptor interaction with their cognate heterotrimeric G proteins (4).

Use of Random Mutagenesis

In contrast to site-directed mutagenesis, random mutagenesis is an unbiased approach that can examine a much broader range of modifications of the receptor protein. Random mutagenesis, therefore, has the potential to reveal unanticipated features of receptor structure and function. Random mutagenesis typically requires functional assay of a much larger number of mutant receptors than analyzed using site-directed mutation. The relatively low throughput inherent to traditional methods of receptor characterization have limited the practical utility of random mutagenesis of mammalian GPCRs. This limitation has become less significant with the recent development of higher throughput functional assays and the successful expression of mammalian GPCRs in more genetically tractable organisms. For example, the budding yeast *Saccharomyces cerevisiae* has been used recently for studying the functional properties of a large number of mutant chemokine receptors in which selected regions of the receptor protein were mutagenized in a nonbiased manner. Analysis of these data identified residues in the receptor protein essential for ligand binding and activation. In addition, this nonbiased screening approach yielded unanticipated information, including the identification of mutations that constitutively activate receptors and the identification of functional mutant receptors predicted to contain fewer than seven transmembrane domains (31).

Use of Biophysical Approaches

Biophysical techniques are essential for detailed examination of protein structure and conformational change. One reason these methodologies have had limited application in the study of GPCRs is that they typically require milligram quantities of receptor, a quantity difficult to acquire from native tissue sources. For many years rhodopsin, purified from retina, was the only GPCR that could be generated in sufficient quantity for biophysical study. Indeed, much of what we know about GPCR structure and conformational change has been elucidated from elegant biophysical studies of rhodopsin. Recently, the development of improved expression and purification strategies have made it possible to obtain other GPCRs in sufficient quantity and purity for biophysical study. Thus it is likely that biophysical approaches will play an increasingly important role in future studies of GPCR structure and activation.

Structural Studies of Rhodopsin

High-resolution structural information can be provided by x-ray diffraction methodologies applied to ordered three-dimensional crystals of pure protein. Rhodopsin, a GPCR mediating phototransduction in the retina, has been a favorite for such studies because it can be purified in sufficient amount and purity to facilitate crystalization. Previous studies using electron diffraction of two-dimensional crystals of rhodopsin obtained structural information to a resolution of approximately 7.5 Å (angstroms; 1 angstrom = 10^{-10}m), revealing the relative orientation of the transmembrane helices in the lipid bilayer (32). Recently x-ray diffraction has

been used to solve the structure of three dimensional crystals of rhodopsin to a resolution of 2.8 Å. This accomplishment is truly a major milestone in the field, revealing for the first time the atomic structure of any GPCR and providing detailed information about specific interactions between this GPCR and retinal, its physiological ligand (33).

Molecular Modeling of GPCR Structure

As the hydrophobic domains predicted to form transmembrane helices are extensively conserved among all GPCRs, it is believed that the general features of rhodopsin's transmembrane structure are relevant to other GPCRs. This has motivated the use of rhodopsin's structure as a "template" on which to predict the stucture of other GPCRs, via computational approaches that identify homologous residues and infer thermodynamically stable conformations of extracellular and cytoplasmic loops. It remains to be determined the degree to which specific features of diverse GPCRs are actually conserved at the level of atomic resolution. Indeed, based on well established differences in the pharmacology of individual GPCRs, one might expect there to be significant limitations of such homology-based predictive methods, at least with respect to structural features involved in drug binding. Nevertheless, the available experimental data leave little doubt that this approach is an important starting point for mechanistic studies and for rational drug development (34).

Biophysical Studies of Conformational Dynamics Involved in GPCR Activation

While crystallographic methods have the potential to provide detailed information regarding the relative positions of all residues of the receptor protein, these methods are inherently limited to reporting on a static structure. Thus additional methods are required to examine dynamic conformational transitions that mediate ligand-dependent signal transduction via GPCRs. Several biophysical approaches have been utilized for this purpose. Classic studies of rhodopsin measured the optical absorbance properties of this photoprotein that are highly sensitive to changes in protein conformation. Sophisticated studies using optical spectroscopy indicate that rhodopsin cycles rapidly through a series of distinct conformational states following photon-induced activation. Many other types of biophysical techniques have been applied to examine specific features of light-induced conformational changes of rhodopsin, as well as to examine ligand-induced conformational changes of other GPCRs (35). Specific residues in the receptor protein can be labeled with a chemical probe, typically using a combination of site-directed mutagenesis and organic chemistry techniques. Spectroscopic methods can then be used to detect conformational changes involving the labeled residue, by measuring changes in the local environment or mobility of the chemical probe. Approaches of this type have been applied to several GPCRs, and have begun to yield interesting new information about the dynamic effects of clinically relevant drugs on GPCR structure (29).

Potential for Rational Drug Design

The availability of increasingly detailed mutational and biophysical data and the development of sophisticated molecular models suggest that it may be possible in the future to design new classes of therapeutically useful drugs based on this information. A precedent for such an approach is the structure-based design of the angiotensin-converting enzyme inhibitor captopril, the first drug on the market that was designed based on its interactions with its target (36). Inferences about the structure of GPCR-ligand interaction are currently used in a limited manner to guide the modification of existing drugs. However, an important goal is to design completely new drugs de novo based on the structural basis of GPCR activation. A clue that this may be possible comes from recent studies of mutant GPCRs, in which histidine residues have been introduced at defined positions in the receptor structure that can be coordinated by certain metal ions. Addition of the metal ion to the receptor, by coordinating histidine residues introduced within specific transmembrane helices, influences the receptor conformation to either activate or inactivate the receptor (37). Thus the metal ion can serve either as an "engineered" agonist or antagonist for certain mutant receptors. While it is unlikely that this strategy will directly yield clinically useful drugs, these exciting studies serve as a proof of the principle motivating further studies of GPCR structure and conformational change.

REGULATION OF RECEPTOR SIGNALING

Part of "22 - Beyond Binding: Molecular and Cell Biological Approaches to Studying G-Protein–Coupled Receptors "

Methods to Examine Regulation of Receptors by Posttranslational Modification: GPCR Phosphorylation

Many different types of posttranslational modification have been implicated in the regulation in of GPCR function, localization or stability. A detailed discussion of this large area of research is beyond the scope of this chapter. Instead, we illustrate the use of specific methods by discussing some aspects of protein phosphorylation, the most extensively characterized type of posttranslational modification that regulates GPCRs.

Work by Edwin G. Krebs and his collaborators in the 1950s demonstrated that enzyme-catalyzed protein phosphorylation and dephosphorylation reactions were involved in the regulation of glycogen phosphorylase and suggested the notion of the phosphate group as a "covalently bound allosteric ligand" (38). Since these seminal studies, phosphorylation has been shown to play a critical role in the

regulation of a wide variety of cellular proteins, including many GPCRs. Phosphorylation of mammalian proteins typically occurs on serine, threonine, or tyrosine residues. Serine/threonine phosphorylation is widely recognized to regulate GPCRs. Tyrosine phosphorylation, a more recently discovered modification that is well established to mediate signaling via non-GPCR growth factor receptors (39), may also play a role in regulating certain GPCRs (40).

A family of enzymes called G-protein–coupled receptor kinases (GRKs) are well known to attenuate GPCR signal transduction and promote the endocytosis of certain GPCRs by clathrin-coated pits. Other kinases, such as the 3′,5′-cyclic adenosine monophosphate (cAMP)-dependent protein kinase (PKA) and protein kinase C can also regulate GPCRs by phosphorylating distinct cytoplasmic serine/threonine residues (41 ,42 ,43 and 44). Certain kinases (such as PKA) typically phosphorylate residues located within a well-defined “consensus sequence,” making it possible to predict potential sites of regulatory phosphorylation simply by examination of the primary structure (polypeptide sequence) of the cytoplasmic domains of the receptor. Residues phosphorylated by other kinases, such as GRKs, are more difficult to predict because they do not conform to a rigidly defined consensus sequence. However, even in the case of enzymes with relatively well-understood substrate specificity *in vitro*, there are major limitations to the use of sequence analysis for predicting phosphorylation sites *in vivo*. Residues conforming to a specific consensus sequence are not always phosphorylated under physiologic conditions, and, conversely, in some cases residues that do not conform to a well-defined consensus sequence can be phosphorylated in the intact cell. Thus it is important to determine the phosphorylation of GPCRs when expressed in the appropriate mammalian cells.

Analysis and Identification of Phosphorylated Proteins In Vivo

There are many ways of detecting phosphorylated proteins. A starting point for many of these methods is resolution of phosphorylated proteins by electrophoresis in sodium dodecyl sulfate polyacrylamide gels (SDS-PAGE). In SDS-PAGE, proteins dissolved in SDS are loaded onto one end of a porous gel and exposed to an electric field, which causes the SDS-coated proteins to move as “bands” in the gel according to differences in relative molecular mass. By using appropriate radiolabeled compounds (such as inorganic phosphate added to the culture medium), it is possible to apply the technique of autoradiography to specifically detect radioactive, phosphorylated proteins resolved by SDS-PAGE. It is also possible to use gel electrophoresis to separate proteins according to relative charge, a property that is modified predictably by certain modifications such as phosphorylation. These types of separation can be combined in the use of two-dimensional gel electrophoresis, which allows high resolution of proteins as “spots” differing in relative size and charge.

Proteins resolved by gel electrophoresis can be transferred to a membrane composed of nitrocellulose or polyvinyl difluoride (PVDF). This allows many manipulations to be performed, such as detection of a specific protein from a complex mixture by the ability of the protein to be bound by a specific antibody. This procedure, called immunoblotting or “Western” blotting, can be used with commercially available antibodies recognizing phosphorylated peptide sequences or phosphorylated amino acids.

GPCRs resolved by gel electrophoresis can also be analyzed by chemical sequencing, typically by a process called Edman degradation, which sequentially cleaves residues from the amino-terminal end of the protein. Phosphorylated residues can be distinguished from their nonphosphorylated counterparts by chromatography or by incorporation of radioactive phosphate, allowing the identification of specific phosphorylated residues in a polypeptide sequence by the order of appearance in the eluate collected after multiple cycles of Edman degradation. A very powerful method for determining amino acid sequence and detecting posttranslational modifications of proteins is via mass spectrometric analysis. For example, with tandem mass spectrometry it is possible to measure the mass of specific protein fragments with an accuracy of one part in 10,000 up to 12,000 daltons and one part in 1,000 up to 25 kd (45). The impressive accuracy of this method makes it possible to detect phosphorylation as well as many other posttranslational modifications, even those that cause subtle changes in the protein size or charge.

Chromatography, which refers to any separation based on differential behavior of a molecule between a stationary phase and a moving phase, offers many ways of identifying protein modifications. High-performance liquid chromatography (HPLC) using reverse-phase (e.g., C18) columns allows the precise separation of peptides derived from proteolytic or chemical fragmentation of GPCRs. By comparing the pattern of peptide fragments derived from the native protein and the modified protein, one can identify specific polypeptide fragments containing the modification of interest. Subsequently, these fragments can be isolated and further analyzed by methods such as Edman degradation or mass spectrometry.

Methods to Examine Regulation of Receptors by Localization and Trafficking

It has been appreciated for many years that a critical parameter that can regulate the strength of functional signal transduction via GPCRs is the actual number of receptors present in target tissues and, in particular, the number of receptors present in the plasma membrane of individual cells. Indeed, disturbances in the regulation of receptor number and/or distribution may be of primary importance in the pathophysiology

of certain neuropsychiatric disorders. For example, long-term administration of dopamine receptor antagonists can induce upregulation of specific receptors, which may contribute to the apparent supersensitivity of dopamine receptors associated with tardive dyskinesia (46). In contrast, prolonged stimulation of certain GPCRs with agonist ligands can lead to a decrease in the number of binding sites available on the cell surface. This phenomenon is termed receptor “down-regulation” and may contribute to the effects of certain antidepressant drugs (47). Studies using radioligand binding and subcellular fractionation techniques provided early evidence that multiple mechanisms are capable of mediating changes in the number of GPCRs present at the cell surface (48). More recently developed molecular and cell biological approaches provide powerful tools for directly visualizing the subcellular localization of GPCRs and for performing biochemical studies of specific receptor trafficking mechanisms.

Immunochemical Methods to Visualize the Subcellular Localization of Receptors

GPCRs can be detected *in situ* in cell or tissue preparations using immunochemical techniques and receptor-specific antibodies. Antibodies that recognize the native receptor protein can be used to examine the localization of endogenously expressed receptors, whereas epitope-tagging methods (see above) can be used to detect mutated versions of the receptor protein or as a means to detect recombinant receptors for which antibodies recognizing the native receptor are not available. In either case the general scheme is as follows: Cells or tissues expressing the receptor of interest are fixed using standard histologic methods. The fixed cells or tissue can be “permeabilized” with a nonionic detergent, to facilitate biochemical access to receptors situated in intracellular membranes, and then specimens are incubated with antibodies recognizing the receptor of interest. After sufficient time has elapsed to allow antibodies to bind to their respective epitopes in the specimen (typically several hours), the specimens are washed extensively to remove nonspecifically associated antibodies. Antibodies bound to the receptor are then detected by incubation of specimens with a “secondary” antibody that binds specifically to the “primary” antibody (the antibody bound initially to the receptor). The secondary antibody is typically coupled to a fluorochrome (such as fluorescein), a recognizable particle (such as colloidal gold), or an enzyme that can produce a localized reaction product (such as horseradish peroxidase) to facilitate direct visualization of the receptor-containing immune complex using various light or electron microscopic techniques (Fig. 22.2).

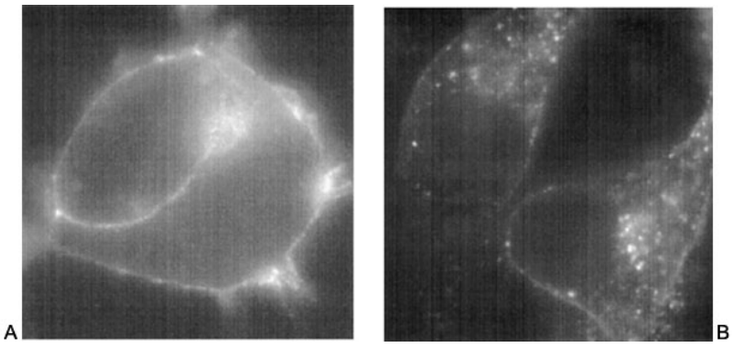


FIGURE 22.2. Visualization of HA epitope-tagged dopamine D1 receptors in transfected cells, using a fluorochrome-labeled secondary antibody and fluorescence microscopy. The ability of this receptor to undergo regulated internalization is indicated by the dopamine-induced redistribution of immunoreactive receptors from the plasma membrane (visualized as linear staining at the cell periphery) to endocytic vesicles (visualized as punctate structures located throughout the cytoplasm). **A:** Untreated cells (no ligand). **B:** Treated with 10 μM dopamine. (Photograph courtesy of Gabriel Vargas.)

Biochemical Methods to Assay Specific Receptor Trafficking Processes

Whereas microscopic imaging can readily provide a great deal of qualitative information about GPCR localization and trafficking, it can be quite challenging to quantitiate

from these data the precise amount of receptor present in a specific subcellular localization or to measure accurately the rate or extent of specific trafficking processes. The importance of these processes has motivated the development of biochemical methods for examining GPCR trafficking. In addition to their utility for receptor localization, antibodies specifically recognizing GPCRs facilitate biochemical studies of GPCR trafficking using techniques adapted from other areas of cell and molecular biology. For example, one method that has been extremely useful for quantitative studies of GPCR endocytosis is cell-surface biotinylation coupled with immunoprecipitation of receptors. Proteins present in the plasma membrane of cells can be specifically labeled by incubating intact cells in the presence of biotin coupled to an activated ester, which is membrane-impermeant and therefore forms a covalent bond only with exposed amine moieties present in plasma membrane proteins. In general, biotinylation in this manner does not adversely affect GPCR function, allowing biotinylation to be used as a chemical tag for surface receptors. The biotin moiety is extremely useful for subsequent detection or purification of surface-tagged proteins because it binds with extremely high affinity (K_a ~10^{15} M^{-1}) to the proteins avidin or strepatavidin. Using variations of this basic biochemistry, it is possible to measure a wide variety of membrane trafficking processes. For example, internalization of GPCRs has been measured by the inaccessibility of biotinylated receptors to a membrane-impermeant reducing agent that “cleaves” the biotin moiety away from tagged proteins (49), and surface biotinylation has been used to measure the rate and extent of proteolytic degradation of receptors after endocytosis (50 ,51).

Methods for Examining Specific Protein Interactions Involved in GPCR Function and Regulation

A salient lesson emerging from recent cell biological studies is that GPCR signal transduction can be viewed, in essence, as a dynamically regulated network of protein–protein interactions that occur in specific subcellular locations. Therefore, an important goal of current and future research is to define these critical protein interactions and elucidate their temporal and spatial regulation in intact cells and tissues. A great deal of effort is presently going into developing and applying novel methods for the study of protein–protein interactions both *in vitro* and *in vivo*, as illustrated by the following examples.

Coimmunoprecipitation Techniques to Examine Defined Protein Interactions with GPCRs in Intact Cells

As discussed above, it is possible to rapidly purify GPCRs from cell or tissue extracts using receptor-specific antibodies or epitope tagging methods. In addition to being extremely useful for examining posttranslational modifications of GPCRs, in some cases it is possible to use these techniques to isolate receptor-containing complexes that presumably reflect protein interactions occurring in the intact cell. The basic idea is to immunopurify a specific GPCR from cell or tissue extracts (or from a partially purified subcellular fraction prepared from a cell or tissue lysate) using an antibody recognizing the native receptor or an engineered epitope tag, and then to analyze proteins specifically associated with this complex using a different antibody. In general, this is accomplished by immunoprecipitation of the receptor followed by analysis of associated proteins in the complex by immunoblotting with the appropriate additional antibody. In some cases, the protein complexes are sufficiently stable that they remain associated through the initial immunopurification of the receptor. In other cases this is not true, and the complexes dissociate before the receptor can be purified from the extract. In this case, various chemical agents can be added prior to cell lysis to physically “cross-link” closely associated proteins with one another by forming covalent bonds that prevent dissociation of the complex. Coimmunoprecipitation has been used to assay GPCR interaction with heterotrimeric G proteins (52) and with ß-arrestins (53), and to examine the regulation of these protein interactions by ligand-induced activation of the receptor.

Use of Coimmunoprecipitation to Examine Oligomerization of GPCRs

The idea that GPCRs may function *in vivo* as higher-order molecular complexes has been suspected for many years. Recent studies provide strong support for this idea and, specifically, provide evidence for homo- and heterodimerization of individual GPCRs *in vivo*. This principle is perhaps best established for receptor tyrosine kinases, where it is well established that oligomerization of receptors is required for appropriate ligand-dependent signal transduction (54). A relatively early hint that GPCRs may also undergo oligomerization came from studies of the $ß_2$-adrenergic receptor using epitope-tagging techniques, where it was observed that receptors tagged with one epitope could specifically coimmunoprecipitate receptors tagged with a distinct epitope (55). Early evidence suggesting a functional role of oligomerization in GPCR signaling came from mutational studies in which structural domains present in distinct, functionally inactive mutant receptors could “complement” one another when coexpressed in cells, suggesting the formation of a functional “hybrid” oligomeric receptor protein (56). More recently, evidence for oligomerization of many GPCRs has been reported. A particularly compelling example of this is the recent observation that distinct subtypes of GABA-B receptor hetero-oligomerize in cells, and that oligomerization is essential for the formation of recombinant receptors possessing the functional properties characteristic

of native GABA-B receptors observed *in vivo* (57 ,58). Recent studies using epitope-tagging and coimmunoprecipitation have demonstrated the formation of homo- and heterodimers of opioid receptors, and suggest that receptor oligomerization may contribute to the remarkable diversity of pharmacologic properties observed in natively expressed opioid receptors (59). There is also emerging evidence that certain GPCRs may associate *in vivo* with completely different classes of receptor protein, such as the dopamine D5 receptor (a GPCR) and the GABA-A receptor (a ligand-gated ion channel). In a recently published study (60), glutathione S-transferase (GST)-fusion proteins encoding the C-terminal tail of the D5 receptor were shown to interact with the GABA-A receptor present in rat hippocampal extracts. Additionally, using an antibody recognizing the dopamine D5 receptor, it was possible to coimmunoprecipitate the GABA-A receptor from cell extracts. Interestingly, this coimmunoprecipitation was detected only when both receptors were stimulated by their respective ligands, suggesting that this heterotypic interaction is regulated in a ligand-dependent manner.

Identification of Novel Protein Interactions with GPCRs

In addition to known proteins that mediate and regulate GPCR signaling (heterotrimeric G proteins, GRKs, arrestins), which were originally identified by functional assays using biochemical purification, cDNA cloning methods have facilitated the identification of additional protein interactions with GPCRs that were completely unanticipated (61). These novel protein interactions, while their functional relevance remains unclear in many cases, are of great interest and potential therapeutic importance as drug targets.

Of the many techniques for identifying novel protein–protein interactions developed over the last 10 years, interaction cloning methods such as the yeast two-hybrid system (62) have been particularly useful for studies of GPCRs. In the yeast two-hybrid system, protein interactions are detected by their ability to reconstitute the activity of a “split” transcriptional activator complex. A transcription factor such as GAL4 can be divided into two domains: a DNA binding domain and a transcriptional activation domain. For the transcription factor to be active, these two domains must be in close proximity to one another, so that the DNA binding domain can bind the promoter sequence in a “reporter” gene and the activation domain can promote gene transcription. A polypeptide sequence for which one wishes to identify putative interacting proteins (such as a sequence derived from a cytoplasmic domain of a GPCR) is cloned into a vector coding for the isolated DNA binding domain from the GAL4 transcription factor, thereby producing a fusion protein containing the GPCR-derived sequence as “bait” with which to search for potential protein interactions. A cDNA library prepared from a tissue of interest is cloned into a cDNA encoding an isolated transcriptional activation domain, producing a large number of fusion proteins containing tissue-derived polypeptide sequences as potential “prey” for protein interaction with the GPCR-derived fusion protein. Both the bait and prey plasmids are transformed into a strain of yeast harboring a “reporter” gene that can be transcribed only in the presence of an “intact” GAL4 transcription factor. Either “half” of the transcription factor is not sufficient to promote efficient transcription of the reporter gene. However, if the fused bait and prey polypeptides form a sufficiently stable protein–protein interaction, they bring their corresponding DNA binding and transcriptional activation domains into close proximity, thus reconstituting transcriptional activation of the reporter gene. Transformed yeast cells containing plasmids encoding the corresponding interacting protein domains can be identified by screening techniques based on GAL4-dependent transcription of reporter genes conferring antibiotic resistance or other selectable metabolic activities, or encoding enzymes that can be detected using colorimetric assays. Assays using the *E. coli*–derived *lacZ* gene, for example, can be used to screen for a characteristic blue reaction product when exposed to 5-bromo-4-chloro-3-indolyl α-D-galactoside.

Protein interactions suggested to occur by the yeast two-hybrid system can be examined using various *in vitro* biochemical techniques, such as affinity chromatography facilitated by GST-fusion proteins. In addition to serving as an independent assay for previously defined candidate interacting proteins, this method can be used to identify novel protein interactions with GPCRs *de novo* (63). In this method a DNA encoding a polypeptide sequence of interest is fused to GST using standard cDNA cloning techniques and expressed as a recombinant protein in *E. coli*. The GST portion of the fusion protein allows the efficient immobilization of the protein by binding to agarose beads covalently derivatized with glutathione. Proteins from a cell or tissue extract that bind to the fusion protein then can be isolated as an immobilized protein complex by affinity chromatography. As an example of the use of these methods, it was shown recently (64) that the third cytoplasmic loop of the dopamine D2 receptor binds specifically to spinophilin, a large cytoskeleton-associated protein that also binds to protein phosphatase-1. This binding was initially identified through use of the yeast two-hybrid system, and then the identification of the specific domains that mediate this protein interaction was accomplished by affinity chromatography using GST-fusion proteins.

Methods for Examining Candidate Protein Interactions in Intact Cells

A major question regarding novel protein interactions with GPCRs, such as those identified using interaction cloning

or protein affinity chromatography, is whether or not they actually occur in an intact cell. Immunocytochemical techniques can provide some insight into this question by determining whether candidate interacting proteins “colocalize” in cells with the appropriate GPCR, as expected if the proteins physically interact *in vivo*. However, even in the event that extensive colocalization is observed, immunocytochemical techniques of this sort do not provide direct evidence for a physical interaction between candidate proteins. Coimmunoprecipitation techniques, as discussed above, provide a useful method for addressing this question. However, demonstrating that a specific protein association can occur *in vivo* is only the first step in the process of assessing the potential physiologic relevance of a novel protein interaction, as this method generally does not provide any information regarding the possible functional activity of a candidate protein interaction. Addressing this question can be a challenging task that involves creative application of diverse techniques and functional assays. Examples of novel protein interactions with GPCRs for which compelling functional data exist include the aforementioned interaction of the D2 dopamine receptor with ABP280 (65) and interaction of the ß$_2$-adrenergic receptor with NHERF/EBP50-family proteins (51 ,63).

EMERGING HORIZONS

Part of "22 - Beyond Binding: Molecular and Cell Biological Approaches to Studying G-Protein–Coupled Receptors "

Unexpected Signaling, Cross-Talk, and Transactivation Involving GPCRs (Fig. 22.3)

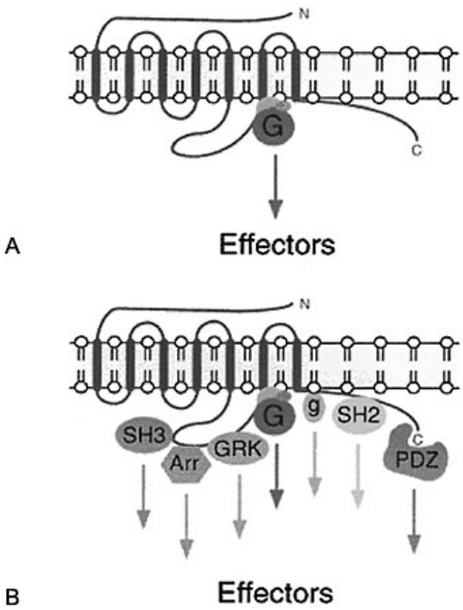


FIGURE 22.3. Schematic diagram of G-protein–coupled receptor (GPCR) signaling. **A:** The G-protein paradigm. Following agonist binding, GPCRs activate heterotrimeric G proteins (G), which then regulate the activity of specific cellular effectors. **B:** Beyond the G-protein paradigm. Following agonist binding, GPCRs can associate with members of diverse families of intracellular proteins, including heterotrimeric G proteins (G), polyproline-binding proteins such as those containing SH3 domains (SH3), arrestins (Arr), G-protein–coupled receptor kinases (GRK), small guanosine triphosphate (GTP)-binding proteins (g), SH2 domain–containing proteins (SH2), and PDZ domain–containing proteins (PDZ). These interactions allow GPCRs to initiate multiple intracellular signaling pathways, with each subtype of receptor likely coupled to a relatively unique set of effectors. (From Hall RA, Premont RT, Lefkowitz RJ. Heptahelical receptor signaling: beyond the G protein paradigm. *J Cell Biol* 1999;145:927–932, with permission.) See color version of figure.

Another line of evidence suggesting the existence of functionally relevant, novel protein interactions involving GPCRs comes from recent work by several labs suggesting that unanticipated functional interactions can occur between GPCRs and receptor tyrosine kinases (RTKs), a distinct family of single-transmembrane receptors involved in growth, differentiation, and oncogenesis (66). The RTK family includes the epidermal growth factor receptor (EGFR), the first receptor shown to have intrinsic tyrosine kinase activity (67 ,68). Whereas the classic pathway for RTK-mediated signaling is initiated by binding of polypeptide growth factors (such as EGF) to the extracellular domain of the RTK, it has been observed recently that certain GPCRs can initiate signaling cascades traditionally thought to be controlled by RTKs. In this situation, the primary signal appears to be through the GPCR, which in turn “transactivates” the RTK. For example, several GPCRs can mediate transactivation of coexpressed EGFRs, thus stimulating mitogenesis by a similar downstream pathway as that initiated by binding of EGF directly to the EGFR (69). One mechanism of GPCR-mediated transactivation involves the activation of a membrane-associated metalloproteinase, which cleaves the EGF precursor protein to generate increased amounts of ligand for the EGFR (70). Another mechanism of cross-talk involves the formation of heteromeric signaling complexes, which include components of both “classical” GPCR and RTK signaling cascades. For example, recent studies suggest that the nonreceptor tyrosine kinase c-Src can associate with the ß$_2$-adrenergic receptor and the ß-arrestin in endocytic membranes, thus mediating mitogenic kinase activation either by a c-Src-mediated phosphorylation of downstream effectors (71) or by c-Src-mediated phosphorylation of co-endocytosed EGFR (72).

Visualization of Protein Localization and Interaction in Living Cells

As discussed above, immunochemical methods are useful for examining the localization of proteins in intact cells.

However, these methods are typically applied to fixed cells because they require disruption of the cell membrane and prolonged incubation of specimens with antibodies used to detect the receptor of interest. The discovery of proteins from certain marine animals that have high levels of intrinsic fluorescence has fostered a revolution in the ability to localize proteins in living cells. These proteins, such as the green fluorescent protein (GFP) isolated from the jellyfish *Aequorea victoria*, are brightly fluorescent molecules that can fold properly in many environments and do not require any additional chromophore for their fluorescence (73 ,74). This allows them to be used to “tag” GPCRs and other important signaling proteins in intact cells. This is accomplished by using site-directed mutagenesis to create a fusion between the GPCR polypeptide sequence and the sequence encoding the GFP tag, analogous to the introduction of an antigenic epitope tag. The localization of the fusion protein can be examined in intact cells using fluorescence microscopy. Examples of this methodology include the visualization of ligand-induced endocytosis of a GFP-tagged $ß_2$-adrenergic receptor in living cells and visualizing the dynamic recruitment of GFP-tagged ß-arrestin from the cytoplasm to the plasma membrane induced by activation of various GPCRs (75 ,76).

While GFP has facilitated the localization of proteins in living cells, localization by itself does not necessarily indicate the occurrence of a physical interaction of a GPCR with a specific protein. The development of mutant versions of GFP, which differ in their excitation and emission spectra, has made it feasible to examine *in vivo* protein interactions using the process of fluorescence resonance energy transfer (FRET) (77). FRET occurs when two suitable fluorophores are present in extremely close proximity so that light produced from one fluorophore can be “transferred” efficiently into exciting the other. FRET can be detected in living cells using sophisticated microscopy, making it possible in principle to detect specific protein interactions with GPCRs and study their localization in real time. FRET imaging has not yet been used extensively for GPCR research but holds great promise for future study of the spatial and temporal dynamics of protein interactions with GPCRs in intact cells and tissues.

SUMMARY AND CONCLUSIONS

Part of "22 - Beyond Binding: Molecular and Cell Biological Approaches to Studying G-Protein–Coupled Receptors "

We have discussed a subset of experimental approaches that have provided powerful new tools for studying GPCR function and regulation. These approaches are responsible, in large part, for the vast explosion of new information about specific mechanisms of GPCR biology that has emerged over the past several years. In many cases these developments have extended directly from seminal observations made originally through classic pharmacologic approaches, which remain of central importance to understanding GPCR function and regulation. Indeed, we view newer molecular and cell biological approaches as complementing, rather than replacing, the sophisticated pharmacologic methods that have been developed over the years since the discovery of receptors as important drug targets.

Molecular cloning techniques have allowed the isolation of cDNAs encoding many G-protein–coupled receptors. The isolation of receptor cDNAs has provided insight into the remarkable structural homology among GPCRs, revealed an unanticipated level of molecular diversity in the GPCR superfamily, allowed functional characterization of defined receptor subtypes in heterologous systems, and made it practical to produce large amounts of receptor protein for pharmacologic, biochemical and biophysical studies.

Structural, biophysical, and molecular modeling approaches hold great promise for ultimately defining the precise atomic determinants of receptor-ligand interaction and for understanding protein conformational changes involved in receptor activation and regulation. Continued progress in this important area may lead to entirely new concepts and methods relevant to therapeutic drug design. Site-directed mutagenesis techniques complement structural and biophysical approaches and have enabled, in the absence of precise structural information, the empirical identification of residues and receptor domains important for ligand binding and activation. Cell biological methods have elucidated mechanisms of signal transduction and regulation in impressive detail, and have revealed a previously unanticipated level of specificity and complexity of crosstalk between signal transduction systems. Emerging technologies for detecting protein interactions in intact cells are suggesting new insights into cell biological mechanisms of GPCR function and regulation, and are beginning to allow real-time examination of the temporal and spatial dynamics of defined protein interactions in living cells.

Based on the new methodologies available today and current pace of progress in using these methods for elucidating GPCR function and regulation, we anticipate that the next several years will see even greater progress in our understanding of the fundamental biology of GPCRs. Indeed, the field of GPCR research is rapidly moving away from a focus on any one set of experimental techniques and has become a vanguard area of integrative structural, molecular, and cell biology. Further developments of these experimental methods, combined with new *in vivo* imaging and genomics approaches that have appeared on the horizon, are likely to fuel continued rapid progress in the field. This exciting progress is fundamentally and directly relevant to the main mission of neuropsychopharmacology: to develop and provide effective therapies for the complex neuropsychiatric disorders that affect our patients.

REFERENCES

1. Langley JN. On nerve endings and on special excitable substances in cells. *Proc R Soc Lond* 1906;B78:170–194.

2. Snyder SH. *Drugs and the brain*. New York: Scientific American Library, WH Freeman, 1996.

3. Barondes SH. *Molecules and mental illness*. New York: WH Freeman, Basingstoke, 1999.

4. Lefkowitz RJ, Kobilka BK, Caron MG. The new biology of drug receptors. *Biochem Pharmacol* 1989;38:2941–2948.

5. Hartman DS, Civelli O. Dopamine receptor diversity: molecular and pharmacological perspectives. *Prog Drug Res* 1997;48:173–194.

6. Mombaerts P. Seven-transmembrane proteins as odorant and chemosensory receptors. *Science* 1999;286:707–711.

7. Old RW, Primrose SB. *Principles of gene manipulation: an introduction to genetic engineering*. Boston: Blackwell Scientific, 1994.

8. Kobilka BK, Dixon RA, Frielle T, et al. cDNA for the human beta 2-adrenergic receptor: a protein with multiple membrane-spanning domains and encoded by a gene whose chromosomal location is shared with that of the receptor for platelet-derived growth factor. *Proc Natl Acad Sci USA* 1987;84:46–50.

9. Bunzow JR, Van Tol HH, Grandy DK, et al. Cloning and expression of a rat D2 dopamine receptor cDNA. *Nature* 1988;336:783–787.

10. Julius D, MacDermott AB, Axel R, et al. Molecular characterization of a functional cDNA encoding the serotonin 1c receptor. *Science* 1988;241:558–564.

11. Kieffer BL, Befort K, Gaveriaux RC, et al. The delta-opioid receptor: isolation of a cDNA by expression cloning and pharmacological characterization. *Proc Natl Acad Sci USA* 1992;89:12048–12052.

12. Evans CJ, Keith DJ, Morrison H, et al. Cloning of a delta opioid receptor by functional expression. *Science* 1992;258:1952–1955.

13. Horn FWJ, Beukers MW, Horsch S, et al. GPCRDB: an information system for G protein-coupled receptors. *Nucleic Acids Res* 1998;26:275–279.

14. Campagne FJR, Reversat JL, Bernassau JM, et al. Visualisation and integration of G protein-coupled receptor related information help the modelling: description and applications of the Viseur program. *J Comput Aided Mol Des* 1999;13:625–643.

15. Van TH, Bunzow JR, Guan HC, et al. Cloning of the gene for a human dopamine D4 receptor with high affinity for the antipsychotic clozapine. *Nature* 1991;350:610–614.

16. Van Tol HH, Wu CM, Guan HC, et al. Multiple dopamine D4 receptor variants in the human population. *Nature* 1992;358:149–152.

17. Asghari V, Schoots O, Van Kats S, et al. Dopamine D4 receptor repeat: analysis of different native and mutant forms of the human and rat genes. *Mol Pharmacol* 1994;46:364–373.

18. Burns CM, Chu H, Rueter SM, et al. Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature* 1997;387:303–308.

19. Bunzow JR, Maneckjee R, Unteutsch A, et al. Pharmacological and functional characterization of a putative member of the opioid receptor family. *Soc Neurosci Abst* 1994;20:744.

20. Chen Y, Fan Y, Liu J, et al. Molecular cloning, tissue distribution and chromosomal localization of a novel member of the opioid receptor gene family. *FEBS Lett* 1994;347(2 and 3):279–283.

21. Reinscheid RK, Nothacker HP, Bourson A, et al. Orphanin FQ: a neuropeptide that activates an opioidlike G protein-coupled receptor. *Science* 1995;270:792–794.

22. Sakurai T, Amemiya A, Ishii M, et al. Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* 1998;92:573–585.

23. Lin L, Faraco J, Li R, et al. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* 1999;98:365–376.

24. Vallar L, Muca C, Magni M, et al. Differential coupling of dopaminergic D2 receptors expressed in different cell types. Stimulation of phosphatidylinositol 4,5-bisphosphate hydrolysis in LtK-fibroblasts, hyperpolarization, and cytosolic-free Ca2+ concentration decrease in GH4C1 cells. *J Biol Chem* 1990;265:10320–10326.

25. Klein U, von Zastrow M. Epitope tagging and detection methods for receptor identification. In Benovic JL. *Regulation of G-protein coupled receptor function and expression*. New York: Wiley-Liss, 2000.

26. Kobilka BK. Amino and carboxyl terminal modifications to facilitate the production and purification of a G protein-coupled receptor. *Anal Biochem*. 1995;231:269–271.

27. Wess J, Blin N, Mutschler E, et al. Muscarinic acetylcholine receptors: structural basis of ligand binding and G protein coupling. *Life Sci* 1995;56:915–922.

28. Cascieri MA, Fong TM, Strader CD. Molecular characterization of a common binding site for small molecules within the transmembrane domain of G-protein coupled receptors. *J Pharmacol Toxicol Methods* 1995;33:179–185.

29. Gether U, Kobilka BK. G protein-coupled receptors. II. Mechanism of agonist activation. *J Biol Chem* 1998;273:17979–17982.

30. Bylund DB. Pharmacological characteristics of alpha-2 adrenergic receptor subtypes. *Ann NY Acad Sci* 1995;763:1–7.

31. Baranski TJ, Herzmark P, Lichtarge O, et al. C5a receptor activation. Genetic identification of critical residues in four transmembrane helices. *J Biol Chem* 1999;274:15757–15765.

32. Unger VM, Hargrave PA, Baldwin JM, et al. Arrangement of rhodopsin transmembrane alpha-helices. *Nature* 1997;389:203–206.

33. Palczewski K, Kumasaka T, Hori T, et al. Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* 2000;289:739–745.

34. Ballesteros JA, Shi L, Javitch JA. Stuctural mimicry in G protein-coupled receptors: implications of the high-resolution stucture of rhodopsin for structure-function analysis of rhodopsin-like receptors. *Mol Pharmacol* 2000;60:1–19.

35. Farrens DL, Altenbach C, Yang K, et al. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* 1996;274:768–770.

36. Opie LH. ACE inhibitors: almost too good to be true. *Scientific American Science and Medicine* 1994;July/August:14–23.

37. Elling CE, Thirstrup K, Holst B, et al. Conversion of agonist site to metal-ion chelator site in the beta(2)-adrenergic receptor. *Proc Natl Acad Sci USA* 1999;96:12322–12327.

38. Krebs EG. The phosphorylation of proteins: a major mechanism for biological regulation. Fourteenth Sir Frederick Gowland Hopkins memorial lecture. *Biochem Soc Trans* 1985;13:813–820.

39. Hunter T, Cooper JA. Protein-tyrosine kinases. *Annu Rev Biochem* 1985;54:897–930.

40. Karoor V, Baltensperger K, Paul H, et al. Phosphorylation of tyrosyl residues 350/354 of the beta-adrenergic receptor is obligatory for counterregulatory effects of insulin. *J Biol Chem* 1995;270:25305–25308.

41. Ferguson SS, Zhang J, Barak LS, et al. Role of beta-arrestins in the intracellular trafficking of G-protein-coupled receptors. *Adv Pharmacol* 1998;42:420–424.

42. Carman CV, Benovic JL. G-protein-coupled receptors: turn-ons and turn-offs. *Curr Opin Neurobiol* 1998;8:335–344.

43. Freedman NJ, Lefkowitz RJ. Desensitization of G protein-coupled receptors. *Recent Prog Horm Res* 1996;51:319–351; discussion 352–313.

44. Lefkowitz RJ, Pitcher J, Krueger K, et al. Mechanisms of beta-adrenergic receptor desensitization and resensitization. *Adv Pharmacol* 1998;42:416–420.

45. Biemann K, Scoble HA. Characterization by tandem mass spectrometry of structural modifications in proteins. *Science* 1987;237:992–998.

46. Meshul CK, Casey DE. Regional, reversible ultrastructural changes in rat brain with chronic neuroleptic treatment. *Brain Res* 1989;489:338–346.

47. Stahl S. 5HT1A receptors and pharmacotherapy. Is serotonin receptor down-regulation linked to the mechanism of action of antidepressant drugs? *Psychopharmacol Bull* 1994;30:39–43.

48. Staehelin M, Simons P. Rapid and reversible disappearance of beta-adrenergic cell surface receptors. *EMBO J* 1982;1:187–190.

49. Vickery RG, von Zastrow M. Distinct dynamin-dependent and -independent mechanisms target structurally homologous dopamine receptors to different endocytic membranes. *J Cell Biol* 1999;144:31–43.

50. Tsao PI, von Zastrow M. Type-specific sorting of G protein-coupled receptors after endocytosis. *J Biol Chem* 2000;275:11130–11140.

51. Cao TT, Deacon HW, Reczek D, et al. A kinase-regulated PDZ-domain interaction controls endocytic sorting of the beta2-adrenergic receptor. *Nature* 1999;401:286–290.

52. Law SF, Manning D, Reisine T. Identification of the subunits of GTP-binding proteins coupled to somatostatin receptors. *J Biol Chem* 1991;266:17885–17897.

53. Lefkowitz RJ. G protein-coupled receptors. III. New roles for receptor kinases and beta-arrestins in receptor signaling and desensitization. *J Biol Chem* 1998;273:18677–18680.

54. de Vos AM, Ultsch M, Kossiakoff AA. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* 1992;255:306–312.

55. Hebert TE, Moffett S, Morello JP, et al. A peptide derived from a beta2-adrenergic receptor transmembrane domain inhibits both receptor dimerization and activation. *J Biol Chem* 1996;271:16384–16392.

56. Maggio R, Vogel Z, Wess J. Coexpression studies with mutant muscarinic/adrenergic receptors provide evidence for intermolecular “cross-talk” between G-protein-linked receptors. *Proc Natl Acad Sci USA* 1993;90:3103–3107.

57. White JH, Wise A, Main MJ, et al. Heterodimerization is required for the formation of a functional GABA(B) receptor. *Nature* 1998;396:679–682.

58. Jones KA, Borowsky B, Tamm JA, et al. GABA(B) receptors function as a heteromeric assembly of the subunits GABA(B)R1 and GABA(B)R2. *Nature* 1998;396:674–679.

59. Jordan BA, Devi LA. G-protein-coupled receptor heterodimerization modulates receptor function. *Nature* 1999;399:697–700.

60. Liu F, Wan Q, Pristupa ZB, et al. Direct protein-protein coupling enables cross-talk between dopamine D5 and gamma-aminobutyric acid A receptors. *Nature* 2000;403:274–280.

61. Hall RA, Premont RT, Lefkowitz RJ. Heptahelical receptor signaling: beyond the G protein paradigm. *J Cell Biol* 1999;145:927–932.

62. Fields S, Song O. A novel genetic system to detect protein-protein interactions. *Nature* 1989;340:245–246.

63. Hall RA, Premont RT, Chow CW, et al. The beta2-adrenergic receptor interacts with the Na+/H+-exchanger regulatory factor to control Na+/H+ exchange. *Nature* 1998;392:626–630.

64. Smith FD, Oxford GS, Milgram SL. Association of the D2 dopamine receptor third cytoplasmic loop with spinophilin, a protein phosphatase-1-interacting protein. *J Biol Chem* 1999;274:19894–19900.

65. Li M, Bermak JC, Wang ZW, et al. Modulation of dopamine D-2 receptor signaling by actin-binding protein (ABP-280). *Mol Pharmacol* 2000;57:446–452.

66. Hunter T. The Croonian Lecture 1997. The phosphorylation of proteins on tyrosine: its role in cell growth and disease. *Philos Trans R Soc Lond [B]* 1998;353:583–605.

67. Cohen S. Epidermal growth factor. *In Vitro Cell Dev Biol* 1987;23:239–246.

68. Carpenter G, Cohen S. Epidermal growth factor. *J Biol Chem* 1990;265:7709–7712.

69. Daub H, Weiss FU, Wallasch C, et al. Role of transactivation of the EGF receptor in signalling by G-protein-coupled receptors. *Nature* 1996;379:557–560.

70. Prenzel N, Zwick E, Daub H, et al. EGF receptor transactivation by G-protein-coupled receptors requires metalloproteinase cleavage of proHB-EGF. *Nature* 1999;402:884–888.

71. Luttrell LM, Ferguson SS, Daaka Y, et al. Beta-arrestin-dependent formation of beta2 adrenergic receptor-Src protein kinase complexes [see comments]. *Science* 1999;283:655–661.

72. Maudsley S, Pierce KL, Zamah AM, et al. The beta(2)-adrenergic receptor mediates extracellular signal-regulated kinase activation via assembly of a multi-receptor complex with the epidermal growth factor receptor. *J Biol Chem* 2000;275:9572–9580.

73. Prasher DC, Eckenrode VK, Ward WW, et al. Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* 1992;111:229–233.

74. Cody CW, Prasher DC, Westler WM, et al. Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. *Biochemistry* 1993;32:1212–1218.

75. Barak LS, Ferguson SS, Zhang J, et al. A beta-arrestin/green fluorescent protein biosensor for detecting G protein-coupled receptor activation. *J Biol Chem* 1997;272:27497–27500.

76. Kallal L, Gagnon AW, Penn RB, et al. Visualization of agonist-induced sequestration and down-regulation of a green fluorescent protein-tagged beta2-adrenergic receptor. *J Biol Chem* 1998;273:322–328.

77. Mendelsohn AR, Brent R. Protein interaction methods—toward an endgame. *Science* 1999;284:1948–1950.

23

Applying Functional Genomics to Neuropsychopharmacology

Michael Brownstein

M. Brownstein: National Institutes of Health, National Institute of Mental Health, Bethesda, Maryland.

> “The time has come,” the Walrus said, “to talk of many things.”
>
> --Lewis Carroll

The time has come indeed. The sequencing of the human genome and the genomes of a number of other species subject to research (1 ,2 ,3 ,4 ,5 and 6) have paved the way for new sorts of studies. Soon researchers will be able to look at the response of every human gene to specific manipulations or developmental events at multiple time points. This will require a new mindset. Researchers will not necessarily be testing specific hypotheses as they have done in the past. Instead, they will rely on the emergence of patterns and systematic features in their data sets (and those of others) to describe the phenomena being examined. Such patterns may hint at functions of collections of genes, the interactions of their products, and their importance in physiologic and pathologic processes. This chapter introduces array technology, discusses the sorts of experiments that can now be done with it, and suggests future advances. Several reviews have already been published on this subject, and the reader should refer to them for additional information (7 ,8 ,9 ,10 ,11 ,12 ,13 ,14 and 15). In addition, university, government, and commercial Web sites are valuable sources of news, background material, reagents, arrays, software, and instrumentation (16 ,17 ,18 ,19 ,20 ,21 ,22 ,23 ,24 ,25 ,26 ,27 ,28 ,29 ,30 ,31 ,32 and 33).

EARLY STUDIES OF GENE EXPRESSION

Part of "23 - Applying Functional Genomics to Neuropsychopharmacology "

The human genome is composed of approximately 3 billion DNA nucleotides encoding more than 100,000 genes (16). Each of these genes must be turned on or off in the right cells at the right time for an individual to develop and prosper. The genes that are ultimately expressed in a particular tissue define it. That is, brain is brain and liver is liver because of the particular collections of transcripts found in their respective cells. Brain, however, is extraordinarily heterogeneous. It has been estimated that nearly half of the genes in the genome are expressed there, distributed among the different neuronal and glial populations.

Genes are made of DNA, a nucleic acid polymer that has deoxyribose as its sugar backbone. Each sugar moiety in the chain has a base (adenine, A; cytosine, C; guanine, G; or thymine, T) attached to it. DNA exists as a double-stranded helix. The two antiparallel strands are bound to one another because their sequences are complementary—that is, the opposing bases are held together by hydrogen bonds, A to T and C to G. Similarly, messenger RNA (mRNA), the transcription product of the coding region of each gene, is complementary to the DNA strand from which it was copied and can bind to it. Northern blotting, the first method developed for detecting single mRNA species in a cellular extract, is based on this phenomenon. In this technique, RNA samples are fractionated by agarose gel electrophoresis, and the RNA bands are transferred (blotted) onto nitrocellulose membranes. Single RNA species can then be detected by hybridizing a radiolabeled DNA to the blot that is complementary to the RNA of interest.

In the past, the responses of cells or organisms to environmental cues were studied on a small scale, one gene or pathway at a time. Initially, Northern blotting was used to examine the abundance of specific mRNA species. Subsequently, other methods were chosen because they were simpler and more sensitive, such as reverse-transcriptase polymerase chain reaction (RT-PCR) (34), or because they were more comprehensive, such as SAGE (serial analysis of gene expression) (35 ,36). These techniques provide useful information, but they are tedious, time-consuming, and expensive to employ.

In the last 5 years, spurred by the availability of large volumes of genomic and cDNA (EST [expressed sequence tag]) sequence data from a variety of organisms, investigators have developed methods to study mRNA profiles in cells and tissues by means of large-scale, high-throughput, parallel methods. In the future, it would be helpful to look at protein and small molecule profiles as well, but the reagents required (panels of antibodies, for example) are difficult to

assemble and utilize. Even though there is not a one-to-one correspondence between the level of a particular transcript in a cell and that of its translation products, a great deal can be learned by performing mRNA expression profiling.

GENE EXPRESSION ARRAYS

Part of "23 - Applying Functional Genomics to Neuropsychopharmacology "

Expression profiling relies on large ordered collections of cDNAs immobilized on glass (microarrays) or synthetic oligonucleotides immobilized on silica wafers or chips (probe arrays). Both of these methods are the conceptual descendants of target nucleic acids immobilized on filters or membranes and detected with complementary radioactive probes. While filter-based systems are commercially available, reasonably priced, and fairly easy to use, it is clear that they will be preempted by glass or chip arrays developed with fluorescent probes. Glass arrays printed on microscope slides are now much cheaper to employ than chips, and many universities and research institutes have already built facilities for printing and probing such arrays. So it is worth discussing the uses to which such arrays have already been put, and the uses to which neuropsychopharmacologists could put them.

Few investigators have used arrays to study brain so far, preferring instead to look at mammalian cell lines and tumors (37 ,38 ,39 ,40 ,41 ,42 ,43 and 44). In addition, many workers have focused on yeast (45 ,46) or prokaryotes (47 ,48) because their genomes are small and have been completely sequenced. Consequently, every protein-encoding gene can be arrayed and examined. This will be true of human and mouse arrays in the not-too-distant future. Meanwhile, experiments can be done with the arrays that are available. These have between a few thousand and a few tens of thousands of elements, and with them we can begin to catalogue genes expressed in regions of the developing and adult nervous system, and to look for alterations in expression patterns associated with pathologic states or physiologic/pharmacologic manipulations (49). Consider, for example, the work that could be done to understand the mechanism(s) of action of selective serotonin reuptake inhibitors (SSRIs) and the reason for their delayed onset of action in depressed patients. As is known, SSRIs increase the availability of serotonin (5-hydroxytryptamine, 5-HT) to presynaptic and postsynaptic receptors, of which there are at least 14 subtypes (50). Among these, 5-HT_{1A} receptors on serotoninergic raphe neurons are thought to play a key role in regulation 5-HT release (51). 5-HT_{1A} agonists, which are used to treat anxiety, inhibit serotonin secretion. Conversely, desensitization of 5-HT_{1A} receptors, which could result from elevated 5-HT levels in the synaptic space following SSRI administration, may have the opposite effect—an increase in 5-HT release by raphe neurons, and chronic stimulation of 5-HT receptors in regions such as the hippocampus, amygdala, and septum. Despite all the research that has been done to date, the identity of the structures and biochemical alterations that are responsible for the antidepressant actions of SSRIs is still moot.

Array experiments will allow investigators to explore the serotoninergic system in a way that is model independent and comprehensive, and the experiments should become easy and cheap enough to perform to permit varying many parameters and comparing many conditions.

Initially, regional responses to a single dose of SSRI at a variety of times in one mouse strain might be examined. Subsequently, mouse strains that differ in their behavioral reactions to SSRIs could be examined; knockout mice known to have altered responses to SSRIs (e.g., 5-HT_{1A} receptor knockouts) could be studied; and drugs that resemble, facilitate, or inhibit the behavioral effects of SSRIs could be investigated. Mice would be better to use for this work than rats as of now because very big mouse arrays are available as are genetically manipulated animals and a variety of well-characterized inbred strains. Unfortunately, mice have small brains, and obtaining samples of minute regions (e.g., raphe nuclei) large enough to make sufficient RNA for labeling is difficult. Help is on the way, though. Better labeling methods, dyes, and detection devices are being developed. In fact, the amount of total RNA needed for an array experiment has already fallen well below 1 µg, and should approach 1 ng shortly.

Each array experiment will let an investigator look simultaneously at thousands of transcripts including those encoding enzymes involved in energy metabolism, receptors, G proteins, second messengers, and ion channels, to name a few. In addition, there will be many species represented on big arrays, the actions of which are unknown. The major task will be to assign them functions (see below).

THE DEVIL IS IN THE DETAILS

Part of "23 - Applying Functional Genomics to Neuropsychopharmacology "

Methods for making and probing arrays and analyzing array data have developed quickly. In spite of this, the supply of arrays has not kept up with demand, and demand should increase dramatically if the goal of using arrays is to compare many conditions and then mine the data systematically for patterns of gene expression. Thus, as stated earlier, costly products are unlikely to gain wide acceptance, and glass slide arrays are likely to be most commonly employed. For this reason, I now discuss their production and use.

Large collections of cDNAs and their sequences are now in the public domain. Some sets of cDNAs have been sequence verified and are ideal to use for preparing arrays; others have not been validated and are less useful. To make arrays, plasmid DNA is prepared from gridded sets of clones to be printed, and (typically) the 3′ end of each cDNA is amplified by PCR. The purified PCR products are then spotted using a robotic arrayer. It is possible to fabricate one's own arrayer (18), but many investigators will prefer

to buy an instrument or obtain arrays from core facilities or commercial vendors. Many thousands of 100-μM spots can be printed on a single glass microscope slide (see ref. 10 for details).

It is important to realize that array experiments do not permit measurement of the amount of each RNA that is present in a sample. This is because the relationship between the amount of transcript in a mixture and the intensity of the fluorescent spot it produces is a complex one—influenced by labeling efficiency, hybridization and wash conditions, and the sequence, quality, and quantity of the printed DNA. Thus, microarrays are typically employed to measure the relative abundances of RNA species in two or more extracts. To achieve this goal, in a two-sample experiment the RNAs are separately labeled with dyes of different colors, and then the products are mixed and hybridized to the arrayed spots (Fig. 23.1). After washing, the slides are scanned with a "reader" or "scanner" and the intensities of the fluorescent signals produced by the two separate dyes are determined spot by spot. Following background subtraction and "normalization" of the signals from the two channels (see below), a ratio of intensities of the two colors is determined for each spot, and the relative abundance of the two input RNAs can then be estimated. Finally, "clustering" methods are used to sort and display the data.

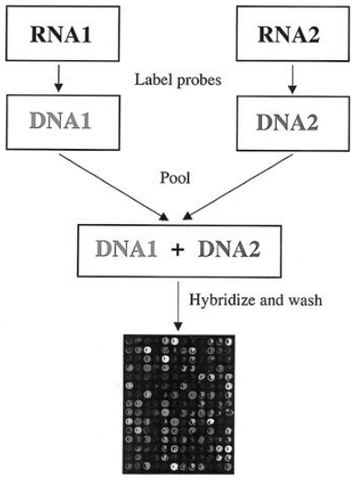


FIGURE 23.1. To perform microarray experiments, RNA is purified from two or more samples of cultured cells or dissected tissues. These RNAs are used to produce labeled probes. In the example given, the dye cy5, which fluoresces red, was used to label probe from sample 1; and the dye cy3, which fluoresces green, was used to label probe from sample 2. The labeled products are mixed and hybridized to the spots on the microarray. Following a wash step, the array is scanned and the signals from the red and green channels are superimposed. If an RNA species is more abundant in sample 1 than 2, the resulting spot will be red; in the reverse case, the spot will be green. When the RNA is equally abundant in the two samples, the spot is yellow. See color version of figure.

Two sorts of experimental paradigms have been defined: type I and type II (42). In the former, two samples are compared to one another; in the latter, multiple samples are compared. To look at multiple samples (e.g., time points, drug doses, developmental ages, brain regions, autopsy specimens), each sample in the set could theoretically be labeled with a different fluorescent dye that could, in turn, be visualized with a different laser. Presently, most commercial readers have only two lasers, but four-color instruments have already appeared on the market. (The number of dyes that it is possible to use for labeling RNA samples is dictated by a reader's ability to resolve the signal from individual dyes and the strength of the signal each dye produces.)

To use a two-color scanner for multiple comparisons, discrete samples must be compared to a reference standard. The ideal standard would have modest amounts of each transcript represented on the array used, because it needs to generate a nonzero denominator for the hybridization ratio. In the case of the mouse, the 17-day embryo and/or the adult brain have been proposed as sources of standard RNA. Pools of cell-line RNAs have been used as standards for human work. It would be useful if a central source of standards existed and if huge batches were prepared.

NORMALIZING RATIOS

Part of "23 - Applying Functional Genomics to Neuropsychopharmacology "

Since it is difficult, if not impossible, to measure the amount of RNA used to produce a labeled probe, normalizing the signals from the source RNAs is essential. To do this, a set of "housekeeping genes" is chosen because their transcription is fairly constant across a range of conditions. The ratio of signals from these genes is set to 1. The housekeeping set needs to be defined empirically, and in looking for candidates to include in such a set, few genes have been found that have constant expression levels. When large arrays are used, this is not a problem; hundreds of genes (or the entire set of genes) can be used for "global normalization" (Fig. 23.2). When small arrays are employed, on the other hand, the size and composition of the gene set used for normalization are very important. Just as a reference standard is urgently needed now, a normalization set supplied by a central site would be quite valuable.

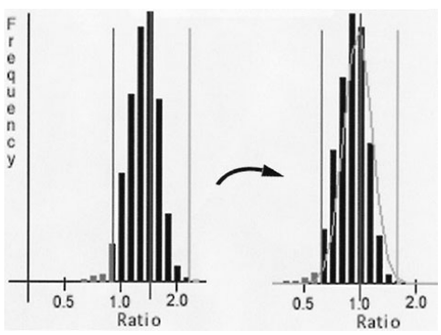


FIGURE 23.2. Normalization. Most genes, especially those with "housekeeping" functions, do not change very much from one experimental condition to another. For this reason, housekeeping genes, or the entire collection of genes on a large array, can be used to set the ratio of signals from the two color channels to 1 as shown here. See color version of figure.

QUALITY CONTROL

Part of "23 - Applying Functional Genomics to Neuropsychopharmacology "

While we have methods to assess the quality of DNA sequence data, for example, there is no generally accepted method for establishing the quality of an array study. In spite of this, there are some controls that can be built into an array. As noted earlier, scientists are arraying DNAs generated by PCR from plasmid templates. It is highly desirable to use sequence-verified cDNA sets. Amplifying these with

specific primers would confirm the identity of each clone, but this is expensive. Assuming we can obtain a reasonably well-validated clone set, the cheapest way to produce the 1.5- to 2.0-kilobase (kb) DNAs for printing is to use vector primers, and to analyze the products on agarose gels. The resulting cDNAs will include T-tails of varying lengths, and repeat sequences. Hybridization of labeled probe to these sequences is prevented by addition of a blocking solution containing oligo (dA) 20-mers, yeast transfer RNA (tRNA), and (for human probes) Cot-1 DNA to the probe solution. To show that the blocking was successful, a number of negative controls should be included among the samples arrayed: spotting buffer, Cot-1 and human genomic DNA, plasmid DNA, oligo (dT), and oligo (dA).

It is useful to array targets for nonmammalian transcripts and to “spike” the samples with the corresponding polyA-tailed RNAs. These RNAs can be added in different concentrations to crude tissue extracts, total RNA samples, or purified polyA-plus RNA to determine extraction efficiencies and detection sensitivities. It would be a mistake to imagine that the added RNA standards can be used to generate figures for absolute amounts of RNA in samples for the reasons given earlier. They should be used exclusively for quality control.

An additional set of spots that have been found useful to array are “landing lights.” These DNAs, which are printed at regular intervals, are used for orientation.

SMALL SAMPLES, FALSE NEGATIVES, AND FALSE POSITIVES

Part of "23 - Applying Functional Genomics to Neuropsychopharmacology "

Failure to detect transcripts or changes in transcripts could result from low-quality arrays or poor labeling methods. Over time, the methods used to make and probe arrays should improve, and false negatives will grow less important. Presently, we can detect RNAs with an abundance of about 1:300,000 in a complex sample. This translates into a few copies per cell if one is studying a homogeneous cell line. Seeing increases in rare transcripts under these circumstances should be simple, but measuring decreases will difficult if not impossible when one can barely detect a weak signal in the first place. Since brain samples are much more heterogeneous than cell lines, the problem of detecting rare mRNAs is even harder. For this reason, it may be necessary to isolate neuronal populations from brain sections by microdissection or to collect single neurons by laser capture methods to enrich and study rare, cell-specific transcripts. To take full advantage of these dissection techniques, methods will have to be developed for isolating and labeling picogram quantities of RNA. (One million cells yield about 5 to 10 µg of total RNA.) It is important to note that labeling methods have to preserve the heterogeneity and relative abundances of the RNAs in the samples to be studied. Care must be taken if PCR is used in the labeling procedure to avoid biasing the sample. Novel labeling methods can be tested using arrays and serially diluted RNA templates.

At present investigators use Northern blotting, the TaqMan system, or *in situ* hybridization histochemistry to weed out false-positive responses of selected mRNAs. When arrays are no longer limiting, this could be accomplished by studying replicate samples, but the argument could be made that one should focus on variations in collections of genes instead of single ones, and that looking at many conditions once may be more powerful than looking at the same condition many times.

BIOINFORMATICS

Part of "23 - Applying Functional Genomics to Neuropsychopharmacology "

Analyzing the earliest, small-scale, array experiments was simply a matter of listing the names of transcripts that appeared to increase or decrease from control levels (Fig. 23.3). As the sizes of arrays increased and labeling methods improved, new algorithms had to be developed that “clustered” the hundreds or even thousands of expression changes found in a typical experiment. Clustering methods permit the classification of genes on the basis of similarities or differences in their patterns of expression across multiple experiments (52 ,53). The output is usually in tabular form. Experimental conditions are listed across the top of a table, and names of genes listed along the side. The response of each gene in each experimental condition is color coded—one color (red) indicating an increase and another (green) a decrease vs. a standard signal (Fig. 23.4). The eye can readily detect patterns in complex images of the sort described, and groups of genes can be identified that parallel one another. Commonly, such genes function in concert.

In the hypothetical SSRI study described above, one group of genes may be increased or decreased in the raphe nuclei following chronic, but not acute, treatment with drug, and a different collection of genes is altered in areas innervated by raphe neurons such as the hippocampus. Different sets of genes might respond to SSRI treatment when behaviorally responsive mouse strains are compared to unresponsive ones, and the pattern of gene expression is different in knockout animals that do not respond to SSRIs as compared with animals that do. Each additional experiment may narrow (or broaden) the list of genes of interest.

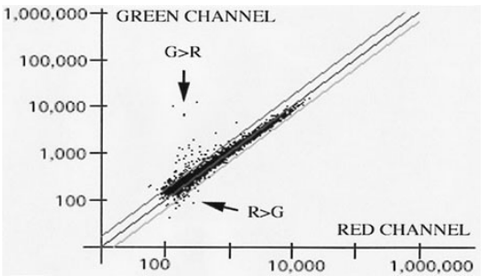


FIGURE 23.3. After normalization, a scatter plot shows that most genes fall in a ratio = 1 space. *Arrows* above and below this space point at genes with altered expression. See color version of figure.

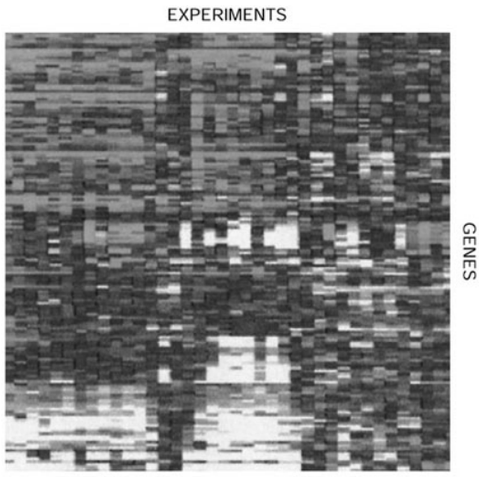


FIGURE 23.4. Clustering. The clustering algorithm has sorted the genes that were studied in a series of experiments according to similarities in their patterns of expression. By convention, red indicates an increase from the standard used, and green a decrease. The collections of genes that are moving up or down in parallel can be readily seen. See color version of figure.

There are ways of evaluating data that are not completely model-independent. Gene responses can be imposed on metabolic charts or on maps of chromosomes. In the former case, increases or decreases in the utilization of certain pathways can be detected; in the latter, deletions or changes in copy number may be recognized, or strong positional candidates identified.

INTERPRETING EXPERIMENTS

Part of "23 - Applying Functional Genomics to Neuropsychopharmacology "

Recent studies of changes in gene expression in yeast associated with nutritional and environmental stresses, the cell cycle, or genetic manipulations are examples of well planned and executed surveys (45 ,46 ,54 ,55 and 56). All of the 6,200 known and predicted protein-coding genes in the yeast genome were arrayed on a single microscope slide, and sufficient numbers of cells were grown to make ample amounts of RNA for labeling. Each experiment gave a richly detailed picture of molecular responses to a physiologic process or perturbation.

Unfortunately, the brain is much more difficult to examine than yeast. It varies with age, and is composed of hundreds of different sorts of cells that express, in aggregate, as many as half of the genes in the genome in a highly regulated manner. To determine the properties of single populations of cells in the context of the intact structure will be difficult, but perhaps not impossible. Initially, it would make sense to identify all the genes expressed in the developing and adult nervous system. This, in fact, is one goal of the Brain Molecular Anatomy Project (BMAP) (57). After this goal is achieved, the regional and cellular localization of “brain genes” will be determined.

In addition to cataloging the transcripts in the brain, it would be helpful to look at the reactions of isolated populations of neurons or glia to specific signals or environmental alterations—e.g., oxidative stress, excitotoxins, neurotransmitters, hormones, and drugs. Some responses may be of a global nature—increases or decreases in energy metabolism or protein biosynthesis—while others may be quite specific to the cell studied or the agent administered.

The availability of transcript maps and collections of “expression motifs” should help us interpret some of the changes observed in human or mouse brain samples. For example, if it were possible to examine the responses of isolated raphe neurons to SSRI treatment *in vitro*, it might be easier to recognize similar responses in tissue samples. Bear in mind, however, that the majority of arrayed genes discovered by sequencing the human genome and large collections of cDNAs are of unknown function. Structural motifs may hint at the function of some gene products (58), but the role of most will remain an enigma. Expression studies may help solve this problem because genes with similar expression profiles often have related jobs. Functional proteomic work will be useful as well. The combination of

two-dimensional gel electrophoresis, ultrasensitive detection methods, and mass spectroscopic analyses will permit researchers to map protein species to specific organelles or macromolecular complexes (59). It is important to remember that most proteins in a cell do not exist or act in isolation. Components of metabolic pathways and regulatory cascades reside in protein communities. Interactions between members of such communities can be detected with yeast two-hybrid methods, and researchers have already begun to examine protein–protein interactions in simple organisms on a genome-wide basis (60 ,61 and 62). Furthermore, yeast "*n*-hybrid" methods have been developed that permit one to look at protein–DNA and protein–RNA interactions (63). Finally, it is worth mentioning that large collections of mice are being produced with random mutations in their genomes. The goal of "saturation mutagenesis" projects is to use multiple screens to identify animals with interesting phenotypes. The goal of investigators who are making insertional mutations in embryonic stem (ES) cell lines, on the other hand, is to determine the insertion site of each cell produced so that knockout animals can be made on demand.

In the future, in analyzing the results of array experiments, the field will benefit from work on animals, proteomics, and earlier expression studies too—but only if everyone adheres to standard formats in archiving and annotating data.

ANNOTATING EXPERIMENTS

Part of "23 - Applying Functional Genomics to Neuropsychopharmacology "

Presently, there are no standards for annotation, but efforts are under way to solve this problem. To create useful and searchable archives, all features of each experiment will have to be described in a standard way using an explicit and unambiguous, "controlled" vocabulary. Some of the required vocabulary already exists. For example, DNAs spotted onto an array can be given and linked to identifiers in public databases. (Unfortunately, different databases sometimes use different identifiers for the same gene. Consequently, the sequence of each DNA on an array should be specified.) Drugs used in array experiments can be referred to by Merck Index number (64), organisms can be described using names in the taxonomy database (65), and mouse strains and mutants can be named according to established rules and guidelines (66). Much of the language needed to describe array experiments has not been standardized, however, and for now databases will have to contain many free-form text fields.

For studies of autopsy samples from psychiatric patients, a good deal of specialized information should be provided. The patient's age at death, gender, diagnosis, genotype (if available), cause of death, postmortem interval, pathology, toxicology screening results, and medication/drug abuse history should be given. The brain region dissected, dissection method used, side of the brain sampled, specimen weight, microdissection procedure and cells selected, RNA extraction method and quality, and labeling method are all essential fields as well. It would be useful to have quality scores for some of these—e.g., the diagnosis and sample condition. Short of a quality score for diagnosis, a detailed description of the key elements of the patient's clinical and laboratory findings should be made available.

Annotating array experiments can be tedious and time-consuming. I am not suggesting that the recommendations above should all be implemented immediately. This would hinder progress. On the other hand, the field would benefit from well-annotated work, and the sooner guidelines are agreed to and implemented, the better. If this chapter serves one useful purpose, it would be to promote this agenda.

REFERENCES

1. http://www.ncbi.nlm.nih.gov/genome/guide/.
2. http://www.ncbi.nlm.nih.gov/UniGene/Hs.stats.shtml.
3. http://www.ncbi.nlm.nih.gov/dbEST/index.html.
4. http://www.celera.com/celerascience/index.cfm.
5. http://www.tigr.org/.
6. http://genome.rtc.riken.go.jp/.
7. Brown PO, Botstein D. Exploring the new world of the genome with DNA microarrays. *Nature Genet* 1999;21(1 suppl):33–37.
8. Brownstein MJ, Trent JM, Boguski MS. Functional genomics. *Trends Guide Bioinformatics* 1998;27–29.
9. Duggan DJ, Bittner M, Chen Y, et al. Expression profiling using cDNA microarrays. *Nature Genet* 1999;21(1 suppl):10–14.
10. Eisen MB, Brown PO. DNA arrays for analysis of gene expression. *Methods Enzymol* 1999;303:179–205.
11. Ferea TL, Brown PO. Observing the living genome. *Curr Opin Genet Dev* 1999;9(6):715–722.
12. Kao CM. Functional genomic technologies: creating new paradigms for fundamental and applied biology. *Biotechnol Prog* 1999;15(3):304–311.
13. Schena M, Shalon D, Davis RW, et al. Quantitative monitoring of gene expression patterns with a complementary DNA microarray [see comments]. *Science* 1995;270(5235):467–470.
14. Schena M, Heller RA, Theriault TP, et al. Microarrays: biotechnology's discovery platform for functional genomics [see comments]. *Trends Biotechnol* 1998;16(7):301–306.
15. Shalon D, Smith SJ, Brown PO. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res* 1996;6(7):639–645.
16. http://www.incyte.com/.
17. http://www.nhgri.nih.gov/DIR/LCG/15K/HTML/.
18. http://cmgm.stanford.edu/pbrown/.
19. http://w95vcl.neuro.chop.edu/vcheung/.
20. http://sequence.aecom.yu.edu/bioinf/funcgenomic.html.
21. http://chroma.mbt.washington.edu/mod_www/.
22. http://www.affymetrix.com/.
23. http://www.biodiscovery.com/.
24. http://tagc.univ-mrs.fr/.
25. http://www.genscan.com/.
26. http://www.genomicsolutions.com/products/bio/index.html.
27. http://www.intelligentbio.com/.
28. http://www.interactiva.de/vlab/xna.html.
29. http://www.apbiotech.com/application/microarray/default.htm.
30. http://www.packardinst.com/.
31. http://www.ultranet.com/˜radius/.

32. http://www.resgen.com/.

33. http://www.stratagene.com/gc/gene.htm.

34. Berchtold MW. A simple method for direct cloning and sequencing cDNA by the use of a single specific oligonucleotide and oligo(dT) in a polymerase chain reaction (PCR). *Nucleic Acids Res* 1989;17(1):453.

35. http://www.ncbi.nlm.nih.gov/sage/.

36. Velculescu VE, Zhang L, Vogelstein B, et al. Serial analysis of gene expression [see comments]. *Science* 1995;270(5235):484–487.

37. Afshari CA, Nuwaysir EF, Barrett JC. Application of complementary DNA microarray technology to carcinogen identification, toxicology, and drug safety evaluation. *Cancer Res* 1999;59(19):4759–4760.

38. MacLachlan TK, Somasundaram K, Sgagias M, et al. BRCA1 effects on the cell cycle and the DNA damage response are linked to altered gene expression. *J Biol Chem* 2000;275(4):2777–2785.

39. Duan Z, Feller AJ, Penson RT, et al. Discovery of differentially expressed genes associated with paclitaxel resistance using cDNA array technology: analysis of interleukin (IL) 6, IL-8, and monocyte chemotactic protein 1 in the paclitaxel-resistant phenotype. *Clin Cancer Res* 1999;5(11):3445–3453.

40. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling [see comments]. *Nature* 2000;403(6769):503–511.

41. Perou CM, Jeffrey SS, van de Rijn M, et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci USA* 1999;96(16):9212–9217.

42. Geiss GK, Bumgarner RE, An MC, et al. Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. *Virology* 2000;266(1):8–16.

43. Ross DT, Scherf U, Eisen MB, et al. Systematic variation in gene expression patterns in human cancer cell lines [see comments]. *Nature Genet* 2000;24(3):227–235.

44. DeRisi J, Penland L, Brown PO, et al. Use of a cDNA microarray to analyse gene expression patterns in human cancer [see comments]. *Nature Genet* 1996;14(4):457–460.

45. Nacht M, Ferguson AT, Zhang W, et al. Combining serial analysis of gene expression and array technologies to identify genes differentially expressed in breast cancer. *Cancer Res* 1999;59(21):5464–5470.

46. Spellman PT, Sherlock G, Zhang MQ, et al. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol Biol Cell* 1998;9(12):3273–3297.

47. Tao H, Bausch C, Richmond C, et al. Functional genomics: expression analysis of *Escherichia coli* growing on minimal and rich media. *J Bacteriol* 1999;181(20):6425–6440.

48. Hayward RE, Derisi JL, Alfadhli S, et al. Shotgun DNA microarrays and stage-specific gene expression in *Plasmodium falciparum* malaria. *Mol Microbiol* 2000;35(1):6–14.

49. Tassone F, Lucas R, Slavov D, et al. Gene expression relevant to Down syndrome: problems and approaches [In Process Citation]. *J Neural Transm Suppl* 1999;57:179–195.

50. Julius D. Serotonin receptor knockouts: a moody subject [comment]. *Proc Natl Acad Sci USA* 1998;95(26):15153–15154.

51. De Vry J. 5-HT1A receptor agonists: recent developments and controversial issues. *Psychopharmacology (Berl)* 1995;121(1):1–26.

52. Eisen MB, Spellman PT, Brown PO, et al. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998;95(25):14863–14868.

53. Zhang MQ. Large-scale gene expression data analysis: a new challenge to computational biologists [published erratum appears in *Genome Res* 1999;9(11):1156]. *Genome Res* 1999;9(8):681–688.

54. DeRisi JL, Iyer VR, Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 1997;278(5338):680–686.

55. Ferea TL, Botstein D, Brown PO, et al. Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proc Natl Acad Sci USA* 1999;96(17):9721–9726.

56. Lashkari DA, DeRisi JL, McCusker JH, et al. Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc Natl Acad Sci USA* 1997;94(24):13057–13062.

57. http://grants.nih.gov/grants/guide/notice-files/not98-052.html.

58. Orengo CA, Todd AE, Thornton JM. From protein structure to function. *Curr Opin Struct Biol* 1999;9(3):374–382.

59. Blackstock WP, Weir MP. Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol* 1999;17(3):121–127.

60. Bartel PL, Roecklein JA, SenGupta D, et al. A protein linkage map of Escherichia coli bacteriophage T7. *Nature Genet* 1996;12(1):72–77.

61. Fromont-Racine M, Rain JC, Legrain P. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens [see comments]. *Nature Genet* 1997;16(3):277–282.

62. Ito T, Tashiro K, Muta S, et al. Toward a protein-protein interaction map of the budding yeast: a comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc Natl Acad Sci USA* 2000;97(3):1143–1147.

63. Vidal M, Legrain P. Yeast forward and reverse 'n'-hybrid systems. *Nucleic Acids Res* 1999;27(4):919–929.

64. http://library.dialog.com/bluesheets/html/bl0304.html#BI.

65. http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/.

66. http://www.informatics.jax.org/mgihome/nomen/table.shtml.

Section III
Emerging Imaging Technologies and Their Application to Psychiatric Research

Robert Desimone
Robert Desimone: National Institute of Mental Health, Intramural Research Program, National Institutes of Health, Bethesda, Maryland.

Emerging Imaging Technologies and Their Application to Psychiatric Research - Introduction

In a series of chapters on advances in neuroimaging techniques, it is ironic that images per se of the brain's structures or of neural activity have actually diminished in importance since publication of the American College of Neuropsychopharmacology's Fourth Generation of Progress. The focus of neuroimaging in cognitive neuroscience and psychiatry is widening beyond initial questions of where normal and pathologic functions are localized in the brain, to begin to include questions of how cognitive operations are carried out and why they sometimes fail. For these more mechanistic questions, images become simply measurements for testing hypotheses and are not an end in themselves.

Structural imaging of the brain with magnetic resonance imaging (MRI) is a good example of a field that is no longer restricted to simple localization of pathology in psychiatric disease. Indeed, there are few psychiatric cases that are characterized by clear pathology that is visible in MRI pictures. By contrast, the new analytic approaches to measuring the size of structures in MRI images described by Evans in this section allow one to track small changes in structures over time, which can potentially reveal abnormal patterns of development in either gray or white matter. Recently, for example, these techniques have been used to track the distribution of gray and white matter during development in child-onset schizophrenia, which is characterized by an abnormal time course of gray matter reductions in several different brain systems. These abnormal trajectories suggest possible concomitant abnormalities in synaptic pruning, which is known to take place throughout development. New diffusion-tensor imaging techniques in MRI, described by Makris et al. in this section, add the ability to track major fiber bundles in the white matter, which can give some insight into cortical connectivity. Makris et al. discuss the current limitations with this technique, including the difficulty in following fiber bundles to their termination.

Another evolutionary change in neuroimaging has been the continued shift from positron emission tomography (PET) to MRI-based techniques for indirect measurement of neural activity. However, as described by Fujita and Innis, PET and single photon emission computed tomography (SPECT) remain the only viable techniques for studying ligand binding in the brain, and the resolution of PET is continuing to increase as new detectors are developed. Fujita and Innis review the status of radiotracer development in PET and SPECT and describe new tracers for measuring postreceptor signal transduction and even gene expression. Another relatively new approach to mapping the distribution and concentration of specific molecules in the brain is magnetic resonance spectroscopy (MRS), described by Rothman et al. in this section. The focus of this chapter is on measurements of metabolites involved in neuroenergetics and amino acid neurotransmission, especially the flux through glutamate/glutamine and ɣ-aminobutyric acid (GABA)/glutamine cycles during neural activity. GABA metabolism, in particular, appears to be sensitive to both psychiatric disorders, such as depression, and to pharmacologic treatment.

Some of the biggest advances in functional MRI (fMRI) methodology for the indirect measurement of neural activity have been in the time domain. Bandettini describes new methods for improving the temporal resolution of fMRI, in particular event-related designs. In more traditional blocked-trial designs, the BOLD signal is averaged for many seconds, typically for several trials of a behavioral task. However, with event-related designs, one can measure BOLD changes for events lasting less than 2 seconds, which allows one to distinguish activity changes in one part of a trial from another, e.g., from the encoding to the retrieval phase of a trial in a memory task.

The specific application of both fMRI and brain-lesion analyses to studies of mood and emotion is the focus of the chapter by Davidson. The brain systems important for the regulation and expression of mood and emotion are highly distributed, and thus it is essential to take a systems approach to imaging and lesion data in this field, rather than a “function per brain structure” approach. Davidson also describes how basic behavioral research lays the necessary groundwork for studying mood and anxiety disorders, and he gives specific examples of basic research into fear and anxiety and its implications for understanding disorders such as social phobia.

Despite improvements in the temporal resolution of fMRI, the technique will never approach the temporal resolution of event-related potential (ERP) methods, which is at the millisecond level. As described by Hillyard and Kutas, new analytic techniques have improved the spatial resolution of ERPs, and there has been considerable progress in combining the spatial resolution of fMRI with the temporal resolution of ERPs. Perhaps even more important than these technologic advances, there have been conceptual advances in understanding the functional components of ERP signals through the application of cognitive theory to ERP paradigms. For example, there are characteristic signals found in visual tasks for the arrival of visual information in a cortical region, for the modulation of this signal by attention, and for the decoding of the visual information into semantic information. With the appropriate task design and with the large base of information acquired on the timing of these cognitive operations in normal subjects, one can then begin to ask how these operations differ in schizophrenia, for example.

Finally, the techniques described in the chapters in this section allow one to create and test models of the functional interactions between different brain structures, but they provide no direct evidence of functional connectivity. One new, direct approach to functional connectivity is an analytic technique known as effective connectivity mapping, described by BÜchel and Friston. Here, one makes use of the moment-to-moment relationships between fMRI signals in different brain regions to create structural equations, which can quantify the contribution of activity in one brain structure to the activity in another. An even more direct approach to measuring connectivity is through the combined use of transcranial magnetic stimulation (TMS) and fMRI, described by George and Bohning. They describe studies in which brain structures are directly stimulated with TMS, and the resulting effects on far-removed brain structures, including in the opposite hemisphere, are measured using fMRI.

The chapters in this section describe an impressive armamentarium of techniques now available to brain imaging researchers, and they outline some promising new directions in the application of these techniques to mental illness. Together, the chapters show that the key to progess in brain imaging studies of pathophysiology will be to see beyond the images.

24

Automated 3D Analysis of Large Brain Mri Databases

Alan C. Evans

Alan C. Evans: Department of Neurology, McGill University, Montreal Neurological Institute, Montreal, Quebec, Canada.

In recent years, the study of gross neuroanatomy and its relationship to behavior and brain function has been reenergized by the advent of imaging techniques and the powerful computational tools with which to analyze high-resolution three-dimensional (3D) brain images (10 ,11 ,22 ,52 ,53). However, such high technology tools often demand that scientific questions be restated and made more amenable to quantitative analysis. Questions such as “How much normal variation is there in the size, shape, or location of an individual brain structure?” or “To what extent does functional architecture of the cortex depend on the anatomic boundaries between anatomic regions?” carry with them the assumption that the borders of individual structures can be specified accurately in any brain. In the past, basic questions of functional neuroanatomy were difficult to address in a systematic way in the living brain. We have learned much from anecdotal reporting of individual patients with various forms of brain lesion or from direct cortical stimulation during brain surgery, but the generalization of individual observation to the wider population has been confounded by the normal variation in brain structure itself. There is then a fundamental interest in understanding the nature of anatomic variability in the population, both for its relationship to functional variability and for the potential of using structural abnormality as a measure of development, normal aging, and disease. For instance, in some degenerative diseases like Huntington's disease and Alzheimer's disease, the sulci become more open and the ventricles become enlarged. Magnetic resonance imaging (MRI)-based measurements of these changes can lead to early diagnosis and treatment, but we need to understand the variation among normal brains first.

Although the study of postmortem neuroanatomy is a long-established science, the ability to accumulate the numbers of brains necessary to make statistically meaningful conclusions about cerebral anatomy is a relatively recent phenomenon. It is still difficult to identify reliably in any single brain the anatomic landmarks, boundaries, and other delimiting features necessary for any subsequent analysis. Thus, we face a new problem posed by this newfound technology and its inflexible demand that anatomic questions be posed in numerical rather than descriptive terms. The tools exist to image large numbers of brains noninvasively with MRI, but we are still struggling with how to extract the anatomic measurements necessary to answer the questions posed above. It is relatively easy to identify the precentral gyrus, but few researchers attempt to define its “top” and “bottom.” Where does the inferior frontal sulcus end? Traditional brain atlases identify brain regions only by pointing to the middle of the region or surface feature, leaving the interfaces between regions unspecified. Neuroanatomists debate the exact boundary of even relatively simple structures such as the thalamus or caudate nucleus. With this context, new initiatives at various laboratories are attempting to standardize and codify the partitioning of the human brain into named regions, not without controversy. Traditional neuroanatomists debate among themselves about what parcellation scheme and nomenclature to use. Computer scientists argue among themselves about whether to use hierarchical, relational, object-oriented, or some other form of database structure to organize the brain parcellation. Both groups tend to misunderstand the importance of the other's concerns. Neurobiologists or physicians are not used to thinking in terms of inclusive sets where, for instance, every structure at one level is wholly included within a higher level organization, where all 3D pixels, i.e., voxels, within the brain space *must* be labeled as one of the structures in the partitioning scheme, or that the cerebrospinal fluid (CSF) ventricular spaces may be declared as being outside of “brain.” Computer scientists tend to ignore the realities that many cortical sulcal features do not exist in every brain, and may be fragmented or have multiple occurrences. Some sophisticated analytic approaches for quantifying anatomic variability assume that a particular landmark can be perfectly identified in any brain when the reality is that errors of 5 to 10 mm typically occur, an error that is about

the same magnitude as the true spatial variation being sought.

Despite heroic efforts in the recent past (2 ,29 ,58 ,60 ,62 ,75 ,79 ,80), manual labeling of many individual MRI data sets in 3D is a labor-intensive effort that is not likely to be widely adopted. Fully automated techniques that produce accurate neuroanatomic segmentation in large numbers of MRI data sets are essential if questions of normal cross-sectional variability, normal longitudinal development, and detection of abnormality in single subjects or in groups are to be answered definitively. Many groups are now engaged in the field of MRI-based quantitative neuroanatomy, and an exhaustive review of the field is beyond the scope of this chapter. A representative sampling of activity by other groups in the field, categorized into the four forms of segmentation discussed in the subsequent Methods section, include the following:

- Tissue classification/voxel morphometry: This refers to MRI intensity-based classification of images into tissue classes and voxel-based statistical analysis of the resulting class maps. In normal brain, the tissue classes are typically gray matter, white matter, and CSF, although there is no reason in principle to restrict to these three tissue types. In these approaches, one or more co-registered MRI images of the same neuroanatomy, obtained using different acquisition protocols [e.g., T1-weighted, T2-weighted, proton density (PD), magnetization transfer), provide the input data. At each voxel the MRI intensity for each of the *N* input images provides an *N*-dimensional "feature vector." Ideally, each tissue class is identified by a unique feature vector. In practice, many confounding factors (e.g., tissue heterogeneity, MRI field distortions, partial volume effects, and image noise) blur the feature space and render it difficult to distinguish accurately even three tissue classes. Many different multivariate statistical methods exist to optimize the class labeling and, for most of them, more independent images (features) help to disentangle overlapping class distributions in feature space.

  Mapping the segmented images into stereotaxic space (69 ,70) allows for group analysis across a population of 3D data sets from different individuals. All of the machinery of random field statistical analysis developed for functional imaging then becomes available for structural analysis (1 ,5 ,30 ,31 ,35 ,54 ,56 ,57 ,81 ,82 and 83).

- Regional parcellation/atlas deformation: Delineation of brain regions within each tissue class (e.g., caudate nucleus in the gray matter class) is not possible using only the information available in the MR image(s) since there is not sufficient differentiation among these regions within the feature space. Some form of prior information on neuroanatomic boundaries is needed, usually in the form of a computerized brain atlas, to assist in 3D brain regional labeling. Regions can be identified by vector boundaries or by labeling of all internal voxels. The atlas or parcellation scheme can be used as a guide to manual segmentation or as the basis for automated regional segmentation in which the atlas space is deformed to match each new 3D brain image. The atlas template is matched to the new MRI volume through a variety of nonlinear deformation techniques, the most successful of which use image similarity criteria to deform one image into another.

  Once delineated in their native space it is possible to map the regional labels into stereotaxic space in much the same way as tissue class maps and to conduct voxel morphometry among groups using the random field statistical analysis (3 ,4 ,6 ,12 ,18 ,18 ,21 ,26 ,32 ,34 ,36 ,39 ,40 and 41 ,50 ,68).

- Surface extraction/cortical unfolding: Regional parcellation is generally quite successful at labeling relatively well-defined 3D brain regions, such as the thalamus, but is typically less successful in identifying cortical gyri. Indeed, the cortex as such is sufficiently important to merit special analytic treatment. Techniques have been developed to "extract" the exterior cortical surface automatically by boundary detection of the intensity interface between gray matter and subarachnoid CSF. To overcome partial effects, some groups have targeted the internal cortical margin at the interface between gray and white matter. Obtaining a measure of the two surfaces simultaneously allows for a measure of cortical thickness at each location over the cortical surface.

  Extraction of the cortical surface has prompted some groups to explore the potential of an "unfolded" cortical surface as a means of studying functional neuroanatomy on a two-dimensional (2D) plane. Arguably, this device reduces the variability of functional areas introduced by cortical folding in three dimensions. The mapping from 3D to 2D is a nontrivial task with many issues surrounding the optimal mapping function, with direct analogies to the well-known cartographic dilemmas of preservation of area, direction, distance etc. (7 ,8 ,20 ,27 ,28 ,38 ,49 ,55 ,77 ,78).

- Sulcal extraction/analysis: The cortical sulci have held a historical position of prominence in functional neuroanatomy, in part because of their utility as approximate landmarks to functional areas. Recent interest has centered upon extracting not just the surface trace of the sulcus as a line but rather the depth of the sulcus as a ribbon. The latter approach provides more information on buried cortex and sulcal shape than a simple line trace, which can be related to genetic and developmental considerations (46 ,51 ,59 ,65 ,76).

In the United States, the Human Brain Project has specifically set out to foster the application of computational techniques, hardware, and algorithms to neuroscience at all spatial scales. We are involved in one of these applications operating at the gross morphology level. The International Consortium for Brain Mapping (ICBM) (52), seeks to create a so-called probabilistic human brain atlas (see below).

This chapter provides an overview of the methods developed by the Brain Imaging Centre (BIC) at the Montreal Neurological Institute for fully automated 3D segmentation of the ICBM database and other MRI databases like it, such as those collected for the creation of normal pediatric development and for evaluation of new pharmaceuticals. A key concept underlying this work is that of the analysis "pipeline," which takes 3D MRI volumes from large numbers of subjects and generates 3D statistical maps of adult brain morphology with no manual intervention. The pipeline concept has also been implemented for clinical trial analysis of MRI data from multiple sites. All data sets, across patients, time points, and pulse sequences, are mapped into a standardized 3D coordinate space for automatic segmentation and statistical analysis.

Once the MRI image has been segmented, each voxel in the 3D image space carries an anatomic label and a measure of the confidence in that label. This information can be used in a variety of ways to detect subtle neuroanatomic or neuropathologic changes:

- Single subject vs. group data for detection subtle of structural abnormality (e.g., misshapen corpus callosum)
- Intergroup cross-sectional comparison (e.g., Alzheimer's disease group vs. normal age-matched controls)
- Longitudinal study in a single subject (e.g., tumor growth, progressive atrophy)
- Longitudinal study in a group [e.g., early development and aging in normal populations, multiple sclerosis (MS) disease progression].

Illustrative example applications of some of these capabilities are described at the end of the chapter.

IMAGE SEGMENTATION METHODS

Part of "24 - Automated 3D Analysis of Large Brain Mri Databases "

Within the BIC image analysis pipeline, MRI data are processed using a series of tools that provide measurements of volume, shape, size, and tissue composition of selected brain regions. These are summarized below. To manage the flow of MRI data through the pipeline, we have developed PCS (Production Control System), which allows the rapid implementation and parallel execution of analysis pipelines for processing large MRI databases. Each processing stage in the pipeline is performed by a single command. PCS allows the user to specify this command with its options, input and output files, and dependencies on other stages in the pipeline using a simple script language. Efficient coarse-grain parallelism is achieved by distributing the individual jobs over a network of workstations. PCS monitors the status of each job and submits a new job when the prerequisites for submission have been satisfied (typically the completion of all stages on whose output data the stage depends). The major elements of this environment include (Fig. 24.1):

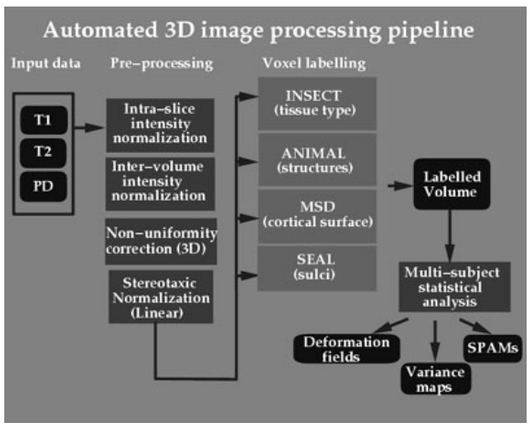


FIGURE 24.1. Brain Imaging Centre (BIC) pipeline environment for magnetic resonance imaging (MRI) processing: major components of pipeline analysis of large ensembles of MRI multispectral data sets. Each multispectral data set yields labeled maps of tissue type, three-dimensional (3D) brain region, and cortical topology.

- Thin-slice MRI data acquisition (typically 1-mm axial sampling, with 1-mm isotropic voxels).
- Multimodal, multidimensional stereotaxic data format (MINC).
- MRI simulator for validation of segmentation algorithms (MRISIM).
- Correction for coil-dependent 3D intensity nonuniformities (N3).
- Within-subject registration of different sequence volumes (MINCTRACC).
- Cross-subject mapping into a standardized "stereotaxic" 3D coordinate space (MRITOTAL).
- Fully automated 3D classification of gray/white/CSF tissue classes (INSECT).
- Fully automated 3D regional segmentation based on prior atlas templates (ANIMAL).
- Fully automated 3D extraction of gray/CSF and gray/white cortical interfaces (MSD, ASP).
- Computer-assisted 3D labeling of individual sulci (SEAL).

Stereotaxic Image Format—MINC

A fundamental aspect of this pipeline environment and its interaction with other sites within ICBM is the MINC image format for intersite data communication. MINC (Medical Image Net CDF), developed at the MNI by Peter Neelin, is a multidimensional, multimodality image file format that supports stereotaxic coordinate representation.

Image volumes can be explored in real time in 3D with continuous update of stereotaxic coordinates. Image files with different native voxel dimensions can be compared directly without regard for the original acquisition sampling grid. This simplifies stereotaxic analysis of MRI data ensembles collected with different voxel dimensions.

MRI Simulation—MRISIM

To assist in the evaluation of these segmentation tools, we created an average MRI data set of a single young normal male, by repeated MRI scanning followed by linear alignment of all volumes. A total of 27 separate MRI scans were collected. The improved signal-to-noise ratio (SNR) in the composite MRI, termed ICBM27, produces a high-definition data set (37), suitable for brain atlas construction, validation of segmentation/mapping algorithms, and MRI simulation. (Note: Since it incorporates the structural idiosyncracies of a single brain, it is not intended for use as a high-definition master data set for stereotaxic normalization.) This data set has been segmented manually to create an accurate digital phantom (17) for use as the source template of an MRI simulator, MRISIM (43).

MRISIM requires as input a set of “fuzzy” structure maps, one for each distinct tissue (or structure) type to be modeled, in which each voxel value is the probability of that voxel containing that tissue (structure) type. Such maps are generated by algorithms like INSECT or ANIMAL (see below) applied to a high-SNR data set. The MRI signal is simulated by solving the Bloch equations for the specified pulse sequence and tissue relaxation characteristics. Noise is modeled from first principles rather than by adding some parametric (e.g., gaussian) noise distribution to the expectation image (42). MRISIM has been used in validation studies for correction of MRI intensity nonuniformity (67) and tissue classification (84). It has been used to create a database of 108 simulate MRI images [3 slice thicknesses × 3 tissue contrasts (T1/T2/PD) × 3 noise levels × 4 levels of radiofrequency (RF) inhomogeneity], available at Web site http://www.bic.mni.mcgill.ca/.

Correction for 3D Intensity Nonuniformity—N3

A major problem for automated MRI image segmentation is the slowly varying change in signal intensity over the image, caused principally by nonuniformities in the radiofrequency field (Fig. 24.2). Apparent signal from any one tissue type is therefore different from one brain area to another, confusing automated segmentation algorithms that assume constant signal for one tissue type. We have developed a fully automated 3D technique for inhomogeneity correction, modeling inhomogeneity as the convolution of the true MRI intensity histogram with a blurring kernel. This effective kernel can be estimated and deconvolved by iterative entropy maximization. The method is applicable to any pulse sequence, field strength, and scanner (66 ,67).

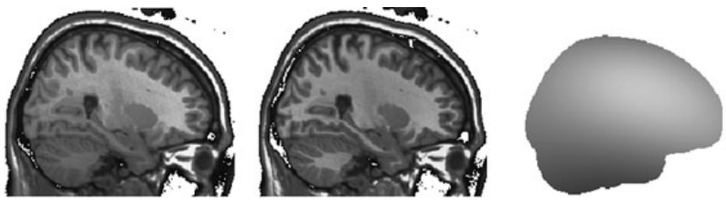


FIGURE 24.2. N3 correction for intensity nonuniformity. MRI image before (*left*) and after (*middle*) correction for nonuniformity field (*right*), estimated using N3. Note the increased uniformity of white matter regions.

Intrasubject Image Alignment—MINCTRACC

Alignment of images from the same subject, either from the same modality at different times in a longitudinal study or from different modalities, is achieved using a linear version of ANIMAL (see below), constrained to a six-parameter (three rotation, three translation) rigid-body transformation (15).

Stereotaxic Transformation—MRITOTAL

Stereotaxic transformation is achieved using a simple nine-parameter linear [three rotation, three translation, three scale, (15)] transformation to match the image volume to a master data set already resident in stereotaxic space. The master data set therefore defines the gross dimensions and orientation of stereotaxic space. We have previously constructed

a composite stereotaxic MRI data set drawn from 305 normal subjects, sampled on a 1-mm voxel grid (24), as that master data set. This mean data set, now termed ICBM305, has been circulated to over 100 international sites and defines the stereotaxic space for the SPM statistical package. That data set was derived from T1-weighted data with 2-mm-thick slice data. More recently, this has been superseded by a composite data set derived from 1-mm-thick data collected within the ICBM project (see below). That latter data set, while exhibiting higher contrast and more anatomic detail than the original ICBM305, was nevertheless mapped into the space of the ICBM305 using the nine-parameter MRITOTAL and is therefore a derivative of that first data set.

Tissue Classification—INSECT

We have developed an algorithm for tissue classification, known as INSECT (Intensity-Normalized Stereotaxic Environment for Classification of Tissue) (25 ,63 ,84). The algorithm operates upon multispectral (typically T1-, T2-, PD-weighted) data sets. In a series of preprocessing steps, each MRI data set is corrected for intensity nonuniformity (67), interslice normalization, and intersubject intensity normalization (Fig. 24.3). Stereotaxic transformation is then performed (15). An artificial neural network (ANN) classifier with one hidden layer is used to assign each voxel to a tissue type (gray/white/CSF) based on its MRI intensity feature space. The algorithm also employs tissue likelihood, based on the spatial location of the voxel in stereotaxic space, as orthogonal prior information to constrain the feature-space assignment. For example, periorbital fat exhibits a similar feature-space signal as white matter and, without consideration of spatial location, would be classified as white matter. Spatial masks expressing the normal distribution of tissue classes in the population (see Fig. 24.8) indicate that the likelihood of finding white matter in the periorbital stereotaxic region is small, and reduce the likelihood of misclassifiation.

INSECT operates on an arbitrary number of input images and generates a user-selected number of output tissue maps.

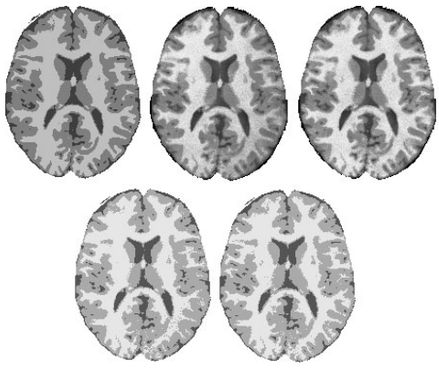


FIGURE 24.3. Classification with and without correction for intensity nonuniformity: tissue classification with INSECT with and without correction for nonuniformity using N3. An idealized 3D digital phantom was created from by segmentation of a high=nsignal-to-noise ratio (SNR) data set (17,37). The initial phantom data (*top left*) contains three classes: cerebrospinal fluid (CSF) (*black*), gray matter (*dark gray*), and white matter (*light gray*). This phantom was used to generate a simulated MRI image with (*top middle*) and without (*top right*) a 20% inhomogeneity running from *top left* to *bottom right* of the image. The INSECT-classified image without prior N3 correction (*bottom left*) exhibits artifactually thicker cortex at *bottom right* and thinner cortex at *top left* of the image, respectively, a consequence of the field inhomogeneity gradient. This artifact is removed in the N3-corrected classification (*bottom right*).

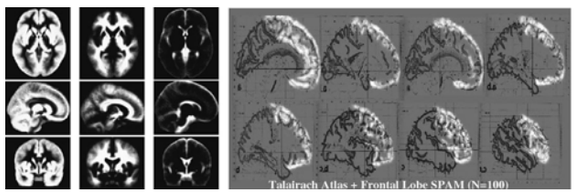


FIGURE 24.8. Tissue probability maps. **Left**: Cuts through INSECT-generated 3D tissue class maps for gray matter, white matter, and cerebrospinal fluid (CSF). **Right**: Serial sagittal sections through Talairach atlas with ANIMAL-generated probabilistic frontal cortex SPAM (statistical probability anatomy map) overlaid. In both cases 100 subjects were used to generate the SPAMs.

Regional Parcellation—ANIMAL

Manual labeling of brain voxels is both time-consuming and subjective. We have previously developed an automated algorithm to perform this labeling in 3D (13). The ANIMAL algorithm (Automated Nonlinear Image Matching and Anatomical Labeling), deforms one MRI volume to match another, previously labeled, MRI volume. It builds up the 3D nonlinear deformation field in a piecewise linear fashion, fitting cubical neighborhoods in sequence using a mutual information residual for parameter optimization (Fig. 24.4). The algorithm is applied iteratively in a multiscale hierarchy. At each step, image volumes are convolved with a 3D gaussian blurring kernel of successively smaller width [32-, 16-, 8-, 4-, and 2-mm full-width at half-maximum (FWHM)]. Anatomic labels are defined in the new volume by interpolation from the original labels, via the spatial mapping of the 3D deformation field. Originally, ANIMAL used 3D gradient magnitude as the image property to be matched. The ridge-tracking Lvv operator is now used to extract additional topologic information on brain shape in each image. Furthermore, the surface trace of major sulci, represented as 3D line segments, can be used as local constraints on image deformation (14 ,16). Both steps increase the correspondence of cortical anatomy across brains.

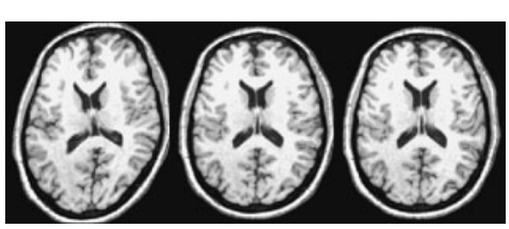


FIGURE 24.4. ANIMAL warping. Slice through a 3D ANIMAL deformation. The *left* image was warped to match the *right*, with the result in the *middle*.

Cortical Surface Segmentation and Unfolding—ASP

We have previously developed a fully automated procedure for unfolding the entire human cortex, using an algorithm that automatically fits a 3D mesh model to the cortical surface extracted from MRI (47). This algorithm, MSD, uses an iterative minimization of a cost function that balances the distance of the deforming surface from (a) the target surface, and (b) the previous iteration surface (Fig. 24.5). Specification of the relative weight of these competing forces allows MSD to range from unconstrained (data-driven) deformation to tightly constrained (model-preserving) deformation. Further shape-preserving constraints to penalize excessive local stretching and bending of the model surface are also employed. The initial mesh surface can be chosen arbitrarily to be a simple geometric object, such as a sphere, an ellipsoid, or two independently fitted hemispheres. The MSD algorithm has formed the basis of cortical analysis at both MNI and UCLA within the ICBM project (71 ,72 and 73). Recently, the algorithm has been extended to allow multiple concentric surfaces to be mapped simultaneously. The new algorithm, Automatic Segmentation

using Proximities (ASP), has the following refinements and capabilities (48), compared with the earlier MSD version:

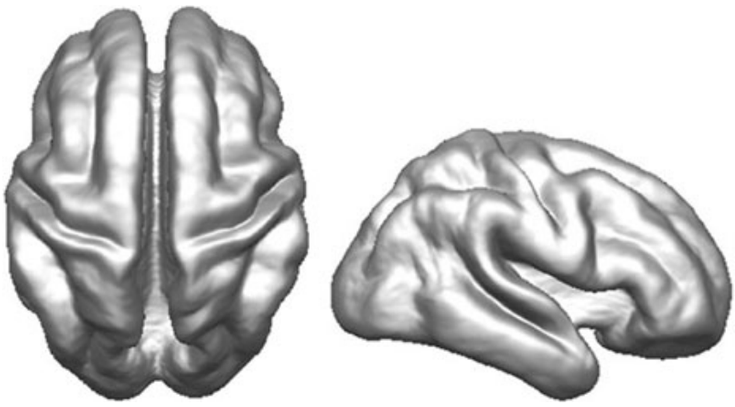


FIGURE 24.5. Average cortical surface. Average of 150 normal cortical surfaces. Note the prominence of the major gyral and sulcal features common to all brains.

- A boundary search along the normal local surface is used to increase the range of attraction of edges.
- The use of proximity constraints with appropriate weights excludes the potential for impossible self-intersecting surface configurations.
- Some arbitrary weights are replaced by more intuitive geometric constraints.
- Multiple surfaces, models, and data sets may be combined into a single objective function.
- Automatic identification of the total cerebral cortical surface from MR images is achieved in a robust way with respect to partial volume effects.
- A preliminary map of cortical gray matter thickness has been produced and related to previous studies.
- A higher resolution average brain surface has been created using the deeper sulcal penetration of ASP compared to earlier versions of this algorithm (47).

As an alternative form of stereotaxy applicable to cortical analysis, ASP also provides a fully automated mapping from 3D to an unfolded surface space. Since ASP iteratively deforms a starting 3D polygonal mesh onto the 3D cortical surface, the inverse mapping projects this fitted surface and topologic feature at each surface vertex back to the model space (47 ,48). Individual anatomic features such as gyral ridges and sulcal valleys are converted to measures of topology, e.g., curvature, mapped on to the model surface. These can be analyzed in terms of 2D variability on the surface of the starting model using a 2D surface coordinate space (Fig. 24.6).

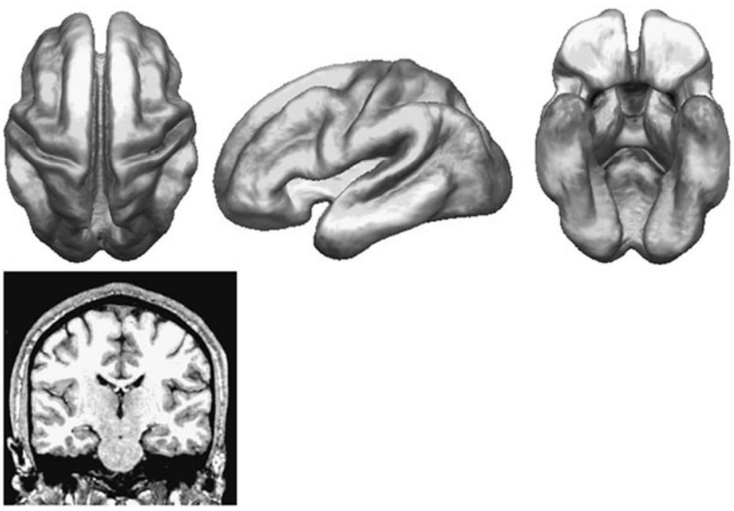


FIGURE 24.6. Cortical thickness. Mean cortical thickness in 150 normal adult brains, color-coded and texture-mapped onto the average cortical surface obtained from the same population.

Sulcal Extraction and Labeling—SEAL

We have implemented an automated sulcal extraction and labeling algorithm (SEAL) (45). At every voxel on the ASP-generated exterior cortical isosurface, SEAL calculates the two principal curvatures: k1, the mean curvature, and k2, the gaussian curvature (g = k1 * k2). Voxels with negative

mean curvature, belonging to sulci, are extracted and pruned to obtain a set of sulcal traces on the cortical surface. SEAL extracts the buried sulcus with an “active ribbon” that evolves in 3D from a superficial trace to the bottom of a sulcus by optimizing an energy function. We have defined a relational graph structure that stores, for each sulcus, its length, depth, and orientation, as well as attributes, e.g., hemisphere, lobe, sulcus type, connecting sulci, etc. Sulcal labeling is performed semiautomatically by tagging a sulcal trace in the 3D graph and selecting from a menu of candidate labels. The menu is restricted to most likely candidates by the use of sulcal probabilistic maps. SEAL identifies the sulci maps that overlap with each selected sulcus with highest likelihood (44 ,45) (Fig. 24.7).

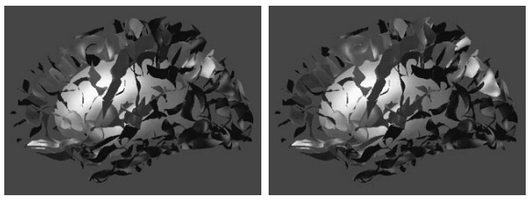


FIGURE 24.7. Use of spatial priors for automatic sulcus labeling within the sulcal extraction and labeling algorithm (SEAL). 3D representation of labeled sulcal folds occurs either automatically with SEAL, using prior probabilities (*left*), or manually labeled by a neuroanatomist (*right*). Different colors represent different sulcal labels, e.g., central sulcus is colored magenta (the smooth object is an average MRI surface, reduced in scale, included only to provide context for the sulcal maps). The automated and manual labeling of the sulci are in broad agreement, although some differences are apparent.

SAMPLE APPLICATIONS

Part of "24 - Automated 3D Analysis of Large Brain Mri Databases "

ICBM: Multicenter Consortium on Statistical Neuroanatomy

The International Consortium for Brain Mapping (ICBM) multicenter initiative was launched in 1993 as part of the Human Brain Project (52). Its overall goal is to create a 3D probabilistic brain atlas, based on MRI volumes from 450 normal adult brains. Within the ICBM project, all scans at all sites were collected with a strictly defined protocol, specifying three MRI volumes per subject (a 1-mm-thick, 1-mm-spaced gradient echo sequence for T1-weighted data and a 2-mm-thick, 1-mm-spaced double-echo sequence for PD and T2-weighted volumes). This database as been segmented using the pipeline environment described above and the variability captured in the form of probability maps as follows. Neuroanatomic variability can be conveniently represented in the form of 3D stereotaxic maps where each voxel expresses the likelihood of finding a particular structure at that location. By labeling one structure, e.g., caudate nucleus, in an ensemble of stereotaxic MRI volumes, a continuous 3D probability field for that structure (0% to 100% at each voxel), termed a statistical probability anatomy map (SPAM), can be constructed and used to test for group difference, e.g., pediatric versus adult brains, or outliers. For visualization purposes, these statistical maps can the thresholded at any level of structural probability to create probability isosurfaces suitable for surface-rendering and 3D display. Example SPAMs are shown for (a) gray/white/CSF tissue classes (INSECT, Fig. 24.8); (b) all major cortical gyri, cerebellum, and deep nuclei (ANIMAL, Fig. 24.8 and Fig. 24.9); and (c) cortical surface (ASP) (23).

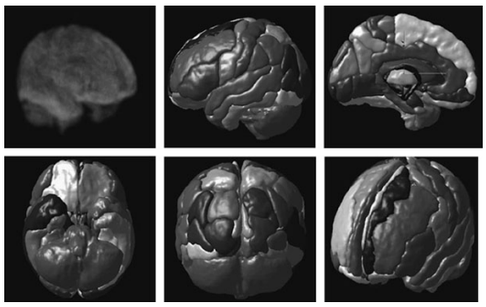


FIGURE 24.9. Rendered probabilistic atlas. Volume rendering (*top left*) and surface renderings (*all others*) of the 3D probabilistic atlas (*N* 100). For the surface renderings, the SPAMs were thresholded at the 40% level to generate regional probability isosurfaces.

Multicenter Clinical Trial Image Analysis

The principles of pipeline analysis described above for large databases of normal brain MRI data are equally applicable for population analysis of neuropathology or for tracking structural change over time, such as the progressive tissue atrophy, which occurs in some degenerative diseases. Indeed, the MRI analysis employed within the ICBM project was originally developed for a multicenter phase III clinical trial of a new pharmaceutical for treatment of multiple sclerosis. In this trial, 14 centers in the U.S. and Canada collected a total of 1,850 data sets, each data set composed of T1, T2, and PD volumes, from 514 subjects. All data collection was coordinated by the BIC clinical trials group, which performed quality control before trial launch and for all data shipped to the BIC for processing. Pipeline analysis of the database was used to generate 3D statistical maps of normal tissues and of MS lesions. In validation studies, the

results obtained with this automated approach for a subset of images were compared with those obtained by totally manual methods at seven established MRI/MS sites in Europe and North America. The results of the comparison indicated no significant differences between the BIC approach and the mean result obtained across the seven sites. They also indicate considerable variability among the sites themselves when analyzing the same data, which emphasizes the importance of the reproducibility of results obtained with a fully automated approach.

After correction for MRI intensity inhomogeneity, interslice and intervolume intensity normalization, and stereotaxic transformation, the multispectral data were tissue classified to identify MS lesion voxels for each patient time point. Figure 24.10 shows a 3D rendering of a probability map for lesion distribution obtained from all data sets. It shows the most likely locations for MS lesions within a population and is a convenient way to distill a large amount of population data into a single entity. Tests of drug effect are reduced to testing for a significant group difference in the overall volume of this distribution above a given threshold when partitioned into drug and placebo groups.

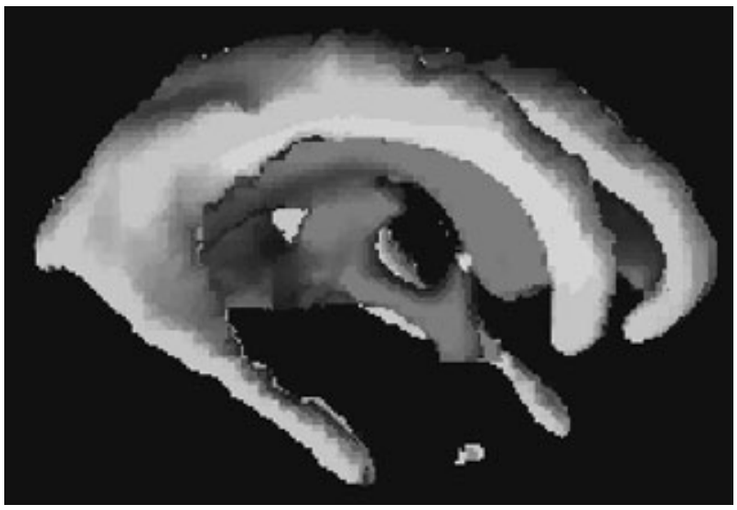


FIGURE 24.10. Multiple sclerosis (MS) lesion probability map. 3D renderings of probability maps for MS lesion (*light region*) and ventricle (*dark region*), obtained from 460 patients.

NIMH Intramural Pediatric Database

As part of an ongoing collaboration with Drs. Jay Giedd and Judy Rapoport at the National Institute of Mental Health (NIMH) Child Psychiatry Branch, the BIC image analysis pipeline has been used to process a large pediatric MRI database collected at the NIMH. Subjects were scanned on a General Electric 1.5 tesla Signa scanner using a 3D SPGR protocol. Approximately 1,800 T1-weighted images with slice thickness of 1.5 to 2.0 mm in the axial plane have

been obtained in approximately 600 children aged 3 to 18 from a number of subgroups:

- Normal development: A subset of this database, 111 normal children aged 4 to 17, was processed using the INSECT algorithm. All data were resampled into stereotaxic space using a simple nine-parameter linear transformation prior to image segmentation. Regression of population mean white matter intensity at each stereotaxic voxel against age yielded a regression map with significant correlation in the left arcuate fasciculus and the bilaterally in the internal capsule (33 ,61). The former tract links the anterior and posterior speech regions, while the latter is part of the corticospinal motor tract. These areas are continuously developing during maturation and it is tempting to interpret the results as increased myelination in these areas during development.

  A subset of the intramural NIMH database has also been analyzed by the ICBM group at UCLA under the direction of Arthur Toga (74). Using MSD-generated surfaces and tensor field analysis, they produced fourdimensional quantitative maps of growth patterns in the developing brain. Serial scanning in children aged 3 to 15 years across time spans of up to 4 years revealed a rostrocaudal wave of growth in the corpus callosum, a fiber system that relays information between brain hemispheres (Fig. 24.11). Peak growth rates, in fibers innervating association and language cortices, were attenuated after puberty, and contrasted sharply with a severe, spatially localized loss of subcortical gray matter. Conversely, at ages 3 to 6 years, the fastest growth rates occurred in frontal networks that regulate the planning of new actions.

- Child-onset schizophrenia: Fifteen patients with childhood-onset schizophrenia and 34 temporally yoked, healthy adolescents, scanned twice with an interval of 4 years, were analyzed using the pipeline (64). Lobar gray and white matter volumes were obtained with INSECT and ANIMAL. A significant decrease in cortical gray matter volume was seen for healthy controls in the frontal (2.6%) and parietal (4.1%) regions. For the childhood-onset schizophrenia group, there was a decrease in volume in these regions (10.9% and 8.5%, respectively) as well as a 7% decrease in volume in the temporal gray matter. Thus, the childhood-onset schizophrenia group showed a distinctive disease-specific pattern, with the frontal and temporal regions showing the greatest between-group differences. Changes in white matter volume did not differ significantly between the two groups. Patients with very early onset schizophrenia exhibit a fourfold greater decrease in cortical gray matter volume during adolescence and a disease-specific pattern of change.
- Attention-deficit/hyperactivity disorder (ADHD): Anatomic studies of boys with ADHD have previously detected volumetric differences in basal ganglia, prefrontal regions, and the cerebellar vermis. This study sought to

replicate those findings in young girls. MRI data from 53 girls with ADHD and 44 healthy matched female controls, ages 5 to 15, were analyzed using ANIMAL. Significantly smaller volumes were observed in prefrontal brain regions, caudate nucleus, globus pallidus, and amygdala bilaterally. The posterior-inferior cerebellar vermis volume and the rostrum of the corpus callosum were also significantly smaller in the ADHD group. Significantly smaller volumes were seen in the same brain regions as previously reported in boys with ADHD. As in boys, ADHD in girls is associated with anatomic deviations in corticostriatal-pallidal-thalamic circuits and in the posterior-inferior cerebellar vermis (9).

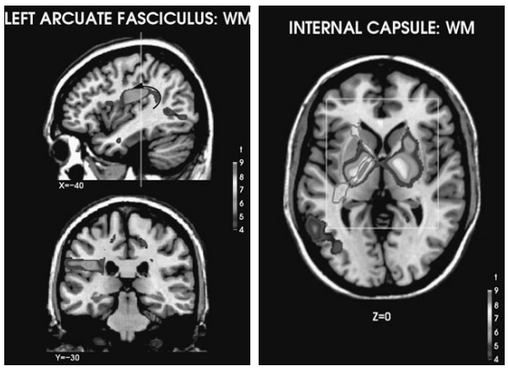


FIGURE 24.11. White matter density changes during pediatric development. Regression maps of white matter density changes over the age range from 4 to 17 (61). These maps show increased white matter density, possibly myelination, in the left arcuate fasciculus (*left*) and internal capsule (*right*), white matter tracts implicated in the development of language and motor skills, respectively.

NIH Extramural Pediatric MRI Database

The NIMH intramural database above has been acquired with only T1-weighted information and sparse behavioral information from a variety of subgroups, including approximately 200 normal children aged 3 to 8. While this database will provide much valuable information on pediatric development, there remains a need to create a more complete database of MRI information from a larger cohort of normal children, well-characterized by behavioral batteries. Therefore, a recent joint initiative by three National Institutes of Health (NIH) agencies (NIMH, NICHD, NINDS) has been launched to create such an MRI database of normal pediatric development in 550 children. This project, drawing upon a clinical trial model, will collect identical imaging and behavioral data at seven U.S. sites. The data will be consolidated into a single database at the BIC for pipeline analysis and eventual dissemination to the community. Each child in the age range of 5 to 18 will be scanned three times over a 6-year period. Behavioral batteries covering the major performance criteria will be collected at each time point. A younger cohort of approximately 100 children, aged 0 to 5, will undergo a more frequent scanning protocol and an age-appropriate behavioral battery. Magnetic resonance spectroscopy (MRS) and diffusion tensor imaging (DTI) information will also be collected at three of the sites to provide information on developmental neurochemistry, myelination, and fiber tract development.

SUMMARY AND FUTURE DIRECTIONS

Part of "24 - Automated 3D Analysis of Large Brain Mri Databases "

This chapter has presented an overview and sample applications of the MRI analysis pipeline environment at the Brain Imaging Centre (BIC) of the Montreal Neurological Institute. The key conceptual elements of this environment are as follows:

1. The use of stereotaxic space for consolidation of large ensembles of MRI data into a common spatial frame for analysis of gross neuroanatomy;
2. Fully automated 3D image preprocessing and segmentation;
3. Statistical analysis using voxel-bases random field theory and general linear models;
4. Incorporation of nonimaging parameters such as behavioral variables, demographic information, and genetic data into the statistical models.

The pipeline is highly modular, allowing for separate development and continued upgrading of the individual elements making up the pipeline. Processing is distributed across the BIC computing infrastructure using the PCS control scripts to optimize the utilization of resources. It has application in a variety of settings from basic neuroscience through clinical research to clinical trials. However, the current environment is focused on gross morphology. Conventional MRI allows us to collect gross anatomic information from a large sample of brains and develop population statistics. Unfortunately, this level of analysis provides no information about the cellular and molecular organization of the brain at a finer scale. A full understanding of functional neuroanatomy links function to macroscopic anatomy via these ultrastructural segregations. High-field MRI offers new possibilities, providing resolution of a few hundred microns over limited volumes. Sectioning, staining, and optical digitization of cadaver brains allow even finer spatial and chemical resolution in limited numbers of brains. A number of sites are bringing together these new acquisition technologies with the concepts of 3D stereotaxic mapping to create probabilistic maps at this finer scale. The advantage of the stereotaxic approach is that information from these many techniques operating at different spatial scales can be consolidated over many years into a systematic description of the whole brain structure and function. Such a rich database of information on both cerebral structure and function, accessible to sophisticated computational and statistical exploration, offers exciting possibilities for future brain research and clinical practice. Quite apart from direct hypothesis testing, such an environment may allow for the detection of hitherto unsuspected patterns of interaction among normal brain elements and the isolation of constellations of measurements that characterize specific disease states.

REFERENCES

1. Abell F, Krams M, Ashburner J, et al. The neuroanatomy of autism: a voxel-based whole brain analysis of structural scans. *Neuroreport* 1999;10(8):1647–1651.

2. Adolphs R, Damasio H, Tranel D, et al. A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *J Neurosci* 2000;20(7):2683–2690.

3. Amunts K, Malikovic A, Mohlberg H, et al. Brodmann's areas 17 and 18 brought into stereotaxic space-where and how variable? *Neuroimage* 2000;11(1):66–84.

4. Ashburner J, Andersson JL, Friston KJ. High-dimensional image registration using symmetric priors. *Neuroimage* 1999;9(6 pt 1):619–628.

5. Ashburner J, Friston KJ. Voxel-based morphometry-the methods. *Neuroimage* 2000;11(6 pt 1):805–821.

6. Ashburner J, Hutton C, Frackowiak R, et al. Identifying global anatomical differences: deformation-based morphometry. *Hum Brain Mapping* 1998;6(5 and 6):348–357.

7. Bakircioglu M, Grenander U, Khaneja N, et al. Curve matching on brain surfaces using frenet distances. *Hum Brain Mapping* 1998;6(5 and 6):329–333.

8. Carman GJ, Drury HA, Van Essen DC. Computational methods for reconstructing and unfolding the cerebral cortex. *Cerebral Cortex* 1995;5(6):506–517.

9. Castellanos FX, Giedd JN, Berquin PC, et al. Quantitative brain magnetic resonance imaging in girls with attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry* 2001;58:289–295.

10. Caviness VS Jr, Lange NT, Makris N, et al. MRI-based brain volumetrics: emergence of a developmental brain science. *Brain Dev* 1999;21(5):289–295.

11. Caviness VS Jr, Makris N, Lange NT, et al. Advanced applications of MRI in human brain science. *Keio J Med* 2000;49(2):66–73.

12. Christensen GE, Joshi SC, Miller MI. Volumetric transformation of brain anatomy. *IEEE Trans Med Imaging* 1997;16(6):864–877.

13. Collins DL, Holmes CJ, Peters TM, et al. Automatic 3D model-based neuroanatomical segmentation. *Hum Brain Mapping* 1995;3:190–208.

14. Collins DL, LeGoualher G, Venegopal R, et al. Cortical constraints for non-linear cortical registration. In: *Proceedings of 4th International Conference on Visualization in Biomedical Computing VBC 1996*. 1996:307–316.

15. Collins DL, Neelin P, Peters TM, et al. Automatic 3D registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr* 1994;18(2):192–205.

16. Collins DL, LeGoualher G, Barillot C, et al. Cortical constraints in fully automated non-linear spatial normalization of 3d brain images. *IEEE Trans Med Imaging* 2000; *submitted*.

17. Collins DL, Zijdenbos AP, Kollokian V, et al. Design and construction of a realistic digital brain phantom. *IEEE Trans Med Imaging* 1998;17(3):463–468.

18. Crespo-Facorro B, Kim JJ, Andreasen NC, et al. Human frontal cortex: an MRI-based parcellation method. *Neuroimage* 1999;10(5):500–519.

19. Csernansky JG, Joshi S, Wang L, et al. Hippocampal morphometry in schizophrenia by high dimensional brain mapping. *Proc Natl Acad Sci USA* 1998;95(19):11406–11411.

20. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. segmentation and surface reconstruction. *Neuroimage* 1999;9(2):179–194.

21. Dale AM, Sereno MI. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. *J Cogn Neurosci* 1993;5:162–176.

22. Duncan JS, Ayache N. Medical image analysis: progress over two decades and the challenges ahead. *IEEE Trans Pattern Analysis Machine Intelligence* 2000;22(1):85–106.

23. Evans AC, Collins DL, Holmes CJ, et al. A 3d probabilistic atlas of normal human neuroanatomy. Proceedings of the 3rd International Conference on Functional Mapping of Human Brain, 1997.

24. Evans AC, Collins DL, Mills SR, et al. 3D statistical neuroanatomical models from 305 MRI volumes. In: *IEEE Conference Record, Nuclear Science Symposium and Medical Imaging Conference*, San Francisco. 1993:1813–1817.

25. Evans AC, Frank JA, Antel J, et al. The role of MRI in clinical trials of multiple sclerosis: comparison of image processing techniques. *Ann Neurol* 1997;41:125–132.

26. Fiez JA, Damasio H, Grabowski TJ. Lesion segmentation and manual warping to a reference brain: intra- and interobserver reliability. *Hum Brain Mapping* 2000;9(4):192–211.

27. Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 1999;9(2):195–207.

28. Fischl B, Sereno MI, Tootell RB, et al. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapping* 1999;8(4):272–284.

29. Frank RJ, Damasio H, Grabowski TJ. Brainvox: an interactive, multimodal visualization and analysis system for neuroanatomical imaging. *Neuroimage* 1997;5(1):13–30.

30. Friston KJ, Frith CD, Liddle PF, et al. Comparing functional (PET) images: the assessment of significant change. *J Cereb Blood Flow Metab* 1991;11:690–699.

31. Friston KJ, Holmes AP, Worsley KJ, et al. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapping* 1995;2:189–210.

32. Gibaud B, Garlatti S, Barillot C, et al. Computerized brain atlases as decision support systems: a methodological approach. *Artif Intell Med* 1998;14(1 and 2):83–100.

33. Giedd JN, Blumenthal J, Jeffries NO, et al. Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neurosci* 1999;2(10):861–863.

34. Goldszal AF, Davatzikos C, Pham DL, et al. An image-processing system for qualitative and quantitative volumetric analysis of brain images. *J Comput Assist Tomogr* 1998;22(5):827–837.

35. Harris G, Andreasen NC, Cizadlo T, et al. Improving tissue classification in MRI: a three-dimensional multispectral discriminant analysis method with automated training class selection. *J Comput Assist Tomogr* 1999;23(1):144–154.

36. Hogan RE, Mark KE, Wang L, et al. Mesial temporal sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients. *Radiology* 2000;216(1):291–297.

37. Holmes CJ, Hoge R, Collins DL, et al. Enhancement of MR images using registration for signal averaging. *J Comput Assist Tomogr* 1998;22(2):324–333.

38. Joshi M, Cui J, Doolittle K, et al. Brain segmentation and the generation of cortical surfaces. *Neuroimage* 1999;9(5):461–476.

39. Kennedy DN, Lange N, Makris N, et al. Gyri of the human neocortex: an MRI-based analysis of volume and variance. *Cerebral Cortex* 1998;8(4):372–384.

40. Kennedy DN, O'Craven KM, Ticho BS, et al. Structural and functional brain asymmetries in human situs inversus totalis. *Neurology* 1999;53(6):1260–1265.

41. Kim JJ, Crespo-Facorro B, Andreasen NC, et al. An MRI-based parcellation method for the temporal lobe. *Neuroimage* 2000;11(4):271–288.

42. Kwan R, Evans AC, Pike GB. An extensible MRI simulator for post-processing evaluation. In: *Proceedings of the 4th International Conference on Visualization in Biomedical Computing, VBC 1996*. Hamburg: 1996:135–140.

43. Kwan RKS, Evans AC, Pike GB. MRI simulation-based evaluation of image processing and classification methods. *IEEE Trans Med Imaging* 1999;18(11):1085–1097.

44. LeGoualher G, Argenti AM, Duyme M, et al. Statistical sulcal shape comparisons: application to the detection of genetic encoding of the central sulcus shape. *Neuroimage* 2000;11(5):564–574.

45. LeGoualher G, Procyk E, Collins DL, et al. Automated extraction and variability analysis of sulcal neuroanatomy. *IEEE Trans Med Imaging* 1999;18(3):206–217.

46. Lohmann G, von Cramon DY, Steinmetz H. Sulcal variability of twins. *Cerebral Cortex* 1999;9(7):754–763.

47. MacDonald D, Avis D, Evans AC. Multiple surface identification and matching in magnetic resonance imaging. Visualization in Biomedical Computing 1994. *Proc SPIE* 1994;2359:160–169.

48. MacDonald D, Kabani N, Avis D, et al. Automated 3d extraction of inner and outer surfaces of cerebral cortex from MRI. *Neuroimage* 2000;12(3):340–356.

49. Magnotta VA, Andreasen NC, Schultz SK, et al. Quantitative in vivo measurement of gyrification in the human brain: changes associated with aging. *Cerebral Cortex* 1999;9(2):151–160.

50. Makris N, Meyer JW, Bates JF, et al. MRI-based topographic parcellation of human cerebral white matter and nuclei ii. rationale and applications with systematics of cerebral connectivity. *Neuroimage* 1999;9(1):18–45.

51. Manceaux-Demiau A, Bryan RN, Davatzikos C. A probabilistic ribbon model for shape analysis of the cerebral sulci: application to the central sulcus. *J Comput Assist Tomogr* 1998;22(6):962–971.

52. Mazziotta JC, Toga AW, Evans AC, et al. A probabilistic atlas of the human brain: theory and rationale for its development. *Neuroimage* 1995;2:89–101.

53. Mazziotta JC, Toga AW, Evans AC, et al. Digital brain atlases [letter]. *Trends Neurosci* 1995;18(5):210–211.

54. Miller M, Banerjee A, Christensen G, et al. Statistical methods in computational anatomy. *Stat Methods Med Res* 1997;6(3):267=n299.

55. Montagnat J, Delingette H, Scapel N, et al. Surface simplex meshes for 3d medical image segmentation. Proceedings, ICRA millennium conference. IEEE International Conference on Robotics and Automation Symposia Proceedings (Cat. No. 00CH37065), IEEE, part vol. 1, 2000.

56. Mummery CJ, Patterson K, Price CJ, et al. A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. *Ann Neurol* 2000;47(1):36=n45.

57. Nopoulos P, Flaum M, O'Leary D, et al. Sexual dimorphism in the human brain: evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. *Psychiatry Res* 2000;98(1):1=n13.

58. Nopoulos P, Swayze V, Flaum M, et al. Incidence of ectopic gray matter in patients with schizophrenia and healthy control subjects studied with MRI. *J Neuropsychiatry Clin Neurosci* 1998;10(3):351=n353.

59. Ono M, Kubik S, Abernathey CD. *Atlas of cerebral sulci*. Stuttgart: Georg Thieme Verlag, 1990.

60. Paus T, Tomaiuolo F, Otaky N, et al. Human cingulate and paracingulate sulci: pattern, variability, asymmetry, and probabilistic map. *Cerebral Cortex* 1996;6:207=n214.

61. Paus T, Zijdenbos A, Worsley K, et al. Structural maturation of neural pathway in children and adolescents. *Science* 1999;283:1908=n1911.

62. Penhune VB, Zatorre RJ, MacDonald JD, et al. Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from MR scans. *Cerebral Cortex* 1996;6(5):661=n672.

63. Pruessner JC, Li LM, Serles W, et al. Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. *Cerebral Cortex* 2000;10(4):433=n442.

64. Rapoport JL, Giedd JN, Blumenthal J, et al. Progressive cortical change during adolescence in childhood-onset schizophrenia: a longitudinal magnetic resonance imaging study. *Arch Gen Psychiatry* 1999;56(7):649=n654.

65. Sastre-Janer FA, Regis J, Belin P, et al. Three-dimensional reconstruction of the human central sulcus reveals a morphological correlate of the hand area. *Cerebral Cortex* 1998;8(7):641=n647.

66. Sled JG, Zijdenbos AP, Evans AC. A comparison of retrospective intensity nonuniformity correction methods for MRI. In: *Information processing in medical imaging*. 1997:459=n464.

67. Sled JG, Zijdenbos AP, Evans AC. A non-parametric method for automatic correction of intensity non-uniformity in MRI data. *IEEE Trans Med Imaging* 1998;17(1):87=n97.

68. Subsol G, Thirion JP, Ayache N. Application of an automatically built 3d morphometric brain atlas: study of cerebral ventricle shape. In: *Visualization in biomedical computing*. 4th International Conference, VBC 1996 Proceedings. Berlin: Springer-Verlag, 1996:373=n382.

69. Talairach J, Szikla G, Tournoux P. *Atlas d'anatomie stereotaxique du telencephale*. Paris: Masson, 1967.

70. Talairach J, Tournoux P. *Co-planar stereotactic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging*. Stuttgart: Georg Thieme Verlag, 1988.

71. Thompson P, MacDonald D, Mega MS, et al. Quantifying and correcting for variable cortical morphology in functional imaging using a deformable probabilistic brain atlas. In: *Human brain mapping*, vol 4. 1997:5423.

72. Thompson PM, Schwartz C, Toga AW. High-resolution random mesh algorithms for creating a probabilistic 3d surface atlas of the human brain. *Neuroimage* 1996;3:19=n34.

73. Thompson PM, Toga AW. A surface-based technique for warping three-dimensional images of the brain. *IEEE Trans Med Imaging* 1996;15(4):402=n417.

74. Thompson PM, Giedd JN, Woods RP, et al. Growth patterns in the developing human brain detected using continuum-mechanical tensor maps. *Nature* 2000;404(6774):190=n193.

75. Tibbo P, Nopoulos P, Arndt S, et al. Corpus callosum shape and size in male patients with schizophrenia [see comments]. *Biol Psychiatry* 1998;44(6):405=n412.

76. Vaillant M, Davatzikos C. Finding parametric representations of the cortical sulci using an active contour model. *Med Image Anal* 1997;1(4):295=n315.

77. Van Essen DC, Drury HA. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *J Neurosci* 1997;17(18):7079=n7102.

78. Van Essen DC, Drury HA, Joshi S, et al. Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. *Proc Natl Acad Sci USA* 1998;95(3):788=n795.

79. Wassink TH, Andreasen NC, Nopoulos P, et al. Cerebellar morphology as a predictor of symptom and psychosocial outcome in schizophrenia. *Biol Psychiatry* 1999;45(1):41=n48.

80. Westbury C, Zatorre RJ, Evans AC. Quantifying variability in the planum temporale: a probability map. *Cerebral Cortex* 1999;9:392=n405.

81. Worsley KJ, Evans AC, Marrett S, et al. A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab* 1992;12:900=n918.

82. Worsley KJ, Marrett S, Neelin P, et al. A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping* 1996;4(1):58=n73.

83. Wright IC, McGuire PK, Poline J-B, et al. A voxel-based method for the statistical analysis of gray and white matter density applied to schizophrenia. *Neuroimage* 1995;2:244=n252.

84. Zijdenbos AP, Evans AC, Riahi F, et al. Automatic quantification of multiple sclerosis lesion volume using stereotaxic space. In: *Proceedings of the 4th International Conference on Visualization in Biomedical Computing*, VBC 1996. 1996:439=n448.

25

In Vivo Magnetic Resonance Spectroscopy Studies of the Glutamate and Gaba Neurotransmitter Cycles and Functional Neuroenergetics

Douglas L. Rothman

Fahmeed Hyder

Nicola Sibson

Kevin L. Behar

Graeme F. Mason

Jun Shen

Ognen A. C. Petroff

Robert G. Shulman

Douglas L. Rothman, Fahmeed Hyder, Kevin L. Behar, Graeme F. Mason, Ognen A. C. Petroff, Robert G. Shulman: Yale University School of Medicine, New Haven, Connecticut.

Nicola Sibson: University of Oxford, Oxford, United Kingdom.

Jun Shen: Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York.

In the last 5 years there has been a renewed interest in the role of metabolism in supporting brain function. Much of this interest is based on the development of functional positron emission tomography (PET) and magnetic resonance imaging (MRI). Although often incorrectly described as directly mapping neuronal activity, both functional PET and MRI actually measure changes in either glucose metabolism or physiologic parameters coupled to glucose metabolism such as blood flow and volume (1). A major limitation in interpreting functional imaging is that the relationship between neuronal activity and the neuroenergetic processes supported by glucose metabolism is poorly defined (2 ,3). The term *neuronal activity* applies to a spectrum of energy-requiring processes including action potential propagation, neurotransmitter release and uptake, vesicular recycling, and maintenance of membrane potentials (4). All of these processes are involved in short-term neuronal information transfer, and the relative distribution of energy among them remains an open question. There is also uncertainty as to how the different classes of neurons in a region contribute to the overall energy consumption. While an increase in the imaging signal is usually assigned to an increase in neuronal excitation, this interpretation is confounded by both inhibitory and excitatory neuronal function requiring energy. Observation of a regional increase or decrease of the functional imaging signal is not sufficient to distinguish these possibilities. Glia also requires energy, and the relationship between its energy demands and neuronal activity remains to be established. Given these uncertainties about the meaning of the signal at a neuronal level, the validity of functional imaging as a tool for studying mental processes has been largely established based on agreement with prior expectation from psychological paradigms (3 ,5).

An alternative approach for imaging brain function, which has the potential of directly measuring metabolic pathways involved in excitatory and inhibitory neurotransmission, is *in vivo* magnetic resonance spectroscopy (MRS). MRS uses technology similar to that of the more familiar MRI. It differs by allowing the measurement of the concentrations and synthesis rates of individual chemical compounds within precisely defined regions in the brain. The basis of its chemical specificity is that the resonance frequency of an MRS active nucleus depends not only on the local magnetic field strength, but also on its chemical environment, a phenomenon referred to as chemical shift. MRS measurements of the ^{1}H nucleus are the most commonly used for *in vivo* studies due to ^{1}H being the most sensitive nucleus present in biological systems. Metabolites that can be measured by ^{1}H MRS include aspartate, γ-aminobutyric acid (GABA), glucose, glutamate, glutamine, and lactate. These metabolites play critical roles in neuroenergetics, amino acid neurotransmission, and neuromodulation. Another nucleus of importance for *in vivo* MRS studies is the

^{13}C nucleus. The natural abundance of the ^{13}C isotope is 1.1% so that in conjunction with the infusion of ^{13}Cenriched substrates the rates of isotopic incorporation into brain metabolites can be measured. Substrates labeled with the nonradioactive, stable, ^{13}C isotope have been employed *in vivo* to study metabolic flux, enzyme activity, and metabolic regulation in the living brain of animals and humans (6 ,7 ,8 ,9 ,10 ,11 ,12 ,13 ,14 ,15 ,16 ,17 ,18 ,19 ,20 ,21 ,22 ,23 ,24 ,25 ,26 ,27 ,28 ,29 ,30 ,31 ,32 ,33 ,34 ,35 ,36 ,37 ,38 and 39). Enhanced sensitivity may be achieved by measuring the ^{13}C enrichment of a molecule through indirect detection through ^{1}H MRS. From these measurements the flux through specific metabolic pathways may be calculated (17 ,18).

This chapter covers the recent development of *in vivo* MRS to study neuronal glutamate and GABA metabolism and the relationship of amino acid metabolism to functional neuroenergetics. The brain pools of GABA, glutamate, and glutamine have been shown to be localized within glutamatergic neurons, GABAergic neurons, and glia, respectively (under nonpathologic conditions). Under nonfasting conditions glucose is the almost exclusive source of energy for the brain. By following the flow of ^{13}C label from glucose into these metabolites, MRS has been used to determine the separate rates of glucose oxidation in these cell types. The metabolism of glutamatergic neurons, GABAergic neurons, and glia is coupled by neurotransmitter cycles. In the glutamate/glutamine cycle, glutamate released from nerve terminals (by either vesicular release or transport reversal) is transported into surrounding glial cells, and converted to glutamine. Glutamine in then transported out of the glia and into the neurons, where it is converted back to glutamate, thereby completing the cycle (Fig. 25.1). By following the flow of ^{13}C label from glutamate into glutamine, the rate of the glutamate/glutamine cycle may be determined using MRS. Through a similar strategy the GABA/glutamine cycle may be measured.

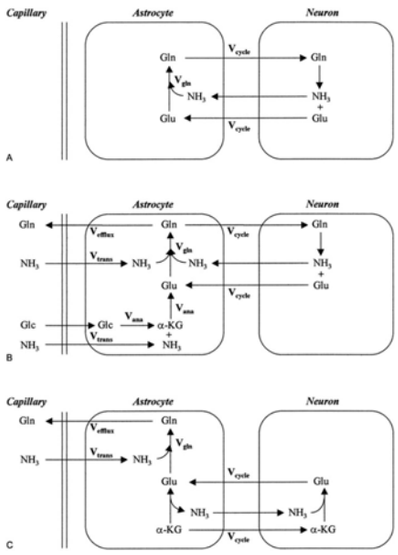


FIGURE 25.1. Schematic representations of the glutamate/glutamine cycle between neurons and astrocytes and the detoxification pathway of glutamine synthesis. **A:** The glutamatine/glutamate cycle between neurons and astrocytes. Released neurotransmitter glutamate is transported from the synaptic cleft by surrounding astrocytic end processes. Once in the astrocyte, glutamate is converted to glutamine by glutamine synthetase. Glutamine is released by the astrocyte, transported into the neuron, and converted to glutamate by phosphate-activated glutaminase (PAG), which completes the cycle. **B:** Including the ammonia detoxification (or anaplerotic) pathway of glutamine synthesis. The net rate of glutamine synthesis reflects both neurotransmitter cycling (V_{cycle}) and anaplerosis (V_{ana}). The stoichiometric relationships required by mass balance between the net balance of ammonia and glutamine and V_{ana} are given in Eq. 2. **C:** An alternative model for neuronal glutamate repletion in which the astrocyte repletes the lost neuronal glutamate by providing the neuron with α-ketoglutarate [or equivalently other tricarboxylic acid cycle (TCA) intermediates] (32,33 and 34). α-Ketoglutarate is converted back to glutamate by neuronal glutamate dehydrogenase. Glc, glucose; a-KG, α-ketoglutarate: V_{trans}, net rate of net ammonia transport into the brain (VNH4 in the text); V_{efflux}, rate of glutamine efflux from the brain; V_{ana}, anaplerotic flux; V_{cycle}, rate of the glutamate/glutamine cycle; V_{gln}, rate of glutamine synthesis. Using [2-^{13}C] glucose (27) and [2-^{13}C] acetate precursors these pathways may now be distinguished.

The application of MRS to study brain glutamate and GABA metabolism and the coupling of neurotransmitter cycling to neuroenergetics have provided several new and controversial insights into the relationship of brain metabolism and function. Contrary to the previous view of a separate metabolic and neurotransmitter pool of glutamate, glutamate release and recycling have been shown to be a major metabolic pathway. MRS studies of GABA metabolism in the rodent and human brain have suggested that there is also an important role of the metabolic pool of GABA in inhibitory function. Another key finding is that the glutamate/glutamine cycle in the cerebral cortex is coupled in a close to 1:1 ratio to neuronal (primarily glutamatergic) glucose oxidation above isoelectricity. This finding, in combination with cellular studies, has led to a model for the coupling between functional neuroenergetics and glutamate neurotransmission. The coupling between neurotransmission and neuroenergetics provides a linkage between the functional imaging signal and specific neuronal processes. This chapter reviews these findings and discusses some of their implications for functional imaging.

MRS is a low spatial resolution method, with a resolution for studying neurotransmitter systems of approximately 1 to 4 mm^3 in animal models and 7 to 40 mm^3 in human brain. Even in the best case the MRS signal is the sum of the signal from a large number of neurons and glia including many different subtypes. Fortunately, nature has localized key enzymes and metabolites involved in neurotransmitter cycling in specific cell types, which greatly simplifies the interpretation of the MRS measurements. The evidence of the cellular compartmentalization of metabolism largely derives from invasive methods with cellular and subcellular resolution, which are reviewed here. As with any new technique there are still uncertainties due to methodologic issues. Studies performed to validate the MRS measurements will be reviewed, and present limits in measurement accuracy and interpretation delineated.

IN VIVO ^{13}C MRS MEASUREMENTS OF THE PATHWAYS OF GLUCOSE OXIDATION: FINDINGS AND VALIDATION

Part of "25 - Magnetic Resonance Spectroscopy Studies of the Glutamate and Gaba Neurotransmitter Cycles and Functional Neuroenergetics "

This section reviews studies in which MRS was used to measure the pathways of glucose oxidation in the cerebral cortex. Glucose oxidation under nonfasting conditions is almost the exclusive source of energy for the brain. The localization of key enzymes involved in GABA and glutamate metabolism in specific cell types provides the capability for MRS to study their separate neuroenergetic requirements. As shown in Fig. 25.2 , which is a ^{13}C MRS spectrum obtained by Gruetter and co-workers (35) at 4 T, the chemical specificity of MRS allows the flow of ^{13}C label from glucose to be followed into several metabolites in the brain coupled to energy metabolism including aspartate, GABA, glutamate, and glutamine. The major finding of these studies is that in normal conditions in nonactivated human cerebral cortex and in rodent models, glucose oxidation in glutamatergic neurons accounts for between 60% and 80% of cerebral cortex energy consumption. The remaining 20% to 40% is primarily distributed between GABAergic neurons and glia.

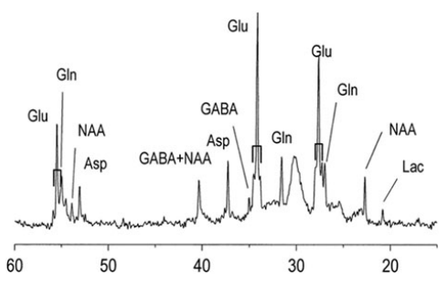


FIGURE 25.2. ^{13}C magnetic resonance spectroscopy (MRS) spectrum in the occipital/parietal lobe at 4 T. The figure shows a 50-minute accumulation ^{13}C MRS spectrum obtained at 4 T approximately 60 minutes after the start of a 1-^{13}C glucose infusion. The spectrum was obtained from a 72-mL volume centered on the midline in the occipital/parietal lobe. The top trace is an expansion of regions of the bottom trace. Labeled resonances include the C2, C3, and C4 positions of glutamate, glutamine, aspartate, and γ-aminobutyric acid (GABA) and the C3 position of lactate. As described in the text (see In Vivo 13C MRS Measurements of the Pathways of Glucose Oxidation: Findings and Validation), the localization of the synthetic enzymes and pools of GABA and glutamine to GABAergic neurons and glia, respectively, and the localization of the majority of the glutamate pool to glutamatergic neurons, allows the relative rates of glucose oxidation in these cell types to be determined from the flow of ^{13}C label into these pools. (From Gruetter R, Seaquist ER, Kim S, et al. Localized in vivo ^{13}C-NMR of glutamate metabolism in the human brain: initial results at 4 tesla. *Dev Neurosci* 1999;20:380–388, with permission.)

MRS Measurement of the Rate of Glucose Oxidation in Glutamatergic Neurons

The initial use of MRS to study brain metabolism was to measure glucose oxidation by following the flow of ^{13}C isotope from [1-^{13}C] glucose into the C4 position of glutamate (2 ,6). Figure 25.3 diagrams the flow of ^{13}C label from a [1-^{13}C] glucose precursor to C4-glutamate and subsequently C4-glutamine. Glucose is metabolized to pyruvate by the glycolytic pathway, which labels C3-pyruvate. The

label is then transferred to the tricarboxylic acid cycle (TCA) by the actions of pyruvate dehydrogenase (PDH) and citrate synthase. When the label reaches C4-α-ketoglutarate it is transferred to the large neuronal glutamate pool by the high activity exchange reactions of the amino acid transaminases and mitochondrial/cytosolic transporters. The large glutamate pool was first identified in ^{14}C tracer studies (40). Based on kinetic and immunohistochemical staining studies, it is believed to correspond to the glutamate pool of glutamatergic neurons (18 ,41 ,42). Due to the rate of these exchange reactions being many times faster than the TCA cycle the glutamate pool acts as a label trap for isotope that enters the neuronal TCA cycle via pyruvate dehydrogenase (17 ,18). ^{13}C MRS may be used to measure the accumulation of ^{13}C label into the trapping glutamate pool, and the kinetic curves analyzed by metabolic modeling to calculate the rate of the neuronal TCA cycle (18). The trapping pool assumption is not essential to determine the rate of the TCA cycle because subsequent labeling in the C3 position of glutamate can be measured to allow calculation of the label exchange rate (17 ,18). Because glucose is the primary fuel for neuronal oxidation, the measurements of the TCA cycle may be converted to measurements of glucose oxidation using known stoichiometries (17 ,18).

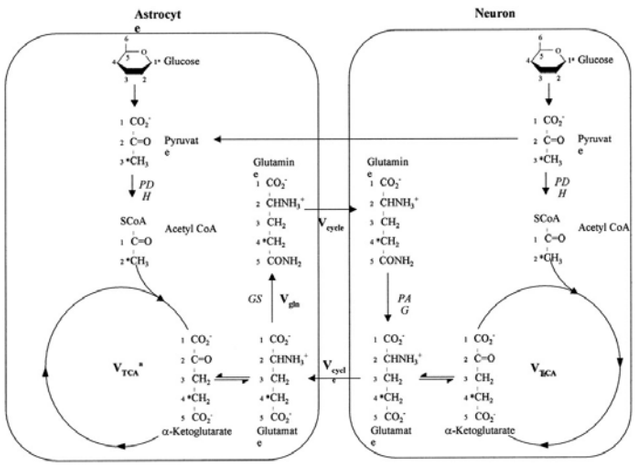


FIGURE 25.3. Isotopic labeling of C4-glutamine by the glutamate/glutamine cycle from a [1-^{13}C] glucose precursor. Infused [1-^{13}C] glucose labels neuronal C3-pyruvate. This label is then incorporated via the combined action of pyruvate dehydrogenase and the TCA cycle into α-ketoglutarate, which is in rapid exchange with glutamate due to the action of several transaminases. The large glutamate pool in the neuron acts as a label trap with [4-^{13}C]-glutamate accumulating at the rate of the neuronal TCA cycle. Released [4-^{13}C] glutamate from the nerve terminal is taken up by glial transport and the ^{13}C label is transferred to [4-^{13}C] glutamine through the action of glutamine synthetase at the rate of the glutamate/glutamine cycle. Interpretation of glutamine labeling is complicated by ^{13}C label entering by the astrocyte pyruvate dehydrogenase reaction. MRS studies using ^{15}N ammonia, [2-^{13}C] glucose, and [2-^{13}C] acetate, as well as comparison with traditional measurements of the uptake of net glutamine precursors, have shown that the majority of labeling in glutamine from [1-^{13}C] glucose is from the glutamate/glutamine cycle (27,36,37,38 and 39).

The rate of neuronal glucose oxidation has been determined in several studies from ^{13}C MRS and ^{1}H-^{13}C MRS measurements of cerebral cortex glutamate turnover from a [1-^{13}C] glucose precursor in animal models (2 ,14 ,15 ,16 and 17 ,21 ,22 ,25 ,26 and 27) and humans (12 ,13 ,18 ,19 ,29 ,31 ,35 ,43 ,44). Comparison of the rates of neuronal glucose oxidation measured in these studies with conventional arteriovenous (AV) difference and PET measurements of total glucose consumption found that the majority (between 70% and 90%) of total glucose oxidation in the rat and human brain is associated with the large glutamate pool, believed to reflect glutamatergic neurons, measured by MRS. In two recent ^{13}C MRS studies of resting awake human occipital parietal cortex, in which other pathways of glucose metabolism were directly measured, a similar range of between 60% (35) and 80% (29) of total glucose oxidation was calculated for the large glutamate pool. The large percentage of cortical synapses that are glutamatergic and the high electrical activity of glutamatergic pyramidal cells (4 ,45) may explain why such a large fraction of total glucose oxidation is associated with glutamatergic neurons.

A caveat to the interpretation of the glutamate turnover measurement is that glutamate is present in all brain cells. Based on the sensitivity limitations of staining methods and kinetic studies in measuring glutamate levels in glia and other neuron types, particularly GABAergic, the assignment of the fraction of glucose oxidation occurring in glutamatergic neurons may be overestimated by up to 20%. In the future, the fraction of glutamate in glia may be measured more accurately through dynamic ^{13}C MRS measurements of glutamate and glutamine labeling during the infusion of labeled acetate that is incorporated into the brain selectively in the glia (28 ,38 ,39).

MRS Measurements of the Rate Glucose Oxidation in GABAergic Neurons

GABA is the major inhibitory neurotransmitter and may represent over 30% of the synapses in the cerebral cortex (4 ,46 ,47). GABA is synthesized from glutamate in GABAergic neurons by the enzyme glutamic acid decarboxylase (GAD). GABA may then be returned to the TCA cycle through successive action of the GABA shunt enzymes, GABA-transaminase, and succinic semialdehyde dehydrogenase, or released from the neuron. Almost all of the brain GABA pool is localized to GABAergic neurons under normal conditions. The labeling of the GABA pool from [1-^{13}C] glucose provides a minimum estimate of the rate of glucose oxidation in the GABAergic neuron. The estimate is a minimum because label may bypass GABA and continue from α-ketoglutarate/glutamate into the TCA cycle directly. *In vitro* MRS analysis of cerebral cortex from extracts of rats infused with [1-^{13}C] glucose has been used to measure the time course of labeling in the GABA and glutamate pools (24 ,48). The isotopic labeling results of the Manor et al. (24)

study were analyzed with a metabolic model to determine the relative rates of glucose oxidation in the glutamate and GABA pool. Under conditions of α-chloralose anesthesia the rate of glucose oxidation in GABAergic neurons was estimated to be between 10% and 20% of total neuronal glucose oxidation. This value is similar to previous estimates obtained using isotopic methods and by inhibiting the degradative enzyme GABA transaminase (24). It should be noted that determination of the rate of GABA synthesis from isotopic methods depends on the assumption that the glutamate precursor pool for GABA is severalfold lower in concentration than GABA, an assumption that is consistent with findings using cellular staining (41 ,42). In the recent studies of human cerebral cortex (13 ,29 ,35), the rate of GABA synthesis, and by inference glucose oxidation in the GABAergic pool, was estimated to be on the order of 10% of total glucose oxidation, although no rates were given. In the future, with the higher sensitivity available using inverse MRS methods in combination with the development of ultrahigh field magnets for human studies, measurements of the rate of glucose oxidation in GABAergic neurons should be possible in humans.

In Vivo MRS Measurements of the Rate of Glucose Oxidation in Glia

A long-term controversy in brain metabolism studies has been the rate of glucose oxidation in glial cells. Early estimates range from 10% to over 50% of glucose oxidation (49). MRS may be used to measure the rate of glial glucose oxidation based on the localization of the enzyme glutamine synthetase in the glia (50). This localization allows the rate of the glial TCA cycle to be calculated from the labeling of glutamine from glial glutamate. The most quantitative early findings were by Van den Berg and co-workers (40), who, using ^{14}C isotopic labeling strategies, assigned a rate to glial pyruvate dehydrogenase, which they referred to as the small glutamate pool, of 15% to 25% of total pyruvate dehydrogenase (neuronal + glial) activity. The pyruvate dehydrogenase rate is equal to the rate of complete glucose oxidation by the TCA cycle plus the rate of net glial anaplerosis. These measurements were performed using extract analysis of whole brains. Two recent ^{13}C MRS measurements of humans have measured glial pyruvate dehydrogenase as accounting for between 8% (29) and 15% (35) of total pyruvate dehydrogenase activity in the occipital parietal lobe. A limitation of these studies is that they did not measure the rate of the glial TCA cycle, only the pyruvate dehydrogenase step, and therefore the total oxidative energy produced in the glia was not calculated. In a preliminary study in human cerebral cortex using [2-^{13}C] acetate, which is exclusively incorporated into the astrocyte (28), as a tracer Lebon et al. (38) found that the glial TCA cycle accounts for approximately 15% of total glucose oxidation.

Summary and Remaining Questions

In vivo MRS measurements of nonactivated cerebral cortex in rats and humans have found that from 60% to over 80% of glucose oxidation is associated with the large glutamate pool, reflecting primarily glutamatergic neurons. The remainder is primarily distributed between GABAergic neurons and glia. The development of new labeling strategies such as [2-^{13}C] acetate and higher sensitivity MRS measurements should allow the contributions of these cell types to be more accurately determined. Within the error of the MRS measurements, and the contribution of glutamate pools in other neuronal classes to the glucose oxidation measurement, there may be up to 10% of total glucose oxidation available for other cell types such as dopaminergic and serotoninergic nerve terminals. An objection that has been raised to these findings is the possibility that small highly metabolically active pools may be missed by the MRS method. Under nonstimulated conditions the good agreement of the MRS measurement with AV difference and PET measurements of total glucose consumption indicate that the contribution of these small pools is not large. MRS measurements of glucose metabolism during cortical activation will be reviewed below (see *In Vivo* MRS Measurements of Neuroenergetics During Functional Activation).

IN VIVO MRS MEASUREMENTS OF THE RATE OF THE GLUTAMATE/GLUTAMINE CYCLE: FINDINGS AND VALIDATION

Part of "25 - Magnetic Resonance Spectroscopy Studies of the Glutamate and Gaba Neurotransmitter Cycles and Functional Neuroenergetics "

The function of the glutamate/glutamine cycle is to prevent depletion of the nerve terminal glutamate pool by synaptic release. Glial cells have a high capacity for transporting glutamate from the synaptic cleft in order to maintain a low ECF (extracellular fluid) concentration of glutamate (50 ,51). *In vivo* and *in vitro* studies indicate that glutamate released by the neuron is taken up by the glia and converted to glutamine by glutamine synthetase (53 ,54), an enzyme found exclusively in glia (52). Glutamine is transported from the glia into the ECF where it is taken up by neurons and converted back to glutamate through the action of phosphate-activated glutaminase (PAG) (55). Based on extensive data from isotopic labeling studies, immunohistochemical staining of cortical cells for specific enzymes, isolated cell, and tissue fractionation studies, it has been proposed that glutamate (as well as GABA) taken up by the glia from the synaptic cleft may be returned to the neuron in the form of glutamine (40 ,56 ,57 and 58). The generally accepted model of the glutamate/glutamine neurotransmitter cycle is shown in Fig. 25.1A .

Despite a wealth of evidence from enzyme localization and isolated cell studies, the rate of the glutamate/glutamine cycle and its importance for brain function have been controversial due to difficulties in performing measurements in the living brain. Because the neurotransmitter glutamate is packaged in vesicles (59 ,60), controversy has arisen about the fraction of glutamate actually involved in the cycle, leading to the concept of a small “transmitter” versus a large “metabolic” glutamate pool. Supporting the concept that glutamate neurotransmitter flux is a small fraction of total glucose metabolism are findings in isolated cells and nonactivated brain slices of a low rate of label incorporation from [1-^{13}C] glucose (61). The concept of a metabolically inactive neurotransmitter pool was brought into question in 1995, when, using ^{13}C nuclear magnetic resonance (NMR), we measured a high rate of glutamine labeling from [1-^{13}C] glucose in the occipital/parietal lobe of human subjects (12). A high rate of glutamine synthesis was calculated from these data (18). At the time of the initial ^{13}C NMR study, the rate of the glutamate/glutamine cycle could not be calculated due to the lack of a model for distinguishing isotopic labeling from this cycle from other sources of glutamine labeling, most significantly removal of cytosolic ammonia produced by metabolism and uptake of plasma ammonia (62). Net ammonia removal requires the de novo glutamine synthesis via the anaplerotic pathway in the glia. In addition, several other pathways, including the glial TCA cycle, have been proposed as providing significant precursors for glutamine

synthesis (61 ,62). To calculate the rate of the glutamate/glutamine cycle, Sibson et al. (25) developed a metabolic model for separating the pathways of glutamine synthesis.

This section reviews ^{13}C MRS measurements of the glutamate/glutamine cycle, emphasizing studies performed to validate the MRS technique. The important and surprising result of these studies is that the glutamate/glutamine cycle is a major metabolic flux, far exceeding de novo glutamine synthesis. The rate of the glutamate/glutamine cycle in the awake resting human cerebral cortex is between 60% and 80% of total glucose oxidation.

Development of a Two-Compartment Metabolic Model of Glutamine Metabolism to Separately Determine the Rate of the Glutamate/Glutamine Cycle and Anaplerotic Glutamine Synthesis

The rate of the glutamate/glutamine cycle is calculated from the time course ^{13}C labeling of glutamine relative to the labeling of its precursor neuronal glutamate. If neuronal glutamate were the only precursor of glutamine, the calculation would be straightforward. Unfortunately, the calculation is complicated by label from [1-^{13}C] glucose entering both the neuronal and glial TCA cycles via pyruvate dehydrogenase. Glutamate in both cell types will be labeled in the C4 position by exchange with α-ketoglutarate in the TCA cycle (Fig. 25.3). The flow of label from C4-glutamate into C4-glutamine is proportional to the total rate of glutamine synthesis. However, unless the relative flow of ^{13}C label into the glial glutamate pool from the glial pyruvate dehydrogenase and neuronal glutamate are distinguished, the fraction of glutamine synthesis due to the glutamate/glutamine cycle cannot be calculated. To determine the rate of the glutamate/glutamine cycle from a [1-^{13}C] glucose precursor, we developed a metabolic model to constrain the rate of glutamine labeling from glial pyruvate dehydrogenase.

Glutamine production via glutamine synthetase requires two substrates, glutamate and ammonia. As shown in the flow diagram of Fig. 25.1A , glutamine synthesis receives precursor glutamate from both glial uptake of released neurotransmitter glutamate and glial anaplerosis. A mathematical model was developed to interpret isotopic data in order to separate these pathways (25 ,27 ,29 ,36). The model extends previous formulations by imposing mass balance constraints on the brain glutamate and glutamine pools that relate the rate of de novo glutamine synthesis to the net uptake of anaplerotic precursors from the blood. Glutamine efflux is the primary source of nitrogen removal from the brain (49 ,62). Nitrogen must be removed from the brain in order to maintain low concentrations of ammonia, which when elevated will interfere with brain function (62). Because at steady state the concentration of glutamine remains constant, loss of glutamine by efflux (V_{efflux}) must be compensated for by de novo synthesis of glutamine by anaplerosis (V_{ana}). For de novo synthesis by anaplerosis, pyruvate derived from glucose is converted by CO_2 fixation (V_{CO2}) to oxaloacetate by the enzyme pyruvate carboxylase, which is active only in the glia (54). Through the action of the TCA cycle oxaloacetate is converted to α-ketoglutarate, which may be converted to glutamate either by ammonia fixation via glia glutamate dehydrogenase or alternatively through transamination with other amino acids (37). Glial glutamate is then converted to glutamine by glutamine synthetase. One or two ammonia molecules are fixed per glutamine molecule synthesized through anaplerosis, depending on the relative fluxes of NH_4^+ fixation versus transamination. Applying nitrogen mass balance constraints leads to the relationship V_{NH4} = (1 to 2)V_{efflux} at steady state. The additional requirement of carbon mass balance leads to the following relationship:

$$V_{ana} = V_{efflux} = V_{CO2} = (\tfrac{1}{2} \text{ to } 1)V_{NH4+} \qquad [1]$$

Total glutamine synthesis is then related to synthesis for ammonia detoxification (V_{ana}) and the glutamate/glutamine cycle (V_{cycle}) by the following expression:

$$V_{gln} = V_{cycle} + V_{ana} \qquad [2]$$

Note that V_{CO2} may be higher than V_{ana} if anaplerosis is needed to replace TCA cycle intermediates lost by oxidative processes or pyruvate recycling (63 ,64).

Examination of Eq. 2 indicates that V_{cycle} may be derived from a measurement of V_{gln} from a ^{13}C MRS experiment in combination with a measurement of any of the rates linked by mass balance considerations to anaplerotic glutamine synthesis. A limitation of isotopic measurements of flux is that isotopic exchange cannot be distinguished from net flux. The linkage between the labeling of glutamine through glial pyruvate dehydrogenase and the brain anaplerosis flux allows the validation of isotopic measurements of glutamine ^{13}C and ^{15}N labeling against traditional AV difference measurements.

The glutamate/glutamine cycle measurement using a [1-^{13}C] glucose precursor also includes contributions from the GABA/glutamine cycle (34 ,57 ,65). GABA is the main inhibitory neurotransmitter, and has been measured by *in vivo* ^{1}H and ^{13}C MRS in animals and humans (12 ,13 ,24 ,29 ,35 ,66) (see In Vivo MRS Studies of GABA Metabolism and the Effects of Disease and Pharmacologic Treatment on Human GABA Metabolism , below). The glutamate/glutamine and GABA/glutamine pathway may be distinguished using [2-^{13}C] glucose and [2-^{13}C] acetate as precursors as described below and in the section In Vivo MRS Studies of GABA Metabolism .

^{13}C NMR Studies of the Glutamate/Glutamine Cycle in Rat Cerebral Cortex

To determine the rate of glutamine synthesis, rats were studied under α-chloralose anesthesia in a 7-T modified Bruker

Biospec spectrometer. A small ^{13}C surface coil was used for transmission and reception. The spectroscopic volume was localized primarily to the motor and somatosensory cortices. The rats were infused with [1-^{13}C] glucose, and the time course of label incorporation into the C4 positions of glutamate and glutamine was measured. The time courses were fitted using differential equations describing the proposed model of glutamate/glutamine cycle. The rate of the neuronal TCA cycle as measured from label incorporation into the C4-glutamate was 0.46 ± 0.12 μmol/g-min [mean ± standard deviation (SD), *n* = 5]. The rate of glutamine synthesis (V_{gln}) was 0.21 ± 0.04 μmol/g-min (*n* = 5), which was nearly half the rate of the TCA cycle (25). These results indicate that glutamine synthesis is a major metabolic pathway in the rat cerebral cortex.

Validation of the Metabolic Model by Comparison of the Increase in the Rate of Glutamine Synthesis During Hyperammonemia with Independent Measures of Net Ammonia Uptake, CO_2 Uptake, and Glutamine Efflux

Elevated plasma ammonia increases the rate of the anaplerotic pathway of glutamine synthesis (62) in order to remove ammonia from the brain. The metabolic model predicts that under conditions of elevated plasma ammonia the increase in the rate glutamine synthesis is stoichiometrically coupled to the increase in the uptake of the anaplerotic substrates CO_2 and ammonia and the efflux of glutamine from the brain ($\Delta V_{gln} = \Delta V_{ana} = \Delta V_{efflux} = \Delta V_{CO2} = ½\Delta V_{NH4s}$). To test the ^{13}C MRS measurement, glutamine synthesis in rat cerebral cortex was measured under normal and elevated plasma ammonia concentrations. Rats were made hyperammonemic (0.35 ± 0.08 mM plasma ammonia vs. basal levels of 0.05 ± 0.01 mM) by a primed continuous infusion of ammonia and studied after 4 hours of hyperammonemia to ensure metabolic steady state. The neuronal TCA cycle rate was not significantly increased under these conditions relative to the control condition, which suggests that brain electrical activity and by inference the glutamate/glutamine cycle were not substantially altered. The rate of glutamine synthesis under hyperammonemic conditions increased by 0.11 ± 0.03 μmol/g-min relative to the rate under normal plasma ammonia levels. The increase in the rate of glutamine efflux (V_{efflux}) measured by AV difference under similar conditions was 0.10 μmol/g-min (67), in good agreement with ΔV_{gln}. Studies that have used ^{14}C isotope to measure the increase in V_{CO2} with hyperammonemia found a rate of ˜0.15 μmol/min-g (68), which is slightly higher than measured by ^{13}C MRS, possibly due to the need for additional incorporation of CO_2 to replace TCA cycle intermediates lost by oxidation. As described below, both AV difference and direct isotope incorporation measurements of ammonia fixation into glutamine under hyperammonemic conditions are also consistent with the predictions of the model. The agreement between the increase in V_{gln} determined by ^{13}C MRS and the increase measured by conventional methods in anaplerotic substrate utilization and glutamine efflux predicted by Eq. 2 provides strong experimental support for the ability to determine V_{cycle} under normal physiologic conditions.

^{15}N MRS Studies to Test the Relationship Between Anaplerotic Glutamine Synthesis and Ammonia Detoxification

^{15}N MRS is a useful method for both *in vitro* and *in vivo* study of cerebral glutamate/glutamine metabolism under hyperammonemic conditions based on the measurement of [5-^{15}N] glutamine and [2-^{15}N] glutamate/glutamine (69 ,70). Incorporation of ^{15}N labeled ammonia into the N5 position of glutamine may be analyzed to calculate the flux through glutamine synthetase. In the absence of label exchange, the rate of incorporation of labeled ammonia into the N2 position of glutamate + glutamine may be analyzed to calculate the rate of glutamate dehydrogenase.

The relationships in Eq. 2, which were used in the modeling of the ^{13}C MRS data to deconvolute ^{13}C labeling in C4-glutamine from neuronal glutamate and glial PDH, are based on mass balance and previous AV difference and ammonia trapping studies. To further test the relationship of Eq. 2 between the rate of ammonia uptake in the cerebral cortex (V_{NH4+}) and anaplerotic glutamine synthesis (V_{ana}), total ammonia uptake was calculated from the time course of the sum of ^{15}N labeled N5 glutamine and N2 glutamate + glutamine in rat cerebral cortex during infusion of ^{15}N ammonia. These were the only compounds into which appreciable ^{15}N label incorporation was observed, which agrees with previous findings that the major flows of ammonia in the brain involve these metabolites (62). The calculated V_{NH4+} from these data was 0.13 ± 0.02 μmol/g-min (*n* = 6). Based on the stoichiometric relationship of the model of $½\Delta V_{NH4} = \Delta V_{ana}$, a rate of anaplerotic glutamine formation of 0.065 ± 0.01 μmol/g-min was predicted. From this measurement an increase in the cerebral glutamine pool during the infusion of 0.065 μmol/min/g × 180 min = 11.7 μmol/g of glutamine was predicted. This calculation is in excellent agreement with the measured increase in glutamine concentration at the end of the study of 11.1 ± 0.4 μmol/g (36).

^{13}C MRS Determination of the Rate of the Glutamate/Glutamine Neurotransmitter Cycle Under Normal Physiologic Conditions

To determine the rate of the glutamate/glutamine cycle from a [1-^{13}C] glucose precursor under physiologic conditions,

Sibson et al. (25) measured V_{gln} and calculated V_{ana} using Eq. 2 and previously published measurements (62 ,71). The value of V_{ana} calculated in this manner ranged from 0.00 to 0.04 μmol/g-min. Comparison with the ^{13}C MRS measurement of V_{gln} of 0.21 ± 0.04 μmol/g-min, yields a V_{cycle} that is 80% to 90% of the rate of glutamine synthesis. A similar high percentage of V_{cycle} was calculated using measurements of the net incorporation of $^{14}CO_2$ into the cerebral cortex (72). The CO_2 measurement is coupled to total brain anaplerosis, which may be higher than anaplerosis used for net glutamine synthesis, and therefore represents the maximum estimate of this flux.

Validation of the Measurement of Glutamine Synthesis by Comparison of Rates Calculated from ^{15}N MRS and ^{13}C MRS Results

To obtain an independent measurement of V_{gln} and V_{ana}, ^{15}N MRS was used to measure the rate of ^{15}N-labeled ammonia incorporation into the N5 position of glutamine and the unresolved resonance of N2 glutamate plus glutamine (36). A mathematical analysis based on the model was used to derive V_{gln} from the MRS measurement of the time course of [5-^{15}N]glutamine and [2-^{15}N] glutamate + glutamine. The labeling in the first hour was almost exclusively within the N5 position of glutamine, which is consistent with the delayed onset of anaplerosis previously reported under these conditions (73) and previous measurements using ^{13}N and ^{15}N labeled ammonia (62 ,69). The low initial rate of anaplerosis allows the rates determined from the ^{15}N NMR study to be compared with the rates measured by ^{13}C NMR under normal physiologic conditions. The measured V_{gln} of 0.20 ± 0.06 μmol/g-min (mean ± SD, *n* = 6) from these studies (36) is in excellent agreement with the results from the ^{13}C NMR measurement of 0.21 ± 0.04 μmol/g-min (25).

Validation of the ^{13}C MRS Measurement of the Glutamate/Glutamine Cycle, and Assessment of Alternate Models of Neuronal/Glial Trafficking, Through Comparison of Results Using [1-^{13}C] Glucose, [2-^{13}C] Glucose, $^{15}NH4$, and [2-^{13}C] Acetate as Precursors

Several alternative models to the glutamate/glutamine cycle (Fig 25.1A ,Fig 25.1B) have been proposed. In one alternative model the ^{13}C labeling of glutamine represents an internal glial glutamate/glutamine cycle as opposed to trafficking between the neuron and glia. Label enters C4-glutamine from [1-^{13}C] glucose in this model through exchange in the glial cell between glutamate and glutamine catalyzed either through the reverse reaction of glutamine synthetase, or alternatively via glial phosphate activated glutaminase. Released neuronal glutamate in this model is taken up directly by the nerve terminal. In another alternate model, diagrammed in Fig. 25.1C , the glia releases α-ketoglutarate, or equivalently citrate or malate, to the neuron to replace the carbon skeleton of released glutamate (32 ,34 ,64). In support of this pathway, which is referred to here as the glutamate/α-ketoglutarate cycle, several TCA cycle intermediates including malate, α-ketoglutarate, and citrate are released from glia in cell culture and may be taken up by synaptosomes and cultured neurons (32 ,33 and 34).

The two pathways of glutamate trafficking shown in Fig. 25.1 cannot be distinguished on the basis of a ^{13}C MRS study using [1-^{13}C] glucose as the label source. As described above, [1-^{13}C] glucose will label both the glial and neuronal glutamate pools directly via pyruvate dehydrogenase. An alternative strategy is to use isotopic precursors that exclusively introduce label into the glia. Analysis of the flow of isotope from the glia into the neuronal glutamate pool yields the rate of total neuronal/glial glutamate trafficking. Comparison with the rate calculated using [1-^{13}C] glucose gives the fraction of neuronal/glutamate trafficking due to the glutamate/glutamine cycle (27 ,36).

The initial use with MRS of the strategy of glial selective precursors to calculate the fraction of glutamate trafficking due to the glutamate/glutamine cycle measurement was by Shen et al. (36), who calculated the relative fraction of the glutamate/glutamine cycle and glutamate/α-ketoglutarate cycle, using ^{15}N MRS measurement of the labeling in glutamine and glutamate from $^{15}NH4^+$. Under hyperammonemic conditions the rate of ^{15}N ammonia incorporation into the N5 and N2 position of glutamine is the same in the glutamate/α-ketoglutarate cycle because only the anaplerotic pathway of glutamine synthesis is present. In contrast, in the glutamate/glutamine cycle, there is additional incorporation of ^{15}N label into the N5 position of glutamine selectively in the glia (due to the localization of glutamine synthetase) due to the cycle. To distinguish these models, the endpoint ^{15}N enrichment of the N2 positions of glutamate and glutamine were calculated relative to the glutamine N5 position for each model using the N5 glutamine labeling curve as an input and compared with experimental values. As shown in Fig. 25.4 , the low ^{15}N fractional enrichment of the N2 position of glutamine and glutamate relative to glutamine N5 at the end of the study strongly supports the glutamate/glutamine cycle as the primary pathway of neuronal glutamate repletion.

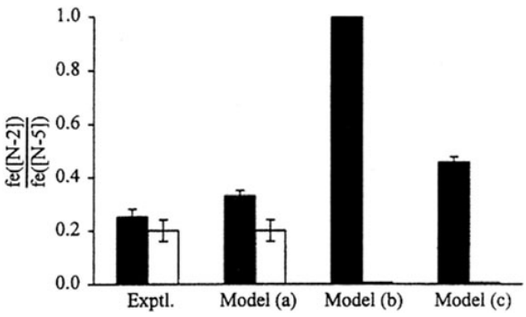


FIGURE 25.4. Calculated $^{15}N2/^{15}N5$ fractional enrichment ratios of glutamine and glutamate for three models of glial glutamine synthesis. Three models of neuronal glutamate completion were compared with experimental results in which the time course of [5-^{15}N] glutamine and [2-^{15}N] glutamine and glutamate were measured by ^{15}N nuclear magnetic resonance (NMR) in the cortex of a rat infused with ^{15}N-labeled ammonia at 7 T (36). The measured ratio at the end of the infusion is in excellent agreement with the ratio predicted if the glutamate/glutamine cycle is the major pathway of astrocytic repletion of released neuronal glutamate (model *a*, which is diagrammed in Fig 25.1A,Fig 25.1B). If instead the cycle was internal to the astrocyte the N2/N5 glutamine relative ^{15}N enrichment would be two times higher than measured and no labeling would have been observed in N2 glutamate (model *b*). If glutamate neurotransmitter repletion took place through the astrocytes providing the neurons with α-ketoglutarate (model *c*, which is diagrammed in Fig 25.1C), the rate of anaplerotic and total glutamine synthesis would be similar and the N5/N2 ratio of glutamine would be close to 1.0 as opposed to the measured ratio of 0.25. A similar labeling strategy has recently been used with [2-^{13}C] glucose and [2-^{13}C] acetate, both substrates that selectively label glutamine through glial specific pathways, and has put an upper limit of the rate of the α-ketoglutarate/glutamate cycle at 20% of the total rate of glutamate trafficking, the remainder being due to the glutamate/glutamine cycle (27,38).

An additional test of the glutamate/glutamine cycle model was recently performed using 2-^{13}C] glucose as an isotopic precursor (27). Label from [2-^{13}C] glucose enters the inner positions of glutamate and glutamine only through pyruvate incorporation into the TCA cycle by pyruvate carboxylase, which is localized to the glia (27 ,74). The initial flow of label from this precursor is into the glial TCA cycle intermediates, and then glial glutamate and glutamine (27).

Subsequently, neuronal/glial cycling moves the label to the neuron where it labels the large glutamate pool. The labeling measured in the glutamate pool is the sum of all trafficking pathways from the glia. In contrast the rate of labeling of glutamine from a [1-^{13}C] glucose precursor is a measure of the glutamate/glutamine cycle. *In vivo* and *in vitro* ^{13}C MRS at 7 T was recently used to measure the labeling time course of glutamate and glutamine in the cerebral cortex of rats under hyperammonemic and normoammonemic conditions during infusion of either [1-^{13}C] or [2-^{13}C] glucose (27). The rate calculated for the neuronal/glial glutamate cycle was similar, with both labels indicating that the glutamate/glutamine cycle is the major pathway of neuronal/glial glutamate trafficking accounting for between 80% and 100% of total glutamate trafficking. A similar conclusion was recently reported for human cerebral cortex using [2-^{13}C] acetate as a precursor (38), which selectively introduces label into glutamate and glutamine through glial pyruvate dehydrogenase.

The Effect of Glutamate Oxidation on the Glutamate/Glutamine Cycle Measurement and Estimates of Its Rate In Vivo

An alternate pathway of neuronal/glial glutamate trafficking is glial glutamate oxidation (10 ,63 ,64). In this pathway glutamate taken up by the glial cell is transaminated into α-ketoglutarate and enters the TCA cycle. Reactions in the TCA cycle convert α-ketoglutarate to oxaloacetate, which is then converted to pyruvate by the action of malic enzyme. The pyruvate formed from glutamate is oxidized in the TCA cycle through the action of pyruvate dehydrogenase. Glutamate lost to the brain by this pathway is then replaced by anaplerosis through pyruvate carboxylase. Evidence of this pathway is derived primarily from isolated cell cultures. It has been proposed that the fraction of glutamate going through this pathway increases with brain electrical activity (64).

The major effect of the glutamate oxidation pathway on the MRS measurement of the glutamate/glutamine cycle is to cause the fraction of glutamine synthesis of net anaplerosis to be overestimated and V_{cycle} to be consequently underestimated, because the labeling of the internal positions of glutamine from the two pathways from [1-^{13}C] and [2-^{13}C] glucose is similar. The unambiguous *in vivo* determination of glutamate oxidation requires the ^{13}C label flow from glutamate to pyruvate to be measured (10 ,63). This measurement is complicated by other metabolic pathways that produce similar labeling patterns, including scrambling of isotopic labeling into other positions of glucose in the liver (23 ,27 ,63 ,75), and as a consequence glutamate oxidation has not been definitively demonstrated *in vivo* under normal physiologic conditions. Suggestive evidence of this pathway is the finding in several studies that the rate of anaplerosis under normal ammonia conditions calculated from labeling of glutamine by ^{13}C labeled glucose is approximately two to three times higher than that predicted from measurements of brain glutamine efflux (27). An alternate possibility is that rather than glutamate oxidation this extra labeling reflects cycling between oxaloacetate and pyruvate to generate reduced nicotinamide adenine dinucleotide phosphate (NADPH) reducing equivalents in the glia, a pathway that has been shown to be highly active in the liver (75).

Validation of the ^{13}C MRS Glutamate/Glutamine Cycle Measurement by Correlation with Brain Electrical Activity

If the ^{13}C labeling measured in glutamine by ^{13}C MRS is due to the glutamate/glutamine cycle, then the calculated rate of this pathway should correlate with brain electrical activity. Neuronal glutamate release is known to increase with neuronal depolarization associated with action potentials.

To test this prediction, ^{13}C MRS was used to measure the rates of neuronal glucose oxidation and the glutamate/glutamine cycle in the rat cerebral cortex at three levels of cortical electrical activity: isoelectric EEG induced by high-dose pentobarbital anesthesia, and at two milder levels of anesthesia (26). During isoelectric conditions, under which minimal glutamate release takes place, almost no glutamine synthesis was measured, consistent with the conclusion that the ^{13}C MRS measurement of glutamine synthesis primarily reflects the glutamate/glutamine cycle. Above isoelectricity, the rates of the glutamate/glutamine cycle and neuronal glucose oxidation both increased with higher electrical activity. The relationship measured in this study between the rate of the glutamate/glutamine cycle and neuronal glucose oxidation is described below (see Determination of the *In Vivo* Coupling Between the Rate of the Glutamate/Glutamine Neurotransmitter Cycle and Neuronal Glucose Oxidation).

^{13}C MRS Measurements of the Rate of the Glutamate/Glutamine Cycle in Human Cerebral Cortex

In 1994 we first demonstrated that *in vivo* ^{13}C NMR may be used to measure the rate of glutamine labeling (12 ,18) from [1-^{13}C] glucose in human occipital/parietal cortex. These studies showed clearly that glutamine is labeled rapidly from [1-^{13}C] glucose in the human cerebral cortex. However, the rate of the glutamate/glutamine cycle was not uniquely determined in the initial experiments due to the inability to distinguish the glutamate/glutamine cycle from other sources of glutamine labeling. To determine whether there is a similar high rate of the glutamate/glutamine cycle in human cerebral cortex as in the rat, we (29) and Gruetter and co-workers (13 ,35) have determined this rate from ^{13}C MRS measurements in the human occipital/parietal lobe.

A time course from the study of Shen and co-workers (29) showing the rapid labeling of C4-glutamine and C4-glutamate from [1-^{13}C] glucose in a single subject is shown in Fig. 25.5 . A best fit of the metabolic model is plotted through the data. A lag is clearly shown in the labeling of C4-glutamine relative to C4-glutamate, which is consistent with the large neuronal glutamate pool being the main precursor for glutamine synthesis. The combination of the metabolic model validated in the rodent and improved MRS sensitivity allowed the rate of the glutamate/glutamine cycle, the neuronal TCA cycle, the glial TCA cycle, and anaplerotic glutamine synthesis to be calculated from the ^{13}C MRS data. The analysis gave a total TCA cycle rate of 0.77 ± 0.05 μmol/min/g (mean ± SD, n = 6), a neuronal TCA cycle rate of 0.71 ± 0.02 μmol/min/g, a glial TCA cycle rate of 0.06 ± 0.02 μmol/min/g, a glutamate-glutamine cycle rate of 0.32 ± 0.04 μmol/min/g (mean ± SD, n = 6), an anaplerotic glutamine synthesis rate of 0.04 ± 0.02, and a glucose oxidation rate of 0.39 ± 0.03 μmol/min/g (mean ± SD, n = 6). In agreement with studies in rat cortex, the glutamate/glutamine cycle is a major metabolic flux in the resting human brain with a rate approximately 80% of the rate of total glucose oxidation.

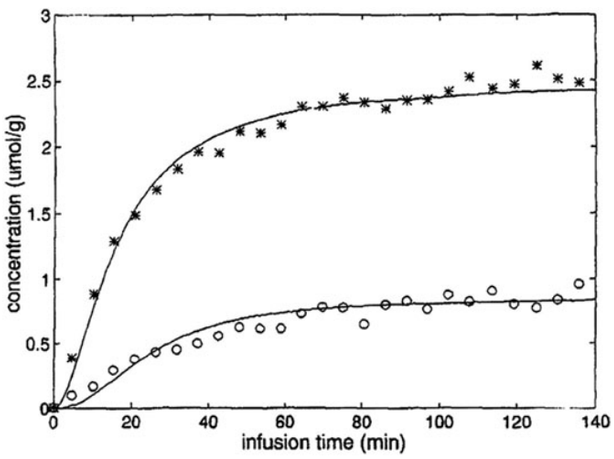


FIGURE 25.5. *In vivo* ^{13}C NMR time course of the human occipital/parietal lobe: the time course from one subject of the concentrations of [4-^{13}C] glutamate and [4-^{13}C]glutamine during a [1-^{13}C] glucose infusion, and the best fit of the two-compartment model to these data. At time 0 on the plot an intravenous infusion of [1-^{13}C] glucose was started. The model is shown to provide an excellent fit to the data. The rise of [4-^{13}C] glutamine is clearly seen to lag the labeling of [4-^{13}C] glutamate, consistent with neuronal glutamate being the main precursor for glutamine synthesis via the glutamate/GABA/glutamine cycle. *, glutamate; ○, glutamine. (From Shen J, Petersen KF, Behar KL, et al. Determination of the rate of the glutamate-glutamine cycle in the human brain by *in vivo* ^{13}C NMR. *Proc Natl Acad Sci USA* 1999;96:8235–8240, with permission.)

In a study performed at 4 T, Gruetter and co-workers (13) measured rapid labeling of glutamate and glutamine from infused [1-^{13}C] glucose using ^{13}C MRS. A high rate of the glutamate/glutamine cycle was measured using a two-compartment model, similar to the model used by Shen and co-workers (29). The improved spectroscopic resolution provided by the higher field strength at 4 T allowed the additional positions of the C2 and C3 resonances of aspartate, glutamate, and glutamine to be incorporated into the modeling. More recently Gruetter and co-workers (35) studied six subjects using localized ^{13}C MRS measurements of a 45-mL volume in the occipital lobe. The main differences from the rates derived from the Shen et al. (29) study are a higher rate of anaplerosis, approximately 25% of total glutamine synthesis as opposed to 11% in the Shen et al. study and a somewhat lower rate of the neuronal TCA cycle of 0.62 ± 0.05 μmol/min/g. The lower calculated neuronal TCA cycle rate was due to a lower rate of neuronal mitochondrial α-ketoglurate/glutamate exchange calculated from the data than in a previous study by Mason and co-workers (18). The lower exchange rate was due to the assignment of a higher concentration of aspartate in the glutamatergic

neuron and glutamate in the astrocyte in the metabolic model of Gruetter and co-workers. The higher anaplerosis rate also reflects differences in which the isotopic data was modeled. In the Shen et al. study anaplerosis was calculated primarily from the labeling kinetics of C4-glutamine and C4-glutamate, whereas in the Gruetter et al. (35) study it was calculated primarily from the measurement of the time course of the differential ^{13}C labeling of the C2, C3, and C4 glutamate and glutamine resonances. Both approaches suffer from needing to deconvolute ^{13}C label entering these carbon positions from pyruvate dehydrogenase from the label entering via pyruvate carboxylase. In the future these differences should be reconcilable by using labeling strategies such as [2-^{13}C] glucose, which labels glutamate and glutamine internal positions only by pyruvate carboxylase. If the anaplerotic pathway is due to glutamate oxidation (see Validation of the 13C MRS Glutamate/Glutamine Cycle Measurement by Correlation with Brain Electrical Activity , above) as opposed to ammonia detoxification, the rate of the glutamate/glutamine cycle reported in both studies is an underestimate by the calculated rate of anaplerosis. The differences in the anaplerotic flux calculated in these studies should not obscure the major point of agreement—that the glutamate/glutamine cycle is major metabolic pathway with a rate accounting for between 60% and 80% of total glucose oxidation in the cerebral cortex.

Cellular and Molecular Evidence that Astroglia have a Major Role in the Uptake of Glutamate Released from Neurons

The high rate of the glutamate/glutamine cycle indicates that astroglial uptake of glutamate and GABA plays a key role in maintaining the low extracellular levels of these neurotransmitters needed for proper receptor-mediated functions. There is considerable evidence from several lines of research that support this conclusion. Overstimulation of glutamate receptors can lead to excitotoxicity (76 ,77). Studies of glutamatergic synapses have shown them to be closely surrounded by glial end processes possessing high densities of glutamate transporters (78). Glutamate and GABA transporters are sodium dependent and electrogenic and are present on both neurons and glia (58 ,78 ,79 and 80). Glutamate transporters have an affinity, K_m, of 1 to 3 μM (80), which is in the range of normal estimated ECF glutamate concentrations. Immunohistochemical studies have showed that the glutamate transporters GLT-1 and GLAST (glutamate astrocytic transporter) are localized primarily in astrocytes (48 ,81 ,82 and 83), whereas EAAC1 is found on neurons (51). Antisense oligonucleotides directed against the astrocytic glutamate transporters GLT-1 or GLAST *in vivo* results in elevated ECF glutamate *in vivo* and excitotoxicity (84 ,85). The majority of glutamate uptake after its release appears to be either postsynaptic or astroglial (86 ,87), although an electrophysiologic study of the hippocampal slice suggests that astroglial uptake dominates (88).

Summary and Remaining Questions

MRS allows the glutamate/glutamine cycle to be measured from the labeling of glutamate and glutamine by ^{13}C and ^{15}N labeled precursors. The major complications in determining the rate of the glutamate/glutamine cycle from isotopic measurements are separating the labeling of glutamine from the glutamate/glutamine cycle from alternate pathways of glutamine synthesis and isotopic exchange, and distinguishing different pathways of neuronal/glial glutamate trafficking. To overcome these obstacles the metabolic modeling of the glutamate/glutamine cycle has been extended to include ammonia detoxification, alternate pathways of glutamate trafficking, and glutamate oxidation (27). The MRS rate measurement has been validated by several strategies including (a) comparison of the rate of glutamine synthesis measured under different ammonia levels with measurements of anaplerotic substrates by AV difference and isotopic trapping methods (25); (b) comparison of the rates of glutamine synthesis and the glutamate/glutamine cycle calculated from the isotopic labeling from [1-^{13}C] glucose, [2-^{13}C] glucose, ^{15}N ammonia, and [2-^{13}C] acetate (27 ,35 ,36 ,39); and (c) measurement of the rate of the glutamate/glutamine cycle as a function of brain electrical activity (26 ,37). The results of these studies indicate that the glutamate/glutamine cycle is the major pathway of glutamine synthesis and neuronal/glial glutamate trafficking under normal conditions, with a rate similar to the rate of neuronal glucose oxidation under conditions of high electrical activity. Measurements in awake nonstimulated human cerebral cortex have found that the rate of the glutamate/glutamine cycle is between 60% and 80% of total glucose oxidative metabolism (29 ,35).

Objections have been raised to the MRS measurement of the glutamate/glutamine cycle for having neglected alternate pathways of glutamate trafficking and the need for comparison with direct measurements of neuronal glutamate release. As described above, isotopic strategies have been developed to assess these pathways and under physiologic conditions they were found to account for less than 20% of glutamate trafficking. However, under pathologic conditions such as seizure the rate of these pathways may be much higher. Glutamate oxidation may have a significant contribution to total neuronal/glial glutamate trafficking. However, the unambiguous *in vivo* measurement of glutamate oxidation will require strategies for eliminating isotopic labeling from other pathways. Although direct measurement of bulk neuronal release of glutamate for comparison with ^{13}C MRS is presently not possible, advances in molecular and cellular methods for studying glutamate transport indicate that neurotransmission is the major, if not exclusive, pathway of glutamate release from glutamatergic neurons and the vast

majority of this flux is taken up by astroglia in the cerebral cortex. Correlation of the MRS glutamate/glutamine cycle with indirect measures of neuronal glutamate release such as microdialysis and nerve terminal labeling would be highly desirable, as would further studies better defining the relevant pool sizes and enzyme distribution in glia and glutamatergic neurons, particularly in regions other than the cerebral cortex.

DETERMINATION OF THE *IN VIVO* COUPLING BETWEEN THE RATE OF THE GLUTAMATE/GLUTAMINE NEUROTRANSMITTER CYCLE AND NEURONAL GLUCOSE OXIDATION

Part of "25 - Magnetic Resonance Spectroscopy Studies of the Glutamate and Gaba Neurotransmitter Cycles and Functional Neuroenergetics "

This section presents evidence from MRS and other studies for a model of the coupling between the glutamate/glutamine cycle and glial glucose uptake and subsequent neuronal oxidation. The model is based on work in cellular systems primarily by Magistretti and co-workers (90) and recent findings, using ^{13}C MRS in rat cortex, that the glutamate/glutamine cycle (a) increases in rate with increasing brain electrical activity in a near 1:1 stoichiometry with neuronal glucose oxidation (26), and (b) is 60% to 80% of the rate of total glucose oxidation in the awake nonstimulated cerebral cortex (13 ,26 ,29 ,35 ,37). Several comprehensive reviews of the evidence from molecular and cellular studies supporting glial localization of glucose uptake related to functional neuroenergetics have been published by Magistretti and co-workers (52 ,89) and are not duplicated here. The focus of this section is on the evidence from *in vivo* studies that support the model and key tests that remain to be performed.

Determination by ^{13}C MRS of the Relationship Between the Glutamate/Glutamine Cycle and Neuronal Oxidative Glucose Consumption

To determine the relationship between the glutamate/glutamine cycle and cerebral cortex neuroenergetics, ^{13}C MRS was used to measure the rate of neuronal glucose oxidation and the glutamate/glutamine cycle in rat cortex under conditions of isoelectric EEG induced by high-dose pentobarbital anesthesia, and at two milder levels of anesthesia (26). The rate of neuronal glucose oxidation and the glutamate/glutamine cycle was calculated using a two-compartment metabolic model from the isotopic turnover of C4-glutamate and C4-glutamine. Under isoelectric conditions, at which minimal glutamate release takes place, almost no glutamine synthesis was measured, consistent with the conclusion that the ^{13}C MRS measurement of glutamine synthesis primarily reflects the glutamate/glutamine cycle. Above isoelectricity, the rates of the glutamate/glutamine cycle and neuronal glucose oxidation both increased with increasing electrical activity. The results, shown in Fig 25.6 , indicate an approximately 1:1 relationship between the increase in the rates of the glutamate/glutamine cycle and neuronal glucose oxidation with brain activity. Under the highest cortical activity studied, the glutamate/glutamine cycle rate was approximately 80% of the rate of neuronal glucose oxidation. A similar ratio of the rate of the glutamate/glutamine cycle to the rate of neuronal glucose oxidation has been reported for measurements of awake nonstimulated human cerebral cortex (13 ,29 ,35).

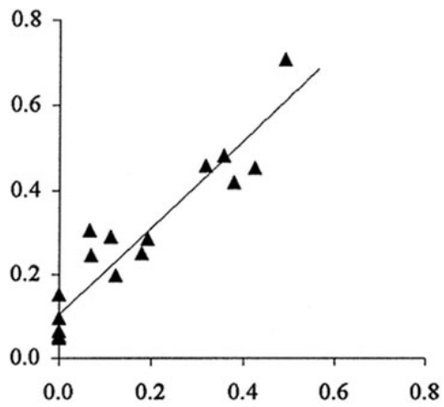


FIGURE 25.6. An approximately 1:1 correlation between the rate of oxidative glucose consumption and the rate of the glutamate glutamine cycle. The rate of neuronal glucose oxidation ($CMR_{glc(ox)}$) and the glutamate/glutamine cycle (V_{cycle}) was measured by ^{13}C MRS at 7 T in the rat somatosensory cortex at different levels of cortical activity induced by anesthesia. A significant positive correlation (p <.001) was found between $CMR_{glc(ox)}$ and V_{cycle}. The regression line shown is $y = 1.04x + 0.10$ with a Pearson product-moment correlation coefficient, r, of 0.94. (From Sibson NR, Dhankhar A, Mason GF, et al. Stoichiometric coupling of brain glucose metabolism and glutamatergic neuronal activity. *Proc Natl Acad Sci USA* 1998;95:316–321, with permission.)

A Model for the Stoichiometric Coupling of the Glutamate/Glutamine Cycle to Neuronal Glucose Oxidation

Figure 25.7 shows a model that provides a mechanistic explanation for the observed ratio of the rates of the glutamate/glutamine cycle to neuronal glucose oxidation (26 ,37 ,90). The model is an extension of the model proposed by Magistretti and co-workers that nonoxidative glial glycolysis is coupled to glutamate uptake due to the preference of the glia to use glycolytic adenosine triphosphate (ATP) to pump out the cotransported three Na^+ ions (52 ,90 ,91). The pyruvate and lactate formed by glial glycolysis would then be transported to the neuron where it is oxidized. Prior to the

in vivo ^{13}C MRS studies, the evidence for the model was primarily from enzyme localization studies and isolated cell studies (see refs. 89 and 90 for reviews of these studies). Both lines of evidence of localization have been criticized based on the presence of the enzymes required for glucose transport and glycolysis in the neurons, and the strong dependence of glutamate-stimulated glial glucose metabolism on cell culture conditions (92).

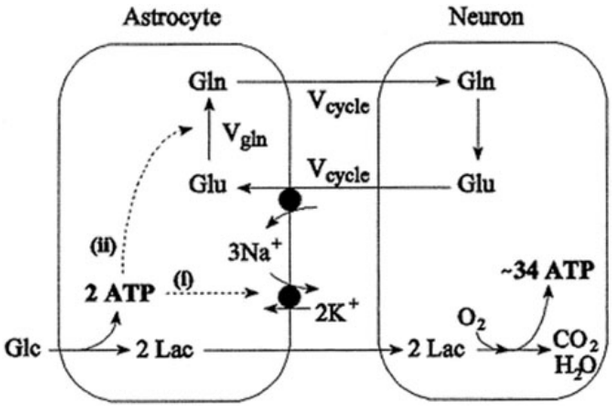


FIGURE 25.7. A metabolic model coupling the glutamate/glutamine cycle to oxidative glucose consumption. In this model the two molecules of adenosine triphosphate (ATP) required by the astrocyte to take up one molecule of glutamate (Glu) and convert it through glutamine synthetase to glutamine (Gln) are provided by nonoxidative glycolysis of one molecule of glucose (Glc). The lactate produced by nonoxidative glycolysis is then released from the astrocyte and taken up by the neuron for oxidative glycolysis. Glc, glucose; Lac, lactate; V_{gln}, rate of glutamine synthesis; V_{cycle}, rate of the glutamate/glutamine cycle. (From Sibson NR, Dhankhar A, Mason GF, et al. Stoichiometric coupling of brain glucose metabolism and glutamatergic neuronal activity. *Proc Natl Acad Sci USA* 1998;95:316–321, with permission.)

Comparison of the In Vivo ^{13}C MRS Results with the Stoichiometry Predicted by the Model

The ambiguities in the determination of the relative rates of metabolic pathways from enzymatic localization and measurements of isolated cells are not unexpected. Metabolic control analysis has shown that the total activity of an enzyme within a metabolic pathway does not determine the flux through the pathway (93). Extrapolation to *in vivo* rates from studies of cell cultures is complicated by the difficulty of reproducing the complex cellular interactions that occur *in vivo* (64). To compare the results of the *in vivo* measurement with the predictions of the model, Sibson et al. (26) calculated the stoichiometric relationship between the glutamate/glutamine cycle and neuronal oxidative glucose consumption. Glutamate is cotransported into the glia with two to three Na^+ ions, with one K^+ ion countertransported (60 ,78 ,94). Transport of three Na^+ ions out of the glia by the Na^+/K^+ adenosine triphosphatase (ATPase) on the glial end process membrane requires approximately one ATP molecule (91). Synthesis of glutamine from glutamate through glutamine synthetase requires one ATP molecule per glutamine molecule synthesized (53). If the ATP for this process were derived entirely from glycolysis, then a 1:1 stoichiometry is predicted between glial nonoxidative glucose consumption and the glutamate/glutamine neurotransmitter cycle. Provided that the lactate formed is released to the neurons for oxidation, then this predicted stoichiometry is in excellent agreement with the *in vivo* ^{13}C MRS findings.

If the model is correct, it may account for a substantial fraction of total glucose consumption in the awake nonstimulated cerebral cortex. Based on the measurements of the rate of the glutamate/glutamine cycle and total glucose oxidation in human cerebral cortex (13 ,29), between 60% and 80% of total brain glucose oxidation may be accounted for by this mechanism.

Additional In Vivo Evidence of the Model

The model (90) has been criticized for not leaving room for other energy consuming processes in the cerebral cortex. Within the error of the measurements, approximately 10% of total glucose oxidation is available for other neuronal systems such as dopaminergic or serotoninergic. The finding of the large majority of glucose oxidation in the cerebral cortex being associated with glutamatergic and GABAergic neurons and their surrounding astrocytes is not surprising, because cell staining studies have shown that the vast majority of synapses and neurons in the cerebral cortex are either glutamatergic or GABAergic (4).

A prediction of the model is that a large fraction of glucose uptake and phosphorylation is localized in the cerebral cortex to the glial end sheaths surrounding the synapses of glutamatergic neurons. In agreement with this prediction studies using ^{14}C-deoxyglucose autoradiography indicate that the majority of brain glucose uptake is used to support synaptic activity. Increased glucose uptake in response to functional stimulation in peripheral neurons and in cortex is primarily localized in dendritic and nerve terminal cortical layers (where there are associated glial end processes) and not in layers associated with cell bodies (1 ,95 ,96 and 97).

The rapid incorporation of ^{13}C label into glutamine by the glutamate/glutamine cycle indicates that the vesicular glutamate pool is rapidly turning over and is in dynamic equilibrium with cytosolic glutamate. This conclusion is in contradiction to the traditional view that the small vesicular pool is metabolically isolated from cellular glutamate metabolism (60 ,61). However, these studies were performed in cellular and tissue preparations, which have a low rate of synaptic metabolism relative to intact cerebral cortex. In support of this conclusion Conti and Minelli (42) showed that inhibition of PAG, which is enriched in nerve terminals (55)

and has been proposed to primarily replete the vesicular pool of glutamate (34), results in a similar rapid depletion of both synaptic and whole cell glutamate in the rat cerebral cortex.

Further support for the coupling between glial glycolysis and the glutamate/glutamine cycle is provided from studies performed looking at glutamine synthesis in mice in which the astrocytic TCA cycle was inhibited by injection with fluoroacetate (28). In these studies mice were given fluoroacetate and injected with a combination of [1,2-^{13}C] acetate and [1-^{13}C] glucose. From measurements of the isotopomer distribution in glutamate and glutamine, the labeling from glucose and acetate was distinguished. The labeling from acetate in glutamate and glutamine was greatly reduced by fluoroacetate administration, which the authors interpreted as resulting from the near-complete inhibition of the glial TCA cycle. Despite this inhibition, there was still a substantial amount of glutamine labeling from [1-^{13}C] glucose, approximately one-third to one-half the labeling found in the control mice. The only mechanism by which this labeling of glutamine from glucose could occur is the glutamate/glutamine cycle, because glutamate labeling in the astrocyte from glucose was completely blocked. The ability to maintain a high glutamate/glutamine cycle flux, despite the near-complete inhibition of glial mitochondrial ATP generation, has been interpreted by Bachelard (98) as supporting the importance of the glutamate/glutamine cycle as well as the potential coupling to glycolytic ATP production: “This singlet labeling of the C4 of glutamine, which can only be derived from [1-^{13}C] glucose metabolism in the neurones, also quite clearly demonstrates that even though the glial TCA cycle is blocked by the toxin, the glia are still capable of participating in the glutamate-glutamine cycle, taking up glutamate from the neurones and converting it to glutamine.”

Another testable prediction of the model is that if another substrate is supplied for neuronal oxidation, the decrease in glucose oxidation will be greater than the decrease in glucose consumption, due to the remaining need for glycolytic ATP to fuel the clearance of glutamate. Consistent with this prediction, an AV difference study of the anesthetized rat brain found that infusion of ß-hydroxybutyrate, which is an alternate fuel for brain oxidative ATP production, led to a two- to threefold greater decrease in glucose oxidation than in glucose consumption, with the difference accounted for by a large increase in the efflux of lactate from the brain (99). Consistent with this finding, Pan et al. (100) measured an increase in brain lactate in 3-day-fasted human subjects with elevated plasma ketone concentrations,

Summary and Remaining Questions

The linear relationship and stoichiometry found using ^{13}C MRS of the rates of the glutamate/glutamine cycle and neuronal glucose oxidation support a direct mechanistic coupling between the glutamate/glutamine cycle and glial glucose uptake. This mechanism may account for between 60% and 80% of the rate of total glucose oxidation in awake nonstimulated human cerebral cortex, and possibly an even larger fraction of the increment in glucose oxidation with stimulation (5 ,90). However, there are alternate potential explanations for the *in vivo* results that need to be tested. Most importantly, it will be necessary to devise strategies for directly distinguishing glial glucose uptake from neuronal glucose uptake and phosphorylation in the intact cerebral cortex. In addition, the stoichiometry between neuronal glucose oxidation and the glutamate/glutamine cycle remains to be measured under conditions of sensory stimulation, and in different brain regions.

IN VIVO MRS STUDIES OF GABA METABOLISM AND THE EFFECTS OF DISEASE AND PHARMACOLOGIC TREATMENT ON HUMAN GABA METABOLISM

Part of "25 - Magnetic Resonance Spectroscopy Studies of the Glutamate and Gaba Neurotransmitter Cycles and Functional Neuroenergetics "

GABA is the major inhibitory neurotransmitter in the cerebral cortex (46 ,47). It is synthesized from glutamate in specialized cells called GABAergic neurons. The release of GABA by a GABAergic neuron inhibits the electrical activity of adjacent neurons. Extensive studies in animals and isolated brain cells and slices have shown that GABAergic function is altered in a variety of models of neurologic and psychiatric disease (46 ,101 ,102). Several antiepileptic and psychiatric drugs are targeted at the GABAergic system. GABA is overlapped in the *in vivo* ^{1}H MRS spectrum by the more intense resonances of macromolecules (103), glutathione, and creatine. The development of ^{1}H MRS spectral editing of GABA in animals and humans (6 ,66 ,104 ,105) has provided a new window on studying GABA metabolism and GABAergic function in animals and humans. Several of the main findings using MRS to study alterations in GABA metabolism in disease and the effect of pharmacologic treatment on GABA metabolism are reviewed below.

MRS Studies of the Effect of the Antiepileptic Drug Vigabatrin on GABA Metabolism

Vigabatrin irreversibly inhibits the enzyme GABA transaminase (GABA-T). GABA-T catalyzes the breakdown of GABA in GABAergic neurons and in astrocytes. By inhibiting GABA-T, the drug leads to an elevation in GABA concentration. The ability of ^{1}H MRS editing to measure GABA elevated by GABA-T inhibitors was first demonstrated in the rat brain (106 ,107). Subsequent MRS editing studies of vigabatrin action on patients have made several new observations relevant to optimum administration of the drug including (a) chronic dosing above 3 g per day

fails to additionally increase GABA concentration (108), (b) GABA concentration reaches a maximum level within 2 hours of initial drug administration (109), (c) the effectiveness of vigabatrin in controlling seizures depends on elevating GABA concentration above the mean level found in nonepileptic subjects (108), (d) GABA concentration is increased by over two times the predrug concentration within 2 hours of an acute dose of 3 g of vigabatrin and remains elevated over 48 hours after a single dose of vigabatrin (109), and (5) there is no down-regulation of GABA-A receptors during chronic dosing with vigabatrin (see refs. 110 and 111 for a review of these studies). Figure 25.8 shows two edited spectra of total GABA obtained from the visual cortex before and after chronic vigabatrin treatment of a patient of epilepsy (108 ,112).

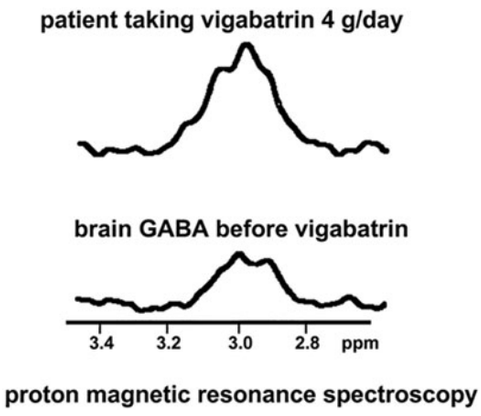


FIGURE 25.8. *In vivo* ^{1}H MRS spectra of total edited GABA from the occipital lobe of a patient with epilepsy before and after treatment with vigabatrin. The ^{1}H MRS spectra were obtained using spectral editing (112) from a 14-cm^{3} volume centered on the midline in the visual cortex. Chronic treatment with vigabatrin led to an over twofold increase in the concentration of total edited GABA, which is the sum of GABA and homocarnosine.

Cortical GABA Synthesis Is Reduced Following Prolonged GABA-Transaminase Inhibition

An initial finding in ^{1}H MRS studies of vigabatrin action in epilepsy patients was that the drug failed to raise GABA above a dose of 3 g/day (108). The enzymatic mechanisms controlling GABA levels *in vivo* are complex, involving short-term regulation of GAD by modulators (e.g., ATP, P_i) that affects the binding of cofactor pyridoxal phosphate, longer-term regulation involving enzyme protein levels, availability of glutamate precursors and their pathways (e.g., GABA-glutamine cycling), and postsynaptic and astroglial catabolic pathways. GAD exists as two major isoforms (GAD_{67} and GAD_{65}) in the brain; each is the product of separate genes (113 ,114) and each has distinct kinetic properties (114 ,115). GAD_{67} is distributed throughout the cytoplasm of GABAergic neurons, whereas GAD_{65} is associated with synaptic terminals. Recently, it was shown that the 67-kd isoform of GAD protein is reduced in response to elevated levels of GABA *in vitro* and *in vivo* (116 ,117). Differential control of the GAD isoforms suggests that they may mediate different fluxes. To investigate the effects of elevated GABA on GABA synthesis and quantitatively assess the role of the GAD isoforms in GABA synthesis, rates of turnover of cortical glutamate and GABA were determined in anesthetized rats during an infusion of [1-^{13}C] glucose after administration of the GABA-transaminase inhibitor vigabatrin (500 mg/kg, i.p.), to increase GABA levels (24). GABA concentration was increased twofold at 24 hours. Tricarboxylic acid cycle flux was not affected by vigabatrin treatment compared to nontreated rats despite the increased GABA level. An analysis of the turnover data revealed a ~70% decrease in the rate of GABA synthesis following vigabatrin-treatment (control, 0.14 μmol/g/min; vigabatrin-treated, 0.04 μmol/g/min). The reduction in GABA synthesis concomitant with the selective inhibition of GAD_{67} suggests that GAD_{67} accounts for the major fraction of GABA synthesis in the rat cerebral cortex under anesthetized nonstimulated conditions. This conclusion is supported by studies of GAD_{67} and GAD_{65} in knockout mice (118 ,119), which have found an order of magnitude greater reduction in GABA concentration in the mice with a GAD_{67} knockout. The isoform composition of human brain is presently unknown.

Elevation of GABA by Other Antiepileptic Drugs

^{1}H MRS studies have found that several antiepileptic drugs, with no known metabolic mechanisms of action on the GABAergic system, also lead to a rapid elevation of GABA concentration. These drugs include GABApentin, topirimate, and lamotragine (120 ,121 ,122 ,123 and 124). The elevation of GABA may be an important mechanism in the effectiveness of these medications as antiepileptic compounds. In addition, these findings provide evidence that the regulation of GABA metabolism is tightly integrated with the regulation of GABAergic function. Figure 25.9 shows a time course of the elevation of GABA after administration of an acute dosage of topirimate. As with vigabatrin, the concentration of GABA reaches a maximum of two times the predrug concentration within 2 hours of drug administration. A similar rapid elevation of GABA was measured after an acute dosage of GABApentin (122 ,124). The ability of ^{1}H MRS to track the response of the GABAergic system to pharmacologic therapy may potentially be useful for developing optimum dosing strategies.

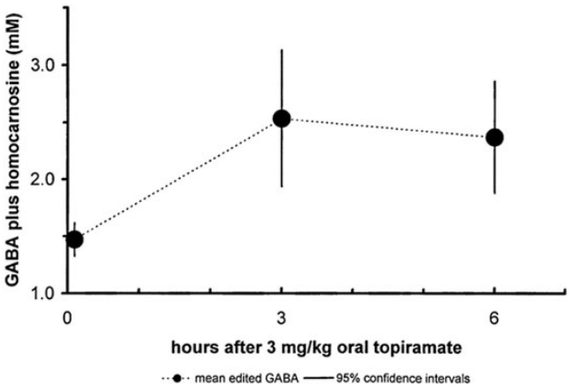


FIGURE 25.9. A time course of GABA plus homocarnosine concentration after administration of topiramate. Topiramate at 3 mg/kg was administered to six volunteers without epilepsy. ^{1}H MRS spectral editing measurements were performed at 4 T (123). GABA levels were measured to peak within 3 hours after administration of topiramate at approximately two times the predrug levels. A study looking at the acute effect of topiramate on GABA levels of epileptic patients found that the increase was almost entirely due to GABA (121). The GABA plus homocarnosine concentrations are normalized to a creatine concentration of 9 μmol/g for comparison with measurements by Petroff and co-workers (111), and the predrug concentrations are in excellent agreement. (Redrawn from Kuzniecky RI, Hetherington HP, Ho S, et al. Topiramate increases cerebral GABA in healthy humans. *Neurology* 1998;51:627–629.)

Effects of Neurologic and Psychiatric Disease on GABA Concentration

Impaired GABAergic function has been implicated in the etiology of epilepsy (101). Consistent with this proposal, ^{1}H MRS editing studies have found decreased GABA in adult epilepsy (111 ,125) and pediatric epilepsy (126). In epilepsy the release of cytosolic GABA has been proposed as an important mechanisms for seizure suppression (127 ,128). The finding that a low GABA concentration was strongly associated with poor seizure control in epilepsy supports this proposal. Further support for a role of cytosolic GABA concentration in inhibiting cortical excitability comes from the finding that the improvement in seizure control of subjects administered vigabatrin chronically depended on the elevation of cerebral cortex GABA concentration (108).

In addition to epilepsy, reduced GABA concentration has been found in unipolar depression (129), alcohol withdrawal, and hepatic encephalopathy (130). These disorders are associated with an alteration in inhibitory GABAergic function. The finding of low GABA associated with these disorders is additional evidence that the brain metabolic GABA pool has an important role in GABAergic function. The finding in unipolar depression appears paradoxical because the condition is not associated with enhanced cortical excitability. A potential explanation of this finding is that the low GABA concentration in this disease is a compensation for the reduction in excitatory glutamatergic activity (129).

Parsing the Edited GABA Resonance

The edited GABA resonance consists of GABA and the GABA derivative homocarnosine, which is a condensation product of GABA and histidine. Homocarnosine is a neuromodulator present in a specific subclass of GABAergic neurons in the primate brain. Short TE ^{1}H MRS with macromolecule suppression may be used to measure the homocarnosine histidine proton resonances in the downfield region of the short TE spectrum (103 ,112). Combining the homocarnosine measurement and the total GABA editing measurements has allowed the separate measurement of homocarnosine and GABA. Through modification of the editing selectivity, the GABA derivative pyrolidinone may also be measured in the edited spectrum (131). It was recently shown that GABA, homocarnosine, and pyrolidinone have different time courses in response to a first-time challenge with vigabatrin (109).

Estimate of the GABA/Glutamine Cycle in Nonstimulated Human Cerebral Cortex

The *in vivo* ^{13}C MRS measurement may potentially be extended to study the rate of GABA synthesis and the GABA/glutamine cycle (34 ,65). However, the ability to directly measure GABA synthesis at 2.1 T is limited by the [2-^{13}C] GABA resonance being overlapped at 2.1 T by the isotopomer sideband of the [4,3-^{13}C] glutamate resonance. Due to the entry of GABA into the glial TCA cycle at the level of succinate, the labeling kinetics of C4-glutamine derived from GABA are indistinguishable from label entering through anaplerosis. From the results of Shen et al. (29), the maximum estimate of the rate of the GABA/glutamine cycle, obtained by assuming that V_{ana}, is entirely due to GABA, would be approximately 10% of the rate of glutamine synthesis (29). In the 4-T studies of Gruetter et al. (13 ,35), the C2 resonance of GABA was resolved (Fig. 25.2), which indicates that direct quantitation of the rate of GABA synthesis and the GABA/glutamine cycle will be possible in human cerebral cortex at higher field strengths.

Summary and Remaining Questions

The ability of ^{1}H MRS to measure regional levels of GABA and GABA derivatives has provided a new window on the GABAergic system in neurologic and psychiatric disease. Reduced levels of cerebral cortex GABA have been found in patients with adult and pediatric epilepsy, depression, and alcohol withdrawal. Studies have found that several of the new generation of antiepileptic drugs raise GABA levels,

and GABA elevation may be related to their effectiveness in seizure depression. The recently demonstrated ability to perform GABA spectroscopic imaging (105) opens up the potential for using regional variations in GABA level diagnostically and to track the effectiveness of drugs targeted at the GABAergic system.

Two important questions are raised by these findings: What is the relationship between GABA levels and the rate of the GABA/glutamine cycle? What is the relationship between the GABA/glutamine cycle and cortical excitability? A preliminary study has recently demonstrated the ability to use isotopic labeling strategies, similar to those developed to measure the glutamate/glutamine cycle, to measure the rate of the GABA/glutamate cycle (39). This strategy, in combination with the manipulation of GABA levels either pharmacologically or through transgenic methods, may provide significant insight into how the regulation of GABA concentration affects GABAergic function.

IN VIVO MRS MEASUREMENTS OF NEUROENERGETICS DURING FUNCTIONAL ACTIVATION

Part of "25 - Magnetic Resonance Spectroscopy Studies of the Glutamate and Gaba Neurotransmitter Cycles and Functional Neuroenergetics "

Under physiologic conditions brain oxygen and glucose consumption are tightly coupled (49), with between 90% and 95% of glucose uptake being completely oxidized. The tightness of this coupling during brain activation was questioned when Fox and co-workers (132) measured by PET a mean increase of 51% in CMR_{glc} in the primary visual cortex of humans during stimulation by a flashing checkerboard pattern accompanied by only a 5% increase in oxygen consumption ($CMRO_2$). This finding was surprising because of the 16- to 18-fold lower ATP production from nonoxidative glycolysis compared with the complete oxidative consumption of glucose. It was concluded from these results that the energy for supporting electrical activity derives primarily from nonoxidative glycolysis as opposed to glucose oxidation (132 ,133). More recently, the greater increase in cerebral blood flow than oxygen consumption that leads to the BOLD (blood oxygenation level dependent) effect has been taken as evidence of the hypothesis of stimulated neuronal activity requiring little energy (134).

The apparently minimal need for energy from glucose oxidation to support stimulated neuronal activity is paradoxical because of the considerable evidence that the majority of energy consumption in the nonstimulated brain, which primarily uses glucose oxidation, is to support neuronal electrical activity. This evidence includes the critical dependence of brain function on oxygen delivery, the 50% to 70% reduction of brain energy requirements under isoelectric conditions (133), and the ^{13}C MRS findings of a high activity of the glutamate/glutamine cycle in the resting awake brain and the linear coupling of this rate to neuronal glucose oxidation (see the above sections In Vivo MRS Measurements of the Rate of the Glutamate/Glutamine Cycle: Findings and Validation , and Determination of the *In Vivo* Coupling Between the Rate of the Glutamate/Glutamine Neurotransmitter Cycle and Neuronal Glucose Oxidation). Furthermore, there is no cellular evidence that stimulated and nonstimulated neuronal activity have different energetic requirements. As discussed below (see Implications of MRS Studies for Understanding Brain Function), the ambiguity created by the variable degree of uncoupling between glucose consumption and oxidation hinders the interpretation of the functional imaging signal quantitatively in terms of changes in neuronal activity.

MRS provides a powerful tool for studying the question of metabolic coupling between glucose and oxygen by allowing measurements of the rates of nonoxidative glycolysis and glucose oxidation. MRS experiments have shown that under stimulated conditions the majority of energy for functional activity is from glucose oxidation (14 ,15 ,135). They have also confirmed the presence of metabolic uncoupling at high levels of brain activity (136 ,137 and 138). A model has been proposed to explain the uncoupling of glucose consumption and oxidation during certain types of stimulation as an extension of the normal energetic processes used to support the glutamate/glutamine cycle (139).

MRS Studies of Lactate Generation and Glucose Oxidation During Sensory Stimulation

A prediction of the presence of uncoupling of the increase of glucose consumption and oxidation during visual activation is that there will be an elevation of lactate in the visual cortex. Several laboratories have found an increase in lactate concentration (136 ,137 and 138) during visual stimulation of the human visual cortex by ^{1}H MRS of approximately 0.2 to 4 μmol/g-min within 2 to 6 minutes of activation. The small increase in lactate is consistent with earlier findings in animal models (40).

The degree of mismatch between the increase in glucose consumption and oxidation during sensory stimulation was studied by Hyder and co-workers (14 ,15) using forepaw stimulation of an anesthetized rat measured the rate of increase in neuronal glucose oxidation from the ^{13}C isotopic turnover of glutamate using POCE (proton observe carbon edit) heteronuclear editing. These studies found a large increase in the rate of glucose oxidation during sensory stimulation (Fig. 25.10). This increase was in good agreement with previous measurements of the increase in total glucose consumption (140). While within the accuracy of the measurements there was room for a significant rate of nonoxidative glycolysis, the contribution of nonoxidative glycolysis to total cerebral ATP production during activation would be minor due to the much greater number of ATP molecules produced by the complete oxidation of glucose (32 ,33 ,34 ,35 and 36) than by nonoxidative glycolysis to lactate (2).

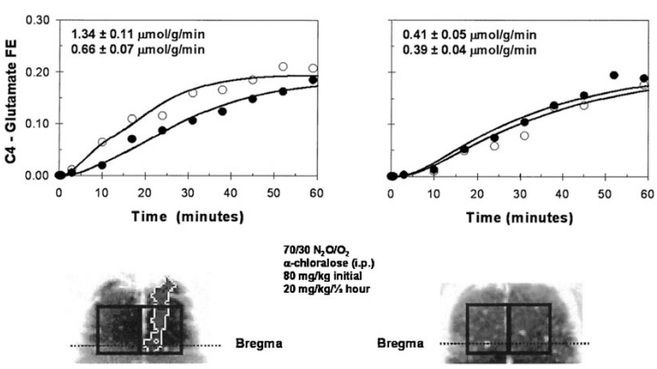


FIGURE 25.10. Time course of C4 glutamate labeling in the ipsi- and contralateral somatosensory cortex of a rat during single forepaw electrical stimulation. *In vivo* ^{1}H-^{13}C MRS spectra were obtained from two 24-μL volumes, positioned in the ipsi- and contralateral somatosensory cortex of rats at 7 T. The spectroscopic volumes are superimposed on a coronal image. The time courses on the right shows the labeling of C4-glutamate during the infusion of a control rat. The time courses on the left were obtained during single forepaw electrical stimulation. The *lighter region* on the image obtained during stimulation is the superposition of the blood oxygenation level dependent (BOLD) functional MRI (fMRI) obtained during the study. The rate of labeling on the contralateral side to the stimulation is observed to increase to approximately two times the rate of the ipsilateral side. The rates on the figure are the calculated rates of the TCA cycle from the group of rats studied in the contra- and ipsilateral volumes. (Redrawn with data from Hyder F, Rothman DL, Mason GF, et al. Oxidative glucose metabolism in rat brain during single forepaw stimulation: a spatially localized $^{1}H[^{13}C]$ NMR study. *J Cereb Blood Flow Metab* 1999;17:1040–1047.)

Calculation of the Relative Contributions of Glucose Oxidation and Nonoxidative Glycolysis to the Incremental Energy Production During Brain Activation

Two recent papers reviewed the increase in glucose oxidation during cognitive and sensory reported in a large number of studies, and concluded that in almost all reports the majority of incremental energy production is from glucose oxidation (131 ,141). Table 25.1 presents the measured fractional change in the rate of total glucose consumption (cerebral metabolic rate of glucose metabolism, CMR_{glc}) and oxidation (cerebral metabolic rate of oxygen, $CMRO_2$) in the human visual cortex during visual stimulation. The studies tabulated used either PET or quantitative functional MRI (fMRI) (142 ,143 ,144 ,145 ,146 ,147 ,148 and 149). In most cases the increase in $CMRO_2$ is greater than reported by Fox et al. (132), and the increase in CMR_{glc} is less. These differences have been attributed to differences in stimulation paradigms, with greater uncoupling from simple stimuli and almost complete coupling from complex stimuli (142 ,149). As shown in the table, even in the most extreme reports of uncoupling the fraction of the increment in total ATP production supplied by glucose oxidation is larger than that supplied by nonoxidative glycolysis.

			Energy Yield (%)		
Stimulation	ΔCMR_{glc} (%)	ΔCMR_{O_2} (%)	(non-ox)CMR_{glc}	(ox)CMR_{glc}	Reference
Visual	51	5	38	62	Fox et al. (132)
	28	28	6	94	Marrett and Gjedde (142)
	29	29	7	93	Marrett and Gjedde (142)
	(31)	16	(10)	90	Davis et al. (143)
	23	(20)	6	(94)	Chen et al.
	24	(20)	6	(94)	Reivich et al. (144)
	(31)	25	(7)	93	Hoge et al. (145)
	(31)	30	(6)	94	Kim et al. (146)
	(31)	5	(27)	73	Kim and Ugurbil (147)
Average	31	20	8	92	
Seizure	400	267	8	92	Borgstrom et al.

Tabulated are the reported increments in CMR_{glc} and CMR_{O_2} from studies using positron emission tomography (PET) or quantitative functional magnetic resonance imaging (MRI) to measure these parameters. The increase in adenosine triphosphate (ATP) production was calculated for each study using a value of 2 ATP molecules produced per glucose molecule consumed in the glycolytic pathway, and 32 additional ATP molecules produced when glucose is completely oxidized. The energy yield is expressed as the percent of the total increase in ATP production from nonoxidative glycolysis [(non-ox)CMR_{glc}] and the oxidative breakdown of glucose in the TCA cycle [(ox)CMR_{glc}]. As shown in the table, even in the most extreme reported cases of uncoupling between CMR_{glc} and CMR_{O_2} the majority of ATP production is from glucose oxidation due to the greater ATP yield. CMR, cerebral metabolic rate.

TABLE 25.1. NEUROENERGETIC YIELD WITH STIMULATION

A limitation of both the PET and quantitative fMRI measurements of oxygen consumption is that they depend on assumptions about blood flow/oxygen delivery coupling in order to derive rates (131). The measurement of glucose oxidation by MRS gets around the requirement of detailed knowledge, or calibration, of this parameter by directly measuring the flow of labeled glucose into the TCA cycle. The recent demonstration of high spatial resolution POCE measurements of glucose oxidation in human visual cortex at 4 T indicate that this method is ready to address this question in the human visual and other sensory systems (31 ,44).

The Glycogen Shunt, a Model of the Mismatch Between Glucose Consumption and Oxidation During Stimulated Neuronal Activity

The results tabulated in Table 25.1 show that the degree of mismatch between the increment in glucose consumption and glucose oxidation is the greatest for stroboscopic stimuli, which require alternate periods of intense activation followed by a quiescent period. For example, in visual stimulation the greatest mismatch was reported for a flashing red dot matrix (132), while no mismatch was found for

the increase from a colored alternating radial checkerboard, which produces a continuous level of stimulation (142 ,149). The largest sustained mismatch between glucose consumption and oxidation occurs in bicuculline-induced status epilepticus where total glucose consumption increases to fourfold the prestatus value, whereas oxidation is increased twofold (49 ,148). In bicuculline-induced status epilepticus, brain cerebral cortex electrical activity is characterized by a burst of intense firing followed by a suppressed period of little electrical activity.

We have proposed a model to explain these observations based on the requirement for ATP from glial glycolysis to supply the ATP needed for glial glutamate clearance and glutamine synthesis (see Determination of the *In Vivo* Coupling Between the Rate of the Glutamate/Glutamine Neurotransmitter Cycle and Neuronal Glucose Oxidation , above). In this model the majority of glucose required to fuel the pumping of glutamate from the synaptic cleft during the intense bursts of neuronal firing induced by sensory stimulation is provided by brain glycogen (150 ,151). Glycogen phosphorylase is kinetically well suited for rapid increases in activity through its regulation by signaling pathways and phosphorylation. There is *in vivo* evidence that brain glycogen may be rapidly mobilized to support function including in status epilepticus (49) and in physiologic brain activation (152 ,153 and 154). In the glycogen shunt model, after an initial period of glycogen depletion during intense stimulation, a steady state is reached in which the glycogen used to rapidly generate ATP for the transport of glutamate into the glial cell and conversion to glutamate during bursts of intense activity is resynthesized during the interim quiescent periods. Only one ATP molecule is produced per glucose moiety used by this pathway, instead of the two produced by glycolysis, and therefore the stoichiometry between the glutamate/glutamine cycle and glial glucose uptake changes to 1:1 from 1:2. The 1:2 ratio is approximately the ratio measured during status epilepticus, in which almost all cortical electrical activity is involved in a burst and suppress pattern. The presence of simultaneous synthesis and breakdown of glycogen has been demonstrated in the exercising muscle (155), and more recently in the cerebral cortex of the stimulated and anesthetized rat (152). Figure 25.11 describes the model schematically.

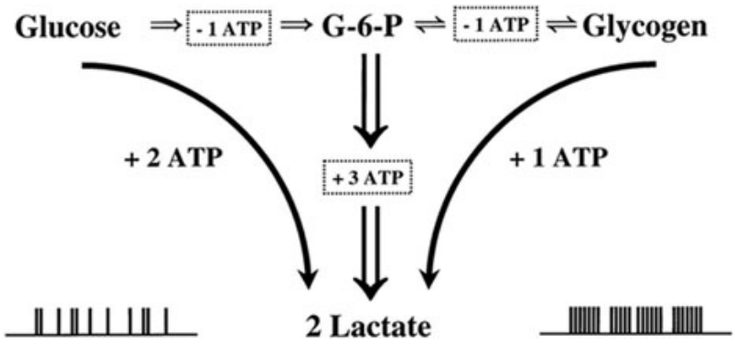


FIGURE 25.11. A schematic diagram of the glycogen shunt model. Glucose taken up by the glia can flow through two pathways after phosphorylation to glucose 6-phosphate. In the standard pathway (*left arrow*), which occurs during normal electrical activity, glucose 6-phosphate is directly converted to lactate by glycolysis producing two ATP molecules per glucose molecule. The stoichiometry between glucose uptake and glutamate transport and conversion to glutamine is 1:1 in this pathway. The majority of lactate is subsequently oxidized in the neuron. In the glycogen shunt pathway (*right arrow*), which occurs during intense repeated bursts of electrical activity, such as in seizures or intense sensory stimulation, glucose is synthesized into glycogen first before being converted to lactate. An ATP molecule is used in the synthesis of glycogen, resulting in a reduction in the energetic yield from the pathway to one ATP molecule per glucose molecule. The stoichiometry between glucose uptake and glutamate transport is increased to 2:1 in this pathway. The extra lactate produced in the shunt pathway is not required for neuronal energy metabolism and eventually leaves the brain.

Summary and Remaining Questions

In vivo MRS studies have made significant contributions to the understanding of functional neuroenergetics. The measurement of the glutamate/glutamine cycle under different levels of electrical activity (see the above sections In Vivo MRS Measurements of the Rate of the Glutamate/Glutamine Cycle: Findings and Validation , and Determination of the *In Vivo* Coupling Between the Rate of the Glutamate/Glutamine Neurotransmitter Cycle and Neuronal Glucose Oxidation) has shown that the majority of brain energy production in even the nonstimulated state supports neuronal activity. Several MRS studies have provided insight into the mismatch between glucose consumption and oxidation during sensory stimulation. Lactate elevation during visual stimulation provided direct validation of the findings using PET (132) of the mismatch between oxygen consumption and glucose consumption (136 ,137 and 138). Although there is a significantly greater increase in glucose consumption than oxidation in certain stimulated states, studies of

the rate of neuronal glucose oxidation during sensory stimulation in the rat cerebral cortex have shown that neuronal glucose oxidation provides the majority of energy production in the stimulated state (14 ,15). The conclusion derived from the MRS studies is consistent with PET measurements of the mismatch when the much greater efficiency of ATP production from glucose oxidation is taken into account (135), as shown in Table 25.1 . A potential explanation of the mismatch has been proposed based on the requirement for rapid glycolytic energy generation to clear glutamate from the synaptic cleft during the bursts of intense neuronal firing associated with stimulated neuronal activity (151). In this glycogen shunt model, the power required is provided by rapid glycogen breakdown. The glycogen is resynthesized during the inter-burst periods resulting in a reduction in the normal stoichiometry between glutamate transport into the glia and glial glucose uptake from 1:1 to 1:2.

Several major questions remain to be addressed on the neuroenergetic support of functional activity. Paramount among these is the need for a measurement of the glutamate/glutamine cycle during sensory stimulation. In addition, although the majority of the increase in energy consumption during stimulation is associated with glutamatergic neurons (15 ,156), the relative contributions of glia and GABAergic neurons are not known. At present there are only minimal data from brain studies supporting the glycogen shunt model of the mismatch between glucose consumption and oxidation. Studies measuring glycogen turnover directly under these conditions (30) may be able to directly test this hypothesis, and better establish the role of glycogen in functional neuroenergetics.

IMPLICATIONS OF MRS STUDIES FOR UNDERSTANDING BRAIN FUNCTION

Part of "25 - Magnetic Resonance Spectroscopy Studies of the Glutamate and Gaba Neurotransmitter Cycles and Functional Neuroenergetics "

The stoichiometry of the rate of the glutamate/glutamine cycle and oxidative glucose metabolism has implications for connecting models of brain function at the macroscopic level, as studied by functional imaging, with neurobiological studies at the level of synapses and networks of neurons. This section reviews work in which this relationship was used to calibrate the PET and fMRI signals and neuroenergetic signals, which are either indirectly or directly measures of functional glucose metabolism, with neurotransmitter cycling (5 ,139 ,157). Some implications of this calibration for the interpretation of brain functional imaging are explored.

Calibration of the Relationship Between the Glutamate/Glutamine Cycle and the PET and BOLD MRI Functional Imaging Signal

At present, functional PET and fMRI are not quantitated in terms of specific neuronal processes involved in information transfer. The functional imaging signal in PET and fMRI is either directly or indirectly coupled to the change in glucose oxidation with activation (1 ,2 ,134). The tight coupling between the glutamate/glutamine cycle and neuronal glucose oxidation in the rat cerebral cortex (26 ,90) provides a relationship for calibrating the functional imaging signal to the specific neuronal process of glutamate release and recycling (5).

The similarity of the ratio of the rates of the glutamate/glutamine cycle to neuronal glucose oxidation in human cerebral cortex and the rat cerebral cortex supports a similar relationship holding for the awake nonstimulated human cerebral cortex. Although studies are needed to establish the exact stoichiometry during sensory or other external stimulation, it is reasonable to extrapolate a positive (and possibly stoichiometric) relationship between changes in the rate of neuronal glucose oxidation and the glutamate/glutamine cycle during activation. Using this relationship, the functional imaging signal may be converted to a first order to changes in the rate of glutamate/glutamine neurotransmitter cycle (5). The advantage of performing this calibration over the direct MRS measurement of the glutamate/glutamine cycle is that several orders of magnitude of higher spatial and temporal resolution is possible with the MRI measurement.

Estimate of the Total Neuroenergetics Used to Support Functional Neuronal Activity During Sensory Stimulation

An assumption in the conventional interpretation of functional imaging is that the increment in neuronal activity during brain activation is sufficient to support the incremental mental processes in a region during sensory stimulation or a cognitive task (5). This interpretation is trivially valid if there is insignificant neuronal activity in the nonstimulated state. The high rate of the glutamate/glutamine neurotransmitter cycle found by ^{13}C MRS in the nonstimulated brain raises the question of whether the incremental neuronal activity is sufficient for mental processing or if the total regional neuronal activity is needed. This question was addressed by analyzing previous studies that measured the change in glucose consumption during stimulation in the sensory cortices of animals stimulated under anesthetized and awake conditions. The studies chosen used anesthetics that interfere minimally with the electrical response to sensory stimulation (157). During anesthesia, the baseline glucose consumption was reduced by as much as two- to threefold. Based on the standard paradigm, a constant increment of neuronal activity, and by inference glucose consumption, during stimulation would be expected regardless of whether the animal was anesthetized or awake. In contrast, if the majority of regional neuronal activity was required for sensory processing, then the glucose consumption required during stimulation would be similar whether the animal was awake or anesthetized. The prediction of these two models is diagrammed in Fig. 25.12 . Results from a number of studies indicated that a similar level of cortical activity was reached during stimulation, independent of the degree of suppression of resting glucose consumption by the anesthesia (139 ,157). These results were supported by the MRS studies that found a large increment in glucose oxidation with somatosensory stimulation under α-chloralose anesthesia (15 ,156). This finding supports the view that during stimulation the total neuronal activity in sensory regions is required to support brain function. Results of this literature survey have recently been reinforced by similar results using quantitative MRI to measure changes in oxygen consumption in the same animal at two different levels of anesthesia (158).

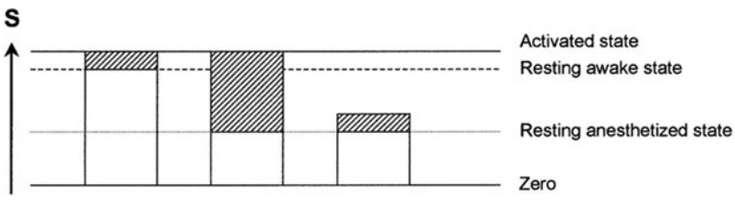


FIGURE 25.12. Schematic representation of possible increase in the functional imaging signal, as measured by neuroenergetics, upon sensory stimulation for an animal that is nonanesthetized (**A**) and anesthetized (**B,C**). The difference in the magnitude of the functional imaging signal, as quantified by glucose oxidation, between stimulated and nonstimulated states is represented by the *shaded rectangles*. In functional imaging these increments are commonly used to identify focally activated regions. The remaining signal, which is represented by *white rectangles*, is removed by fMRI analysis methods. ^{13}C MRS studies have shown that a large fraction of the total signal is from neuroenergetic processes coupled to neuronal activity. If the neuronal activity needed to perform the task was the same as the increment from the awake state, then the increase in the signal upon stimulation would be the same for the anesthetized or awake states (compare **A** and **C**). If instead a large fraction of the total neuronal activity in the region supports the sensory processing, then the incremental signal from anesthesia would be much larger (compare **A** and **B**). A survey of results in the literature showed that in most cases, when anesthetics such as α-chloralose that do not block stimulated electrical activity are used, the total glucose consumption and oxidation rises to the same absolute level during stimulation independent of the initial awake state. These results support the magnitude of the neuronal activity required to support a task (**B**) being substantially larger than the increment in neuronal activity over the resting awake state (**C**). (From Shulman RG, Rothman DL, Hyder F. Stimulated changes in localized cerebral energy consumption under anesthesia. *Proc Natl Acad Sci USA* 1999;96:3245–3250, with permission.)

Implications of the Calibration of the Functional Imaging Signal on the Standard Interpretations of Functional Imaging Studies

The goal of many functional imaging studies is to determine the anatomic location of brain regions involved in performing mental processes. To achieve this goal, subjects are given cognitive or motor tasks to perform, or exposed to sensory stimulation, while being scanned. The degree of involvement of a region in the performance of a task is determined by the increment in the magnitude of the imaging signal relative to the signal when the subject is in a resting state, or performing some other task. An implicit assumption in this analysis often used is that the size of the increment of the signal is proportional to the total neuronal activity recruited by these mental processes (3 ,5 ,159).

As described above, MRS studies have shown that the total neuronal activity in a region, as quantified by the glutamate/glutamine cycle, is much larger than the incremental increase with functional activation. The impact on interpretation of knowing the total magnitude of, as opposed to the incremental, neuronal activity associated with the processing of a stimulus or task may be illustrated through a simple example. Consider a hypothetical experiment in which two subjects perform a cognitive task. In one subject the regional increment in the functional imaging signal in the frontal lobe, quantified as the change in the rate of glucose oxidation, is 1% of the resting rate of total glucose oxidation. In the second subject the same task induces a signal/rate increment of 2%. In the standard interpretation, the second subject recruited twice the neuronal activity to perform the task as the first subject. If instead these increments are calibrated as increments in the glutamate/glutamine cycle the relative difference in neuronal activity is only

a few percent. This example shows that knowing the total size of the signal associated with neuronal activity is important in cases where inferences are being made about differences in the level of neuronal activity, such as when functional imaging is used to study the effects of drugs or disease on brain function (139 ,160). It is also important in the interpretation of functional imaging data to locate a mental process or function (5).

Implications for Studies of Brain Function

The prevailing theory used to interpret functional imaging studies, particularly of cognitive processes, is based on cognitive psychology (3 ,139 ,159 ,161). In the cognitive psychology model of the brain, complex mental processes are broken down using information theory into component processes, sometimes called modules. Functional imaging is used to locate these postulated modules. The search for specialized functional areas has been facilitated by analysis methods such as statistical parametric mapping. These methods identify regional changes in the imaging signal based on the statistical significance of the temporal correlation of the signal changes with the mental processes postulated by the investigator (which are expressed in the design matrix) involved in performing a task (162). In the statistical representation of brain activity, the magnitude of the regional neuronal activity is often ignored in favor of a binary representation as “active” and “inactive.” Statistical methods have been widely adopted to analyze fMRI data. A key assumption often used in the statistical representation of brain function is that the regional neuronal activity used for performing a mental process is independent and separable from the activity of other regions (162 ,163), as well as the neuronal activity recruited by other processes being supported within the same region.

The use of MRS to calibrate neuroimaging provides the potential for examining complex regional brain interactions that do not fulfill the strict modular criteria of independence. Several lines of experiments have shown that parallel regional brain functions interact, and alter the magnitude of the neuronal activity used in processing a stimulus or task (163 ,164 and 165). An example are the studies by Desimone's group (164) on the effect of attention upon the neuronal activity, as measured by the distribution of single neuron firing rates in a region, associated with visual perception. An illustrative set of experiments used two closely spaced visual signals within the same receptor field of a particular region of the striate cortex. The change in neuronal firing rate obtained from one stimulus was found to depend critically on the degree of attention paid to the nearby stimulus. Consistent with this result, an effect of attention on the magnitude fMRI BOLD signal in the human extrastriate cortex has been reported (165 ,166). The calibration of the functional imaging signal in terms of the glutamate/glutamine cycle will extend these studies by allowing these interactions between regions to be described quantitatively in terms of neuronal activity changes, as is presently is done only in electrophysiology studies of animal cerebral cortex. This quantitation should allow the exploration of these complex interactions in much finer detail in humans than is presently possible.

In addition to providing enhanced capability to understand horizontal interactions between brain regions, the calibration of neuroimaging by MRS also allows a vertical dimension of neuronal activity to be explored. The MRS finding of a high rate of the glutamate/glutamine cycle even under nonstimulated conditions is consistent with recent experimentally based proposals that maintaining a constant high level of neuronal activity is critical for brain function. Two recent experiments support this hypothesis. The need for substantial unfocused neuronal activity for the service of even sensory responses was suggested by a recent experiment of Grinvald's group (167). Starting with the recognition that “cortical neurons are spontaneously active in the absence of external input even in primary sensory areas,” the authors studied the correlation between single-unit recordings and real-time optical imaging, which provided a measure of total neuronal activity in the region. They concluded by suggesting that “in the absence of stimulation the cortical network wanders through various states represented by coherent firing of different neuronal assemblies,” and that a stimulus pushes the network into a state representing the stimulus. Analogously, Singer (168) measured the temporal synchronization of neuronal responses and concluded, “Of the many responses of V1 those that become synchronized best will be particularly effective in influencing neurons in higher areas.” These studies both support the presence of a large amount of neuronal activity in the unstimulated state and suggest a role for this activity in brain function. The results from MRS studies provide quantitative measures of the total amount of stimulated and unstimulated activity in a region, and thereby can provide a quantitative basis for analysis.

SUMMARY AND CONCLUSIONS

Part of "25 - Magnetic Resonance Spectroscopy Studies of the Glutamate and Gaba Neurotransmitter Cycles and Functional Neuroenergetics "

Below is a summary of the major findings of MRS studies of the glutamate/glutamine cycle, GABA/glutamine cycle, and functional neuroenergetics, and some of the implications of these findings for understanding brain function.

Approximately 60% to 80% of total glucose oxidation (and energy consumption) in the nonstimulated cerebral cortex is by glutamatergic neurons, with most of the remainder in GABAergic neurons and glia (13 ,18 ,24 ,26 ,27 ,29 ,35 ,38).

The energetic needs of glutamatergic and GABAergic neurons and glia dominate cerebral cortex energy requirements.

In the awake nonstimulated cerebral cortex of humans

and rats, the rate of the glutamate/glutamine cycle is 60% to 80% of total glucose oxidation (13 ,26 ,29 ,35):

1. Glutamate release and recycling is a major metabolic pathway.
2. Glutamate metabolism and neurotransmission can no longer be conceptually separated.
3. The nonstimulated awake brain has a high level of neuronal activity, most likely greater than the increment in activity with external stimulation.

The rate of the glutamate/glutamine cycle increases linearly with neuronal glucose oxidation in a close to 1:1 stoichiometry (26):

1. Energy metabolism in cortical glutamatergic neurons is tightly coupled to glutamate release and recycling.
2. The stoichiometry supports a model in which astrocyte glucose uptake is coupled mechanistically to the glutamate/glutamine cycle (90) through the need for glycolytic ATP to transport glutamate into the astrocyte and synthesize glutamine.
3. The increase in glucose consumption measured during functional activation may be directly coupled to the glutamate/glutamine cycle, providing a calibration for the functional imaging signal.

The GABA level in human cerebral cortex is reduced in epilepsy, alcohol withdrawal, and depression and is raised by several pharmacologic treatments (111 ,129):

1. The concentration of the metabolic pool of brain GABA may play a critical role in inhibitory GABAergic function.
2. Measuring cerebral cortex GABA level provides a useful index of brain GABAergic function and the effectiveness of certain antiepileptic drugs.

Reduction in the activity of GAD_{67} by elevation of GABA leads to a major reduction in the rate of GABA synthesis under nonstimulated conditions (24):

1. GAD_{67} is the major enzyme controlling nonstimulated GABA synthesis in the rate cerebral cortex.
2. Through regulation of GABA concentration GAD_{67} may play a key role in the etiology and pharmacology of epilepsy and other neurologic and psychiatric disorders.
3. The ability of ^{13}C MRS to measure the rate of GABA synthesis in combination with GABA/glutamine neurotransmitter cycling (39) may allow the functional roles of GAD_{65} and GAD_{67} isoforms to be distinguished quantitatively.

The glycogen shunt model provides a mechanistic explanation for the apparent uncoupling of glucose consumption and oxidation during sensory stimulation (151).

The majority of energy to support incremental and total neuronal activity during sensory stimulation is provided by neuronal oxidative glucose metabolism (14 ,131 ,156):

1. The total as opposed to incremental neuronal activity is required to support brain function during sensory stimulation (143).
2. A large amount of unfocused neuronal activity in the nonstimulated state is required for brain function.

ACKNOWLEDGMENTS

Part of "25 - Magnetic Resonance Spectroscopy Studies of the Glutamate and Gaba Neurotransmitter Cycles and Functional Neuroenergetics "

We gratefully acknowledge grant support from the National Institute of Health—DK-27121 (R.G.S.), NS-32126 (D.L.R.), HD-32573 (K.L.B.), NS-34813 (K.L.B.), NS-37203 (F.H.), RO1-NS032518, PO1-NS06208 (O.A.C.P.)—and the National Science Foundation—DBI-9730892 (F.H.). We benefited greatly from discussions with Dr. James Lai on the localization and kinetics of key enzymes in glycolysis and the glutamate/glutamine cycle and the careful reading of the manuscript by Dr. Vincent Lebon.

REFERENCES

1. Sokoloff L. Relationship between functional activity and energy metabolism in the nervous system: whether, where and why? In: Lassen NA, Ingvar DH, Raichle ME, et al., eds. *Brain work and mental activity*. Copenhagen: Munksgaard, 1991:52–64.

2. Fitzpatrick SM, Rothman DL. New approaches to functional neuroenergetics. *J Cognit Neurol* 1999;11:4,467–471.

3. Raichle ME. Behind the scenes of functional brain imaging: a historical and physiological perspective. *Proc Natl Acad Sci USA* 1998;95:576–772.

4. Shephard GM. *The synaptic organization of the brain*. Oxford: Oxford University Press, 1994.

5. Shulman RG, Rothman DL. Interpreting functional imaging studies in terms of neurotransmitter cycling. *Proc Natl Acad Sci USA* 1998;95:11993–11998.

6. Rothman DL, Behar KL, Hetherington HP, et al. Homonuclear 1H-double resonance difference spectroscopy of the rat brain in vivo. *Proc Natl Acad Sci USA* 1984;81:6330–6334.

7. Rothman DL, Novotny EJ, Shulman GI, et al. ^{1}H-[^{13}C] NMR measurements of [4-^{13}C]-glutamate turnover in human brain. *Proc Natl Acad Sci USA* 1992;89:9603–9606.

8. Behar KL, Petroff OAC, Prichard JW, et al. (1986) Detection of metabolites in rabbit brain by ^{13}C NMR spectroscopy following administration of [1-^{13}C]glucose. *Magn Reson Med* 1986;3:911–920.

9. Fitzpatrick SM, Hetherington HP, Behar KL, et al. The flux from glucose to glutamate in the rat brain in vivo as determined by ^{1}H-observed, ^{13}C-edited NMR spectroscopy. *J Cereb Blood Flow Metab* 1990;10:170–179.

10. Cerdan S, Künnecke B, Seelig J. Cerebral metabolism of [1,2-$^{13}C_2$] acetate as detected by in vivo and in vitro ^{13}C NMR. *J Biol Chem* 1990;265:12916–12926.

11. Gruetter R, Novotny EJ, Bouleware SD, et al. Direct measurement of brain glucose concentrations in humans by ^{13}C NMR spectroscopy. *Proc Natl Acad Sci USA* 1992;89:1109–1112.

12. Gruetter R, Novotny EJ, Boulware SD, et al. Localized ^{13}C. Neurochem NMR spectroscopy in the human brain of amino acid labeling from ^{13}C glucose. *J Neurochem* 1994;63:1377–1385.

13. Gruetter R, Seaquist ER, Kim S, et al. Localized in vivo ^{13}C-NMR of glutamate metabolism in the human brain: initial results at 4 Tesla. *Dev Neurosci* 1998;20:380–388.

14. Hyder F, Chase JR, Behar KL, et al. Increased tri-carboxylic acid cycle flux in rat brain during forepaw stimulation detected with ^{1}H-[^{13}C] NMR. *Proc Natl Acad Sci USA* 1996;93:7612–7617.

15. Hyder F, Rothman DL, Mason GF, et al. Oxidative glucose metabolism in rat brain during single forepaw stimulation: a spatially localized ^{1}H[^{13}C] NMR Study *J Cereb Blood Flow Metab* 1997;17:1040–1047.

16. Hyder F, Renken R, Rothman DL. In vivo carbon-edited detection with proton echo-planar spectroscopic imaging (ICED PEPSI): [3,4-$^{13}CH_2$]glutamate/glutamine tomography in rat brain. *Magn Reson Med* 1999;42:997–1003

17. Mason GF, Rothman DL, Behar KL, et al. NMR determination of the TCA cycle rate and alpha-ketoglutarate/glutamate exchange rate in rat brain. *J Cereb Blood Flow Metab* 1992;12:434–447.

18. Mason GF, Gruetter R, Rothman DL, et al. Simultaneous Determination of the rates of the TCA cycle, glucose utilization, alpha-ketoglutarate/glutamate exchange, and glutamine synthesis in human brain by NMR. *J Cereb Blood Flow Metab* 1995;18:12–25.

19. Mason GF, Pan JW, Chu WJ, et al. Measurement of the tricarboxylic acid cycle rate in human gray and white matter in vivo by 1H-13C magnetic resonance spectroscopy at 4.1T. *J Cereb Blood Flow Metab* 1999;19:1179–1188.

20. Kunnecke B, Cerdan S, Seelig J. Cerebral metabolism of [1,2-^{13}C] glucose and [U-^{13}C] 3 hydroxybutyrate in rat brain as detected by ^{13}C NMR spectroscopy. *NMR Biomed* 1993;6:264–277.

21. Van Zijl PMC, Chesnick AS, DesPres D, et al. In vivo proton spectroscopy and spectroscopic imaging of [1-^{13}C] glucose and its metabolic products. *Magn Reson Med* 1993;30:544–551.

22. Van Zijl PMC, Chesnick AS, DesPres D, et al. Determination of cerebral glucose transport and metabolic kinetics by dynamic MR spectroscopy. *Am J Physiol* 1997;273:E1216–1217.

23. Lapidot A, Gopher A. Cerebral metabolic compartmentation. Estimation of glucose flux via pyruvate carboxylase/pyruvate dehydrogenase by ^{13}C NMR isotopomer analysis of D-[U-^{13}C] glucose metabolites. *J Biol Chem* 1994;269:27198–27208.

24. Manor D, Rothman DL, Mason GF, et al. The rate of turnover of cortical GABA from [1-^{13}C] glucose is reduced in rats treated with the GABA-transaminase inhibitor vigabtrin (γ-vinyl GABA). *Neurochem Res* 1996;12:1031–1041.

25. Sibson NR, Dhankhar A, Mason GF, et al. In vivo ^{13}C NMR measurement of cerebral glutamine synthesis as evidence for glutamate-glutamine cycling, *Proc Natl Acad Sci USA* 1997;94:2699–2704.

26. Sibson NR, Dhankhar A, Mason GF, et al. Stoichiometric coupling of brain glucose metabolism and glutamatergic neuronal activity. *Proc Natl Acad Sci USA* 1998;95:316–321

27. Sibson NR, Mason GF, Shen J, et al. In vivo ^{13}C NMR measurement of neurotransmitter glutamate cycling, anaplerosis and TCA cycle flux in rat brain during [2-^{13}C]glucose infusion. *J Neurochem* 2001;76(4):975–989.

28. Hassel B, Bachelard H, Jones P, et al. Trafficking of amino acids between neurons and glia in vivo, effects of inhibition of glial metabolism by fluoroacetate. *J Cereb Blood Flow Metab* 1997;17:1230–1238.

29. Shen J, Petersen KF, Behar KL, et al. Determination of the rate of the glutamate-glutamine cycle in the human brain by in vivo ^{13}C NMR. *Proc Natl Acad Sci USA* 1999;96:8235–8240.

30. Choi I-Y, Tkac I, Ugurbil K. Noninvasive measurements of [1-^{13}C] glycogen content and metabolism in rat brain in vivo. *J Neurochem* 1999;73:1300.

31. Pan JW, Stein DT, Mason GF, et al. Gray and white matter metabolic rate in human brain by spectroscopic imaging. *Magn Reson Med* 2000;44:673–679.

32. Shank RP, Campbell GL. (1984) Alpha-ketoglutarate and malate uptake and metabolism by synaptosomes: further evidence for an astrocyte to neuron metabolic shuttle. *J Neurochem* 1984;42:1153–1161.

33. Schousboe A, Westergaard N, Sonnewald U, et al. Glutamate and glutamine metabolism and compartmentation in astrocytes. *Dev Neurosci* 1993;15:359–366.

34. Peng L, Hertz L. Utilization of glutamine and TCA cycle constituents as precursors for transmitter glutamate and GABA. *Dev Neurosci* 1998;15:367–377.

35. Gruetter R, Seaquist ER, Ugurbil K. In vivo studies of neurotransmitter and amino acid metabolism in human brain. *J Neurochem* 2000;74(suppl):S44.

36. Shen J, Sibson NR, Cline G, et al. ^{15}N NMR spectroscopy studies of ammonia transport and glutamine synthesis in the hyperammonemic rat brain. *Dev Neurosci* 1998;20:438–443.

37. Sibson NR, Shen J, Mason GF, et al. Functional energy metabolism: in vivo ^{13}C NMR evidence for coupling of cerebral glucose consumption and glutamatergic neuronal activity. *Dev Neurosci* 1998;20:321–330.

38. Lebon V, Petersen KF, Cline GW, et al. Measurement of total neuronal/astroglial glutamate-glutamine trafficking and astroglial TCA cycle flux in human cerebral cortex using ^{13}C NMR spectroscopy during the infusion of [2-^{13}C] acetate. *J Neurochem* 2001;*in press*.

39. Patel AB, de Graaf RA, Mason GF, et al. GABAergic and glutamatergic neurotransmitter cycling in the rat cerebral cortex. *Proc Int Soc Magn Reson Med* 2001;*in press*.

40. Van den Berg CJ, Garfinkel D. A simulation study of brain components. Metabolism of glutamate and related substances in mouse brain. *Biochem J* 1971;123:211–218.

41. Ottersen OP, Zhang N, Walberg F. Metabolic compartmentation of glutamate and glutamine: morphological evidence obtained by quantitative immunocytochemistry in rat cerebellum. *Neuroscience* 1992;46:519–534.

42. Conti F, Minelli A. Glutamate immunoreactivity in rat cerebral cortex is reversibly abolished by 6-diazo-5-oxo-L-norleurcine (DON), an inhibitor of phosphate-activated glutaminase. *J Histochem Cytochem* 1994;42:717–726.

43. Chen W, Novotny EJ, Boulware SD, et al. Quantitative measurements of regional TCA cycle flux in visual cortex of human brain using 1H-{13C} NMR spectroscopy. *Proc Soc Magn Reson* 1994;63.

44. Chen W, Zhu X-H, Gruetter R, et al. Study of oxygen utilization changes in human visual cortex during hemifield stimulation using 1H-13C MRS and fMRI. *Proc Int Soc Magn Reson Med* 2000;438.

45. Nowak LG, Sanchez-Vives MV, McCormick DA. Influence of low and high frequency inputs on spike timing in visual cortical neurons. *Cereb Cortex* 1997;7:487–501.

46. Roberts E. Failure of GABAergic inhibition: a key to local and global seizures. *Adv Neurol* 1986;44:319–341.

47. Roberts E. The establishment of GABA as a neurotransmitter. In: Squires RF, ed. *GABA and benzodiazepine receptors*, vol 1. Boca Raton: CRC Press, 1988:1–21.

48. Shank RP, Leo GC, Zielke HR. Cerebral metabolic compartmentation as revealed by nuclear magnetic resonance analysis of D-[1-^{13}C]glucose metabolism. *J Neurochem* 1993;61:315–323.

49. Siesjo BK. *Brain energy metabolism*. New York: Wiley, 1978.

50. Rothstein JD, Martin L, Levey AI, et al. Localization of neuronal and glial glutamate transporters. *Neuron* 1994;13:713–725.

51. Tsacopoulos M, Magistretti PJ. Metabolic coupling between glia and neurons. *J Neurosci* 1996;16:877–885.

52. Martinez-Hernandez A, Bell KP, Norenberg MD. Glutamine synthetase: glial localization in brain. *Science* 1977;195:1356–1358.

53. Meister A. Glutamine synthetase from mammalian tissues. *Methods Enzymol* 1985;113:185–199.

54. Wiesinger H. Glia-specific enzyme systems. In: Kettenmann H, Ransom BR, eds. *Neuroglia*. Oxford: Oxford University Press, 1995:488–499.

55. Kvamme E, Torgner IAA, Svenneby G. Glutaminase from mammalian tissues. *Methods Enzymol* 1985;113:241–256.

56. Erecinska M, Silver IA. Metabolism and role of glutamate in mammalian brain. *Prog Neurobiol* 1990;35:245–296.

57. Sonnewald U, Westergaard N, Schousboe A, et al. Direct demonstration by [13C]NMR spectroscopy that glutamine from astrocytes is a precursor for GABA synthesis in neurons. *Neurochem Int* 1993;22:19–29.

58. Schousboe A, Westergaard N. Transport of neuroactive amino acids in astrocytes. In: Kettenmann H, Ransom BR, eds. *Neuroglia*. New York: Oxford University Press, 1995:246–258.

59. Maycox PR, Hell JW, Jahn R. Amino acid neurotransmission: spotlight on synaptic vesicles. *Trends Neurosci* 1990;13:83–87.

60. Nicholls DG, Attwell D. The release and uptake of excitatory amino acids. *Trends Pharmacol Sci* 1990;11:462–468.

61. Badar-Goffer RS, Ben-Yoseph O, Bachelard HS, et al. Neuronal-glial metabolism under depolarizing conditions. *Biochem J* 1992;282:225–230.

62. Cooper AJL, Plum F. Biochemistry and physiology of brain ammonia. *Physiol Rev* 1987;67:440–519.

63. Haberg A, Qu Hong, Bakken IJ, et al. In vitro and ex vivo 13C-NMR spectroscopy studies of pyruvate recycling in brain. *Dev Neurosci* 1998;20:389–398.

64. Hertz L, Yu AC, Kala G, et al. Neuronal-astrocytic and cytosolic-mitochondrial metabolite trafficking during brain activation, hyperammonemia and energy deprivation. *Neurochem Int* 2000;37:83–102.

65. Reubi JC, Van den Berg C, Cuenod M. Glutamine as a precursor for the GABA and glutamate transmitter pools. *Neurosci Lett* 1978;10:171–174.

66. Rothman DL, Petroff OAC, Behar KL, et al. Localized ^{1}H NMR measurements of GABA levels in human brain in vivo. *Proc Natl Acad Sci USA* 1993;90:5662–5666.

67. Dejong CHC, Kampman MT, Deutz NEP, et al. Cerebral cortex ammonia and glutamine metabolism during liver insufficiency-induced hyperammonemia in the rat. *J Neurochem* 1992;59:1071–1079.

68. Berl S, Takagaki G, Clarke DD, et al. Carbon dioxide fixation in the brain. *J Biol Chem* 1962;237:2570–2573.

69. Kanamori K, Ross BD. ^{15}N n.m.r. measurement of the in vivo rate of glutamine synthesis and utilization at steady state in the brain of the hyperammonaemic rat. *Biochem J* 1993;293:461–468.

70. Kanamori K, Ross BD. In vivo activity of glutaminase in the brain of hyperammonaemic rats measured in vivo by ^{15}N nuclear magnetic resonance. *Biochem J* 1995;305:329–336.

71. Hawkins RA, Miller AL, Nielsen RC, et al. The acute action of ammonia on rat brain metabolism in vivo. *Biochem J* 1973;134:1001–1008.

72. Cheng S. CO_2 fixation in the nervous tissue. *Int Rev Neurobiol* 1971;14:125–157.

73. Fitzpatrick SM, Hetherington HP, Behar KL, et al. Effects of acute hyperammonemia on cerebral amino acid metabolism and pHi in vivo, measured by ^{1}H and ^{31}P NMR. *J Neurochem* 1989;2:741–479.

74. Kanamatsu T, Tsukada Y. Effects of ammonia on the anaplerotic pathway and amino acid metabolism in the brain: an ex vivo 13C NMR spectroscopic study of rats after administering [2-13C]glucose with or without ammonium acetate. *Brain Res* 2000;841:11–19.

75. Jones JG, Naidoo R, Sherry AD, et al. Measurement of gluconeogenesis and pyruvate recycling in the rat liver: a simple analysis of glucose and glutamate isotopomers during metabolism of [1,2,3-(13)C3]propionate. *FEBS Lett* 1997;412:131–137.

76. Schousboe A, Westergaard N, Sonnewald U, et al. Regulatory role of astrocytes for neuronal biosynthesis and homeostasis of glutamate and GABA. *Prog Brain Res* 1992;94:199–211.

77. Yudkoff M, Nissim I, Daikhin Y, Lin ZP, et al. Brain glutamate metabolism: neuronal-astroglial relationships. *Dev Neurosci* 1993;15:343–350.

78. Danbolt NC, Storm-Mathisen J, Kanner BI. An [Na+-K+] coupled L-glutamate transporter purified from rat brain is located in glial cell processes. *Neuroscience* 1992;51:295–310.

79. Kanner BI, Sharon I. Active transport of L-glutamate by membrane vesicles isolated from rat brain. *Biochemistry* 1978;17:3949–3953.

80. Pines G, Kanner BI. Counterflow of L-glutamate in plasma membrane vesicles and reconstituted preparations from rat brain. *Biochemistry* 1990;29:11209–11214

81. Chaudhry FA, Lehre KP, van Lookeren Campagne M, et al. Glutamate transporters in glial plasma membranes: highly differentiated localizations revealed by quantitative ultrastructural immunocytochemistry. *Neuron* 1995;15:711–720.

82. Lehre KP, Levy LM, Ottersen OP, et al. Differential expression of two glial glutamate transporters in the rat brain: quantitative and immunocytochemical observations. *J Neurosci* 1995;15:1835–1853.

83. Milton ID, Banner SJ, Ince PG, et al. Expression of the glial glutamate transporter EAAT2 in the human CNS: an immunohistochemical study. *Mol Brain Res* 1997;52:17–31.

84. Rothstein JD, Jin L, Dykes-Hoberg M, et al. Chronic inhibition of glutamate uptake produces a model of slow neurotoxicity. *Proc Natl Acad Sci USA* 1993;90:6591–6595.

85. Rothstein JD, Dykes-Hoberg M, Pardo CA, et al. Knockout of glutamate transporters reveal a major role for astroglial transport in excitotoxicity and clearance of glutamate. *Neuron* 1996;16:675–686.

86. Ginsberg SD, Martin LJ, Rothstein JD. Regional deafferentation down-regulates subtypes of glutamate transporter proteins. *J Neurochem* 1995;65:2800–2803.

87. Ginsberg SD, Rothstein JD, Price DL, et al. Fimbria-fornix transections selectively down-regulate subtypes of glutamate transporter and glutamate receptor proteins in septum and hippocampus. *J Neurochem* 1996;67:1208–1216.

88. Bergles DE, Jahr CE. Synaptic activation of glutamate transporters in hippocampal astrocytes. *Neuron* 1997;9:1297–1308.

89. Pellerin L, Pellegri G, Bittar PG, et al. Evidence supporting the existence of an activity dependent astrocyte neuron lactate shuttle. *Dev Neurosci* 1998;20:291–299.

90. Magistretti PJ, Pellerin L, Rothman DL, et al. Energy on demand. *Science* 1999;283:496–497.

91. Pellerin L, Magistretti PJ. Glutamate uptake into astrocytes stimulates aerobic glycolysis: a mechanism coupling neuronal activity to glucose utilization. *Proc Natl Acad Sci USA* 1994;91:10625–10629.

92. Hertz L, Swanson RA, Newman GC, et al. Can experimental conditions explain the discrepancy over glutamate stimulation of aerobic glycolysis? *Dev Neurosci* 1998;20:339–347.

93. Fell DA. Metabolic control analysis: a survey of its theoretical and experimental development. *Biochem J* 1997;286:313–330.

94. Martin DL. The role of glia in the inactivation of neurotransmitters. In: Kettenmann H, Ransom BR, eds. *Neuroglia*. Oxford: Oxford University Press, 1995:732–745.

95. Kadekaro M, Crane AM, Sokoloff L. Differential effects of electrical stimulation of sciatic nerve on metabolic activity in spinal cord and dorsal root ganglion in the rat. *Proc Natl Acad Sci USA* 1985;82:6010–6013.

96. Kennedy C, Des Rosiers MH, Sakurada O. Metabolic mapping of the primary visual system of the monkey by means of the autoradiographic [14C] deoxyglucose technique. *Proc Natl Acad Sci USA* 1976;73:4230–4234.

97. Yarowsky P, Kadekaro M, Sokoloff L. Frequency-dependent activation of glucose utilization in the superior cervical ganglion by electrical stimulation of cervical sympathetic trunk. *Proc Natl Acad Sci USA* 1983;80:4179–4183.

98. Bachelard H. Landmarks in the application of 13C-Magnetic Resonance Spectroscopy to studies of neuronal/glial relationships. *Dev Neurosci* 1998;20:277–188.

99. Ruderman NB, Ross PS, Berger M, et al. Regulation of glucose and ketone-body metabolism in brain of anaesthetized rats. *Biochem J* 1974;138:1–10.

100. Pan JW, Rothman DL, Behar KL,. et al. Human brain beta-hydroxybutyrate and lactate increase in fasting induced ketosis. *J Cereb Blood Flow* 2000;*in press*..

101. Meldrum BS. GABAergic mechanism in the pathogenesis and treatment of epilepsy. *Br J Clin Pharmacol* 1989;27:3S–11S.

102. Olsen RW, Avoli M. GABA and epileptogenesis. *Epilepsia* 1997;38:399–407.

103. Behar KL, Rothman DL, Spencer DD, et al. Analysis of macromolecule resonances in ^{1}H NMR spectra of human brain. *Magn Reson Med* 1994;32:294–302.

104. Hetherington HP, Newcomer BR, Pan JW. Measurement of human cerebral GABA at 4.1 T using numerically optimized editing pulses. *Magn Reson Med* 1998;39:6–10.

105. Shen J, Shungu DC, Rothman DL. In vivo chemical shift imaging of γ-aminobutyric acid in the human brain. *Magn Reson Med* 1999;41:35–42.

106. Behar KL, Boehm D. Measurement of GABA following GABA-transaminase inhibition by gabaculine: a ^{1}H and ^{31}P NMR spectroscopic study of rat brain in vivo. *J Cereb Blood Flow Metab* 1991;11(suppl 2):S783.

107. Preece NE, Williams SR, Jackson G, et al. ^{1}H-NMR studies of vigabatrin-induced increases in cerebral GABA. *Proc Soc Mag Reson Med* 1991;1000.

108. Petroff OAC, Rothman DL, Behar KL, et al. Human brain gamma-aminobutyric acid (GABA) levels and seizure control following initiation of vigabatrin therapy. *J Neurochem* 1996;67:2399–2404.

109. Petroff OAC, Behar KL, Rothman DL. New NMR measurements in epilepsy. Measuring brain GABA in patients with complex partial seizures. *Adv Neurol* 1999;79:945–951.

110. Petroff OAC, Rothman DL. Measuring human brain GABA in vivo: effects of GABA-transaminase inhibition with vigabatrin. *Mol Neurobiol* 1998;16:97–121.

111. Petroff OAC, Hyder F, Rothman DL, et al. Functional imaging in the epilepsies proton MRS: GABA and glutamate. *Adv Neurol* 2000;83:263–272.

112. Rothman DL, Behar KL, Prichard JW, et al. Homocarnosine and the measurement of neuronal pH in patients with epilepsy. *Magn Reson Med* 1997;32:924–929.

113. Martin DL, Rimvall K. Regulation of γ-aminobutyric acid synthesis in the brain. *J Neurochem* 1993;60:395–407.

114. Martin DL, Liu H, Martin SB, et al. Structural features and regulatory properties of the brain glutamate decarboxylases. *Neurochem Int* 2000;37:111–119.

115. Kaufman DL, Houser CR, Tobin AJ. Two forms of the gamma-aminobutyric acid synthetic enzyme glutamate decarboxylase have distinct intraneuronal distributions and cofactor interactions. *J Neurochem* 1991;56:720–723.

116. Rimvall K, Martin DL. The level of GAD_{67} protein is highly sensitive to small increases in intraneuronal gamma-aminobutyric acid levels. *J Neurochem* 1994;62:1375–1381.

117. Sheikh SN, Martin DL. Elevation of brain GABA levels with vigabatrin (gamma-vinylGABA) differentially affects GAD_{65} and GAD_{67} expression in various regions of rat brain. *J Neurosci Res* 1998;52:736–741.

118. Asada H, Kawamura Y, Maruyama K, et al. Mice lacking the 65 kDa isoform of glutamic acid decarboxylase (GAD_{65}) maintain normal levels of GAD67 and GABA in their brains but are susceptible to seizures. *Biochem Biophys Res Commun* 1996;229:891–895.

119. Asada H, Kawamura Y, Maruyama K, et al. Cleft palate and decreased brain γ-aminobutyric acid in mice lacking the 67-kDa isoform of glutamic acid decarboxylase. *Proc Natl Acad Sci USA* 1996;94:6496–6499.

120. Petroff OAC, Rothman DL, Behar KL, et al. Low brain GABA level is associated with poor seizure control. *Ann Neurol* 1996;40:908–911.

121. Petroff OAC, Hyder F, Mattson RH, et al. Topiramate increases brain GABA, homocarnosine, and pyrrolidinone in patients with epilepsy. *Neurology* 1999;52:473–478.

122. Petroff OAC, Hyder F, Rothman DL, et al. Effects of gabapentin on brain GABA, homocarnosine, and pyrrolidinone in epilepsy. *Epilepsia* 2000;41:675–680.

123. Hetherington HP, Kuzniecky RI, Pan JW. Acute and chronic alterations in human cerebral GABA levels in response to topiramate, lamotrigine, and gabapentin. *Proc Soc Magn Reson Med* 1998;1:542.

124. Kuzniecky RI, Hetherington HP, Ho S, et al. Topiramate increases cerebral GABA in healthy humans. *Neurology* 1998;51:627–629.

125. Petroff OAC, Rothman DL, Behar KL, et al. The effect of gabapentin on brain gamma-aminobutyric acid in patients with epilepsy. *Ann Neurol* 1996;39:95–99.

126. Novotny EJ, Hyder F, Rothman DL. GABA changes with vigabatrin in the developing human brain. *Epilepsia* 1998;40:462–466.

127. Kocsis JD, Mattson RH. GABA levels in the brain: a target for new antiepileptic drugs. *Neuroscientist* 1996;6:326–334.

128. Richerson GB, Gaspary HL. Carrier-mediated GABA release: Is there a functional role? *Neuroscientist* 1987;3:151–157.

129. Sanacora G, Mason GF, Rothman DL, et al. Preliminary evidence of reduced cortical GABA levels in depressed patients assessed using 1H-magnetic resonance spectroscopy. *Arch Gen Psychiatry* 1999;56:1043.

130. Behar KL, Rothman DL, Petersen KF, et al. Preliminary evidence of reduced cortical GABA levels in localized 1-H-MR spectra of alcohol-dependent and hepatic encephalopathy patients. *Am J Psychiatry* 1999;156:952–954.

131. Hyder F, Petroff OAC, Mattson RH, et al. Localized ¡H NMR measurements of 2-pyrrolidinone in human brain in vivo. *Magn Reson Med* 1999;41:889–896.

132. Fox PT, Raichle ME, Mintun MA, et al. Nonoxidative glucose consumption during focal physiologic neural activity. *Science* 1988;241:462–464

133. Gjedde A. The energy cost of neuronal depolarization. In: Gulyas B, Ottoson D, Roland PE, eds. *Functional organisation of the human visual cortex*. Wenner-Gren International Symposium Series, vol 61. Oxford: Pergamon Press, 1993:291–306.

134. Ogawa S, Menon RS, Kim SG, et al. On the characteristics of functional magnetic resonance imaging of the brain. *Annu Rev Biophys Biomol Struct* 1998;27:447–474.

135. Hyder F, Shulman RG, Rothman DL. Regulation of cerebral oxygen delivery. *Adv Exp Med Biol* 1999;471:99–109

136. Prichard JW, Rothman DL, Novotny EJ, et al. Lactate rise detected by ^{1}H NMR in human visual cortex during physiological stimulation. *Proc Natl Acad Sci USA* 1991;88:5829–5831

137. Sappey-Marinier D, Calabrese G, Fein G, et al. Effect of photic stimulation on human visual cortex lactate and phosphates using 1H and 31P magnetic resonance spectroscopy. *J Cereb Blood Flow Metab* 1992;12:584–592.

138. Frahm J, Kruger G, Merboldt KD, et al. Dynamic uncoupling and recoupling of perfusion and oxidative metabolism during focal brain activation in man. *Magn Reson Med* 1996;35:143–148.

139. Shulman RG. Functional imaging studies: linking mind and basic neuroscience. *Am J Psychiatry* 2001;158:11–20.

140. Ueki M, Linn F, Hossman KA. Functional activation of cerebral blood flow and Metabolism before and after global ischemia of rat brain. *J Cereb Blood Flow Metab* 1988;8:486–494.

141. Hyder F, Kennan RP, Kida I, et al. Dependence of oxygen delivery on blood flow in rat brain: a 7 Tesla nuclear magnetic resonance study. *J Cereb Blood Flow Metab* 2000;20:485–498.

142. Marrett S, Gjedde A. Changes of blood flow and oxygen consumption in visual cortex of living humans. *Adv Exp Med Biol* 1997;413:205–208.

143. Davis TL, Kwong KK, Weisskoff RM, et al. Calibrated functional MRI: mapping the dynamics of oxidative metabolism. *Proc Natl Acad Sci USA* 1998;95:1834–1839.

144. Reivich M, Alavi A, Gur RC. Positron emission tomographic studies of perceptual tasks. *Ann Neurol* 1984;15:S61–S65.

145. Hoge RD, Atkinson J, Gill B, et al. Linear coupling between cerebral blood flow and oxygen consumption in activated human cortex. *Proc Natl Acad Sci USA* 1999;96:9403–9408.

146. Kim SG, Rostrup E, Larsson HB, et al. Determination of relative CMRO2 from CBF and BOLD changes: significant increase of oxygen consumption rate during visual stimulation. *Magn Reson Med* 1999;41:1152–1161.

147. Kim SG, Ugurbil K. Comparison of blood oxygenation and cerebral blood flow effects in fmri: estimation of relative oxygen consumption change. *Magn Reson Med* 1997;38:59–65.

148. Borgstrom L, Chapman AG, Siesjo BK. Glucose consumption in the cerebral cortex of rat during bicuculline-induced status epilepticus. *J Neurochem* 1976;27:971–973.

149. Marret S, Fujita H, Meyer E, et al. Stimulus specific increase of oxidative metabolism in human visual cortex. In: Uemura K, Lassen NA, Jones T, eds. *Quantification of brain function: tracer kinetics and image analysis in brain PET*. Amsterdam: Elsevier, 1993:217–224.

150. Sibson NR, Rothman DL, Behar KL, et al. The glycogen shunt and brain energetics. *Proc Int Soc Magn Reson Med* 1999;1:730.

151. Shulman RG, Rothman DL. The “glycogen shunt” in exercising muscle: a novel role for glycogen in muscle energetics and fatigue. *Proc Natl Acad Sci USA* 2001;98:457–461.

152. Dienel GA, Wang RY, Cruz NF. Channeling of ^{14}C glucose into tricarboxylic acid cycle derived metabolic pools in neurons and glia during and after sensory stimulation. Proceedings of the International Society for Neurochemistry 3rd International Conference on Brain Energy Metabolism, 1997.

153. Swanson RA, Morton MM, Sagar SM, et al. Sensory stimulation induces local cerebral glycogenolysis: demonstration by autoradiography. *Neuroscience* 1992;51:451–461.

154. Magistretti PJ, Sorg O, Martin J-L. Regulation of glycogen metabolism in astrocytes: physiological, pharmacological and pathological aspects. In: Murphy S, ed. *Astrocytes: pharmacology and function*. San Diego: Academic Press, 1993:243–265.

155. Price TB, Taylor R, Mason GM, et al. Turnover of human muscle glycogen during low-intensity exercise. *Med Sci Sports Exerc* 1994;26(8):983–991.

156. Hyder F, Shulman RG, Rothman DL. A model for the regulation of cerebral oxygen delivery. *J Appl Physiol* 1998;85:554–564.

157. Shulman RG, Rothman DL, Hyder F. Stimulated changes in localized cerebral energy consumption under anesthesia. *Proc Natl Acad Sci USA* 1999;96:3245–3250.

158. Hyder F, Shulman RG, Rothman DL, et al. Localized energetic changes with brain activation from anesthesia II: relative BOLD changes at 7T. *Proc Int Soc Magn Reson Med* 2000;8:968.

159. Posner MI, Raichle ME. *Images of mind*. New York: Freeman, 1994.

160. Buchsbaum M, Hazlett EA. Positron emission tomography studies of abnormal glucose metabolism in schizophrenia. *Schizophr Bull* 1998;24(3):343–364.

161. Smith EE. Infusing cognitive neuroscience into cognitive psychology. In: Solso RL, ed. *Mind and brain sciences in the 21st century*. A Bradford Book, MIT Press, 1997:71–89.

162. Frackowiack RSJ, Fristion KJ, Frith CD, et al. *Human brain function*. San Diego: Academic Press, 1997:25–41.

163. Jennings JM, Mcintosh AR, Kapur S, et al. Cognitive subtractions may not add up: the interaction between semantic processing and response mode. *Neuroimage* 1997;5:229–239.

164. Luck SJ, Cheazzi L, Hillyard SA, et al. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol* 1997;77(1):24–42.

165. Kastner S, DeWeerd P, Desimone R, et al. Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 1998;282:108–111.

166. Chawla D, Rees G, Friston KJ. The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neurosci* 1999;2:671–676.

167. Tsodyks M, Kenet T, Grinvald A, et al. Linking spontaneous activity of single cortical neurons and the underlying functional architecture. *Science* 1999;286:1943–1946.

168. Singer W. Putative functions of temporal correlations in neocortical processing. In: Koch C, David JL, eds. *Large scale neuronal theories of the brain*. MIT Press, 1994:201–238.

26

The Spatial, Temporal, and Interpretive Limits of Functional MRI

Peter A. Bandettini

Peter A. Bandettini: Unit on Functional Imaging Methods, Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, Maryland.

Since the inception of functional magnetic resonance imaging (fMRI) in 1991, an explosive growth in the number of users has been accompanied by steady widening of its range of applications. A recent search of the National Library of Medicine database for articles with *fMRI* or *BOLD* (blood oxygenation-dependent) in the title revealed more than 1,000 citations. Improvements continue in pulse sequence design, data processing, data interpretation, and the tailoring of cognitive paradigms to the unique advantages and limits of the technique. This chapter describes the receding limits of fMRI. Specifically, the limits of spatial resolution, temporal resolution, interpretability, and implementation are discussed. The goal is to give the reader a perspective of the evolution of fMRI in the past 9 years and a sense of excitement regarding its ultimate potential.

A user of fMRI primarily is interested in extracting at least one of three types of neuronal information: where neuronal activity is happening, when it is happening, and the degree to which it is happening. To extract this information optimally, an understanding of the basics of some of the more esoteric details is necessary, which are presented in this chapter. First, the basics of fMRI contrast are discussed. Second, the key of fMRI interpretation, the *neuronal–hemodynamic transfer function*, is described. Third, an overview of methods by which neuronal activation is played out in fMRI subjects and subsequently measured is provided. In this section, the popular technique of *event-related fMRI* (ER-fMRI) is described in detail, along with emerging methods of neuronal information extraction. Fourth, the issues of temporal and spatial resolution are discussed. Fifth, the limits of interpretation and the potential for further neuronal–hemodynamic information extraction are discussed. Lastly, some implementation limits are finally given as a practical guideline.

CONTRAST IN fMRI

Part of "26 - The Spatial, Temporal, and Interpretive Limits of Functional MRI "

Several types of physiologic information can be mapped with fMRI. This information includes baseline cerebral blood volume (1 ,2 and 3), changes in blood volume (4), baseline and changes in cerebral perfusion (5 ,6 ,7 ,8 ,9 and 10), and changes in blood oxygenation (11 ,12 ,13 ,14 ,15 ,16 and 17). Recent advances in fMRI pulse sequence and experimental manipulation have allowed quantitative measures of cerebral metabolic rate of oxygen ($CMRO_2$) changes and dynamic, noninvasive measures of blood volume with activation to be extracted from fMRI data (18 ,19 and 20).

Blood Volume

In the late 1980s, the use of rapid MRI allowed tracking of transient signal intensity changes over time. One application of this utility was to follow the T2*- or T2-weighted signal intensity as a bolus of an intravascular paramagnetic contrast agent passed through the tissue of interest (2). As it passed through, susceptibility-related dephasing increased then decreased as the bolus washed out. The area under these signal attenuation curves is proportional to the relative blood volume. In 1990, Belliveau and colleagues (4) took this technique one step further and mapped blood volume during rest and during activation. The first maps of brain activation obtained with fMRI were demonstrated with this technique. As soon as the technique was demonstrated, it was rendered obsolete (for brain activation imaging) by the use of an endogenous and oxygen-sensitive contrast agent—hemoglobin.

Blood Oxygenation

As early as the 1930s, it was known that hemoglobin is paramagnetic and deoxyhemoglobin is diamagnetic (21). In 1982, it was discovered that changes in blood oxygenation change the T2 of blood, but it was not until 1989 that

this knowledge was used to image *in vivo* changes in blood oxygenation (22). Blood oxygenation-dependent contrast, coined *BOLD contrast* by Ogawa et al. (23), was used to image the activated brain for the first time in 1991. Interestingly, Ogawa et al. predicted its utility for functional brain imaging; however, they predicted a signal *decrease* rather than a signal *increase*, as implied by some earlier positron emission tomography (PET) results by Fox and Raichle (24) suggesting that the oxygen extraction fraction decreased during activation. The first results of the use of BOLD contrast were published in 1992 (13 ,15 ,23). Because of its sensitivity and ease of implementation, the contrast used to observe susceptibility changes with changes in blood oxygenation is the most commonly used functional brain imaging contrast, and this is the technique primarily discussed in this chapter. The cascade of events that follow brain activation and lead to BOLD signal changes is shown in Fig. 26.1 .

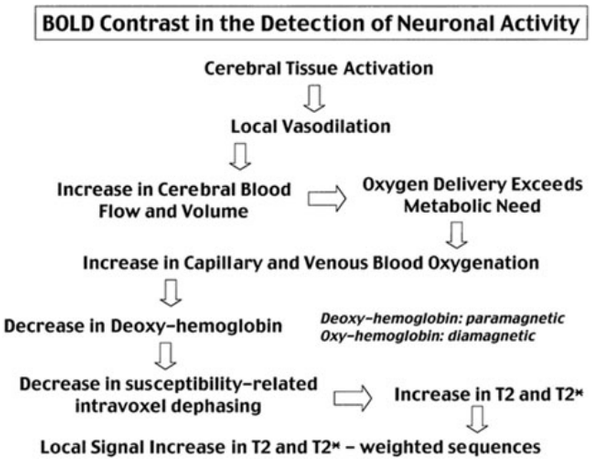


FIGURE 26.1. The cascade of hemodynamic and magnetic resonance imaging events that occur following brain activation.

Blood Perfusion

An array of new techniques exist for mapping cerebral blood *perfusion* in humans. Arterial spin labeling-based perfusion mapping MRI techniques are similar to those applied in other modalities, such as PET and single-photon emission computed tomography (SPECT); in-flowing blood is tagged and then allowed to flow into the imaging plane. The radiofrequency (RF) tagging pulse is usually a 180-degree pulse that "inverts" the magnetization.

Generally, these techniques can be subdivided into those that use continuous arterial spin labeling, which involves continuously inverting blood flowing into the slice, and those that use pulsed arterial spin labeling, which periodically inverts a block of arterial blood and measures the arrival of that blood into the imaging slice. Examples of these techniques are *echo-planar imaging with signal targeting and alternating RF* (EPISTAR) (25) and *flow-sensitive alternating inversion recovery* (FAIR) (10 ,26). Recently, a pulsed arterial spin-labeling technique known as *quantitative imaging of perfusion using a single subtraction* (QUIPSS) has been introduced (27).

Hemodynamic Specificity

With each of the above-mentioned techniques for imaging volume, oxygenation, and perfusion changes, the precise type of observable cerebrovascular information can be more finely delineated. Although this information is typically more than the cognitive neuroscientist requires, it is useful to give an abbreviated summary of how specific MRI can actually be. Regarding susceptibility contrast imaging, spin-echo sequences are more sensitive to small susceptibility compartments (capillaries and red blood cells), and gradient-echo sequences are sensitive to susceptibility compartments of all sizes (28 ,29 ,30 and 31). Outer-volume RF saturation removes in-flowing spins (32), thereby reducing non–susceptibility-related inflow changes when short-repetition time (with high flip angle) sequences are used. Diffusion weighting or "velocity nulling," involving the use of $b > 50$ s/mm^2, reduces the intravascular signal (33), thereby reducing, but not eliminating, large-vessel effects (intravascular effects are removed but extravascular effects remain) in gradient-echo fMRI and all large-vessel effects in spin-echo fMRI. Performing BOLD contrast fMRI at high field strengths has the same effect as diffusion weighting in the context of susceptibility-based contrast because the T2* and T2 of venous blood becomes increasingly shorter than the T2* and T2 of gray matter as field strength increases; therefore less signal arises from venous blood at higher field strengths. (34). Figure 26.2 is a schematic diagram summarizing the pulse sequence selectivity of the specific aspects of the vasculature.

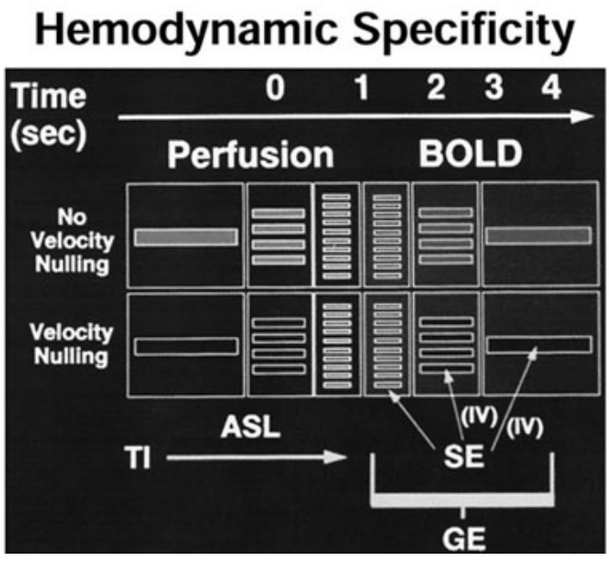


FIGURE 26.2. The vascular tree, including arteries *(left)* and arterioles, capillaries, and veins *(right)*. If the inside of the vessel drawing is filling in, the signal has an intravascular contribution. Arterial spin labeling (ASL) is differentially sensitive to the arterial–capillary region of the vasculature, depending on the inversion time (TI) used and whether or not velocity nulling (otherwise called *diffusion weighting*) gradients are used. A small amount of velocity nulling and a TI of about 1 s make ASL techniques selectively sensitive to capillaries. Susceptibility-based techniques, including gradient-echo and spin-echo, are also differentially sensitive to specific aspects of the vasculature. Gradient-echo techniques are sensitive to susceptibility perturbers of all sizes; therefore, they are sensitive to all intravascular and extravascular effects. Spin-echo techniques are sensitive to susceptibility perturbers about the size of a red blood cell or capillary, so that they are sensitive to intravascular effects in vessels of all sizes and to extravascular capillary effects. Velocity nulling makes gradient-echo sequences sensitive to extravascular capillary-to-vein effects, and makes spin-echo sequences selectively sensitive only to capillary effects. See color version of figure.

Cerebral Metabolic Rate of Oxygen

Recently, advances in mapping activation-induced changes in the $CMRO_2$ with fMRI have been developed (18 ,20 ,35 ,36 and 37). The basis for such measurement is that BOLD and perfusion contrast can be explained by the combination of a handful of parameters. The key, then, is either to constrain the contrast or manipulate the physiologic state such that the number of parameters reduces to one or two. Normalization by means of a hypercapnia has evolved as a method for mapping changes in $CMRO_2$ (18). The basic idea is that when the brain is activated, increases in flow, volume, and oxygenation are accompanied by an increase in $CMRO_2$. When a subject at rest is undergoing a hypercapnic stress (5% CO_2), the cerebral flow, volume, and oxygenation increase without an accompanying increase in $CMRO_2$; therefore,

less oxygen is extracted from the blood stream, so that the blood oxygenation change, relative to the perfusion change, is greater than with brain activation. By comparing the ratio of the (simultaneously measured) perfusion and BOLD signal changes during hypercapnia and during brain activation, $CMRO_2$ information can be derived.

HEMODYNAMIC TRANSFER FUNCTION

Part of "26 - The Spatial, Temporal, and Interpretive Limits of Functional MRI "

The hemodynamic transfer function is referred to here as the combined effect on the fMRI signal change by the spatial and temporal variation in neuronal–vascular coupling, blood volume, blood flow, blood oxygenation, hematocrit, and vascular geometry, among other things. A goal of fMRI method development is to characterize this transfer function completely (i.e., its spatial, temporal, pulse sequence, subject, physiologic, and pharmacologic dependencies), so that more precise inferences can be made about underlying neuronal activation location, magnitude, and timing. The ultimate limits of fMRI depend on this characterization. This goal is particularly relevant in the context of understanding pharmacologic effects on brain function.

After the onset of activation, or rather after the neuronal firing rate has passed an integrated temporal–spatial threshold, either direct neuronal, metabolic, or neurotransmitter-mediated signals reach arteriole sphincters and cause dilation. The time for this initial process to occur is likely to be less than 100 ms. After vessel dilation, the blood flow rate increases by 10% to 200%. The time for blood to travel from arterial sphincters through the capillary bed to pial veins is about 2 to 3 s. This transit time determines how rapidly the blood oxygenation saturation increases in each part of the vascular tree. As shown in Fig. 26.2 , depending on the pulse sequence used, different aspects of the hemodynamics are manifested.

Location

In resting state, hemoglobin oxygen saturation is about 95% in arteries and 60% in veins. The increase in hemoglobin saturation with activation is largest in veins, changing from about 60% to 90%. Likewise, capillary blood saturation changes from about 80% to 90%. Arterial blood, already saturated, shows no change. This large change in saturation is one reason why the strongest BOLD effect is usually seen in draining veins.

The second reason why the strongest BOLD effect is seen in draining veins is that activation-induced BOLD contrast is highly weighted by blood volume in each voxel. Because capillaries are much smaller than a typical imaging voxel, most voxels, regardless of size, likely contain about 2% to 4% capillary blood volume. In contrast, because the size and spacing of draining veins are on the same scale as most imaging voxels, it is likely that veins dominate the relative blood volume in any voxel that they pass through. Voxels that pial veins pass through can have 100% blood volume, whereas voxels that contain no pial veins may have only 2% blood volume. This stratification in blood volume distribution strongly determines the magnitude of the BOLD signal.

As suggested in Fig. 26.2 , different pulse sequence weightings can give different locations of activation. For example, Fig. 26.3 shows the activation in the motor cortex with two different functional MRI contrast weightings collected in the same plane—perfusion and BOLD. Although much overlap is seen, the hot spots vary by as much as 10 mm. The perfusion change map is sensitive primarily to *capillary* perfusion changes, whereas the BOLD contrast activation map is weighted mostly by veins. A potential worry regarding fMRI location is that venous blood, flowing away from the activated area, may maintain its elevated oxygen

saturation as far as a centimeter away. When brain activation is observed on a scale of centimeters, this has not been a major concern. Nevertheless, this issue is discussed in detail later in the chapter.

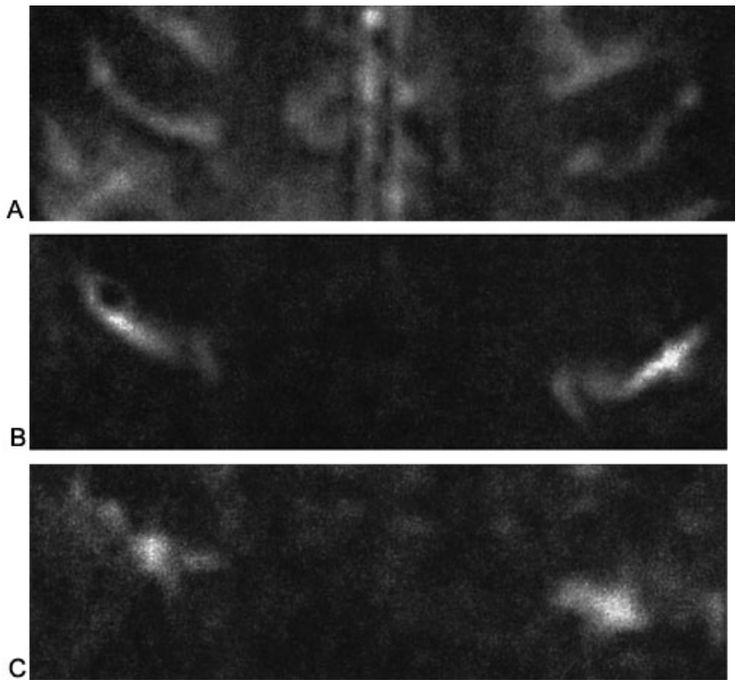


FIGURE 26.3. Comparison of activation-induced signal changes in perfusion and BOLD (blood oxygenation-dependent) measurements. Both measurements were obtained at 3 T. Perfusion measurements were obtained by using FAIR-EPI (flow-sensitive alternating inversion recovery echo-planar imaging) with an inversion time of 1,400 ms. BOLD (blood oxygenation-dependent) measurements were obtained by using gradient-echo EPI with an echo time of 30 ms.

Latency

One of the first observations made regarding fMRI signal changes is that after activation, the BOLD signal takes about 2 to 3 s to begin to deviate from baseline (16 ,38). Because the BOLD signal is highly weighted toward venous oxygenation changes, with a flow increase, the time for venous oxygenation to begin to increase is about the time that it takes blood to travel from arteries to capillaries and draining veins, which is 2 to 3 s. The hemodynamic “impulse response” function has been effectively used to characterize much of the BOLD signal change dynamics (39 ,40 and 41). This function has been derived empirically by performing very brief and well-controlled stimuli. In addition, it can be derived by deconvolving the neuronal input from the measured hemodynamic response (42 ,43). This type of analysis assumes that the BOLD response behaves in a manner that can be completely described by linear systems analysis, which is still an open issue. Regardless, observed hemodynamic response to any neuronal activation can be predicted with a reasonable degree of accuracy by convolving expected neuronal activity timing with the BOLD “impulse response” function. This function has typically been mathematically described by a γ variate function (39).

If a task onset or duration is modulated, the accuracy to which one can correlate the modulated input parameters to the measured output signal depends on the variability of the signal within a voxel or region of interest. In a study by Savoy et al. (44) addressing this issue, variability of several temporal sections of an activation-induced response was determined. Six subjects were studied, and for each subject, 10 activation-induced response curves were analyzed. The relative onsets were determined by finding the latency with which the correlation coefficient was maximized with each of three reference functions representing three parts of the response curve: the entire curve, the rising section, and the falling section. The standard deviations of the whole curve, rising phase, and falling phase were found to be 650, 1,250, and 450 ms, respectively.

Across-region differences in the onset and return to baseline of the BOLD signal during cognitive tasks have been observed. For example, during a visually presented event-related word stem completion task, Buckner et al. (45) reported that the signal in visual cortex increased about 1 second before the signal in the left anterior prefrontal cortex. One might argue that this observation makes sense from a cognitive perspective because the subject first observes the word stem and then, after about a second, generates a word to complete this task. Others would argue that the neuronal onset latencies should not be more than about 200 ms. Can inferences of the cascade of brain activation be made on this time scale from fMRI data? Without a method to constrain or work around the intrinsic variability of the onset of BOLD signal over space, such inferences should not be made in temporal latency differences below about 4 s.

Lee et al. (46) were the first to observe that the fMRI signal change onset within the visual cortex during simple visual stimulation varies from 6 to 12 s. These latencies were also shown to correlate with the underlying vascular structure. The earliest onset of the signal change appeared to be in gray matter, and the latest onset appeared to occur in the largest draining veins. Similar latency dispersions in motor cortex have been observed. In one study, latency differences, detected in visual cortex with the Hilbert transform, did not show a clear correlation of latency with evidence for draining veins (47).

Figure 26.4 is a summary of the sources of temporal variability. Figure 26.4A shows a plot of the average time course from the motor cortex as a result of 2-second finger tapping. As mentioned, the first source of variability is the intrinsic noise in the time series signal. The standard deviation of the signal is on the order of 1%. The second source of variability is that of the hemodynamic response. As mentioned, this ranges from 450 to 1,250 ms, depending on whether one is observing the rising phase of the signal or

the falling phase. The third source of variability is the latency spread over space.

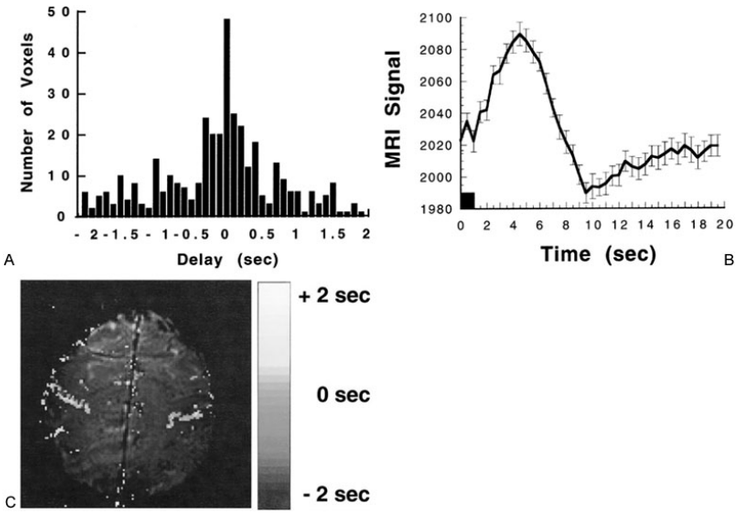


FIGURE 26.4. Demonstration of several of the limits of functional magnetic resonance imaging temporal resolution. Echo-planar imaging was performed at 3 T by using a Bruker Biospec 3T/60 equipped with a local head gradient coil. A time course series of axial images (matrix size = 96 × 96, field of view = 20 cm, echo time = 40 ms, repetition time = 500 ms, flip angle = 80 degrees) through the motor cortex was obtained. Bilateral finger tapping was performed for 2 s, followed by 18 s of rest. These figures demonstrate that the upper temporal resolution is determined by the variability of the signal change in time and space. **A:** Time course of the signal elicited by tapping fingers for 2 s. The standard deviation at each point is in the range of 1% to 2%. The standard deviation of the hemodynamic change, in time, is in the range of 450 to 650 ms. **B:** Map of the dot product (a measure of the activation-induced signal change magnitude) and the relative latencies or delays of the reference function (the plot in **A** was used as the reference function) at which the correlation coefficient was maximized. The spatial distribution of hemodynamic delays has a standard deviation of about 900 ms. The longest delays approximately match the regions that show the highest dot product and the area where veins are shown as *dark lines* in the T2*-weighted anatomic image. **C:** Histogram of relative hemodynamic latencies. This was created from the latency map in **(B)**.

The plot in Fig. 26.4A was used a reference function for correlation analysis and allowed to shift ± 2 s. Figure 26.4B is a histogram of a number of voxels in an activated region that demonstrated a maximum correlation with the reference function at each latency (relative to the average latency) to which the reference function was shifted. The spread in latencies is more than 4 s. Figure 26.4C includes a map of dot product (measure of signal change magnitude) and latency; the regions showing the longest latency roughly correspond to the regions that show the largest signal changes. Although these largest signal changes are likely downstream draining veins, it is important to note that this approximate correlation between latency and magnitude is extremely weak. Many very small signal changes show very long latencies. It is also interesting to note that the inverse, that many large signal changes show short latencies, is typically not true. This implies that many downstream vessels may be almost fully diluted back to resting state oxygenation, therefore showing only a small signal change but still a large latency. Again, work is ongoing to characterize this effect better.

Magnitude

As previously discussed, the magnitude of the fMRI signal change is influenced by many variables across subjects, neuronal systems, neuronal systems, and voxels. Making a complete and direct correlation between neuronal activity and fMRI signal change magnitude in a single experiment will remain impossible until all the variables can be characterized on a voxel-related basis. Because of these primarily physiologic variables, the magnitude of BOLD signal changes on brain activation maps typically ranges from 1% to 5% [at, say, 1.5 tesla (T), gradient-echo sequence, echo time of 40 ms]. The picture is not that bleak, though. In the past several years, considerable progress has been made in characterizing

the magnitude of the fMRI signal changes with underlying neuronal activity.

The progressive series of studies was as follows: First, as mentioned previously, it was clear that areas that showed significant BOLD signal change were in the appropriate neuronal area corresponding to specific, well-characterized tasks. Second, inferred neuronal modulation was carried out by systematically varying some aspect of the task. Clear correlations between BOLD signal change magnitude and visual flicker rate, contrast, word presentation rate, and finger tapping rate were observed (13 ,48 ,49 and 50). This parametric experimental design represented a significant advance in the manner in which fMRI experiments were performed, enabling more precise inferences, not about the BOLD signal change with task modulation. Nevertheless, of course, the degree of neuronal activation (i.e., integrated neuronal firing over a specified area) was still inferred.

Recently, several more intriguing studies have emerged correlating measured neuronal firing rate with well-known stimuli in animals (51) and humans (52 ,53) and demonstrating a remarkably high correlation between BOLD signal change and electrophysiologic measures.

Linearity

Related to the topic of signal change magnitude is that of BOLD signal change linearity. It has been found that, with very brief stimulus durations, the BOLD response shows a larger signal change magnitude than expected if one assumes that it behaves as a linear system (54 ,55). This greater than expected BOLD signal change is generally specific to stimulus durations below 3 s. Reasons for nonlinearities in the event-related response can be neuronal, hemodynamic, or metabolic in nature. The neuronal input may not be a simple boxcar function. Instead, an increased neuronal firing rate at the onset of stimulation (neuronal “bursting”) may cause a slightly larger amount of vasodilation that later plateaus at a lower steady-state level. The amount of neuronal bursting necessary to change the hemodynamic response significantly, if a linear neuronal-hemodynamic coupling is assumed, is quite large. For example, to account for the almost double functional contrast for the experimental relative to the linear convolution-derived single-event responses, the integrated neuronal response greater than 2 s must double. If it is assumed that neuronal firing is at a higher rate only for about the first 50 ms of brain activation, the neuronal firing rate must be 40 times greater than steady state for this duration.

BOLD contrast is highly sensitive to the interplay of blood flow, blood volume, and oxidative metabolic rate. If, with activation, any one of these variables changes with a different time constant, the fMRI signal may show fluctuations until a steady state is reached (56 ,57). For instance, an activation-induced increase in blood volume would slightly reduce the fMRI signal because more deoxyhemoglobin would be present in the voxel. If the time constant for blood volume changes were slightly longer than that of flow changes, then the activation-induced fMRI signal would first increase, then be reduced as blood volume later increased. The same could apply if the time constant of the oxidative metabolic rate were slightly slower than that of flow and volume changes. Evidence for an increased oxidative metabolic rate after 2 min of activation is given by Frahm et al. (57), but no evidence suggests that the time constant of the increase in oxidative metabolic rate is only seconds longer than the flow increase time constant—as would be required to be applicable only to relatively high-amplitude single-event responses. These hemodynamics, which may also differ on a voxel-related basis, remain to be characterized fully.

SCANNER-RELATED ISSUES

Part of "26 - The Spatial, Temporal, and Interpretive Limits of Functional MRI "

A complete discussion of all scanner-related issues and potential solutions is beyond the scope of this chapter. A rudimentary yet necessary description of the most basic problems and solutions is presented in the following sections. Most practitioners of functional MRI typically undergo a painful, frustrating, and prolonged period of learning about all scanner-related limitations and issues. Some give up hope completely. Those who are determined usually emerge hopeful again, and with a much better “feel” for what can and cannot be done in regard to brain imaging. This learning process also applies to understanding the physiology of the signal, but typically the greatest anguish arises in the context of MRI pulse sequences and hardware.

In the first place, all the categories listed below are more or less linked. In this section, an attempt is made to walk the reader informally through the trade-offs and issues involved in performing an fMRI experiment.

Acquisition Rate

Image acquisition rate is ultimately limited by how fast the signal can be digitized and by how rapidly the imaging gradients can be switched. MRI can be logically divided into single-shot and multishot techniques. In single-shot imaging, spins are excited with a single excitation pulse and all the data necessary to create an image are collected at once. Echo-planar imaging (EPI) is the most common single-shot technique; with one “echo,” a single “plane” is acquired. Multi-shot techniques are most commonly used for high-resolution anatomic imaging. In clinical scanning (with multi-shot imaging), a single “line” of raw data is acquired with each RF excitation pulse. Because of the relatively long time it takes for the longitudinal magnetization to return to equilibrium (characterized by the T1 of the tissue), a certain amount of time, between 50 and 500 ms, is spent waiting between shots; otherwise, soon no signal would be

left. It would be “saturated.” Because of this necessary waiting time, multishot techniques are typically slower than single-shot techniques. For a 150-ms “waiting time,” or repetition time (TR), an image with 128 lines of raw data would take 150 ms multiplied by 128, or 19.2 s.

In the case of EPI, the entire data set for a plane is typically acquired in about 20 to 40 ms. In the context of performing a BOLD experiment, the echo time (TE) is about 20 to 40 ms. Along with some additional time for applying other necessary gradients, the total time for an image to be acquired is about 60 to 100 ms, so that 10 to 16 images can be acquired in a second. Improvements in digital sampling rates and gradient slew rates will allow small improvements, but essentially, this is about the upper limit for imaging humans.

In the context of an fMRI experiment with EPI, the typical image acquisition rate is determined by how many slices can temporally fit into a TR. For whole-brain imaging, approximately 20 slices (5-mm thickness) are required to cover the entire brain. This allows a TR of about 1.25 to 2 s at minimum. This image collection rate is more than adequate to capture most of the details of the slow and dispersed hemodynamic response.

Spatial Resolution

The spatial resolution is also primarily determined by gradient strength, digitizing rate, and time available. For multishot imaging, as high a resolution as desired can be achieved if one is willing to wait. One can keep on collecting lines of data with more RF pulses. For EPI, the signal decay rate (described by T2* with gradient-echo EPI and by T2 with spin-echo EPI) plays a significant role in determining the resolution. One can sample for only so long before the signal has completely decayed away. For this reason, the resolution of EPIs is generally much lower than that of multishot images. To get around this problem, two further strategies are commonly used. The first is multishot EPI, in which the full EPI acquisition is used multiple times (but many fewer times than in typical clinical multishot imaging) and interleaved to increase the resolution. The second is to perform an operation called *conjugate synthesis*, which basically makes use of the fact that, in raw data space, half of the data is redundant. This allows at most twice the resolution, with some cost in signal to noise and image quality. An example of multishot EPI is shown in Fig. 26.5 .

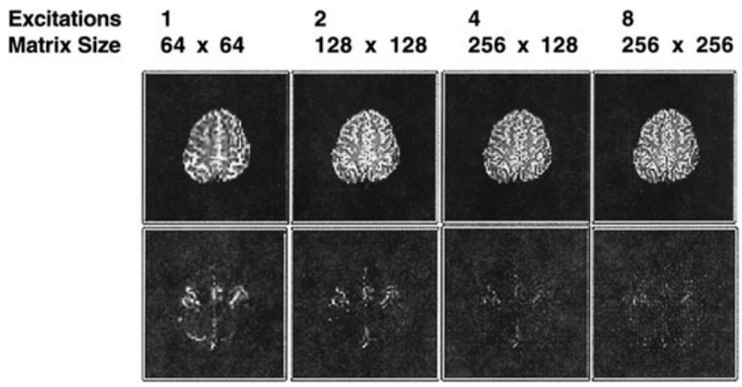


FIGURE 26.5. An example of multishot echo-planar imaging. Number of excitations ranged from 1 to 8. The image resolution increases but the signal to noise and functional contrast to noise decrease. In addition, instabilities are introduced into the time course by the use of multishot imaging.

Signal to Noise

The signal to noise and the functional contrast to noise are influenced by many variables. These include, among others, voxel volume, TE, TR, flip angle, receiver bandwidth, field strength, and RF coil used. Not considering fMRI for a moment, the image signal to noise is increased with larger voxel volume, shorter ET, longer RT, larger flip angle (to 90 degrees), narrow receiver bandwidth, higher field strength, and smaller RF coil. That said, in the context of fMRI, the functional contrast to noise is optimized with a voxel volume equaling the size of the activated area, TE ≈ gray matter T2*, short TR (optimizing samples per unit time), flip angle = Ernst angle = Cos (-TR/T1), narrow receiver bandwidth, high field strength, and smaller RF coil. Of course, all of these variables come at some expense to others.

Stability

Theoretically, the noise, if purely thermal in nature, should propagate similarly over space and across time. In fMRI, this is not at all the case because each image is essentially captured in 40 ms and the time series is collected in minutes. Stability is much more of an issue on the longer time scale.

Flow and motion are correlated in many areas with cardiac and respiratory cycles. Subject movement and scanner instabilities also contribute. Single-shot techniques have generally better temporal stability than multishot techniques because, with multishot techniques, the image is collected over a larger time scale; instabilities on a longer time scale enter into the image creation itself. This leads to nonrepeatable ghosting patterns that generally decrease temporal signal to noise ratio. Work is ongoing to characterize and reduce temporal instabilities for both single-shot and multishot imaging techniques (58 ,59). Correction techniques include cardiac gating (60), navigator pulses (61 ,62), and raw data reordering (63 ,64).

Image Quality

Image quality issues that are the most prevalent are image warping and signal dropout. Although books can be written on this subject, the description here is limited to the bare essentials.

Image warping is fundamentally caused by either or both of two factors, B_o-field inhomogeneities and gradient nonlinearities. A nonlinear gradient causes nonlinearities in spatial encoding, so that the image is distorted. This is primarily a problem when local, small-gradient coils are used that have a small region of linearity that drops off rapidly at the base and top of the brain. With the growing prevalence of whole-body gradient coils for performing EPI, this problem is no longer a major issue. If the B_o field is inhomogeneous, as is typically the case with imperfect shimming procedures, particularly at higher field strengths, the protons are processing at frequencies different from those expected in their particular location. This causes image deformation in the areas of poor shim, particularly with the long readout window or acquisition time of EPI. A solution is either to shim better (65 ,66) or map the B_o field and perform a correction based on this map (67).

Signal dropout is related to the problem described above in that it is also caused by localized B_o-field inhomogeneities, typically at interfaces of tissues having different susceptibilities. If within a voxel, because of the B_o inhomogeneities, spins are of different frequencies, their signals cancel each other out. Several strategies exist for reducing this problem. One is, again, to shim as well as possible at the desired area. Because of still imperfect shimming procedures, this is usually not satisfactory. The other is to reduce the voxel size (increase the resolution), so that stratification of different frequencies is reduced within a voxel. The third is to choose the slice orientation such that the smallest voxel dimension (in many studies, the slice thickness is greater than the in-plane voxel dimension) is perpendicular to the largest B_o gradient. For this reason, many studies are performed with the use of sagittal or coronal slice orientations.

As with many of the topics discussed, much more can be said, but the goal here is simply to provide an introduction and references to additional reading material.

BEST RESULTS SO FAR

Part of "26 - The Spatial, Temporal, and Interpretive Limits of Functional MRI "

The primary discussion up to this point has focused on the limits imposed by the scanner and the hemodynamics. In this section, some of the most successful, thought-provoking, and innovative fMRI studies, from a methodologic perspective, performed as of September 2000 are discussed. The best results in temporal and spatial resolution are presented.

Temporal Resolution

As explained in previous sections, MR images can be acquired at an extremely rapid rate; therefore, scanner-related limits are not the prime determinant of the upper limits of temporal resolution in fMRI. The key to increasing the temporal resolution in fMRI is either to characterize the hemodynamic response better or to work around its limits. The results described here are examples of this work during the past few years.

To obtain information about relative onsets of cascaded neuronal activity from hemodynamic latency maps, it is possible to determine *relative* latency changes on modulation of the task timing. In a study of Savoy et al. (68), activation onset latencies of 500 ms were discernible when they used a visual stimulation paradigm in which the left and right hemifields were stimulated at relative delays of 500 ms. First, the subject viewed a fixation point for 10 s. Then, the subject's left visual hemifield was activated 500 ms before the right. Both hemifields were activated for 10 s, then the left hemifield stimulus was turned off 500 ms before the right.

Although with careful choice of region of interest, from which the time course plot is made, these onset differences can be shown, maps of latency cannot reveal the onset differences because, as mentioned, the variability over space, which is about 4 s, dominates the inserted 500-ms variability from left to right hemifield.

To *map* the relative latency differences between hemifields, it is necessary to modulate the relative stimulation timing. As an extension of their results, the left—right onset order was switched so that, in the second run, the right hemifield was activated and turned off 500 ms before the left. Latency maps were made for each onset order and subtracted from each other to reveal a clear delineation between the right and left hemifields that was not apparent in each of the individual maps. This operation and the resulting *relative* latency map is shown in Fig. 26.6 . These maps are of the change in onset of one area relative to another, not of absolute latency. It is also useful to note that the standard deviation of these maps is reduced simply to the standard

deviation of the latencies in each voxel, not the standard deviation of the latencies over space. Maps such as these can be extremely useful in determining which regions of activation are modulated relative to other areas with a specific task timing modulation.

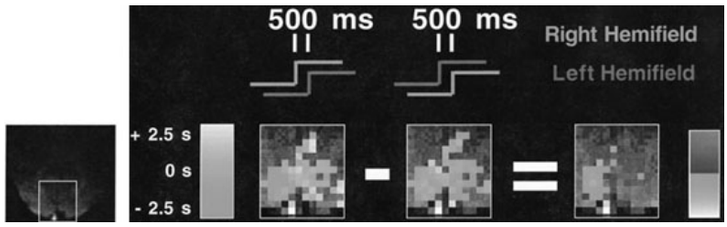


FIGURE 26.6. The use of latency maps and task modulation to extract relative latencies. Activation within a region of visual cortex is shown. In one condition *(left)*, the right visual hemifield stimulation precedes the left by 500 ms. In the other condition *(middle)*, the left visual hemifield precedes the right hemifield stimulation by 500 ms. Latency maps from both these conditions show an across-voxel spread of ± 2.5 s, which is too large to identify clearly the relative latencies across hemifields. However, once the data are normalized for this intrinsic variance by directly comparing the hemodynamic response from the two different lags within individual voxels, the offset between the left and right hemifields can be observed *(right)*. This demonstration suggests that normalization of the hemodynamic lag can allow small relative temporal offsets to be identified. These normalized offsets can then be compared across regions to make inferences about neuronal delay. For this experiment, the repetition time was 400 ms. See color version of figure.

Published work by Menon et al. (69), Kim et al. (70), Richter et al. (71), and Bandettini (72) has explored the temporal resolution limits of fMRI. The results of Menon et al. (69), similar to those mentioned above, indicate that relative brain activation timings on the order of 50 ms can be discerned.

In the study of Richter et al. (71), a parametrically varied event-related mental rotation task was used. Each mental rotation task was presented individually. A high correlation between task duration and event-related width was demonstrated. The longer the task took to accomplish (larger rotation angle), the wider the event-related response was shown to be in specific parietal locations.

Spatial Resolution

The hemodynamic point spread function was first considered and characterized by Engel et al. (73 ,74 and 75). Localization to 1.1 mm was determined.

The first successful mapping of ocular dominance columns in humans was published by Menon et al. (76). Their intriguing results show that the optimal way to pull out differences in activation across closely spaced units is to perform very brief stimuli so as not to reduce the dynamic range of the oxygenated blood that is flowing away beyond the unit of activation.

In regard to MRI pulse sequence, it is important to note that mapping cortical columns multishot imaging is required (76). Performance of multishot imaging requires either navigator echoes or shot-to-shot phase-correction schemes. If these are not performed, temporal stability is seriously compromised.

Many have argued that some aspects of the BOLD signal change dynamics are more spatially localized to neuronal activity. Specifically, the evasive “pre-undershoot” has been indicated as such (77 ,78). The rationale is that this transient “dip” in the signal that occurs 0.5 to 2 s after the onset of activation and quickly gives way to the much larger signal increase is secondary to direct extraction of oxygen from the blood by adjacent activated tissue. Recently published work has demonstrated the utility of such an approach for mapping cortical columns in animals (79 ,80 and 81).

The highest-resolution fMRI performed with single-shot EPI was obtained by Jesmanowicz et al. (82). Here, a partial k-space strategy was used to obtain a presumed 256 × 256 resolution. The actual resolution achieved is debatable because T2* effects reduced the resolution below that implied by the matrix size.

NEURONAL ACTIVATION INPUT STRATEGIES

Part of "26 - The Spatial, Temporal, and Interpretive Limits of Functional MRI "

Much of this chapter has been devoted to the basics and some esoteric issues regarding fMRI. This section provides an overview of the types of neuronal input strategies with which fMRI has been used to extract information about what the brain is doing. Given a question of brain function, what are the available strategies that one can use to design their paradigm? A schematic summary of these strategies appears in Fig. 26.7 .

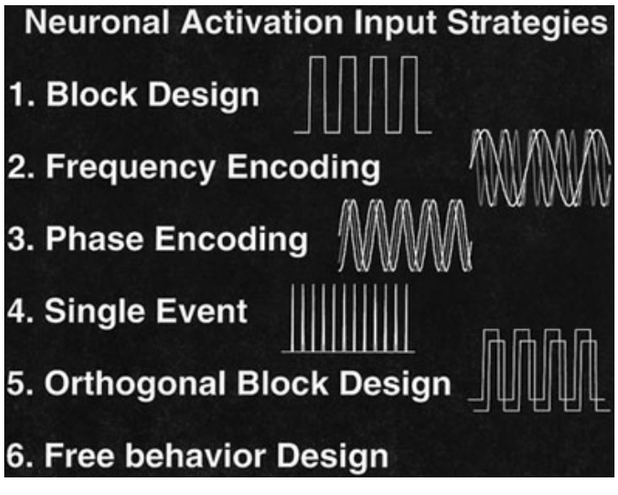


FIGURE 26.7. Overview and schematic depiction of types of neuronal input strategies available in functional magnetic resonance imaging. In addition, parametric designs, which involve systematically varying some aspect of the intensity of the neuronal input, can be applied to any of the design strategies.

Block Design

A block design paradigm was the first used in fMRI and is still the most prevalent neuronal input strategy. Borrowed

from previous PET studies, it involves having a subject alternately perform a task for at least 10 s, then a control task for a similar time, so that the hemodynamic response reaches a steady state in each condition. This is a useful technique in that it is easy to implement, and standard statistical tests can be used to compare each condition.

Phase and Frequency Encoding

Phase encoding of a stimulus involves varying some aspect of the stimulus in a continuous and cyclic manner. This strategy has been most successfully used in retinotopic mapping (75 ,83 ,84). In this type of study, the eccentricity of a visual stimulus ring is continuously varied; then, after the most extreme eccentricity is reached, the cycle is repeated. The data are typically subjected to Fourier analysis, and the areas that show a signal change temporal phase that correlates with the stimulus phase are mapped. This is a powerful technique because it makes use of the entire time series, in that there are no “off” states, and lends itself to Fourier analysis. This method has also been used for somatotopic mapping (85) and tonotopic mapping (86).

Frequency encoding is much less common but can be achieved for certain types of stimuli. The method is to designate a specific on–off frequency for each type of stimulus used. Again, Fourier analysis reveals the most power under a spectral peak corresponding to the brain area specific to the particular on/off frequency. The utility of this method has been demonstrated in mapping left and right motor cortex by cueing the subject to perform a finger-tapping task at different on–off rates for each hand (87).

Orthogonal Designs

Orthogonal task design is a powerful extension of block design. The basic concept is that if one designs two different task timings to create BOLD responses that are orthogonal to each other, then these tasks can be performed simultaneously during a single time series collection with no cross-task interference, so that comparison is much more precise. This technique was pioneered by Courtney et al. (88). In their study, six orthogonal tasks were designed into a single time series. This type of design also lends itself to event-related studies.

Parametric Designs

As mentioned in the section on magnitude, parametric designs are powerful in that more precise statements can be made about relative neuronal activity. A parametric task design simply involves systematically varying some aspect of the task during the time series. This may be a finger-tapping rate, stimulus contrast or flicker rate, cognitive load, or attention demand, and instead of mapping the magnitude of the change with a single task, the slope of the change corresponding to a task is mapped. In this manner, relative brain activation magnitude may be teased out of the time series.

Event-Related Designs

Before 1995, a critical question in event-related fMRI was whether a transient cognitive activation could elicit a significant and usable fMRI signal change. In 1996, Buckner et al. (45) demonstrated that event-related fMRI in fact lends itself quite well to cognitive neuroscience questions. In their study, a word stem completion task was performed; a block design and an event-related strategy were used. Robust activation in the regions involved with word generation were observed in both cases.

Given the substantial number of recent reports of event-related fMRI (40 ,41 and 42 ,65 ,89 ,90 ,91 ,92 ,93 ,94 ,95 ,96 ,97 ,98 ,99 ,100 ,101 ,102 ,103 ,104 ,105 ,106 ,107 ,108 ,109 ,110 ,111 and 112), it can probably be said that this is one of the more exciting developments in fMRI since its discovery.

The advantages of event-related activation strategies are many (113). These include the ability to randomize task types in a time series more completely (114 ,115 and 116), the ability to analyze fMRI response data selectively based on measured behavioral responses to individual trials (111), and the option to incorporate overt responses into a time series. Separation of motion artifact from BOLD changes is possible by use of the temporal response differences between motion effects and the BOLD contrast-based changes (91).

When a constant ISI is used, the optimal interstimulus interval (ISI) is about 10 to 12 s. Dale and Buckner (43) have shown that responses to visual stimuli, presented as rapidly as once every 1 s, can be adequately separated by

using overlap correction methods. Overlap correction methods are only possible if the ISI is varied during the time series. These results appear to demonstrate that the hemodynamic response is sufficiently linear that deconvolution methods can be applied to extract overlapping responses. Burock et al. (95) have demonstrated that remarkably clean activation maps can be created with an average ISI of 500 ms. If one assumes that the hemodynamic response is essentially a linear system, there appears to be no obvious minimum ISI in attempts to estimate the hemodynamic response. Dale has suggested that an exponential distribution of ISIs with a mean as short as psychophysically possible is optimal for estimation (100). Of course, the speed at which one can present stimuli depends on the study being performed. Many cognitive tasks may require a slightly longer average presentation rate.

Future work in event-related experimental optimization rests on what further information can be derived from these responses. Between-region, between-voxel, between-subject, and stimulus-dependent variations in amplitude, latency, shape, and responsivity of the event-related fMRI responses are still relatively uncharacterized. Reasons for these differences are also still unclear.

Free Behavior Designs

With many types of cognitive neuroscience questions, it is not possible to constrain the timing or performance of a task precisely. It is necessary to allow the subject to perform the task “freely” and obtain a continuous measurement of the performance, then use the measurement as a reference function for subsequent time series analysis. Examples of this type of design are emerging. As an example, skin conductance changes are difficult to predict or control. In a study by Patterson et al. (117), skin conductance was simultaneously recorded during an array of tasks and during “rest.” The skin conductance signal change was then used as a reference function in the fMRI time series analysis. In several cortical and subcortical regions, signal changes were observed that were highly correlated with the skin conductance changes. Without the use of this measurement, such signal changes would have appeared as noise. It is thought that this type of design will become more prevalent as methods to monitor subject performance or state precisely become more sophisticated.

CONCLUSION

Part of "26 - The Spatial, Temporal, and Interpretive Limits of Functional MRI "

This chapter has attempted to combine a review of the fundamentals of fMRI with a glimpse of the state of the art. Starting with the basics of fMRI contrast, the discussion moved on to hemodynamic transfer function—the basis of understanding fMRI signal change. Characteristics related to the hemodynamic transfer function include location, latency, magnitude, and linearity. Then, perhaps less provocative but still important issues of working with an MRI scanner and understanding some practical limitations were discussed. A sampling of best results, those successfully bringing into play many of the features of experimental design and analysis already mentioned, was presented. The chapter ended with a brief overview of neuronal input strategies, or rather, ways in which one can activate the brain in the context of an fMRI experiment.

Functional MRI is about 9 years old and apparently still at the beginning of its growth curve in terms of users and applications. Clinical applications are just beginning, whereas cognitive neuroscience applications are in full swing. The field of fMRI continues to develop along intertwining paths of understanding signals, creating tools, and refining the questions being asked.

REFERENCES

1. Rosen BR, Belliveau JW, Aronen HJ, et al. Susceptibility contrast imaging of cerebral blood volume: human experience. *Magn Reson Med* 1991;22:293–299.

2. Rosen BR, Belliveau JW, Chien D. Perfusion imaging by nuclear magnetic resonance. *Magn Reson Q* 1989;5:263–281.

3. Moonen CTW, vanZijl PCM, Frank JA, et al. Functional magnetic resonance imaging in medicine and physiology. *Science* 1990;250:53–61.

4. Belliveau JW, Kennedy DN, McKinstry RC, et al. Functional mapping of the human visual cortex by magnetic resonance imaging. *Science* 1991;254:716–719.

5. Williams DS, Detre JA, Leigh JS, et al. Magnetic resonance imaging of perfusion using spin-inversion of arterial water. *Proc Natl Acad Sci U S A* 1992;89:212–216.

6. Detre JA, Leigh JS, Williams DS, et al. Perfusion imaging. *Magn Reson Med* 1992;23:37–45.

7. Edelman RR, Sievert B, Wielopolski P, et al. Noninvasive mapping of cerebral perfusion by using EPISTAR MR angiography. *J Magn Reson Imaging* 1994;4(P):68(abst).

8. Kwong KK, Chesler DA, Weisskoff RM, et al. *Proceedings of the second annual meeting of the Society of Magnetic Resonance*, San Francisco, 1994:1005(abst).

9. Wong EC, Buxton RB, Frank LR. Implementation of quantitative perfusion imaging techniques for functional brain mapping using pulsed arterial spin labeling. *Nucl Magn Reson Biomed* 1997;10:237–249.

10. Kim S-G. Quantification of relative cerebral blood flow change by flow-sensitive alternating inversion recovery (FAIR) technique: application to functional mapping. *Magn Reson Med* 1995;34:293-301.

11. Ogawa S, Lee TM, Kay AR, et al. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci U S A* 1990;87:9868–9872.

12. Turner R, LeBihan D, Moonen CTW, et al. Echo-planar time course MRI of cat brain oxygenation changes. *Magn Reson Med* 1991;22:159–166.

13. Kwong KK, Belliveau JW, Chesler DA, et al. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci U S A* 1992;89:5675–5679.

14. Ogawa S, Lee TM. Functional brain imaging with physiologically sensitive image signals. J Magn Reson Imaging 1992;2(P):S22(abst).

15. Bandettini PA, Wong EC, Hinks RS, et al. Time course EPI of human brain function during task activation. *Magn Reson Med* 1992;25:390–397.

16. Frahm J, Bruhn H, Merboldt K-D, et al. Dynamic MR imaging of human brain oxygenation during rest and photic stimulation. *J Magn Reson Imaging* 1992;2:501–505.

17. Haacke EM, Lai S, Reichenbach JR, et al. *In vivo* measurement of blood oxygen saturation using magnetic resonance imaging: a direct validation of the blood oxygen level-dependent concept in functional brain imaging. *Hum Brain Mapping* 1997;5:341–346.

18. Davis TL, Kwong KK, Weisskoff RM, et al. Calibrated functional MRI: mapping the dynamics of oxidative metabolism. *Proc Natl Acad Sci U S A* 1998;95:1834–1839.

19. Kim S-G, Ugurbil K. Comparison of blood oxygenation and cerebral blood flow effects in fMRI: estimation of relative oxygen consumption change. *Magn Reson Med* 1997;38:59–65.

20. vanZijl PCM, Eleff SM, Ulatowski JA, et al. Quantitative assessment of blood flow, blood volume, and blood oxygenation effects in functional magnetic resonance imaging. *Nat Med* 1998;4:159–160.

21. Pauling L, Coryell CD. The magnetic properties and structure of hemoglobin, oxyhemoglobin, and carbon monoxyhemoglobin. *Proc Natl Acad Sci U S A* 1936;22:210–216.

22. Thulborn KR, Waterton JC, Matthews PM, et al. Oxygenation dependence of the transverse relaxation time of water protons in whole blood at high field. *Biochim Biophys Acta* 1982;714:265–270.

23. Ogawa S, Tank DW, Menon R, et al. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proc Natl Acad Sci U S A* 1992; 89:5951–5955.

24. Fox PT, Raichle ME. Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proc Natl Acad Sci U S A* 1986;83:1140–1144.

25. Edelman R, Siewert B, Darby D. Qualitative mapping of cerebral blood flow and functional localization with echo planar MR imaging and signal targeting with alternating radiofrequency (EPISTAR). *Radiology* 1994;192:1–8.

26. Kwong KK, Chesler DA, Weisskoff RM, et al. MR perfusion studies with T1-weighted echo planar imaging. *Magn Reson Med* 1995;34:878–887.

27. Wong EC, Buxton RB, Frank LR. Quantitative imaging of perfusion using a single subtraction (QUIPSS and QUIPSS II). *Magn Reson Med* 1998;39:702–708.

28. Ogawa S, Menon RS, Tank DW, et al. Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging: a comparison of signal characteristics with a biophysical model. *Biophys J* 1993;64:803–812.

29. Boxerman JL, Hamberg LM, Rosen BR, et al. MR contrast due to intravascular magnetic susceptibility perturbations. *Magn Reson Med* 1995;34:555–566.

30. Kennan RP, Zhong J, Gore JC. Intravascular susceptibility contrast mechanisms in tissues. *Magn Reson Med* 1994;31:9–21.

31. Bandettini PA, Wong EC. Effects of biophysical and physiologic parameters on brain activation-induced R2* and R2 changes: simulations using a deterministic diffusion model. *Int J Imaging Systems Technol* 1995;6:134–152.

32. Duyn JH, Moonen CTW, vanYperen GH, et al. Inflow versus deoxyhemoglobin effects in BOLD functional MRI using gradient-echoes at 1.5 T. *Nucl Med Reson Biomed* 1994;7:83–88.

33. Boxerman JL, Bandettini PA, Kwong KK, et al. The intravascular contribution to fMRI signal change: Monte Carlo modeling and diffusion-weighted studies *in vivo. Magn Reson Med* 1995;34:4–10.

34. Menon RS, Ogawa S, Tank DW, et al. 4-Ttesla gradient recalled echo characteristics of photic stimulation-induced signal changes in the human primary visual cortex. *Magn Reson Med* 1993;30:380–386.

35. Kim SG, Rostrup E, Larsson HB, et al. Determination of relative $CMRO_2$ from CBF and BOLD changes: significant increase of oxygen consumption rate during visual stimulation. *Magn Reson Med* 1999;41:1152–1161.

36. Hoge RD, Atkinson J, Gill B, et al. Investigation of BOLD signal dependence on cerebral blood flow and oxygen consumption: the deoxyhemoglobin dilution model. *Magn Reson Med* 1999;42:849–863.

37. Hoge RD, Atkinson J, Gill B, et al. Stimulus-dependent BOLD and perfusion dynamics in human V1. *Neuroimage* 1999;9(6 Pt 1):573–585.

38. Bandettini PA. *MRI studies of brain activation: dynamic characteristics. Functional MRI of the brain*. Berkeley, CA: Society of Magnetic Resonance in Medicine, 1993:143.

39. Cohen MS. Parametric analysis of fMRI data using linear systems methods. *Neuroimage* 1997;6:93–103.

40. Josephs O, Turner R, Friston K. Event-related fMRI. *Hum Brain Mapping* 1997;5:243–248.

41. Friston KJ, Josephs O, Rees G, et al. Nonlinear event-related responses in fMRI. *Magn Reson Med* 1998;39:41–52.

42. Glover GH. Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 1999;9:416–429.

43. Dale AM, Buckner RL. Selective averaging of rapidly presented individual trials using fMRI. *Hum Brain Mapping* 1997;5:329–340.

44. Savoy RL, O'Craven KM, Weisskoff RM, et al. Exploring the temporal boundaries of fMRI: measuring responses to very brief visual stimuli. *Book of abstracts of the twenty-fourth annual meeting of the Society for Neuroscience*, Miami, 1994:1264 (abst).

45. Buckner RL, Bandettini PA, O'Craven KM, et al. Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *Proc Natl Acad Sci U S A* 1996;93:14878–14883.

46. Lee AT, Glover GH, Meyer CH. Discrimination of large venous vessels in time-course spiral blood-oxygen-level-dependent magnetic-resonance functional neuroimaging. *Magn Reson Med* 1995;33:745–754.

47. Saad ZS, DeYoe EA. Time delay estimates of FMRI signals: efficient algorithm and estimate variance. *Proceeding of the nineteenth annual international conference of the IEEE/EMBS*, Chicago, 1997:460–463.

48. Tootell RB, Reppas JB, Kwong KK, et al. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci* 1995;15:3215–3230.

49. Binder JR, Rao SM, Hammeke TA, et al. Effects of stimulus rate on signal response during functional magnetic resonance imaging of auditory cortex. *Cogn Brain Res* 1994;2:31–38.

50. Rao SM, Bandettini PA, Binder JR, et al. Relationship between finger movement rate and functional magnetic resonance signal change in human primary motor cortex. *J Cereb Blood Flow Metab* 1996;16:1250–1254.

51. Disbrow EA, Slutsky DA, Roberts TP, et al. Functional MRI at 1.5 tesla: a comparison of the blood oxygenation level-dependent signal and electrophysiology [In Process Citation]. *Proc Natl Acad Sci U S A* 2000;97:9718–9723.

52. Rees G, Friston K, Koch C. A direct quantitative relationship between the functional properties of human and macaque V5. *Nat Neurosci* 2000;3:716–723.

53. Heeger DJ, Huk AC, Geisler WS, et al. Spikes versus BOLD: what does neuroimaging tell us about neuronal activity? [News; Comment]. *Nat Neurosci* 2000;3:631–633.

54. Boynton GM, Engel SA, Glover GH, et al. Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 1996;16:4207–4221.

55. Vazquez AL, Noll DC. Nonlinear aspects of the BOLD response in functional MRI. *Neuroimage* 1998;7:108–118.

56. Buxton RB, Wong EC, Frank LR. A biomechanical interpretation of the BOLD signal time course: the balloon model. *Proceedings of the fifth annual meeting of International Society of Magnetic Resonance Medicine*, Vancouver, 1997.

57. Frahm J, Krüger G, Merboldt K-D, et al. Dynamic uncoupling and recoupling of perfusion and oxidative metabolism during focal activation in man. *Magn Reson Med* 1996;35:143–148.

58. Noll DC, Genovese CR, Vazquez AL, et al. Evaluation of respiratory artifact correction techniques in multishot spiral functional MRI using receiver operator characteristic analyses. *Magn Reson Med* 1998;40:633–639.

59. Jezzard P. Physiological noise: strategies for correction. In: Moonen CTW, Bandettini PA, eds. *Functional MRI*. Berlin: Springer-Verlag, 1999:173–181.

60. Guimaraes AR, Melcher JR, Talavage TM, et al. Imaging subcortical auditory activity in humans. *Hum Brain Mapping* 1998;6:33–41.

61. Lee CC, Jack CR, Jr., Grimm RC, et al. Real-time adaptive motion correction in functional MRI. *Magn Reson Med* 1996;36:436–444.

62. Hu X, Kim S-G. Reduction of signal fluctuations in functional MRI using navigator echoes. *Magn Reson Med* 1994;31:495–503.

63. Le TH, Hu X. Retrospective estimation and correction of physiological artifacts in fMRI by direct extraction of physiological activity from MR data. *Magn Reson Med* 1996;35:290–298.

64. Wowk B, McIntyre MC, Saunders JK. k-Space detection and correction of physiological artifacts in fMRI. *Magn Reson Med* 1997;38:1029–1034.

65. Constable RT, Carpentier A, Pugh K, et al. Investigation of the human hippocampal formation using a randomized event-related paradigm and z-shimmed functional MRI. *Neuroimage* 2000;12:55–62.

66. Glover GH. 3D z-shim method for reduction of susceptibility effects in BOLD fMRI. *Magn Reson Med* 1999;42:290–299.

67. Jezzard P, Balaban RS. Correction for geometric distortion in echo planar images from Bo field distortions. *Magn Reson Med* 1995;34:65–73.

68. Savoy RL, Bandettini PA, Weisskoff RM, et al. Pushing the temporal resolution of fMRI: studies of very brief visual stimuli, onset variability and asynchrony, and stimulus-correlated changes in noise. *Proceedings of the third annual meeting of the Society of Magnetic Resonance*, Nice, 1995:450(abst).

69. Menon RS, Luknowsky DC, Gati JS. Mental chronometry using latency-resolved functional MRI. *Proc Natl Acad Sci U S A* 1998;95:10902–10907.

70. Kim SG, Richter W, Ugurbil K. Limitations of temporal resolution in functional MRI. *Magn Reson Med* 1997;37:631–636.

71. Richter W, Somorjai R, Summers R, et al. Motor area activity during mental rotation studied by time-resolved single-trial fMRI. *J Cogn Neurosci* 2000;12:310–320.

72. Bandettini PA. The temporal resolution of functional MRI. In: Moonen CTW, Bandettini PA, eds. *Functional MRI*. Berlin: Springer-Verlag, 1999:205–220.

73. Engel SA, Rumelhart DE, Wandell BA, et al. fMRI of human visual cortex. *Nature* 1994;369:525 [published erratum appears in *Nature* 1994;370:106].

74. Engel SA, Rumelhart DE, Wandell BA, et al. fMRI of human visual cortex [Letter] [published erratum appears in *Nature* 1994;370:106]. *Nature* 1994;369:525.

75. Engel SA, Glover GH, Wandell BA. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb Cortex* 1997;7:181–192.

76. Menon RS, Thomas CG, Gati JS. Investigation of BOLD contrast in fMRI using multi-shot EPI. *Nucl Magn Reson Biomed* 1997;10:179–182.

77. Hu X, Le TH, Ugurbil K. Evaluation of the early response in fMRI in individual subjects using short stimulus duration. *Magn Reson Med* 1997;37:877–884.

78. Yacoub E, Hu X. Detection of the early negative response in fMRI at 1.5 tesla. *Magn Reson Med* 1999;41:1088–1092.

79. Duong TQ, Kim DS, Ugurbil K, et al. Spatiotemporal dynamics of the BOLD fMRI signals: toward mapping submillimeter cortical columns using the early negative response [In Process Citation]. *Magn Reson Med* 2000;44:231–242.

80. Grinvald A, Slovin H, Vanzetta I. Non-invasive visualization of cortical columns by fMRI [News; Comment]. *Nat Neurosci* 2000;3:105–107.

81. Kim DS, Duong TQ, Kim SG. High-resolution mapping of iso-orientation columns by fMRI [see Comments]. *Nat Neurosci* 2000;3:164–169.

82. Jesmanowicz A, Bandettini PA, Hyde JS. Single-shot half NEX 256 × 256 resolution EPI at 3 tesla. *Proceedings of the fifth annual meeting of the International Society of Magnetic Resonance Medicine*, Vancouver, 1997.

83. Sereno MI, Dale AM, Reppas JR, et al. Functional MRI reveals borders of multiple visual areas in humans. *Science* 1995;268:889–893.

84. DeYoe EA, Carman G, Bandettini P, et al. Mapping striate and extrastriate areas in human cerebral cortex. *Proc Natl Acad Sci U S A* 1996;93:2382–2386.

85. Servos P, Zacks J, Rumelhart DE, et al. Somatotopy of the human arm using fMRI. *Neuroreport* 1998;9:605–609.

86. Talavage TM, Ledden PJ, Sereno MI, et al. Preliminary fMRI evidence for tonotopicity in human auditory cortex. *Neuroimage* 1996;3:S355.

87. Bandettini PA, Jesmanowicz A, Wong EC, et al. Processing strategies for time-course data sets in functional MRI of the human brain. *Magn Reson Med* 1993;30:161–173.

88. Courtney SM, Ungerleider LG, Keil K, et al. Transient and sustained activity in a distributed neural system for human working memory. *Nature* 1997;386:608–611.

89. Bandettini PA, Cox RW. Event-related fMRI contrast when using constant interstimulus interval: theory and experiment. *Magn Reson Med* 2000;43:540–548.

90. Belin P, Zatorre RJ, Hoge R, et al. Event-related fMRI of the auditory cortex. *Neuroimage* 1999;10:417–429.

91. Birn RM, Bandettini PA, Cox RW, et al. Event-related fMRI of tasks involving brief motion. *Hum Brain Mapping* 1999;7:106–114.

92. Buckner RL. Event-related fMRI and the hemodynamic response. *Hum Brain Mapping* 1998;6:373–377.

93. Buckner RL, Koutstaal W, Schacter DL, et al. Functional-anatomic study of episodic retrieval. II. Selective averaging of event-related fMRI trials to test the retrieval success hypothesis. *Neuroimage* 1998;7:163–175.

94. Buckner RL, Goodman J, Burock M, et al. Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron* 1998;20:285–296.

95. Burock MA, Buckner RL, Woldorff MG, et al. Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *Neuroreport* 1998;9:3735–3739.

96. Clare S, Humberstone M, Hykin J, et al. Detecting activations in event-related fMRI using analysis of variance. *Magn Reson Med* 1999;42:1117–1122.

97. D'Esposito M, Postle BR, Ballard D, et al. Maintenance versus manipulation of information held in working memory: an event-related fMRI study. *Brain Cogn* 1999;41:66–86.

98. D'Esposito M, Postle BR, Jonides J, et al. The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related functional MRI. *Proc Natl Acad Sci U S A* 1999;96:7514–7519.

99. D'Esposito M, Zarahn E, Aguirre GK. Event-related functional MRI: implications for cognitive psychology. *Psychol Bull* 1999;125:155–164.

100. Dale AM. Optimal experimental design for event-related fMRI. *Hum Brain Mapping* 1999;8:109–114.

101. Davis KD, Kwan CL, Crawley AP, et al. Event-related fMRI of pain: entering a new era in imaging pain. *Neuroreport* 1998;9:3019–3023.

102. Friston KJ, Fletcher P, Josephs O, et al. Event-related fMRI: characterizing differential responses. *Neuroimage* 1998;7:30–40.

103. Friston KJ, Zarahn E, Josephs O, et al. Stochastic designs in event-related fMRI. *Neuroimage* 1999;10:607–619.

104. Josephs O, Henson RN. Event-related functional magnetic resonance imaging: modelling, inference and optimization. *Philos Trans R Soc Lond B Biol Sci* 1999;354:1215–1228.

105. Kang AM, Constable RT, Gore JC, et al. An event-related fMRI study of implicit phrase-level syntactic and semantic processing. *Neuroimage* 1999;10:555–561.

106. Kiehl KA, Liddle PF, Hopfinger JB. Error processing and the rostral anterior cingulate: an event-related fMRI study. *Psychophysiology* 2000;37:216–223.

107. McCarthy G. Event-related potentials and functional MRI: a comparison of localization in sensory, perceptual and cognitive tasks. *Electroencephalogr Clin Neurophysiol Suppl* 1999;49:3–12.

108. Pinel P, Le Clerc HG, van de Moortele PF, et al. Event-related fMRI analysis of the cerebral circuit for number comparison. *Neuroreport* 1999;10:1473–1479.

109. Postle BR, D'Esposito M. “What”-Then-Where” in visual working memory: an event-related fMRI study. *J Cogn Neurosci* 1999;11:585–597.

110. Rosen BR, Buckner RL, Dale AM. Event-related functional MRI: past, present, and future. *Proc Natl Acad Sci U S A* 1998;95:773–780.

111. Schacter DL, Buckner RL, Koutstaal W, et al. Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. *Neuroimage* 1997;6:259–269.

112. Schad LR, Wiener E, Baudendistel KT, et al. Event-related functional MR imaging of visual cortex stimulation at high temporal resolution using a standard 1.5 T imager. *Magn Reson Imaging* 1995;13:899–901.

113. Zarahn E, Aguirre G, D'Esposito M. A trial-based experimental design for fMRI. *Neuroimage* 1997;6:122–138.

114. Clark VP, Maisog JM, Haxby JV. fMRI study of face perception and memory using random stimulus sequences. *J Neurophysiol* 1998;79:3257–3265.

115. Dale A, Buckner R. Selective averaging of individual trials using fMRI. *Proceedings of the third international conference on functional mapping of the human brain*, Copenhagen, 1997:S47.

116. McCarthy G, Luby M, Gore J, et al. Infrequent events transiently activate human prefrontal and parietal cortex as measured by functional MRI. *J Neurophysiol* 1997;77:1630–1634.

117. Patterson J, Bandettini P, Ungerleider LG. Simultaneous skin conductance measurement and fMRI during cognitive tasks: correlations of skin conductance activity with ventromedial prefrontal cortex (PFC) and orbitofrontal cortex (OFC) activity. *Nero Image* 2000;11:235.

27
Diffusion Tensor Imaging

Nikos Makris

G. M. Papadimitriou

A. J. Worth

Bruce G. Jenkins

L. Garrido

A. Gregory Sorensen

V. J. Wedeen

David S. Tuch

O. Wu

Merit E. Cudkowicz

V. S. Caviness Jr.

B. R. Rosen

David N. Kennedy

Nikos Makris, G. M. Papadimitriou, A. J. Worth, Bruce G. Jenkins, L. Garrido, A. Gregory Sorensen, V. J. Wedeen, David S. Tuch, O. Wu, Merit E. Cudkowicz, V. S. Caviness, Jr., B. R. Rosen, and David N. Kennedy: Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts.

The central nervous system (CNS) spans across different levels of organization, covering the gamut from genes to behavior. It is the aim of neuroscience to elucidate all these levels at a gross and an ultrastructural resolution and to define the relationships among them. One of the highest levels of organization is the systems level, which includes the motor, sensory, and central neural systems. Among the central systems are included those related to cognitive function, such as attention, memory, language, and executive function. Each system can be considered as a set of interconnected processors or centers constituted by nerve cells. Their physical connections are composed of axons of different lengths that can form fascicles as they run from origin to destination. Within the neocortex, these connections are selective and architectonic (1). How cytoarchitecture, connections, and function relate within the neocortex is a fundamental problem in neuroscience. This question addresses basic organizational principles of the nervous system and aims to elucidate the mechanisms through which the cerebrum mediates behavior (2 ,3). Behavior, to a large extent the product of cerebral function and at the same time a modulator of this function, is certainly the key to solving the problem of meaning related to any neural system. To explore thoroughly the behavioral dimension of an organism, it appears necessary to perform studies *in vivo*. If experiments were designed in such manner that information related to structure was gathered along with behavioral information, we would have an ideal setting for structural–functional or, in the case of disease, anatomic–clinical correlation. This can be done in the domain of neuroimaging by using currently available technology in an unprecedented way. In the past two centuries, functional–structural correlations were derived mostly from experimental nonhuman material, whereas anatomic–clinical correlations were derived principally from human behavioral and, eventually, postmortem lesion analyses. With the tremendous development of magnetic resonance imaging (MRI) technology—both structural and functional MRI (fMRI) and magnetic resonance spectroscopy (MRS)—the study of the structure, function, and metabolism of the living human is an ongoing reality.

One of the latest advancements of MRI technology has been *diffusion tensor imaging* (DTI), a technique capable of measuring the diffusivity of water molecules and rendering visible the preferential orientation of their movement. As water molecules diffuse within the brain, their movement is constrained by the structural “fingerprint” of the tissue. If the tissue is equally distributed in all directions, like light in a room, then a “random walk” of water molecules occurs and their diffusion is *isotropic* (i.e., “equally behaving” in all directions). On the other hand, within a strongly oriented tissue such as a white matter tract, water diffusion is not equal in all directions but is instead *anisotropic*, specifically predominant along the direction of the tissue. The strongly parallel axonal arrangement within a fiber bundle creates a highly oriented and anisotropic environment for

water molecules. Because DTI can measure the anisotropy and orientation of a tissue, and because brain white matter can be characterized to a large extent by its orientation and anisotropy, the detection of white matter fiber pathways has become feasible in the living brain.

Tractography (i.e., the "writing" or tracing of tracts) has been considered to be the most difficult task in neuroanatomy (4). To delineate neuroanatomic connections exactly, white matter fiber tracts have been traced systematically in postmortem experimental material during the last decades. It has emerged from studies in nonhuman primates that fiber pathways establish connections between distinctive architectonic cortical areas, and that they constitute fundamental components of neural systems that subserve specific functions (1 ,5 ,6). Extensive research during the past two centuries in higher brain function (and aphasiology in particular) has demonstrated the tight relationship between damage of commissural or associational fiber tracts and breakdown in cognitive human behavior (7 ,8 and 9). In the light of brain organization at a systems level in terms of architecture and its connections, it becomes evident how relevant the precise knowledge of neocortical pathways is to the understanding of human behavior in normal and disease states. Although the delineation of fiber tracts has been accomplished precisely and comprehensively in postmortem nonhuman primates, this goal has not been met in humans satisfactorily. Moreover, the study of fiber tracts *in vivo* is only beginning and is largely based on MRI techniques. DTI is one avenue that may provide solutions to the connectional structure of the CNS. Even though the identification and reconstruction of major fiber bundles has been accomplished with the use of DTI and computational techniques, the basic problems related to the biological sources of the DTI signal remain to be clarified. Thus, to achieve a comprehensive understanding of the sensitivity and specificity of the DTI technique, studies addressing both the gross and ultrastructural level of the white matter are necessary.

In this chapter, we overview the connectional composition of the cerebrum, emphasizing the morphology and architectonic structure of its pathways. Subsequently, we overview the DTI technique and illustrate its use at an ultrastructural and a gross neuroanatomic level. In this perspective, we elaborate on three representative white matter fiber pathways of the brain and emphasize how they can be mapped in terms of computational MRI neuroanatomy. Additionally, we discuss the utility, potential, and limitations of the DTI technique in the context of its applications and its integration within a larger neurofunctional MRI examination. For this purpose, we present a case of amyotrophic lateral sclerosis that we studied with MRI, DTI, MRS, and fMRI.

OVERVIEW OF WHITE MATTER PATHWAYS

Part of "27 - Diffusion Tensor Imaging "

Anatomic Connections

The white matter of the CNS contains axons serving corticocortical, commissural, cortico-subcortical, and cerebellar connections. These connections can be categorized in three principal classes—namely, associational, commissural, and projectional. Intrahemispheric associational corticocortical connectivity in particular is accomplished in general by (a) short U fibers that constitute the local circuitry within a gyrus, (b) intermediate-range fibers within the extent of a lobe, and (c) long association pathways that connect different lobes (10). A tract can be described at a morphologic level in terms of a set of descriptors (i.e., its stem, splay, and origins and terminations), whereas the set or map of its architectonic connections is its principle structural descriptor.

Morphologic Descriptors of a Fiber Pathway

Conceptually, a white matter fiber pathway is a group of axons that originate from a set of neuronal bodies and end on one or more sets of target neurons. Morphologically, we can distinguish three different components in a fiber bundle—namely (a) a compact portion or "stem" or "stalk" (11), where the axons run together to form a fascicle; (b) a zone of divergence/convergence called *spray* (12) or *splay*, where the axons fan out; and (c) the distal, peripheral *origins* and *terminations*, where the axons either originate or end. However, in current MRI research, the term *extreme periphery* seems more practical than the terms *origin* and *termination* for a couple of reasons. First, DTI does not allow the identification of this portion of a fiber pathway clearly enough to justify the application of such specific terminology. Second, both the corticocortical associational and the commissural connections are generally bidirectional, meaning that within the stem of a tract fibers are running in both directions; thus, origins and terminations pertaining to the same fiber tract occur in adjacent locations. The differentiation of these features is well beyond the current spatial resolution capabilities of DTI. Therefore, at a morphologic level, a comprehensive delineation of a fiber pathway should include its stem, splay, and extreme periphery. A description of the extreme periphery of an association fiber tract is related to a cortical field of origin or termination (Fig. 27.1).

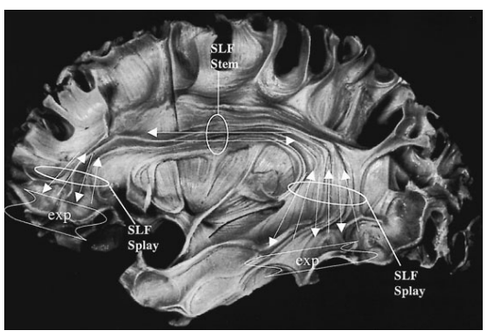


FIGURE 27.1. Morphologic description of a fiber pathway. In this lateral preparation of Ludwig and Klingler (53), the sample tract is the human superior longitudinal fasciculus, which is a long, associational, corticocortical connection. Its three components are (a) the stem, where axons run compactly and in a bidirectional fashion; (b) the splays or sprays, where the axons diverge or converge; and (c) the extreme peripheries, the cortical regions within which the axons originate or terminate. *Arrows* denote the bidirectionality of axons within the stem and the splays of the fiber tract as they run from origin to termination.

Architectonic Connections

The detailed characterization of a fiber pathway in terms of its cytoarchitectonic correlates is a fundamental step for the understanding of cerebral structural and functional organization. These issues have been addressed at the architectonic level in experimental animals and in human postmortem material. In particular, cerebral cytoarchitecture has been described precisely in both the human and the experimental animal, such as the macaque. In the experimental nonhuman primate, the white matter fiber pathways have been delineated precisely and comprehensively (i.e., in terms

of their stems, splays, and origins and terminations). Although this is currently the case for the nonhuman primate, the status of research in human brain anatomic connectivity is very different.

Studies in Human Postmortem Material

In human postmortem material, traditional techniques such as myelin stain, bichromate fixation, and gross anatomic dissection allow visualization of the stems of these fiber bundles (10). With very few exceptions, in which dyes such as the carbocyanine dye (DiI) were used for very short connections (13), histologic description of human fiber pathways is incomplete because it does not provide a detailed understanding of their origins and terminations, and no technique is available that can identify with certainty the origin and distal terminations of a fiber pathway satisfactorily in the human. The closest inferences at this level of description are obtained from white matter degeneration studies of brains with specific neurologic damage. Most of these studies deal with cortico-subcortical connections and are not specific because the cortical lesions that cause the remote degeneration are very large (10 ,14).

Studies in the Experimental Nonhuman Primate

Experimental approaches with available techniques have addressed the problem of origin and termination of fiber pathways in the monkey. The injection of radioactively labeled amino acids (15) into nonhuman primate brain with appropriate histologic processing techniques permits the accurate and reliable interpretation of the origins, trajectories, and terminations of the subcomponents of the various fiber pathways of the CNS. With this technique, a number of studies of the fiber connections between the different lobes of the cerebral cortex, between the two hemispheres, and between the cerebral cortex and subcortical regions have been carried out in the monkey. These studies have demonstrated that the various fiber bundles are distinct and occupy unique trajectories from origin to termination.

Extrapolation from Animal Experimental Material to Human Material

In the monkey, it is known how different pathways correlate with radioactively labeled material, so that their origin and termination can be delineated. Because the stems of the major pathways are similar in the monkey and the human, one can extrapolate the origins and terminations of the observed pathways in the human to correlate specific fiber pathways with cortical architectonic fields. Drawing these inferences, we can formulate for each individual fiber bundle a specific map that characterizes the tract in terms of its connections.

Maps of Anatomic Connectivity

A map of anatomic connectivity (MAC) is a *set of neuroanatomic regions interconnected by a particular white matter fiber pathway*, and we can symbolize it as $PATHWAY_{MAC}$. For instance, the MAC for cingulum bundle (CB) would be CB_{MAC} (Fig. 27.2). Whereas in the human we have precise knowledge only of the stem of this bundle, the architectonic connections in the nonhuman primate are well documented. A comprehensive description of maps for anatomic cerebral connectivity has been formulated, derived from anatomic studies in the human and by extrapolation from experimental material, and has been integrated in the context of a methodology for topographic characterization and quantification of human forebrain white matter (16 ,17). Because of their versatility, MACs could be relevant in the formulation of more sophisticated tractographic experiments and in interpreting neurofunctional data. Integrating the neuroanatomic knowledge of fiber tracts within an fMRI experiment is an additional challenge, but at the same time it seems to be key if we are to study behavior in normal and clinical conditions, which implies the study of fiber pathways *in vivo*.

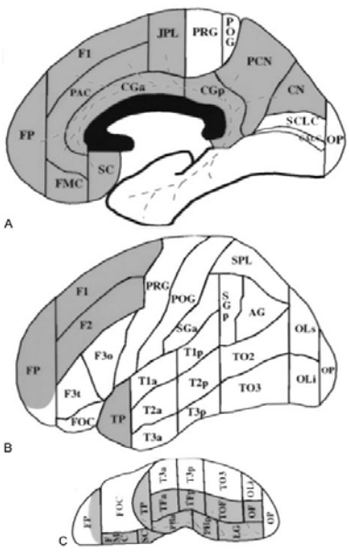


FIGURE 27.2. Map of cortical anatomic connectivity (MAC) for the cingulum bundle (CB), or CB_{MAC}. The connections of the CB are represented in the mesial **(A)**, lateral **(B)**, and ventral **(C)** views of the human brain on a cortical parcellation system (46). The *shaded area* in blue within the frontal pole in the ventral view corresponds approximately to the rostral part of the frontoorbital cortex that is anterior to the transverse orbital sulcus. AG, angular gyrus; CALC, intracalcarine cortex; CGa, cingulate gyrus, anterior; CGp, cingulate gyrus, posterior; CN, cuneiform cortex; CO, central operculum; F1, superior frontal gyrus; F2, middle frontal gyrus; F3o, inferior frontal gyrus, pars opercularis; F3t, inferior frontal gyrus, pars triangularis; FMC, frontal medial cortex; FO, frontal operculum; FOC, frontal orbital cortex; FP, frontal pole; H1, Heschl gyrus; INS, insula; JPL, juxtaparacentral cortex; LG, lingual gyrus; OP, occipital pole; OF, occipital fusiform gyrus; OLi, lateral occipital cortex, inferior; OLs, lateral occipital cortex, superior; PAC, paracingulate cortex; PCN, precuneus; PHa, parahippocampal gyrus, anterior; PHp, parahippocampal gyrus, posterior; PO, parietal operculum; POG, postcentral gyrus; PP, planum polare; PRG, precentral gyrus; PT, planum temporale; SC, subcallosal cortex; SCLC, supracalcarine cortex; SGa, supramarginal gyrus, anterior; SGp, supramarginal gyrus, posterior; SPL, superior parietal lobule; T1a, superior temporal gyrus, anterior; T1p, superior temporal gyrus, posterior; T2a, middle temporal gyrus, anterior; T2p, middle temporal gyrus, posterior; T3a, inferior temporal gyrus, anterior; T3p, inferior temporal gyrus, posterior; TFa, temporal fusiform, anterior; TFp, temporal fusiform, posterior; TO2, middle temporal gyrus, temporooccipital; TO3, inferior temporal gyrus, temporooccipital; TOF, temporooccipital fusiform gyrus; TP, temporal pole. See color version of figure.

In Vivo Analysis of Fiber Pathways

The capability of studying tracts in the living human brain opens up a new window in structural–functional and anatomic–clinical relationships. Currently, the detection of

fiber tracts *in vivo* has been addressed by MRI techniques such as DTI. DTI analysis enables us to characterize a white matter fiber pathway in terms of its orientation, location, and size. To date, tractography has been performed in two different ways. Using *manual* or *model-independent methods*, we can derive the trajectory of the fiber bundle and approximate its extreme peripheries (12). Using mathematically driven *model-based methods*, we can also trace a fiber pathway (18 ,19 and 20). In the section on applications, we give examples in which both methods are used and different tracts are visualized in two and three dimensions. Although the field of DTI-based brain tractography is expanding rapidly with impressive results, it has to be pointed out that certain basic conceptual obstacles still need to be overcome. For instance, in this stage, it has not been demonstrated that we are able to delineate completely and precisely a fiber tract in the brain by means of any DTI analysis technique. At the most, we can identify and characterize the stems of the major fiber tracts (12 ,21); however, the problem of elucidating the splays and extreme peripheries of the bundles remains to be solved reliably. Therefore, when we use the term *pathway*, *tract*, or *bundle*, we currently refer to its stem.

OVERVIEW OF DIFFUSION TENSOR IMAGING

Part of "27 - Diffusion Tensor Imaging "

Self-diffusion of molecules has been studied with magnetic resonance methodologies for several decades (22 ,23). For comprehensive reviews of the use of diffusion in nuclear magnetic resonance, we refer the reader to other sources (24 ,25 and 26). With respect to the physical principles underlying diffusion (also known as *brownian motion*), water in tissues with an oriented structure tends to diffuse more *along* the orientation of the tissue structure (Fig. 27.3). The incoherent motion of the diffusing water, when it occurs in the presence of a magnetic field gradient, leads to dephasing of the MR signal. This dephasing produces signal attenuation (SA), which is related to the magnitude of diffusivity of the water along the direction and magnitude of the applied gradient in an exponential fashion. For anisotropic gaussian diffusion, the SA is proportionate to

$$q_1 = \text{atan}(z/x) \;/\; n/2$$

$$q_2 = \text{atan}(y/h) \;/\; p/2$$

$$\text{where } h = \text{sqrt}(z^2 + x^2).$$

For isotropic diffusion, this reduces to the Stejskal–Tanner relation: $SA = SA_0\, e^{-bD}$, where D is the diffusion coefficient and b is the diffusion sensitivity factor. Note that $b = \gamma^2 g^2 \delta^2 (\Delta - \delta/3)$, where the values of g, δ, and Δ correspond to the values of the gradient amplitude, duration, and spacing, respectively, and γ is the hydrogen gyromagnetic ratio (27). The diffusion process can be parameterized by a 3 × 3 symmetric tensor, which can be represented by an ellipsoid, as shown in Fig. 27.3 .

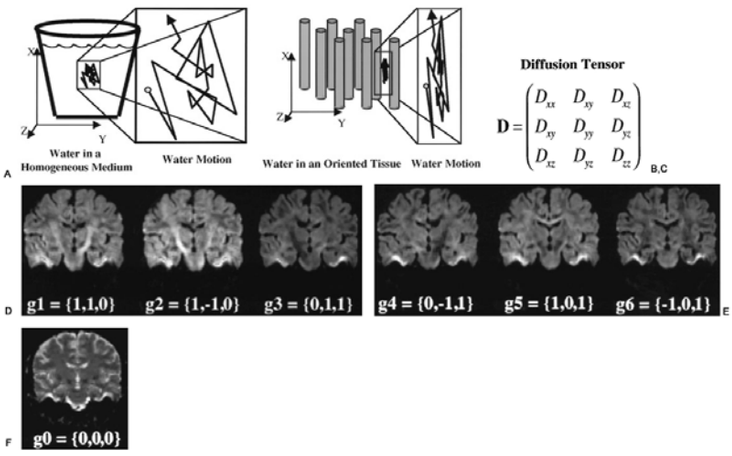


FIGURE 27.3. Schematic representation of water diffusion in the presence of **(A)** a nonoriented tissue structure and **(B)** an oriented tissue structure. Water mobility is highest along the direction with least interference—that is, along the direction of the fibers. **D,E:** Overview of the procedural steps involved in diffusion tensor acquisition. Six images with gradient sampling directions of g_1 = {1, 1, 0}, g_2 = {1, -1, 0}, g_3 = {0, 1, 1}, g_4 = {0, -1, 1}, g_5 = {1, 0, 1}, and g_6 = {-1, 0, 1}, and an additional baseline acquisition with no diffusion gradients, g_0 = {0, 0, 0}, are acquired. The six first data sets (gradients g_1 through g_6) are analyzed relative to the baseline acquisition with no diffusion encoding (g_0). This set of observations is sufficient to define the symmetric diffusion tensor representation of water self-diffusion, shown in **(C)**. See color version of figure.

HOW DO WE “BUILD” A DIFFUSION IMAGE?

Part of "27 - Diffusion Tensor Imaging "

The acquisition of the DTI requires the acquisition of six directionally weighted samples of the effect of the diffusion process relative to the axes of the imaging system. Specifically, the magnitude of the diffusion attenuation in MR signal along the *x*, *y*, and *z* axes themselves, as well as in the *xy*, *xz*, and *yz* directions, must be measured. The attenuation of MR signal in the presence of gradients in each of these directions is calculated relative to an image acquired with no diffusion encoding (baseline). Hence, seven (six directions, one baseline) acquisitions for each slice level are required. Once the tensor is sampled, the magnitude calculated from the trace expresses the *total* (no directionality) diffusivity at each voxel location (Fig. 27.4).

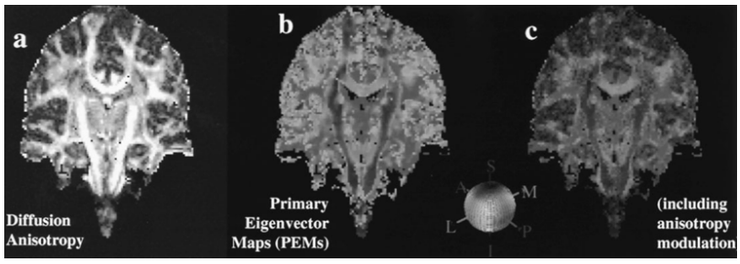


FIGURE 27.4. The tensor representation can be visualized in many ways. **A:** *Diffusion anisotropy*, defined as the relative magnitude of the major axis of the diffusion ellipsoid in comparison with the minor axes, can be visualized; in this figure, regions of high anisotropy are bright, yielding an observable substructure within the cerebral white matter. **B:** A primary eigenvector map (PEM) can be generated to observe the orientation of the major axis of the diffusion ellipse in three-dimensional space; red indicates medial–lateral, blue indicates superior–inferior, and green indicates anterior–posterior orientation, respectively. **C:** The PEM can include anisotropy modulation if the intensity of the color is made proportional to the degree of anisotropy present in each voxel. Regions of high anisotropy can be colored with intense color, whereas regions of low anisotropy have a pale coloring, so that the underlying anatomic image can be viewed. See color version of figure.

The directionality of diffusion is assessed by an eigen decomposition of the diffusion tensor. The largest eigenvalue corresponds to the major axis of the diffusion ellipsoid and so represents the major directionality of diffusion at that location.

Color Coding

A color is assigned for each voxel location by using the primary eigenvector (corresponding to the largest eigenvalue) of the diffusion tensor. At each voxel, the absolute values of the *x*, *y*, and *z* components are used as the red, green, and blue color values, respectively, such that a red voxel in the image means the vector points left–right (or right–left), green means anteroposteriorly (or posteroanteriorly), and blue represents superoinferiorly (or inferosuperiorly). For instance, if a vector points mostly in the red direction, then the *x* value of the vector will be large and the color will be pure red; otherwise, the color will be a mixture of red, green, and blue, depending on the magnitudes of the vector components (i.e., on the direction of the vector). In Fig. 27.4 , this color coding scheme is shown with the appropriate color painted onto a sphere, and the principal eigenvector map (PEM) is the result of color-coding a tensor image (Fig. 27.4).

Modulation by Anisotropy

To further distinguish white matter fiber pathways from other regions, the color is modulated by a measure of anisotropy of the voxel. This emphasizes the stems (i.e., the compact portions of the fiber tracts) by diminishing the brightness of everything else. Here, *anisotropy* can be either lattice (21 ,28 ,29) or fractional (30) anisotropy. Figure 27.4C shows the effect of anisotropy modulation.

“ZORRO”

Part of "27 - Diffusion Tensor Imaging "

In the section on applications, we illustrate tractographic representations of different pathways that have been created with use of a novel tool for DTI data analysis. This tool, which we call *zorro* (for its capability to create masks), is described here.

Zorro is a program for visualization and quantitative measurement of diffusion-weighted MR tensor scans. It was written with the “visualization toolkit” (31). Its main purpose is to create “masks” interactively that designate regions of voxels in the three-dimensional (3D) data. The masks are then used to make 3D visualizations and quantitative measurements, such as volume, anisotropy, and direction. The program loads both tensor files and their registered nonattenuated baseline echo-planar imaging (T2-EPI) files. Anisotropy values are calculated from the tensor data (e.g., fractional, lattice), or they are loaded if they were previously calculated. A rough segmentation into brain, background (air), and cerebrospinal fluid is performed with use of the T2-EP image, and this facilitates visualization of the data by providing an anatomic context. Zorro can display all three colorized diffusion eigenvectors, anisotropy images, the T2-EP image, the segmented image, and all mask images. A mouse click prints all numeric data for a given voxel (the full tensor, anisotropy values, and mask values). The brightness and contrast can be adjusted for both color (vector) and grayscale (scaler) images, and there are options to show and hide the background and cerebrospinal fluid in the color images and to enable and disable the anisotropy modulation of colors. Any number of different voxel “masks” can be created by clicking on particular voxels. One click adds the voxel to the mask being edited, and another turns it off. Once chosen, various statistics may be calculated from any value associated with every voxel in the mask. Different masks may be combined in binary operations (and, or, xor).

Masks may be semiautomatically created by using “region growing.” For this, a mouse click starts the region at a “seed” voxel, and then each of the seed';s neighbors in 3D is added to the region if the neighbor is similar enough to the seed voxel. A voxel is “similar enough” if, for instance, its anisotropy and primary diffusion direction differ from the seed's by less than some given threshold. Once a neighbor is included, it becomes a seed, and its neighbors are checked to see if they should be included in the result. In zorro, the three kinds of region growing are *direction of interest*, *change in angle*, and *flow*. For all types, one can also specify an anisotropy threshold, so that if the anisotropy of the neighbor is below this threshold, it will not be included in the result. Region growing can also be 2D instead of 3D and can be prevented in cerebrospinal fluid regions or the background region.

Flow Region Growing

Flow region growing involves three vectors: the seed voxel's primary diffusion vector, the neighbor's primary diffusion

vector, and the displacement vector (a vector from the seed voxel location to the neighbor voxel location). A neighbor voxel is included if the displacement vector matches both the seed and neighbor voxels' primary diffusion vectors. This match is determined by comparing a threshold to the product of the dot products of the displacement vector with the seed and the neighbor's primary diffusion vectors. This means that both primary diffusion vectors are pointing in a direction similar to that of the displacement vector.

Change-in-Angle Region Growing

Here, a neighbor is included in the mask if its direction is close enough to the seed direction. This match is determined by comparing a threshold to the acrcosine of the dot product of the seed's and neighbor's primary diffusion vectors. As the region grows, the angle is defined by the new seeds.

Direction of Interest

Here, an absolute direction is specified and each neighbor is included if its direction is close enough to a user-specified direction of interest. This match is determined by comparing a threshold to the acrcosine of the dot product of the user-specified direction of interest and the neighbor's primary diffusion vector. Again, growing can be limited in each of these region-growing methods by providing a threshold for anisotropy and by choosing not to grow out of the 2D slice or into the background or cerebrospinal fluid.

The problem with region growing is that noisy data may cause the region to “escape” and grow out of a region where it might be expected to remain. Besides allowing the interactive examination of tensor and anisotropy data, zorro also produces three kinds of output: mask statistics, angular histograms, and 3D visualization with the use of “boxels.”

Mask Statistics

After all voxels in a mask have been chosen by manual editing, region growing, or a combination thereof, zorro prints the average and variance of fractional and lattice anisotropies taken over all voxels in the given mask.

Angular Histogram

The primary diffusion direction (an angle in 3D) can be represented by two angles (q_1, q_2). An angular histogram then can be presented as a 2D image in which each pixel is a bin of the histogram and the brightness of a pixel represents the height of that bin. The size of the image represents the number of bins chosen to characterize each angle, n. Angles are calculated and then scaled to a range of $\pm n/2$. For a given primary diffusion vector (x,y,z):

$q_1 = \text{atan}(z/x) / n/2$

$q_2 = \text{atan}(y/h) / p/2$

where $h = \text{sqrt}(z^2 + x^2)$.

To create the histogram, for each voxel of the mask, the bin (pixel location) indicated by these two scaled angles for the primary diffusion vector is incremented.

3D Visualization by “Boxels”

“Boxels” are a method for visualizing DTI data in which all tensor information can be seen at the same time. Along with color coding and modulation by anisotropy, boxes are drawn to indicate the directions and relative lengths of all three vectors. The orientation and lengths of the three parallel faces are proportional to the eigenvectors and eigenvalues, respectively.

APPLICATIONS AND DISCUSSION

Part of "27 - Diffusion Tensor Imaging "

Diffusion tensor imaging tractography has opened up the capability to study white matter fiber pathways in the living brain in clinical conditions and in normality, and can be integrated within more general structural–functional, clinical, and behavioral paradigms. However, basic problems remain to be elucidated. At an anatomic level, fiber tracts can be delineated precisely only at the level of their stems, not at their splays and extreme peripheries. This is a challenging problem that should be solved if DTI is to enable us to delineate tracts in their entirety. At a signal analysis level, a fundamental question relates to the source of the DTI signal. For instance, the relative contributions of the intracellular and extracellular compartments to the diffusion signal are not known with certainty (32 ,33 ,34 and 35). One powerful approach to the solution of this problem is study at the ultrastructural level.

Ultrastructural Studies

To acquire a better understanding of *in vivo* diffusion MR measurements in the nervous system, we investigated the diffusion of water in excised sciatic nerve of mice. Fixed and freshly excised nerves were placed straight in a tube and immersed in Fluorinert (Sigma, St. Louis, Missouri). The sample was positioned in a solenoid with its axis coinciding with the nerve axis and perpendicular to B_0. Also, it was possible to exploit the particular geometry of the sample to align the nerve axis with one of the field gradients (in these experiments, the g_y gradient, or 0, 1, 0). The diffusion MRI experiments were performed at 2.0 tesla (T) (proton frequency

of 84.74 MHz) on a SISCO (Varian Associates, Palo Alto, California) system equipped with an 18-cm horizontal-bore superconducting magnet (Nalorac, Martinez, California) and a set of coils capable of producing 120-mT/m field gradients. Proton MR diffusion-weighted images were obtained by using a spin-echo sequence with an echo time of 50 ms and a repetition time of 1 s. The measurements were performed at room temperature (21°C). Typically, 12 scans were averaged. The *b* values used for diffusion weighting were 0, 200, 400, 800 and 1,600 s • mm^2. The slice thickness was 2 mm, and the field of view was 7.5 × 7.5 mm^2 with an in-plane resolution of 256 × 128 pixels. Figure 27.5A shows a transaxial proton image of a nerve obtained with a *b* value of 0.

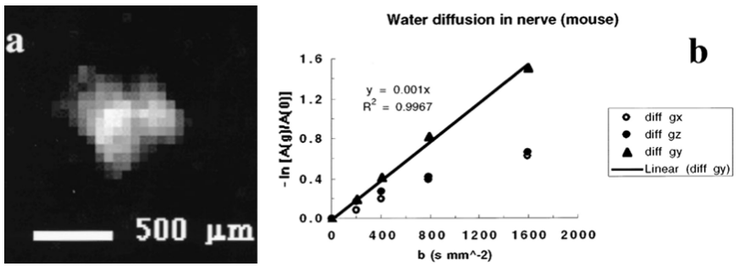


FIGURE 27.5. A: Transaxial proton image of the sciatic nerve of a mouse. **B:** Graph illustrating the variation of the magnetic resonance signal as a function of the orientation of the diffusion gradient. The *circles* (open and closed) correspond to measurements with the diffusion gradients applied in the plane perpendicular to the nerve axis. The *triangles* correspond to diffusion measurements obtained by using the gradient parallel to the nerve axis. The anisotropy of the restriction to the diffusion of water molecules in the cellular compartments is clearly shown.

The effect of restriction to molecular diffusion within the nerve is shown in Fig. 27.5B . Diffusion gradients applied along directions g_x = (1,0,0) and g_z = (0,0,1), which were perpendicular to the nerve axis g_y = (0,1,0) in these experiments, show that the presence of more barriers (i.e., cell membranes and cytoplasmic fibers) hinder molecular diffusion across rather than along the nerve.

Macroscopic Studies

In a sense, the cerebral white matter can be considered as a finite set of discrete and topographically organized fiber pathways or connections that convey connectivity within the brain, and the physical connections in their entirety would make up *anatomic connectivity*. On the other hand, the physiologic outcome of activities of neuronal assemblies, which is coherent in nature and is expressed in the correlated time structure of the firing pattern of its member neurons, would be the basis of *functional connectivity* (i.e., the temporal correlation between remote neurophysiologic events) (36). In the context of a functional experiment, functional connectivity would allow us an array of possible interpretations. The simplest of these solutions that can replicate the observed functional connectivity describes the interactions and connections that are sufficiently active to be detectable at the time of observation. This simplest solution is the *effective connectivity*, which accounts for the interaction that one neural system exerts on another (37).

In vivo DTI-based fiber tract analysis is relevant for the study of *structural-functional* and *anatomic-clinical* relationships. The details of systems neuroanatomy are critical for studies of lesion analysis and for the analysis and interpretation of metabolic and functional neuroimaging data. Damage to specific fiber pathways correlates with the decreased cerebral metabolism pattern observed in stroke patients. Functional activation studies in which PET and fMRI are used contribute to the knowledge of the spatial distribution of cortical and subcortical processing elements. In the creation of effective connectivity models, actual knowledge of brain anatomy is an integrative part because it is utilized for the design of the *a priori* model used in structural equation modeling (38 ,39 ,40 and 41). Use of actual individual information regarding *in vivo* white matter fiber pathway topography and volumetry may optimize and increase the predictive power of these models. These anatomic and neuroimaging studies of the constituents of the fiber tracts and the connections of the human cerebral cortex will be important in acquiring an understanding of the distributed neural circuits that subserve normal brain function. They will also pave the way for future morphometric studies of the white matter fiber systems in normal populations, and for lesion–deficit correlations in patients with focal brain lesions.

The significance of this method is underscored in the study of normal white matter neuroanatomy and also of several *diseases that affect the cerebral white matter*, such as stroke, head trauma, spinal cord injury, and neurodegenerative diseases. Applications of the DTI technique in white matter pathway analysis in normal persons and in cases of *amyotrophic lateral sclerosis* are illustrated. Three normal tracts—the corticospinal projection, the corpus callosum, and the cingulum fiber system—have been selected for the illustration because of their different orientations and also because they represent distinctive classes of connection (i.e., projectional, commissural, and associational, respectively). We have used "zorro" in these examples to show the various ways in which tractography can be performed. Specifically, the corpus callosum and the corticospinal tracts were handpicked, whereas the cingulum fiber system was reconstructed automatically in three dimensions. The results are shown in Fig. 27.6 and Fig. 27.7 .

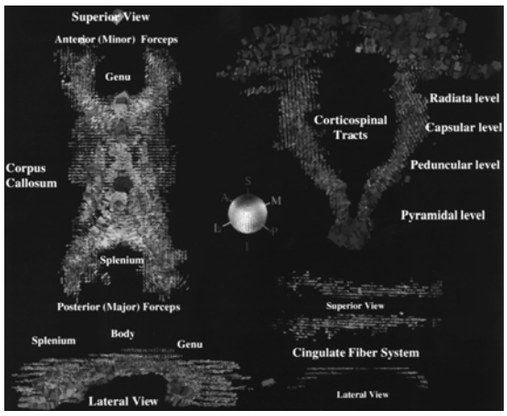


FIGURE 27.6. Three-dimensional reconstruction of stems of individual pathways of a living normal young adult human subject. Note that scaling of tracts is unequal. Red indicates medial–lateral, blue indicates superior–inferior, and green indicates anterior–posterior orientation, respectively, as shown in the color-coded sphere in the center of the figure. For this caption, three representative fiber tracts were selected based on class of pathway provenience and orientation. The corpus callosum, a commissural tract with general mediolateral orientation, is shown in superior and lateral views. Specifically, its genu, body, and splenium are colored in red because they run mediolaterally, whereas its anterior (forceps minor) and posterior (forceps major) forcipes are colored in green because they are oriented anteroposteriorly. The corticospinal tract, a projectional pathway with superior–inferior orientation, is depicted in blue in a frontal view. Finally, the cingulum bundle, a long, associational corticocortical fiber tract with anterior–posterior orientation, is colored green above the body of the corpus callosum for its major trajectory, whereas it is colored blue along its vertically oriented portions in front of the genu and behind the splenium of the corpus callosum. See color version of figure.

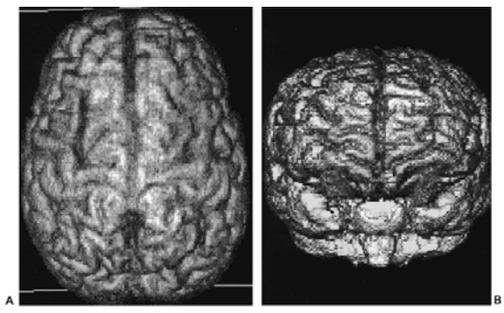


FIGURE 27.7. Three-dimensional (3D) reconstruction of fiber tracts within the three-dimensionally rendered brain of a normal young adult. **A:** A combined 3D representation of the cingulum bundle (green) and the corpus callosum (red). **B:** A composite 3D rendition of the corticospinal projection (blue) and the corpus callosum (red). Red indicates medial-lateral, blue superior-inferior, and green anterior-posterior orientation, respectively. See color version of figure.

Amyotrophic Lateral Sclerosis

In amyotrophic lateral sclerosis (ALS), as degeneration of the upper motor neuron occurs, a progressive damage to the corticospinal tract has been documented in the caudorostral direction (42 ,43 ,44 and 45). We used DTI in two patients (cases 1 and 2) to document pronounced bilateral reduction in the size of the corticospinal tract at the level of the medulla oblongata. Case 1 illustrates an integrated MRI examination of structure, function, and metabolism. A change in morphology of the corticospinal tract can be appreciated in conventional MR images or in the T2-EP images at its ventral surface shown in Fig. 27.8 A,D. However, these MRI techniques do not allow an accurate definition of the borders of the tract. With the aid of tractographic DTI, the corticospinal tract is readily visualized as it courses rostrocaudally. In addition, bordering structures can be visualized, such as the inferior olivary nucleus; thus, a more precise definition of the caudal borders of the corticospinal tract allows quantification of the fiber pathway at this anatomic level. Fractional and lattice anisotropy indices were measured in a 13-voxel region (voxel size = $2 \times 2 \times 2$ mm^3) for the corticospinal tract in the medulla, as shown in Fig. 27.8C . The average fractional anisotropy index was .49 ± .02, and the lattice anisotropy index was .30 ± .01. In this case, in addition to DTI, we collected single-voxel MRS measurements in medulla, and the data indicated metabolic changes—loss of NAA (*N* = acetylaspartate) and possibly elevated glutamate/glutamine (glx), shown in Fig. 27.8H . Similarly, we acquired MRS measurements in the right and left primary motor cortex, shown in Fig. 27.8H . Of note is that the observed metabolic pattern showed a slightly decreased NAA/Cr (creatine) ratio in the left motor cortex in comparison with the contralateral counterpart. Finally, fMRI was acquired while flexion was performed (frequency of 1.5 Hz) with either the left or right hand. An expanded area of activation was observed in the left hemisphere that corresponded with this patient's motor symptomatology (i.e., right-hand weakness much greater than left-hand weakness) (Fig. 27.8I).

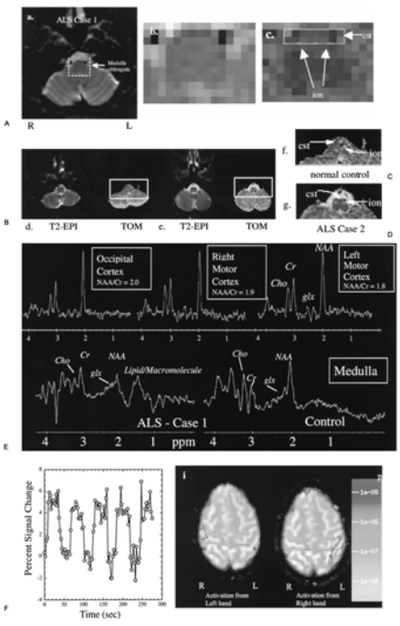


FIGURE 27.8. Comparison of diffusion tensor imaging in the brainstem (medulla oblongata) of a normal control **(D,F)** and two patients with amyotrophic lateral sclerosis, ALS case 1 **(A-C)** and ALS case 2 **(E,G)**. A T2 echo-planar image (T2-EPI) and tensor orientation map (TOM) are included at each level. In this image, green voxels indicate anteroposteriorly, red indicate mediolaterally, and blue indicate superoinferiorly oriented diffusion directions. **H:** Four magnetic resonance spectra from an ALS patient (case 1). The spectra in the top row come from a point resolved spectroscopy sequence (PRESS) with TR (repetition time)/TE (echo time) 2,000/144 ms in either left motor cortex (hand area), right motor cortex (hand area), or occipital cortex. The occipital cortex serves as a control region that is essentially unaffected by ALS. A slight asymmetry in the NAA/Cr and glx/Cr values between left and right motor cortex perhaps reflects the large motor asymmetry in this patient. Also shown on the bottom are PRESS spectra from medulla in the same patient and from a normal control (TR/TE 2,000/30 ms). Note the lower NAA/Cr values in the patient and the higher intensity in the glx region. **I:** This figure shows activation in the hand motor cortex after performance of hand flexion (frequency of 1.5 Hz) with either the left or right hand. Note the expanded area of activation in the left brain, which correlates with the asymmetry of symptoms. The left hand is essentially normal, whereas the right hand is much more severely affected. No difference was noted in frequency of performance of the task with the left or right hand. Also shown is a time course from a region of interest in left motor cortex showing the response of the BOLD (blood oxygenation-dependent) signal to the off and on periods. AG, angular gyrus; CALC, intracalcarine cortex; CGa, cingulate gyrus, anterior; CGp, cingulate gyrus, posterior; CN, cuneiform cortex; CO, central operculum; F1, superior frontal gyrus; F2, middle frontal gyrus; F3o, inferior frontal gyrus, pars opercularis; F3t, inferior frontal gyrus, pars triangularis; FMC, frontal medial cortex; FO, frontal operculum; FOC, frontal orbital cortex; FP, frontal pole; H1, Heschl gyrus; INS, insula; JPL, juxtaparacentral cortex; LG, lingual gyrus; OP, occipital pole; OF, occipital fusiform gyrus; OLi, lateral occipital cortex, inferior; OLs, lateral occipital cortex, superior; PAC, paracingulate cortex; PCN, precuneus; PHa, parahippocampal gyrus, anterior; PHp, parahippocampal gyrus, posterior; PO, parietal operculum; POG, postcentral gyrus; PP, planum polare; PRG, precentral gyrus; PT, planum temporale; SC, subcallosal cortex; SCLC, supracalcarine cortex; SGa, supramarginal gyrus, anterior; SGp, supramarginal gyrus, posterior; SPL, superior parietal lobule; T1a, superior temporal gyrus, anterior; T1p, superior temporal gyrus, posterior; T2a, middle temporal gyrus, anterior; T2p, middle temporal gyrus, posterior; T3a, inferior temporal gyrus, anterior; T3p, inferior temporal gyrus, posterior; TFa, temporal fusiform, anterior; TFp, temporal fusiform, posterior; TO2, middle temporal gyrus, temporooccipital; TO3, inferior temporal gyrus, temporooccipital; TOF, temporooccipital fusiform gyrus; TP, temporal pole (46). See color version of figure.

Integrated Magnetic Resonance Neurologic Examination

Diffusion tensor imaging can be an important part of a comprehensive neurofunctional MRI examination. Quantitative assessment of anatomic volume as measured by morphometric MRI (46), of fiber pathway anisotropy and size by DTI, of metabolic state by MRS, and of functional state by fMRI provides a highly multidimensional data space for the elucidation of structural, functional, behavioral, and clinical correlates. For instance, the relative contributions of each of these imaging modalities can be assessed for any developmental, aging, or degenerative process with the use of principal component analysis. This leads to an optimized, temporally efficient MRI neurologic examination that captures the salient features of the structural, metabolic, and functional states as they change over time. Monitoring these manifestations enables us to elucidate the neurobiological underpinnings of normal brain development and aging in addition to the endpoints of etiology, natural history, and therapeutic intervention in disease states (47). The ALS case 1 is an example of this application. We have proposed using this technique to assess the changes in both motor cortical “remapping” and recruitment of additional cortical areas that may occur progressively in ALS. We successfully performed fMRI in this same patient (Fig. 27.8A), in addition to MRS and DTI. The data indicated larger areas of activation in the left than in the right hemisphere and reflected the motor asymmetry of this patient. The pattern of activation in the right hemisphere was essentially normal, whereas that in the left hemisphere showed recruitment of posterior parietal and association areas. This pattern of “expansion” of cortical activation has also been observed in other neurodegenerative conditions, such as Parkinson disease (48 ,49) and stroke (50 ,51 and 52).

CONCLUSION

Part of "27 - Diffusion Tensor Imaging "

An understanding of the human cerebral white matter, specifically its fiber pathways, is needed. Historically, this objective was achieved to a certain extent in postmortem material, and the findings were therefore of limited practical value. The advancement of neuroimaging technology with such techniques as DTI has made it possible to study human white matter fiber pathways *in vivo*, and therefore in clinical conditions. The DTI method opens up a new approach, *tractography*, for studying the various white matter fiber pathways that are particularly involved in normal cognitive

processing and in certain disease states, such as language, developmental, psychiatric, and demyelinating disorders. In addition, this imaging technique may allow a better understanding of the state of white matter pathways during development, aging, and recovery following brain damage. Finally, the integration of DTI in a more comprehensive MR neurologic examination would allow us to monitor the evolution of a pathologic condition and thus elucidate the endpoints of etiology, natural history, and therapeutic intervention in disease states.

ACKNOWLEDGMENTS

Part of "27 - Diffusion Tensor Imaging "

We thank the following persons for assistance in these projects: D. N. Pandya, E. Kaplan, J. P. Vonsattel, T. G. Reese, E. Kraft, D. N. Caplan, and R. H. Brown, Jr. This work was supported in part by grants from the National Institutes of Health (NS27950, DA09467, and NS34189 as part of the Human Brain Project), and from the National Alliance for Research on Schizophrenia and Depression (NARSAD), the Fairway Trust, the Giovanna Armenise-Harvard Foundation for Advanced Scientific Research, and the Amyotrophic Lateral Sclerosis Association (ALSA).

Dr. Sorensen receives research support from, is a consultant for, or has spoken on behalf of the following companies within the last year: Siemens Medical Systems, General Electric Medical Systems, Neurocrine Biosciences Inc., Matrix Pharmaceuticals Inc., Millennium Pharmaceuticals Inc., Vertex Pharmaceuticals Inc., and Berlex Laboratories.

In addition, Dr. Sorensen has an equity position in and holds the position of Medical Director at EPIX Medical Inc., a specialty pharmaceutical company based in Cambridge, Massachusetts, engaged in developing targeted contrast agents for cardiovascular MRI.

REFERENCES

1. Pandya DN, Yeterian E. Architecture and connections of cortical association areas. In: Peters, Jones EG, eds. *Cerebral cortex*. New York: Plenum Publishing, 1985.

2. Mountcastle VB. *The mindful brain: part I*. Cambridge, MA: MIT Press, 1978.

3. Shepherd GM. *Neurobiology*. New York: Oxford University Press, 1994.

4. Cajal SR. *Histology of the nervous system of man and vertebrates*. New York: Oxford University Press, 1995.

5. Pandya DN, Kuypers HGJM. Cortico-cortical connections in the rhesus monkey. *Brain Res* 1969;13:13–36.

6. Jones EG, Powell TPS. An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain* 1970;93:793–820.

7. Sperry RW. Cerebral organization and behavior. *Science* 1961;133:1749–1757.

8. Geschwind N. Disconnexion syndromes in animals and man. *Brain* 1965;88:584–644.

9. Mesulam M-M. *Principles of behavioral neurology*. Philadelphia: FA Davis Co, 1985.

10. Dejerine J. *Anatomie des centres nerveux*. Paris: Rueff et Cie, 1895.

11. Krieg WJS. *Connections of the cerebral cortex*. Chicago: Brain Books, 1963.

12. Makris N, Worth AJ, Sorensen AG, et al. Morphometry of *in vivo* human white matter association pathways with diffusion-weighted magnetic resonance imaging. *Ann Neurol* 1997;42:951–962.

13. Friedman DI, Johnson JK, Chorsky RL, et al. Labelling of human retinohypothalamic tract with the carbocyanine dye, DiI. *Brain Res* 1991;560:297–302.

14. Clarke S, Miklossy J. Occipital cortex in man: organization of callosal connections, related myelo- and cytoarchitecture, and putative boundaries of functional visual areas. *J Comp Neurol* 1990;298:188–214.

15. Cowan WM, Gottlieb DI, Hendrickson AE, et al. The autoradiographic demonstration of axonal connections in the central nervous system. *Brain Res* 1972;37:21–51.

16. Makris N, Meyer J, Bates J, et al. MRI-based topographic parcellation of cerebral central white matter and nuclei: II. Rationale and applications with systematics of cerebral connectivity. *Neuroimage* 1999;9:18–45.

17. Meyer J, Makris N, Bates J, et al. MRI-based topographic parcellation of the human cerebral white matter: I. Technical foundations. *Neuroimage* 1999;9:1–17.

18. Mori S, Crain B, Chacko VP, et al. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol* 1999;45:265–269.

19. Conturo TE, Lori NF, Cull TS, et al. Tracking neuronal fiber pathways in the living human brain. *Proc Natl Acad Sci U S A* 1999;96:10422–10427.

20. Wedeen VJ, Davis TL, Lautrup BE, et al. Diffusion anisotropy and white matter tracts. *Proceedings of the second international conference on functional mapping of the human brain* 1996;3: S146.

21. Pierpaoli CJ, Basser PJ, Barnett A, et al. Diffusion tensor MR imaging of the human brain. *Radiology* 1996;201: 637–648.

22. Hahn E. Spin echoes. *Phys Rev* 1950;80:580–594.

23. Carr HY, Purcell EM. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys Rev* 1954;94:630–638.

24. Le Bihan D, Breton E, Lallemand D, et al. MR imaging of intravoxel incoherent motions: applications to diffusion and perfusion in neurologic disorders. *Radiology* 1986;161:401–407.

25. Filley CM. The behavioral neurology of cerebral white matter. *Neurology* 1998;50:1535–1540.

26. Rye DB. Tracking neural pathways with MRI. *Trends Neurosci* 1999;22:373–374.

27. Stejskal EO, Tanner JE. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J Chem Phys* 1965;42: 288–292.

28. Pierpaoli C, Basser PJ. Toward a quantitative assessment of diffusion anisotropy. *Magn Reson Med* 1996;36:893–906.

29. Pierpaoli C, Basser PJ. Toward a quantitative assessment of diffusion anisotropy [Erratum]. *Magn Reson Med* 1996;37:972.

30. Basser PJ, Pierpaoli C. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J Magn Reson B* 1996;111:209–219.

31. Schroeder W, Martin K, Lorensen B. *The visualization toolkit*. Upper Saddle River, NJ: Prentice Hall, 1998.

32. Nicholson C, Sykove E. Extracellular space structure revealed by diffusion analysis. *Trends Neurosci* 1998;21:207–215.

33. Beaullieu C, Fenrich F, Allen PS. Multicomponent water proton transverse relaxation and T2-discriminated water diffusion in myelinated and nonmyelinated nerve. *Magn Reson Imaging* 1998;16:1201–1210.

34. Norris DG, Niendorf T. Interpretation of DW-NMR data: dependence on experimental conditions. *Nucl Magn Reson Biomed* 1995;8:280–288.

35. Peled SH, Gudbjartsson H, Westin C-F, et al. Magnetic resonance imaging shows orientation and asymmetry of white matter fiber tracts. *Brain Res* 1998;780:27–33.

36. Aertsen A, H. Preissl H. Dynamics of activity and connectivity in physiological neuronal networks. In: Schuster, ed. *Nonlinear dynamics and neuronal networks*. New York: VCH Publishers, 1991.

37. Friston K. Functional and effective connectivity in neuroimaging: a synthesis. *Hum Brain Mapping* 1994;2:56–78.

38. McIntosh AR, Gonzalez-Lima F. Structural equation modelling and its application to network analysis in functional brain imaging. *Hum Brain Mapping* 1994;2:2–23.

39. Gonzalez-Lima F, McIntosh AR. Neural network interactions related to auditory learning analyzed with structural equation modelling. *Hum Brain Mapping* 1994;2:23–44.

40. Horwitz B, Tagamets M-A, McIntosh AR. Neural modeling, functional brain imaging, and cognition. *Trends Cogn Sci* 1999;3:91–98.

41. Buchel C, Coull JT, Friston KJ. The predictive value of changes in effective connectivity for human learning. *Science* 1999;283:1538–1541.

42. Colmant H-J. Die myatrophische Lateralsklerose. *Handbuch der speziellen pathologischen Anatomie und Histologie* 1958;XIII/2 Bandteil B:2624–2692.

43. Rowland L. Natural history and clinical features of amyotrophic lateral sclerosis and related motor neuron diseases. In: Calne, ed. *Neurodegenerative diseases*. Philadelphia: WB Saunders, 1994.

44. Martin J, Swash M, Schwartz M. New insights in motor neuron disease. *Neuropathol Appl Neurobiol* 1990;16:97–110.

45. Brown R. Amyotrophic lateral sclerosis. *Arch Neurol* 1997;54:1246–1250.

46. Caviness VSJ, Makris N, Meyer J, et al. MRI-based parcellation of human neocortex: an anatomically specified method with estimate of reliability. *J Cogn Neurosci* 1996;8:566–588.

47. Jenkins BG. An integrated strategy for evaluation of metabolic and oxidative defects in neurodegenerative illness using magnetic resonance techniques. *Ann N Y Acad Sci* 1999;893:214–242.

48. Brooks D. The role of the basal ganglia in motor control: contributions from PET. *J Neurol* Sci 1995;128:1–13.

49. Brooks DJ. Motor disturbance and brain functional imaging in Parkinson's disease. *Eur Neurol* 1997;38[Suppl 2]:26–32.

50. Cramer SC, Nelles G, Benson RR, et al. A functional MRI study of subjects recovered from hemiparetic stroke. *Stroke* 1997;28:2518–2527.

51. Rossini PM, Caltagirone C, Castriota-Scanderbeg A, et al. Hand motor cortical area reorganization in stroke: a study with fMRI, MEG and TCS maps. *Neuroreport* 1998;9:2141–2146.

52. Thulborn KR, Carpenter PA, Just MA. Plasticity of language-related brain function during recovery from stroke. *Stroke* 1999;30:749–754.

53. Ludwig E, Klingler J. *Atlas cerebri humani*. Boston: Little, Brown and Company, 1956.

28

Activation Paradigms in Affective and Cognitive Neuroscience: Probing the Neuronal Circuitry Underlying Mood and Anxiety Disorders

Richard J. Davidson

Richard J. Davidson: Laboratory for Affective Neuroscience, University of Wisconsin, Madison, Wisconsin.

Virtually all forms of psychopathology are associated with disturbances in various aspects of affect and cognition. Although most clinical research has relied on relatively coarse phenomenologic descriptions of symptoms, recent work in neuroimaging with behavioral activation paradigms offers a new and more penetrating look at specific cognitive and affective processes in psychopathology. This new trend is predicated on the view that we must go beyond phenomenology to understand the brain circuitry that is associated with complex mood and anxiety disorders. Advances in our understanding of these conditions will emerge from research that is designed to examine more specific cognitive and affective processing abnormalities. This work holds promise in revealing additional targets for therapeutic intervention, both behavioral and pharmacologic. It also will be important in helping to expose the heterogeneity of these disorders and in offering more meaningful ways in which to parse various subtypes. Finally, by examining the impact of particular therapeutic interventions on functional brain activity elicited in the context of activation paradigms, a better understanding of the impact of these interventions on specific subcomponents of the brain circuitry underlying affect and cognition is likely to emerge.

In this chapter, some key elements of the circuitry that is most relevant to understanding mood and anxiety disorders are first reviewed. The role of individual differences in the functional activity of this circuitry is then considered. The next section reviews key approaches and findings of activation paradigms that have been used in this area. The chapter concludes with a summary and a discussion of future trends in this rapidly developing area.

CIRCUITRY OF AFFECT AND COGNITION IN MOOD AND ANXIETY DISORDERS

Part of "28 - Activation Paradigms in Affective and Cognitive Neuroscience: Probing the Neuronal Circuitry Underlying Mood and Anxiety Disorders "

The review presented in this section of the key components of the circuitry underlying aspects of emotion and cognition that are most relevant to mood and anxiety disorders is gleaned mostly from studies of lesions experimentally produced in animals, the human lesion literature, and neuroimaging studies in normal humans. The review focuses on various territories of the prefrontal cortex, amygdala, hippocampus, and anterior cingulate cortex. Collectively, these studies provide important clues regarding the types of activation paradigms that are most promising for use in patients with mood and anxiety disorders to probe the underlying circuitry of affect and cognition. Research in which activation paradigms with neuroimaging are applied in patients with mood and anxiety disorders is reviewed in a subsequent section.

Prefrontal Cortex

A large corpus of data at both the animal and human levels implicates various sectors of the prefrontal cortex (PFC) in both cognition and emotion. The PFC is not a homogeneous zone of tissue; rather, it has been differentiated on the basis of both cytoarchitectonic and functional considerations. The three subdivisions of the primate PFC that have been consistently distinguished are the dorsolateral, ventromedial, and orbitofrontal sectors of the PFC. In addition, it appears that important functional differences exist between the left and right sides within some of these sectors.

The role of the PFC in cognitive control has recently been reviewed (1), so it is not extensively considered here other than to underscore that a major function of the PFC in general is “to extract information about the regularities

across experiences and so impart rules that can be used to guide thought and action" (1). One of the principal roles of the PFC is to represent goal-relevant information, a key component of both complex thought and emotion. As many studies at the nonhuman primate level have now documented, reward-related information plays a key role in modulating the activity of PFC neurons. Activity in both lateral and ventromedial zones of the PFC is associated with the identity and size of expected rewards (2). This component of PFC activity is likely governed by a dopaminergic input from the ventral tegmental area of the midbrain (see ref. 1 for review). Our notion of the role of the PFC in pre-goal attainment positive affect is based on this corpus of research, which is discussed below (3 ,4).

The case for the differential importance of left and right PFC sectors in emotional processing was first made systematically in a series of studies of patients with unilateral cortical damage (5 ,6 and 7). Each of these studies compared the mood of patients with unilateral left or right-sided brain damage and found a greater incidence of depressive symptoms following left-sided damage. In most cases, the damage was fairly gross and likely included more than one sector of the PFC and often other brain regions as well. The general interpretation that has been placed on these studies is that depressive symptoms are increased following left-sided anterior PFC damage because this brain territory participates in certain forms of positive affect, particularly pre-goal attainment positive affect; damage leads to deficits in the capacity to generate this form of positive affect, a hallmark feature of depression (8). It should be noted that not all studies support this conclusion. In a recent metaanalysis of lesion studies, Carson et al. (9) failed to find support for this hypothesis. Davidson (10) has previously reviewed many of these studies and has addressed a number of critical methodologic and conceptual concerns in this literature. The most important of these issues is that according to the diathesis stress model of anterior activation asymmetry proposed by Davidson and colleagues (11 ,12 and 13), individual differences in anterior activation asymmetry, whether lesion-induced or functional, represent a diathesis. As such, they alter the probability that specific forms of emotional reactions will occur in response to the requisite environmental challenge. In the absence of such a challenge, the pattern of asymmetric activation will simply reflect a propensity but will not necessarily culminate in differences in mood or symptoms. In a recent study of mood sequelae in patients with unilateral lesions with the largest sample size to date (*n* = 193), Morris et al. (14) found that among stroke patients, it was only in those with small lesions that the relation between left PFC damage and depressive symptoms was observed. It is likely that larger lesions intrude on other brain territories and mask the relation between left PFC damage and depression.

A growing corpus of evidence in normal intact humans is consistent with the findings derived from the lesion studies. Davidson and colleagues have reported that induced positive and negative affective states shift the asymmetry in prefrontal brain electrical activity in lawful ways. For example, film-induced negative affect increases relative right-sided prefrontal and anterior temporal activation (15), whereas induced positive affect elicits an opposite pattern of asymmetric activation. This general pattern has been replicated by others using similar measures (16 ,17). In positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) studies, with considerably better spatial resolution, similar PFC activations have been reported, although many important methodologic details must be considered in interpreting the findings (see ref. 4 for review). The most important of these is considered in a later section. In addition, a body of evidence supports the conclusion that individual differences in baseline levels of asymmetric activation in these brain regions are lawfully related to variations in dispositional affective style (18).

The ventromedial PFC has been implicated in the anticipation of future positive and negative affective consequences. Bechara and colleagues (19) have reported that patients with bilateral lesions of the ventromedial PFC have difficulty anticipating future positive or negative consequences, although immediately available rewards and punishments do influence their behavior. Such patients show decreased levels of electrodermal activity in anticipation of a risky choice in comparison with controls, whereas controls exhibit such autonomic change before they explicitly know that a choice is risky (20 ,21 and 22).

The findings from the lesion method when effects of small unilateral lesions are examined and from neuroimaging studies in normal subjects and patients with anxiety disorders converge on the conclusion that increases in right-sided activation in various sectors of the PFC are associated with increased negative affect. Less evidence is available for the domain of positive affect, in part because positive affect is much harder to elicit in the laboratory and because of the negativity bias (23 ,24). This latter phenomenon refers to the general tendency of organisms to react more strongly to negative than to positive stimuli, perhaps as a consequence of evolutionary pressures to avoid harm. The findings of Bechara et al. (19) on the effects of ventromedial PFC lesions on the anticipation of future positive *and* negative affective consequences are based on studies of patients with bilateral lesions. It will be of interest in the future to examine patients with unilateral ventromedial lesions to ascertain whether valence-dependent asymmetric effects are also present, although most lesions in this PFC territory are bilateral.

Systematic studies designed to disentangle the specific role played by various sectors of the PFC in emotion are lacking, although a growing corpus of work illustrates the functional differentiation among different sectors of the PFC in different aspects of cognitive control (25). Many theoretical accounts of emotion assign it an important role

in guiding action and organizing behavior toward the acquisition of motivationally significant goals (26 ,27). This process requires that the organism have some means of representing affect in the absence of immediately present rewards and punishments and other affective incentives. Such a process may be conceptualized as a form of affective working memory. It is likely that the PFC plays a key role in this process (28). Damage to certain sectors of the PFC impairs an individual's capacity to anticipate future affective outcomes and consequently results in an inability to behave in an adaptive fashion. Such damage is not likely to disrupt an individual's response to immediate cues for reward and punishment, only the anticipation before and maintenance after an affective cue has been presented. This proposal can be tested with current neuroimaging methods (e.g., fMRI) but has not yet been rigorously evaluated. With regard to the different functional roles of the dorsolateral, orbitofrontal, and ventromedial sectors of the PFC, Davidson and Irwin (4) suggested on the basis of both human and animal studies that the ventromedial sector is most likely involved in the representation of elementary positive and negative affective states in the absence of immediately present incentives. The orbitofrontal sector has most firmly been linked to rapid learning and unlearning of stimulus-incentive associations and has been particularly implicated in reversal learning (29). As such, the orbitofrontal sector is likely key to understanding aspects of emotion regulation (30). One critical component of emotion regulation is the relearning of stimulus-incentive associations that may have been previously maladaptive, a process likely requiring the orbitofrontal cortex. The dorsolateral sector is most directly involved in the representation of goal states toward which more elementary positive and negative states are directed.

Amygdala

A large corpus of research at the animal (mostly rodent) level has established the importance of the amygdala in emotional processes (31 ,32 and 33). Because many reviews of the animal literature have appeared recently, a detailed description of these studies is not presented here. LeDoux and colleagues have marshaled a large corpus of compelling evidence to suggest that the amygdala is necessary to establish conditioned fear. Whether the amygdala is necessary to express that fear following learning and whether the amygdala is the actual locus where learned information is stored is still a matter of some controversy (34 ,35). The classic view of amygdala damage in nonhuman primates (resulting in major affective disturbances as expressed in the Kluver–Bucy syndrome, in which the animal exhibits an abnormal approach, hyperorality and hypersexuality, and little fear) is now thought to be a function of damage elsewhere in the medial temporal lobe. When very selective excitotoxic lesions of the amygdala are made that preserve fibers of passage, nothing resembling the Kluver–Bucy syndrome is observed (36). This diverse array of findings suggests a more limited role for the amygdala in certain forms of emotional learning, although the human data imply a more heterogeneous contribution.

Although the number of patients with discrete lesions of the amygdala is small, they have provided unique information about the role of this structure in emotional processing. A number of studies have now reported specific impairments in the recognition of facial expressions of fear in patients with restricted amygdala damage (37 ,38 ,39 and 40). Recognition of facial signs of other emotions have been found to be intact. In a study that required subjects to make judgments about the trustworthiness and approachability of unfamiliar adults based on facial photographs, patients with bilateral amygdala damage judged the unfamiliar persons to be more approachable and trustworthy than did control subjects (41). Recognition of vocalic signs of fear and anger was found to be impaired in a patient with bilateral amygdala damage (42), which suggests that this deficit is not restricted to facial expressions. Other researchers demonstrated an impairment of aversive autonomic conditioning in a patient with amygdala damage despite the fact that the patient demonstrated normal declarative knowledge of the conditioning contingencies (43). Collectively, these findings from patients with selective bilateral destruction of the amygdala suggest specific impairments on tasks that tap aspects of negative emotion processing. Most of the studies have focused on perception; the data clearly show the amygdala to be important in recognizing cues of threat or danger. The conditioning data also indicate that the amygdala may be necessary for acquiring new implicit autonomic learning of stimulus–punishment contingencies. In one of the few studies to examine the role of the amygdala in the expression of already learned emotional responses, Angrilli and colleagues (44) described a patient with a benign tumor of the right amygdala who underwent an emotion-modulated startle study. Among control subjects, they observed the well-known effect of startle potentiation during the presentation of aversive stimuli. In the patient with right amygdala damage, no startle potentiation was observed in response to aversive versus neutral stimuli. These findings suggest that the amygdala may be necessary for the expression of an already learned negative affect.

Hippocampus and Anterior Cingulate Cortex

In this section, the contributions of the hippocampus and anterior cingulate cortex (ACC) to emotion and cognition are briefly mentioned. A more extensive discussion of the contributions of this circuit to emotional and cognitive processing can be found in several recent reviews (4 ,45 ,46 and 47).

The hippocampus has been implicated in various aspects of memory (47), particularly declarative memory of the sort we experience when we consciously recall an earlier episode. The contribution of the hippocampus to emotion and affective

style has only recently begun to be gleaned from the available corpus of animal studies on its role in context-dependent memory (48). This literature has generally supported a role for the hippocampus in the learning of context. For example, when an animal is exposed to a procedure in which a discrete cue is paired with an aversive outcome, in addition to learning the specific cue–punishment contingency, the animal learns to associate the context in which the learning occurs with the aversive outcome. Lesions to the hippocampus abolish this context-dependent form of memory but have no effect on learning of the cue–punishment contingency. The fact that the hippocampus has a very high density of glucocorticoid receptors and participates in the regulation of the hypothalamic–pituitary–adrenal axis is particularly germane to the importance of this structure in regulating emotion. Basic research at the animal level has demonstrated the powerful impact of glucocorticoids on hippocampal neurons (32 ,49). Data indicate that the exogenous administration of hydrocortisone to humans impairs explicit memory that is presumably hippocampus-dependent (50), although other data that suggest that in more moderate amounts, cortisol may facilitate memory (Abercrombie, *unpublished doctoral dissertation*, Department of Psychology, University of Wisconsin at Madison, 2000). A number of investigators using MRI-based measures have reported that hippocampal volume is significantly decreased in patients with several stress-related disorders, including posttraumatic stress disorder (PTSD) (51) and depression (52 ,53), although several failures to replicate these findings have also been reported (54). In the studies in which hippocampal atrophy has been found, the implication is that excessively high levels of cortisol associated with the stress-related disorder cause hippocampal cell death and result in the hippocampal atrophy seen on MRI. Although virtually all these studies have focused on the effects of hippocampal changes on cognitive function, particularly declarative memory, we have proposed that the hippocampus also plays a key role in the context modulation of emotional behavior (55). Moreover, we have suggested that it is in the affective realm that the impact of hippocampal involvement in psychopathology may be most apparent, and that in persons with compromised hippocampal function, the normal context-regulatory role of this brain region is impaired, so that they consequently display emotional behavior in inappropriate contexts. This argument holds that what may be particularly abnormal in disorders such as PTSD and depression is not the display of “abnormal emotion” but rather the display of perfectly normal emotion in inappropriate contexts. For example, in the case of PTSD, extreme fear and anxiety were likely very adaptive in the original traumatic context. This extreme emotional response probably plays an important role in facilitating an organism's withdrawal from a threatening situation. However, in PTSD, this response is elicited in inappropriate situations. The patient with PTSD behaves like the animal with a hippocampal lesion in failing to modulate emotional responses in a context-appropriate manner. These suggestions are only inferential at the present time. Neuroimaging studies are needed to document the role of the hippocampus in this process in normal and disordered populations. In addition, further study is needed to understand how and why the hippocampus may preferentially extract and process information about context. Finally, some research (56) indicates that other structures with direct connections to the hippocampus (e.g., the bed nucleus of the stria terminalis) play a role similar to that of the hippocampus. More work is needed to understand the differential contributions of the various components of this circuitry.

Many studies that have used neuroimaging methods to probe patterns of brain activation during the arousal of emotion have reported that the ACC activates in response to emotion. Several investigators (45 ,57) have recently distinguished between cognitive and affective subdivisions of the ACC based on where activations lie in response to tasks that are purely cognitive versus those that include aspects of emotion. The various tasks used to make these inferences are described in a subsequent section. Based on the model of Carter et al. (58) of the role of the ACC in conflict monitoring in the cognitive domain, we have proposed that the affective subdivision of the ACC may play a similar role in emotion (4). When emotion is elicited in the laboratory, something of a conflict arises because social norms dictate certain rules for participant behavior that do not usually include the display of strong emotion. Thus, the very process of activating emotion in the unfamiliar context of a laboratory environment might activate the ACC. Carter et al. (58) have suggested that ACC activation results in a call for further processing by other brain circuits to address the conflict that has been detected. In most individuals, automatic mechanisms of emotion regulation are likely invoked to dampen strong emotion that may be activated in the laboratory. The initial call for the processes of emotional regulation may result from ACC activation.

PROBING THE NEURAL CIRCUITRY OF AFFECT AND COGNITION IN PATIENTS WITH MOOD AND ANXIETY DISORDERS: CONCEPTUAL AND METHODOLOGIC CONUNDRA

Part of "28 - Activation Paradigms in Affective and Cognitive Neuroscience: Probing the Neuronal Circuitry Underlying Mood and Anxiety Disorders "

In this section, some of the key conceptual and methodologic issues in the use of activation paradigms to probe dysfunctions in the underlying neural circuitry of cognition and affect in patients with mood and anxiety disorders are considered. Issues specific to the study of dysfunctions in the circuitry of emotion in children are considered in a recent review by Davidson and Slagter (59). A key issue that is often neglected in the design of activation studies is the specification of how deficits in the process that is being

studied may account for the symptoms of the disorder. For example, many of the early PET studies in patients with various types of psychopathology used easy continuous performance tasks in which behavioral differences between groups were not expected to occur, or they used unilateral somatosensory stimulation (see ref. 60 for review of early studies). Just what the hypothesized relation was between abnormalities in activation patterns in response to such tasks and symptoms of the disorder being studied was most often not specified in these earlier studies. The better the conceptual link between task performance and symptomatology, the more useful an activation paradigm will be for revealing the underlying deficits in the disorder in question. Several examples of strong conceptual connections between specific task-related deficits and symptomatology in both the cognitive and affective domains are available and can be consulted by the interested reader (see ref. 61 for an example in the cognitive domain and ref. 30 for an example in the affective domain).

The use of tasks that require active performance on the part of subjects poses a host of methodologic issues that are crucial for studies of psychopathology. One of the most important of these is matching the difficulty of an experimental task with that of a control task. This is an issue with a long history in experimental research in psychopathology (62), although the neuroimaging field has yet to appreciate its significance fully. When performance on two tasks is compared between groups, it is imperative that the difficulty of the two tasks be matched. If one task is more difficult than the other task in normal subjects, than a differential deficit on one versus the other task may be a consequence of differences in task difficulty and not specific to the processes that are putatively required for performance of the task. Chapman and Chapman (62) have provided many examples of such artifactual group differences that are products of variation in task performance rather than reflections of differential deficit. It is therefore essential in neuroimaging studies for activation tasks to be matched in this way. If the tasks that are being compared in imaging studies are not matched, then any difference found in activation between tasks may arise as a consequence of differences in the difficulty level of the tasks. Unfortunately, the neuroimaging literature is replete with task comparisons for tasks that do indeed differ in the level of difficulty and thus are particularly problematic for comparisons between groups. The challenge is to design control conditions that are matched to the experimental conditions in regard to basic stimulus and response components, in addition to task difficulty. In one of the few studies to have addressed this potential source of confound, Barch et al. (63), using fMRI, found that the sustained PFC increases in working memory tasks were a function of specific task requirements when they compared such tasks to control tasks that were matched in level of difficulty but did not require working memory.

In studies with patients, investigators frequently wish to examine changes over time with treatment. In this way, effects that may be specifically associated with the symptoms of the disorder can be disentangled from those associated with vulnerability to the disorder. The latter class of effects may also arise as a consequence of scarring—effects produced by having once had the disorder. In experimental designs that require subjects to be scanned and administered tasks on two or more occasions, it is imperative to have data on the test–retest stability of the effects in question. If the effects do not show stability over time, it becomes difficult to interpret group differences in change over time in task activations. We have strongly advocated the psychometric assessment of both psychophysiologic (64) and neuroimaging (65) measures. Such assessments can turn up important surprises. Resting regional glucose metabolism measured with PET is frequently used to assess baseline differences in regional brain activation in various forms of psychopathology. Using MRI coregistration and regions of interest, we recently examined the test–retest stability across a 6-month period of such baseline measures of glucose metabolism in subcortical regions implicated in affective processing. We found that all the regions we examined showed good test–retest stability, including the left and right hippocampus, left and right anterior caudate region, left and right thalamus, and the left amygdala, but not the right amygdala (65). The right amygdala apparently varied over time, in part because metabolic rate in this region was more affected by the stress of the first scan in comparison with activation elsewhere.

Emotional pictures are frequently used to provoke changes in affect in imaging studies (66). When these pictures are used to compare patients and controls over time, it is again important to establish that the effects produced are stable over time in normal subjects. We used startle to probe the test–retest stability of the potentiation produced by negative pictures and the attenuation produced by positive pictures, and we found poor stability when the same pictures were used on both occasions. It was only when different pictures were used, matched on valence and arousal characteristics to the original set, that we found better stability (64). These data underscore the importance of not assuming that effects will be stable over time and the utility of actually measuring the test–retest stability of both task performance and activation changes in normal subjects before conducting a longitudinal study of changes in patients.

The final issue I wish to raise here pertains to studies in which emotion is provoked by specific task manipulations, such as pleasant and unpleasant pictures, guided imagery, monetary rewards and punishments, and symptom provocation with the use of actual feared objects, pictures of objects, or imagined objects. When such paradigms are used, it is imperative for the investigator to verify independently the presence of the intended affective state. Ideally, such verification should include more than self-report measures. For example, peripheral biological indices (e.g., emotion-modulated

startle, electrodermal activity) can often be effective when utilized in imaging studies to provide an independent index of the effects of the intended emotion. Moreover, when such measures are used, correlations between activations produced by the task in question and changes in the peripheral biological index can be computed and are often revealing. For example, Furmark et al. (67) found that subjects showing larger conditioned electrodermal changes in a classic conditioning task showed greater increases in blood flow in the right amygdala during conditioning.

ACTIVATION STUDIES IN PATIENTS WITH MOOD AND ANXIETY DISORDERS

Part of "28 - Activation Paradigms in Affective and Cognitive Neuroscience: Probing the Neuronal Circuitry Underlying Mood and Anxiety Disorders "

Most of the extant imaging studies of patients with mood disorders have been performed with PET while the subjects are in a baseline state. These findings have been recently reviewed elsewhere (68). Recent studies using these methods have reported associations between the severity of particular symptom clusters and patterns of regional blood flow or metabolism (69 ,70 and 71). These studies have underscored the importance of differentiating among various symptoms of depression and illustrate the lawful relations that can be gleaned by examining associations between specific symptoms and patterns of regional brain activity. The few studies using activation paradigms that have been conducted in patients with mood disorders have utilized complex cognitive tasks designed to activate the PFC and ACC. Several studies from Dolan's group (72 ,73 and 74) assessed the relationship of regional blood flow to performance on complex planning tasks during depressed mood in normal subjects and unipolar depressives. Depressed subjects failed to show normal task-related increases in blood flow in regions of the PFC, ACC, basal ganglia, and thalamus.

Several reports have been published of deficits in task performance in depressed patients in which tasks were used that have been extensively studied in previous neuroimaging or neuropsychological research. For example, Merriam et al. (75) studied Wisconsin Card Sorting performance in a large group of patients with major depression who had been without medication for at least 28 days. They found significant deficits on various indices of the Wisconsin Card Sorting task in these patients in comparison with controls. Moreover, patients with more severe depression, reflected in the Hamilton Depression Scale, performed more poorly. Merriam et al. (75) interpreted their data as consistent with suggestions of a dysfunction in prefrontal function in depression.

Other investigators have suggested that in addition to prefrontal deficits, right-sided parietal dysfunction can also contribute to specific symptoms of depression (76). Henriques and Davidson (77), using extremely carefully psychometrically matched verbal and spatial tasks chosen to reflect left- and right-sided posterior cortical function, found a selective deficit on the spatial cognitive task (dot localization) in depressed subjects in comparison with controls. Moreover, in this study, measures of brain electric activity paralleled the performance data and revealed deficits in activation in the right posterior scalp.

We have begun using positive and negative emotional pictures to probe affective processing in depressed patients and controls and to examine changes over time with treatment (see ref. 78 for early preliminary findings). In more recent work with this same paradigm, we have found that patients show a reduction in MR signal intensity in the amygdala in response to negative versus neutral pictures with treatment, whereas controls tested at the same points in time do not. Moreover, the magnitude of MR signal change in the amygdala predicts treatment response (55).

A unique strategy used in research on mood disorders is the short-term depletion of tryptophan among remitted depressed patients maintained on selective serotonin reuptake inhibitors. The depletion of tryptophan, which reduces the presynaptic availability of serotonin, often results in depressive relapse. Thus, this method can be powerfully harnessed to examine activation patterns during the production of depressive relapse in mood-disordered patients. Bremner at al. (79) examined regional metabolic rate with PET during tryptophan depletion and placebo. When they compared subjects who showed a depletion-induced relapse in symptoms with those without relapse, they found that tryptophan depletion resulted in decreases in regional metabolism in the dorsolateral PFC, thalamus, and orbitofrontal cortex in patients who relapsed, but not in patients without relapse. Furthermore, patients who relapsed had a higher baseline (i.e., placebo) metabolism in several areas, including the dorsolateral PFC, orbitofrontal cortex, hippocampus, and amygdala, than those who did not relapse, which possibly suggests that increased basal activity in these structures increases vulnerability to depressive relapse.

We are currently using a task designed to elicit anticipatory positive affect, a form of positive affect that, as noted earlier in this article, is probably implemented at least in part in the dorsolateral PFC. We have hypothesized that this form of positive affect is abnormally decreased in patients with depression (80 ,81). The task we designed is a computerized “lottery” task in which subjects are required to choose digits that may or may not match the digits displayed by a computer after a 10-second delay during which the digits spin like a slot machine. We have found reliable attenuation of startle magnitude at selected points in time during this task (82), and we are now studying a variant of this task in the scanner with fMRI in both normal persons and patients with depression.

Many more studies have been performed in patients with various anxiety disorders (see ref. 83 for recent review). In general, most studies that have used either symptom provocation or other procedures designed to activate the amygdala have found greater activation in this region in response to

such stimuli in anxious patients than in controls. For example, in two studies using script-driven imagery and PET to assess regional blood flow, increased activation was found in the amygdala of patients with PTSD (84 ,85). In a more recent study comparing patients with PTSD and controls, Rauch et al. (86) reported an increased activation of the amygdala in the PTSD patients in response to masked facial expressions of fear versus masked expressions of happiness.

Right-sided activation in various territories of the PFC has been found as a general characteristic of anxiety when symptoms are provoked in patients with several different anxiety disorders (e.g., obsessive-compulsive disorder, simple phobia, and PTSD) (87). In a series of studies that used PET to measure regional cerebral blood flow, Fredrikson and colleagues (88; see ref. 89 for review) reported increases in secondary visual associative regions in patients with snake phobia in response to the presentation of phobia-relevant visual stimuli (e.g., pictures of snakes) versus control visual stimuli. Interestingly, in a separate group of patients with arachnophobia, this pattern did not change after the administrative of diazepam when the subjects were rescanned (90). Using fMRI, Birbaumer et al. (91) explored activation of the amygdala of patients with social phobia relative to that in healthy controls as they were exposed to slides of neutral faces and aversive odor stimuli. The subjects in this study were all male; seven had been given a DSM-IV diagnosis of social phobia and five were healthy controls matched for age, sex, and education. Neutral faces, which do not lead to amygdala activation in nonpsychopathologic humans (92), and aversive odors, which are significantly associated with amygdala activation in comparison with a no-odorant control condition (93), were presented to all the subjects. Birbaumer et al. (91) compared activation in the thalamus and amygdala in the two groups. In both groups, odors elicited greater bilateral activation in the amygdala than in the thalamus. In contrast, the social phobics responded to the faces with significantly greater bilateral amygdala activation than did the controls. However, no difference in regional activation of the thalamus was found between the two groups in response to the neutral faces. Interestingly, although significant amygdala activation was noted in the social phobics, their subjective ratings of the faces did not differ from those of the controls.

SUMMARY AND CONCLUSIONS

Part of "28 - Activation Paradigms in Affective and Cognitive Neuroscience: Probing the Neuronal Circuitry Underlying Mood and Anxiety Disorders "

This chapter began with discussion of some key components of the circuitry underlying affect and cognition that are most relevant to an understanding of affective and cognitive dysfunction in patients with mood and anxiety disorders. Emphasis was placed on the PFC, amygdala, hippocampus, and ACC. Next, some important conceptual and methodologic problems that plague research in this area were considered. The relevance of the task chosen in activation studies to the underlying symptoms of the disorder should be made explicit in this type of research. Several psychometric problems were then considered, including the issues of matching experimental and control tasks according to level of difficulty and of establishing the reliability of tasks before using them in longitudinal studies of patients in whom changes produced by treatment are being examined. Finally, in studies of emotion, the importance of independent verification of elicitation of the intended emotion was emphasized.

Recent activation studies in patients with mood and anxiety disorders were reviewed. It should be apparent from this review that studies using this strategy are currently lacking despite its obvious importance in revealing the abnormalities in circuitry that underlie basic cognitive and affective processes. It is imperative that the next generation of clinical investigators be trained in the methods and techniques of affective and cognitive neuroscience, the area where such activation paradigms are typically first developed.

It is also imperative that the results of burgeoning research on cognitive and affective information-processing deficits in mood and anxiety disorders (see ref. 94 for review) be used to develop new tasks that can be applied with neuroimaging to probe the circuitry associated with specific types of processing anomalies. For example, an extensive corpus of literature has now documented biases in forms of explicit memory in depression and biases in attention in various types of anxiety disorders. This information can be used to design activation paradigms that are more closely linked to the various hypothesized underlying information-processing deficits. Such research should help to uncover abnormalities in the circuitry underlying the processing of emotion and cognition in patients with mood and anxiety disorders, and should also provide new targets for novel therapeutic approaches.

REFERENCES

1. Miller EK. The prefrontal cortex and cognitive control. *Nat Rev Neurosci* 2000;1:59–65.

2. Leon MI, Shadlin MN. Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque monkey. *Neuron* 1999;24:415–425.

3. Davidson RJ. Affective style and affective disorders: perspectives from affective neuroscience. *Cogn Emotion* 1998;12:307–320.

4. Davidson RJ, Irwin W. The functional neuroanatomy of emotion and affective style. *Trends Cogn Sci* 1999;3:11–21.

5. Gainotti G. Emotional behavior and hemispheric side of lesion. *Cortex* 1972;8:41–55.

6. Robinson RG, Starr LB, Price TR. A two-year longitudinal study of mood disorders following stroke: prevalence and duration at six-months follow-up. *Br J Psychiatry* 1984;144:256–262.

7. Sackeim HA, Weiman AL, Gur RC, et al. Pathological laughter and crying: functional brain asymmetry in the expression of positive and negative emotions. *Arch Neurol* 1982;39:210–218.

8. Watson D, Clark LA, Weber K, et al. Testing a tripartite model: II. Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. *J Abnorm Psychol* 1995;104:15–25.

9. Carson AJ, MacHale S, Allen K, et al. Depression after stroke and lesion location: a systematic review. *Lancet* 2000;356:122–126.

10. Davidson RJ. Cerebral asymmetry and emotion: conceptual and methodological conundrums. *Cogn Emotion* 1993;7:115–138.

11. Davidson RJ. Cerebral asymmetry, emotion and affective style. In: Davidson RJ, Hugdahl K, eds. *Brain asymmetry*. Cambridge, MA: MIT Press, 1995:361–387.

12. Davidson RJ. Anterior electrophysiological asymmetries, emotion and depression: conceptual and methodological conundrums. *Psychophysiology* 1998;35:607–614.

13. Henriques JB, Davidson RJ. Left frontal hypoactivation in depression. *J Abnorm Psychol* 1991;100:535–545.

14. Morris PLP, Robinson RG, de Carvalho ML, et al. Lesion characteristics and depressed mood in the stroke data bank study. *J Neuropsychiatry Clin Neurosci* 1996;8:153–159.

15. Davidson RJ, Ekman P, Saron C, et al. (1990). Approach/withdrawal and cerebral asymmetry: emotional expression and brain physiology. *J Pers Soc Psychol* 1990;58:330–341.

16. Ahern GL, Schwartz GE. Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis. *Neuropsychologia* 1985;23:745–755.

17. Jones NA, Fox NA. Electroencephalogram asymmetry during emotionally evocative films and its relation to positive and negative affectivity. *Brain Cogn* 1992;20:280–299.

18. Davidson RJ. The neuroscience of affective style. In: Gazzaniga MS, ed. *The cognitive neurosciences*, second ed. Cambridge, MA: MIT Press, 2000:1149–1159.

19. Bechara A, Damasio AR, Damasio H, et al. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 1994;50:7–15.

20. Bechara A, Tranel D, Damasio H, et al. Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cereb Cortex* 1996;6:215–225.

21. Bechara A, Damasio H, Tranel D, et al. Deciding advantageously before knowing the advantageous strategy. *Science* 1997;275:1293–1295.

22. Bechara A, Damasio H, Damasio AR, et al. Different contributions of the human amygdala and ventromedial prefrontal cortex to decision making. *J Neurosci* 1999;19:5473–5481.

23. Cacioppo JT, Gardner WL. Emotion. *Annu Rev Psychol* 1999;50:191–214.

24. Taylor SE. Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychol Bull* 1991;110:67–85.

25. Owen AM, Evans AC, Petridies M. Evidence for a two-stage model of spatial working memory processing within the lateral prefrontal cortex: a positron emission tomography study. *Cereb Cortex* 1996;6:31–38.

26. Frijda NH. Emotions are functional, most of the time. In: Ekman P, Davidson RJ, eds. *The nature of emotion: fundamental questions*. Oxford University Press:1994:112–122.

27. Levenson RW. Human emotion: a functional view. In: Ekman P, Davidson RJ, eds. *The nature of emotion: fundamental questions*. Oxford University Press:1994:123–126.

28. Watanabe M. Reward expectancy in primate prefrontal neurons. *Nature* 1996;382:629–632.

29. Rolls ET. *The brain and emotion*. New York: Oxford University Press, 1999.

30. Davidson RJ, Putnam KM, Larson CL. Dysfunction in the neural circuitry of emotion regulation—a possible prelude to violence. *Science* 2000;289:591–594.

31. LeDoux JE. *The emotional brain: the mysterious underpinnings of emotional lift*. New York: Simon & Schuster, 1996.

32. Cahill L, McGaugh JL. Mechanisms of emotional arousal and lasting declarative memory. *Trends Neurosci* 1998;21:273–313.

33. Aggleton JP. The contribution of the amygdala to normal and abnormal emotional states. *Trends Neurosci* 1993;16:328–333.

34. Cahill L, Weinberger NM, Roozendaal B, et al. Is the amygdala a locus of "conditioned fear"? Some questions and caveats. *Neuron* 1999;23:227–228.

35. Fanselow MS, LeDoux JE. Why we think plasticity underlying pavlovian fear conditioning occurs in the basolateral amygdala. *Neuron* 1999;23:229–232.

36. Kalin NH, Shelton SE, Kelley AE, et al. The primate amygdala mediates acute fear but not the behavioral and physiological components of anxious temperament. *J Neurosci* 2001;21:2067-2074.

37. Adolphs R, Damasio H, Tranel D, et al. Fear and the human amygdala. *J Neurosci* 1995;15:5879–5891.

38. Adolphs R, Damasio H, Tranel D, et al. Cortical systems for the recognition of emotion in facial expressions. *J Neurosci* 1996;16:7678–7687.

39. Broks P, Young AW, Maratos EJ, et al. Face processing impairments after encephalitis: amygdala damage and recognition of fear. *Neuropsychologia* 1998;36:59–70.

40. Calder AJ, Young AW, Rowland D, et al. Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. *Cogn Neuropsychol* 1996;13:699–745.

41. Adolphs R, Tranel D, Damasio AR. The human amygdala in social judgment. *Nature* 1998;393:470–474.

42. Scott SK, Young AW, Calder AJ, et al. Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature* 1997;385:254–257.

43. Bechara A, Tranel D, Damasio H, et al. Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* 1995;269:1115–1118.

44. Angrilli A, Mauri A, Palomba D, et al. Startle reflex and emotion modulation impairment after a right amygdala lesion. *Brain* 1996;119:1991–2000.

45. Bush G, Luu P, Posner MI. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci* 2000;4:215–222.

46. Eichenbaum H. A cortical-hippocampal system for declarative memory. *Nat Rev Neurosci* 2000;1:41–50.

47. Zola SM, Squire LR. The medial temporal lobe and the hippocampus. In: Tulving E, Craig FIM, eds. *The Oxford handbook of memory*. New York: Oxford University Press, 2000:485–500.

48. Faneslow MS. Contextual fear, gestalt memories, and the hippocampus. *Behav Brain Res* 2000;110:73–81.

49. McEwen BS. Protective and damaging effects of stress mediators. *N Engl J Med* 1998;338:171–179.

50. Kirschbaum C, Wolf OT, May M, et al. Stress- and treatment-induced elevations of cortisol levels associated with impaired declarative memory in healthy adults. *Life Sci* 1996;58:1475–1483.

51. Bremner JD. The lasting effects of psychological trauma on memory and the hippocampus (*anonymous*, 1999).

52. Sheline, Wang, Gado, et al. Hippocampal atrophy in major depression. *Proc Nat Acad Sci* 1996;93:3908-3913.

53. Bremner JD, Randall P, Scott TM, et al. MRI-based measurement of hippocampal volume in patients with combat-related posttraumatic stress disorder. *Am J Psychiatry* 1995;152:972-981.

54. Vakili K, Pillay SS, Lafer B, et al. Hippocampal volume in primary unipolar major depression: a magnetic resonance imaging study. *Biol Psychiatry* 2000;47:1087–1090.

55. Davidson RJ, Jackson DC, Kalin NH. Emotion, plasticity, context and regulation: perspectives from affective neuroscience. *Psychol Bull* 2000;126:890–909.

56. Davis M, Lee YL. Fear and anxiety: possible roles of the amygdala and bed nucleus of the stria terminalis. *Cogn Emotion* 1998;12:277–306.

57. Whalen P, Bush G, McNally RJ, et al. The emotional counting Stroop paradigm: a functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biol Psychiatry* 1998;44:1219–1228.

58. Carter CS, Botvinick MM, Cohen JD. The contribution of the anterior cingulate cortex to executive processes in cognition. *Rev Neurosci* 1999;10:49–57.

59. Davidson RJ, Slagter HA. Probing emotion in the developing brain: functional neuroimaging in the assessment of the neural substrates of emotion in normal and disordered children and adolescents. *Ment Retard Dev Disabil Res Rev* 2000;6:166–170.

60. Guze BH, Baxter LR, Szuba MP, et al. Positron emission tomography and mood disorders. In: Hauser P, ed. *Brain imaging in affective disorders*. Washington, DC: American Psychiatric Press, 1991:63–87.

61. Braver TS, Barch DM, Cohen JD. Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function. *Biol Psychiatry* 1999;46:312–328.

62. Chapman LJ, Chapman JP. Problems in the measurement of cognitive deficits. *Psych Bull* 1973;79:380-385.

63. Barch DM, Braver TS, Nystrom LE, et al. Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia* 1997;35:1373–1380.

64. Larson CL, Ruffalo D, Nietert J, et al. Temporal stability of the emotion-modulated startle response. *Psychophysiology* 2000;37:92–101.

65. Schaefer SM, Abercrombie HC, Lindgren KA, et al. Six-month test-retest reliability of MRI-defined PET measures of regional cerebral glucose metabolic rate in selected subcortical structures. *Hum Brain Mapping* 2000;10:1–9.

66. Irwin W, Davidson RJ, Lowe MJ, et al. Human amygdala activation detected with echo-planar functional magnetic resonance imaging. *Neuroreport* 1996;7:1765–1769.

67. Furmark T, Fischer H, Wicke JD, et al. The amygdala and individual differences in fear conditioning. *Neuroreport* 1997;8:3957–3960.

68. Davidson RJ, Abercrombie HC, Nitschke JB, et al. Regional brain function, emotion and disorders of emotion. *Curr Opin Neurobiol* 1999;9:228–234.

69. Abercrombie HC, Schaefer SM, Larson CL, et al. Metabolic rate in the right amygdala predicts negative affect in depressed patients. *Neuroreport* 1998;9:3301–3307.

70. Drevets WC, Videen TO, Price JL, et al. A functional anatomical study of unipolar depression. *J Neurosci* 1992;12:3628–3641.

71. Bench CJ, Friston KJ, Brown RG, et al. Regional cerebral blood flow in depression measured by positron emission tomography: the relationship with clinical dimensions. *Psychol Med* 1993;23:579–590.

72. Baker SC, Frith CD, Dolan RJ. The interaction between mood and cognitive function studied with PET. *Psychol Med* 1997;27:565–578.

73. Elliot R, Baker SC, Rogers RD, et al. Prefrontal dysfunction in depressed patients performing a complex planning task: a study using positron emission tomography. *Psychol Med* 1997;27:931–942.

74. Elliot R, Sahakian BJ, Michael A, et al. Abnormal neural response to feedback on planning and guessing tasks in patients with unipolar depression. *Psychol Med* 1998;28:559–571.

75. Merriam EP, Thase ME, Haas GL, et al. Prefrontal cortical dysfunction in depression determined by Wisconsin Card Sorting test performance. *Am J Psychiatry* 1999;156:780–782.

76. Finset A. Depressed mood and reduced emotionality after right-hemisphere brain damage. In: Kinsbourne M, ed. *Cerebral hemisphere function in depression*. Washington, DC: American Psychiatric Press, 1988:49–64.

77. Henriques JB, Davidson RJ. Brain electrical asymmetries during cognitive task performance in depressed and nondepressed subjects. *Biol Psychiatry* 1997;42:1039–1050.

78. Kalin NH, Davidson RJ, Irwin W, et al. Functional magnetic resonance imaging studies of emotional processing in normal and depressed patients: effects of venlafaxine. *J Clin Psychiatry* 1997;58:32–39.

79. Bremner JD, Innis RB, Salomon RM, et al. Positron emission tomography measurement of cerebral metabolic correlates of tryptophan depletion-induced depressive relapse. *Arch Gen Psychiatry* 1997;54:364–374.

80. Henriques JB, Glowacki JM, Davidson RJ. Reward fails to alter response bias in depression. *J Abnorm Psychol* 1994;103:460–466.

81. Henriques JB, Davidson RJ. Decreased responsiveness to reward in depression. *Cogn Emotion* 2000;15:711–724.

82. Skolnick AJ, Davidson RJ. Level of expectation of reward modulates eyeblink startle response. *Psychophysiology* 2000;37:S91.

83. Malizia AL. What do brain imaging studies tell us about anxiety disorders? *J Psychopharmacol* 1999;13:372–378.

84. Rauch SL, van der Kolk BA, Fisler RE, et al. A symptom provocation study of posttraumatic stress disorder using positron emission tomography and script-driven imagery. *Arch Gen Psychiatry* 1996;53:380–387.

85. Shin LM, Kosslyn SM, McNally RJ, et al. Visual imagery and perception in posttraumatic stress disorder. *Arch Gen Psychiatry* 1998;54:233–241.

86. Rauch SL, Whalen PJ, Shin LM, et al. Exaggerated amygdala response to masked facial stimuli in postraumatic stress disorder: a functional MRI study. *Biol Psychiatry* 2000;47:769–776.

87. Rauch SL, Savage CR, Alpert NM, et al. A study of three disorders using positron emission tomography and symptom provocation. *Biol Psychiatry* 1997;42:446–452.

88. Fredrikson M, Wik G, Greitz T, et al. Regional cerebral blood flow during experimental phobic fear. *Psychophysiology* 1993;30:126–130.

89. Fredrikson M, Fischer H, Wik G. Cerebral blood flow during anxiety provocation. *J Clin Psychiatry* 1997;58:16–21.

90. Fredrikson M, Wik G, Annas P, et al. Functional neuroanatomy of visually elicited simple phobic fear: additional data and theoretical analysis. *Psychophysiology* 1995;32:43–48.

91. Birbaumer N, Grodd W, Diedrich O, et al. fMRI reveals amygdala activation to human faces in social phobics. *Neuroreport* 1998;9:1223–1226.

92. Davidson RJ, Irwin W. Functional MRI in the study of emotion. In: Moonen CTW, Bandettini PA, eds. Functional MRI. Berlin: Springer-Verlag, 1999:487–499.

93. Zald DH, Pardo JV. Emotion, olfaction and the human amygdala: amygdala activation during aversive olfactory stimulation. *Proc Natl Acad Sci U S A* 1997;94:4119–4124.

94. Mineka S, Rafaeli-Mor E, Yovel I. Cognitive biases in emotional disorders: information processing and social-cognitive perspectives. In: Davidson RJ, Goldsmith HH, Scherer K, eds. *Handbook of affective science*. New York: Oxford University Press *(in press)*.

29

Interactions Among Neuronal Systems Assessed with Functional Neuroimaging

Christian Büchel

Karl Friston

Christian Büchel: Cognitive Neuroscience Laboratory, Department of Neurology, Hamburg University, Hamburg, Germany.

Karl Friston: Wellcome Department of Cognitive Neurology, London, United Kingdom.

In the late nineteenth century, the early investigations of brain function were dominated by the concept of functional segregation. This approach was driven largely by the data available to scientists of that era. Patients with circumscribed lesions were found who were impaired in one particular ability while their other abilities remained largely intact. Indeed, descriptions of patients with different kinds of aphasia (an impairment of the ability to use or comprehend words), made at this time, have left a permanent legacy in the contrast between Broca's and Wernicke's aphasia. These syndromes were thought to result from damage to anterior or posterior regions of the left hemisphere, respectively. In the first part of the twentieth century, the idea of functional segregation fell into disrepute and the doctrine of “mass action” held sway, according to which higher abilities depended on the function of the brain “as a whole” (1). This doctrine was always going to be unsatisfactory. However, with the resources available at the time, it was simply not possible to make any progress studying the function of the “brain as a whole.” By the end of the twentieth century, the concept of functional segregation had returned to domination.

The doctrine is now particularly associated with cognitive neuropsychology and is enshrined in the concept of double dissociation (2). A double dissociation is demonstrated when neurologic patients can be found with “mirror” abnormalities. For example, many patients have been described who have severe impairments of long-term memory but whose short-term memory is intact. In 1969, Warrington and Shallice (3) described the first of a series of patients who had severe impairments of phonologic short-term memory but no impairments of long-term memory. This is a particularly striking example of double dissociation. It demonstrates that different brain regions are involved in short- and long-term memory. Furthermore, it shows that these regions can function in a largely independent fashion. This observation caused major problems for theories of memory, extant at the time, according to which inputs to long-term memory emanated from short-term memory systems (4).

Functional brain imaging avoids many of the problems of lesion studies, but, here too, the field has been dominated by the doctrine of functional segregation. Nevertheless, it is implicit in the subtraction method that brain regions communicate with each other. If we want to distinguish between brain regions associated with certain central processes, for example, then we design an experiment in which the sensory input and motor output are the same across all conditions. In this way, activity associated with sensory input and motor output will cancel out. The early studies of reading by Petersen et al. (5) and Posner et al. (6) are still among the best examples of this approach. The design of these studies was based on the assumption that reading goes through a single series of discrete and independent stages; visual shapes are analyzed to form letters, letters are put together to form words, the visual word form is translated into sound, the sound form is translated into articulation, and so on. By a comparison of suitable tasks (e.g., letters vs. false font, words vs. letters), each stage can be isolated and the associated brain region identified. Although subsequent studies have shown that this characterisation of the brain activity associated with reading is a considerable oversimplification, the original report still captures the essence of most functional imaging studies; a number of discrete cognitive stages are mapped onto discrete brain areas. Nothing is revealed about how the cognitive processes interact or how the brain regions communicate with each other. If word recognition really did depend on the passage of information through a single series of discrete stages, we

would at least like to know the temporal order in which the associated brain regions are engaged. Some evidence comes from encephaloelectrographic and myoelectrographic studies. In fact, we know that word recognition depends on at least two parallel routes—one of meaning and one of phonology (7). Given this model, we would like to be able to specify the brain regions associated with each route and have some measure of the strengths of the connections between these different regions.

In this chapter, we show how new methods for measuring effective connectivity allow us to characterize the interactions between brain regions that underlie the complex interactions among different processing stages of functional architectures.

DEFINITIONS

Part of "29 - Interactions Among Neuronal Systems Assessed with Functional Neuroimaging "

In the analysis of neuroimaging time series (i.e., signal changes in a set of voxels, expressed as a function of time), functional connectivity is defined as the *temporal correlations between spatially remote neurophysiological events* (8). This definition provides a simple characterization of functional interactions. The alternative is effective connectivity, *the influence one neuronal system exerts over another* (9). These concepts originated in the analysis of separable spike trains obtained from multiunit electrode recordings (10 ,11). Functional connectivity is simply a statement about the observed correlations; it does not comment on how these correlations are mediated. For example, at the level of multiunit microelectrode recordings, correlations can result from *stimulus-locked transients*, evoked by a common afferent input* , or reflect *stimulus-induced oscillations*, phasic coupling of neural assemblies mediated by synaptic connections (12). Effective connectivity is closer to the notion of a connection, either at a synaptic (cf synaptic efficacy) or cortical level. Although functional and effective connectivity can be invoked at a conceptual level in both neuroimaging and electrophysiology, they differ fundamentally at a practical level. This is because the time scales and nature of neurophysiologic measurements are very different (seconds vs. milliseconds and hemodynamic vs. spike trains). In electrophysiology, it is often necessary to remove the confounding effects of stimulus-locked transients (that introduce correlations *not* causally mediated by direct neural interactions) to reveal an underlying connectivity. The confounding effect of stimulus-evoked transients is less problematic in neuroimaging because propagation of signals from primary sensory areas onward is mediated by neuronal connections (usually reciprocal and interconnecting). However, it should be remembered that functional connectivity is not necessarily a consequence of effective connectivity (e.g., common neuromodulatory input from ascending aminergic neurotransmitter systems or thalamocortical afferents), and when it is, effective influences may be indirect (e.g., polysynaptic relays through multiple areas). In this chapter, we focus only on effective connectivity. More details about functional connectivity can be found in Friston et al. (8).

EFFECTIVE CONNECTIVITY

Part of "29 - Interactions Among Neuronal Systems Assessed with Functional Neuroimaging "

A Simple Model

Effective connectivity depends on two models: a mathematical model, describing “how” areas are connected, and a neuroanatomic model, describing “which” areas are connected. We shall consider linear and nonlinear models. Perhaps the simplest model of effective connectivity expresses the hemodynamic change at one voxel as a weighted sum of changes elsewhere. This can be regarded as a multiple linear regression, in which the effective connectivity reflects the amount of rCBF (regional cerebral blood flow) variability, at the target region, attributable to rCBF changes at a source region. As an example, consider the influence of other areas *M* on area *V1*. This can be framed in a simple equation:

$$V1 = M c + e \qquad [1]$$

where *V1* is an $n \times 1$ column vector with n scans, M is an $n \times m$ matrix with m regions and n observations (scans), c is an $m \times 1$ column vector with a parameter estimate for each region, and e is a vector of error terms.

Implicit in this interpretation is a mediation of the influence among brain regions by neuronal connections with an effective strength equal to the (regression) coefficient c. This highlights the fact that the linear model assumes that the connectivity is constant over the whole range of activation and does not depend on input from other sources.

Experience suggests that the linear model can give fairly robust results. One explanation is that the dimensionality (the number of things that are going on) of the physiologic changes can be small by experimental design. In other words, the brain responds to simple and well-organized experiments in a simple and well-organized way. Generally, however, neurophysiologic interactions are nonlinear, and the adequacy of linear models must be questioned (or at least qualified). Consequently, we focus on a nonlinear model of effective connectivity (13).

Structural Equation Modeling

The simple model above is sufficient to analyze effective connectivity to one region at a time (e.g., *V1* or *V2*). We will now introduce structural equation modeling as a tool allowing for more complicated models comprising many

regions of interest and demonstrate how nonlinear interactions are dealt with in this context. The basic idea behind structural equation modeling differs from the usual statistical approach of modeling individual observations. In multiple regression or ANCOVA (analysis of covariance) models, the regression coefficients derive from the minimization of the sum of squared differences of the predicted and observed dependent variables (i.e., activity in the target region). Structural equation modeling approaches the data from a different perspective; instead of variables being considered individually, the emphasis lies on the variance–covariance structure.* Thus, models are solved in structural equation modeling by minimizing the difference between the observed variance–covariance structure and the one implied by a structural or path model. In the past few years, structural equation modeling has been applied to functional brain imaging. For example, McIntosh et al. (14) demonstrated the dissociation between ventral and dorsal visual pathways for object and spatial vision by using structural equation modeling of positron emission tomographic (PET) data in the human. In this section, we focus on the theoretic background of structural equation modeling and demonstrate this technique with the use of functional magnetic resonance imaging (fMRI).

In terms of neuronal systems, a measure of covariance represents the degree to which the activities of two or more regions are related (i.e., functional connectivity). The study of variance–covariance structures here is much simpler than in many other fields; the interconnection of the dependent variables (regional activity of brain areas) is anatomically determined, and the activation of each region can be directly measured with functional brain imaging. This represents a major difference from “classic” structural equation modeling in the behavioral sciences, in which models are often hypothetical and include latent variables denoting rather abstract concepts, such as intelligence.

As mentioned above, structural equation modeling minimizes the difference between the observed or measured covariance matrix and the one that is implied by the structure of the model. The free parameters (path coefficients or connection strengths; *c* above) are adjusted to minimize the difference between the measured and modeled covariance matrix.† (See ref. 15 for details.)

An important issue in structural equation modeling is the determination of the participating regions and the underlying anatomic model. Several approaches to this issue can be adopted. These include categoric comparisons between different conditions, statistical images highlighting structures of functional connectivity, and nonhuman electrophysiologic and anatomic studies (16).

With respect to anatomic connectivity in humans, the advent of new MR techniques promises a better characterization of neuronal connectivity in humans. Diffusion tensor imaging measures the anisotropy of diffusion in the brain. The main anisotropy exists in the white matter because the orientation of neuronal fibres (axons) allows molecules to diffuse more easily along the fiber than in other directions. Therefore, the main direction of the diffusion tensor reflects the underlying orientation of white matter tracts. Through tracing algorithms, it is now possible to infer the connectivity of individual regions (e.g., activations derived from an fMRI study) in an individual brain (17) (Fig. 29.1).

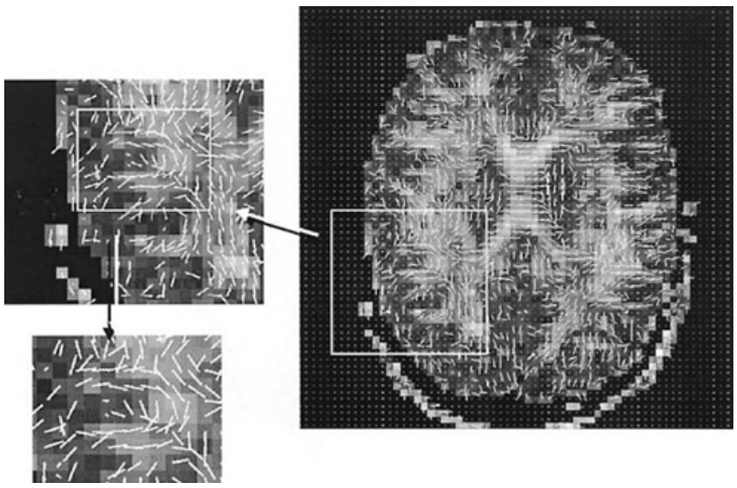


FIGURE 29.1. Axial diffusion tensor image, obtained by using a TurboSTEAM diffusion sensitized pulse sequence on a Siemens Vision 1.5T MR scanner. Voxel size 3 × 3 × 3 mm. Average of 20 replications. *Needles* in each voxel show the largest eigenvector of the tensor (i.e., the main orientation of diffusion within this voxel). In white matter, the major axis of diffusion is constrained by the orientation of white matter tracts and therefore provides a good estimate of the direction of fiber bundles (17). As expected, the corpus callosum in the center of the image shows predominantly horizontal fibers connecting both hemispheres. In the occipital cortex, parts of the optic radiation with a predominantly anterior–posterior fiber orientation can be seen. The precision of the method is highlighted by the demonstration of corticocortical U fibers, magnified in the small image. (From Nolte U, Finsterbusch J, Frahm J. Rapid whole brain diffusion mapping without susceptibility artifacts using diffusion-weighted single-shot STEAM MRI. *Proceedings of the eighth annual meeting of the International Society of Magnetic Resonance in Medicine*, Denver, 2000:807.)

A model is always a simplification of reality; exhaustively correct models either do not exist or are too complicated to understand. In the context of effective connectivity, one has to find a compromise between complexity, anatomic accuracy, and interpretability. Mathematical constraints on the model also exist; if the number of free parameters exceeds the number of observed covariances, the system is underdetermined and no single solution exists.

Each estimated model can be analyzed to give an overall

goodness-of-fit measure for use when different models are compared with each other. A "nested model" approach can be used to compare different models (e.g., data from different groups or conditions) in the context of structural equation modeling. A so-called null model is constructed in which the estimates of the free parameters are constrained to be the same for both groups. The alternative model allows free parameters to differ between groups. The significance of the differences between the models is expressed by the difference of the goodness-of-fit statistic. Consider the following hypothetical example. Subjects are scanned under two different conditions (e.g., *attention* and *no attention*). The hypothesis might be that within a system of regions *A, B, C*, and *D*, the connectivity between *A* and *B* is different under the two attentional conditions. To determine whether the difference in connectivity is statistically significant, we estimate the goodness-of-fit measure for two models. Model 1 allows the connectivity between *A* and *B* to take different values for both conditions. Model 2 constrains the path coefficient between *A* and *B* to be equal for *attention* and *no attention*. If the change of connectivity between *attention* and *no attention* for the connection of *A* and *B* is negligible, the constrained model (model 2) should fit the data as well as the free model (model 1). We can now infer whether the difference of the two goodness-of-fit measures is significant. Nonlinear models can also be accommodated in the framework of structural equation modeling by introducing additional variables containing a nonlinear function (e.g., $f(x) = x^2$) of the original variables (18). Interactions of variables can be incorporated in a similar fashion, wherein a new variable, containing the product of the two interacting variables, is introduced as an additional influence. We will now demonstrate these ideas with an example. More details of structural equation modeling, including the operational equations, can be found in ref. 15 .

Example: Learning

In the first example, we were interested in changes in effective connectivity over time as expected during paired-associates learning (19). In the case of object-location memory, several functional studies have demonstrated activation of ventral occipital and temporal regions during the retrieval of object identity and, conversely, increased responses in dorsal parietal areas during the retrieval of spatial location (20). These results suggest domain-specific representations in posterior neocortical structures that are closely related to those involved in perception, a finding that accords with the segregation of ventral and dorsal pathways in processing categoric or spatial stimulus features, respectively. Another phenomenon observed in some learning studies is a decrease of neural responses (i.e., adaptation) to repeated stimulus presentations. This repetition suppression has been replicated consistently in primate electrophysiologic and human functional imaging studies (21). For object-location learning, it is intuitively likely that two specialized systems need to interact to establish an association. Domain-specific representations or repetition suppression is not sufficient to account for this associative component. In other words, functional segregation and localized response properties cannot account for associative learning alone.

In our fMRI experiment, decreases in activation during learning, indicative of repetition suppression, were observed in several cortical regions in the ventral and dorsal visual pathway. Within the framework of repetition suppression, it has been hypothesized that decreases in neural responses are a secondary result of enhanced response selectivity (22). By analogy to the development and plasticity of cortical architectures, this refined selectivity is likely to be a consequence of changes in effective connectivity within the system at a synaptic level. We explicitly addressed this notion by characterising time-dependent changes in effective connectivity during learning.

The experiment was performed on a 2-tesla (T) MRI system equipped with a head volume coil. fMRI images were obtained every 4.1 seconds with echo-planar imaging (48 slices in each volume). Six subjects had to learn and recall the association between 10 simple line drawings of real-world objects and 10 locations on a screen during fMRI. Each learning trial consisted of four conditions: *encoding, control, retrieval*, and *control* (Fig. 29.2A). The behavioral data acquired during *retrieval* demonstrated that all six subjects were able to learn the association between object identity and spatial location, for all 10 objects, within eight learning blocks, as indicated by the ensuing asymptotic learning curves (Fig. 29.2B).

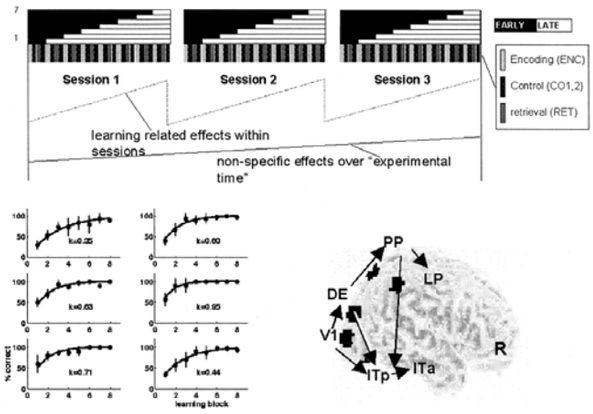


FIGURE 29.2. Changes in effective connectivity over time in paired-associates learning. **A:** Design of the study. Blocks of *encoding* and *retrieval* were alternated with control conditions. Subjects had to complete three individual learning sessions to avoid the confounding effect of time. **B:** Behavioral performance data for each of the six subjects averaged across all three learning sessions. **C:** Anatomic model. Processing of object identity is mainly a property of the ventral visual pathway, whereas object location is a property of the dorsal stream. We focused on the interstream connections (mainly posterior parietal cortex to posterior inferotemporal cortex) based on the hypothesis that learning the association of object identity and spatial location leads to an increase in effective connectivity between the ventral and dorsal streams. (From Büchel C, Coull JT, Friston KJ. The predictive value of changes in effective connectivity for human learning. *Science* 1999;283:1538–1541, with permission.)

The structural model used in the analysis embodies connections within and across ventral and dorsal visual pathways and was based on anatomic studies in primates (Fig. 29.2C). Primary visual cortex was modeled as the origin of both pathways. In addition to "interstream" connections between dorsal extrastriate cortex and the fusiform region and between the posterior parietal cortex and the posterior inferotemporal cortex, we included direct connections based on a hierarchic cortical organization. Given our hypothesis relating to changes in effective connectivity between dorsal and ventral pathways, the path analysis focused on the connection between posterior parietal cortex (PP, dorsal stream) and posterior inferotemporal cortex (ITp, ventral stream). We divided each learning session into EARLY (first part) and LATE (second part) observations and estimated separate path coefficients for each partition.

The path coefficient between PP and ITp increased significantly during learning in the group ($p < .05$) and was confirmed by an analysis of individual subjects showing an increase in effective connectivity between PP and ITp of 0.27. In contrast to the connections between streams, connections within the dorsal pathway decreased over time.

The estimated change in connectivity from PP to ITp

clearly depended on the cutoff point between EARLY and LATE. To establish unequivocally a relationship between neurophysiologically mediated changes in connectivity and behavioral learning, we examined the relationship between the temporal pattern of effective connectivity changes and learning speed for all sessions and subjects. We estimated the differences in effective connectivity for seven EARLY and LATE partitions by successively shifting the cutoff. The cutoff time at which the connectivity change peaked was used as a temporal index of changes in effective connectivity (i.e., plasticity). The significant regression of *k*, a measure of learning speed* , on this plasticity index indicated that for sessions showing fast learning (i.e., high value of *k*), the maximum difference in path coefficients between PP and ITp was achieved earlier in the session (i.e., EARLY comprises fewer scans relative to LATE) (Fig. 29.3). In other words, the temporal pattern of changes in effective connectivity strongly predicted learning or acquisition.

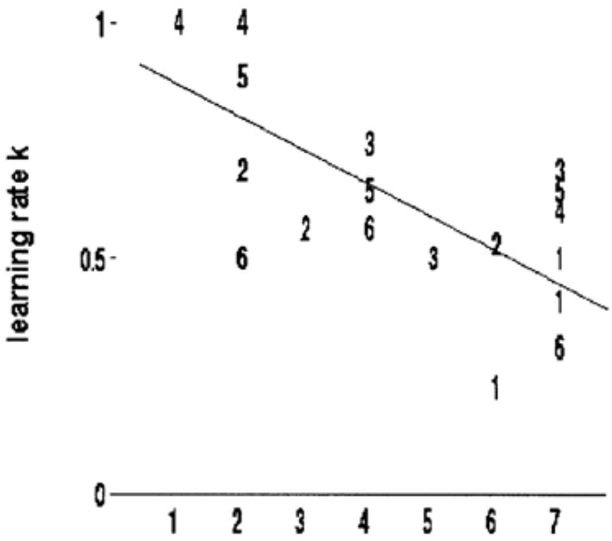


FIGURE 29.3. Changes in effective connectivity predict learning. This graph shows the correlation between the temporal index of changes in effective connectivity and learning. The temporal index is defined as the time of a maximum increase in effective connectivity between posterior parietal cortex and posterior inferotemporal cortex. For example, a temporal index of 3 indicates that the maximum increase in effective connectivity occurred between the third and fourth blocks. The *numbers* denote the subject from which this temporal index of effective connectivity was obtained. Each subject was scanned during three independent learning sessions; therefore, each number appears three times. A negative slope means that the maximum increase in effective connectivity occurs earlier in fast learning. (From Büchel C, Coull JT, Friston KJ. The predictive value of changes in effective connectivity for human learning. *Science* 1999;283:1538–1541, with permission.)

Example: Attention

Electrophysiologic and neuroimaging studies have shown that attention to visual motion can increase the responsiveness of the motion-selective cortical area (*V5*) (23 ,24) and the PP (25). Increased or decreased activation in a cortical area is often attributed to attentional modulation of the cortical projections to that area. This leads to the notion that attention is associated with changes in connectivity.

Here we present fMRI data from an individual subject, scanned under identical visual motion stimulus conditions while only the attentional component of the tasks employed was changed. First, we identify regions that show differential activations in relation to attentional set. In the second stage,

changes in effective connectivity to these areas are assessed with structural equation modeling. Finally, we show how these attention-dependent changes in effective connectivity can be explained by the modulatory influence of parietal areas by using a nonlinear extension of structural equation modeling. The specific hypothesis we addressed was that parietal cortex could modulate the inputs from *V1* to *V5*.

The experiment was performed on a 2-T MRI system equipped with a head volume coil. fMRI images were obtained every 3.2 seconds with echo-planar imaging (32 slices in each volume). The subject was scanned during four different conditions: *fixation, attention, no attention*, and *stationary*. Each condition lasted 32 seconds to give 10 volumes per condition. We acquired a total of 360 images. During all conditions, the subjects looked at a fixation point in the middle of a screen. In this section, we are interested only in the two conditions with visual motion (*attention* and *no attention*), in which 250 small white dots moved radially from the fixation point, in random directions, toward the border of the screen at a constant speed of 4.7 degrees per second. The difference between *attention* and *no attention* lay in the explicit command given to the subject shortly before the condition: *just look* indicated *no attention*, and *detect changes* indicated the *attention* condition. Both visual motion conditions were interleaved with *fixation*. No response was required.

Regions of interest were defined by categoric comparisons with use of an output statistical image (*SPM{Z}*) comparing *attention* with *no attention* and comparing *no attention* with *fixation*. As predicted, given a stimulus consisting of radially moving dots, we found activation of the lateral geniculate nucleus, primary visual cortex (*V1*), motion-sensitive area (*V5*), and posterior parietal complex. For the subsequent analysis of effective connectivity, we defined regions of interest with a diameter of 8 mm centered around the most significant voxel as revealed by the categoric comparison. A single time series, representative of this region, was defined by the first eigenvector of all the voxels in the region of interest (15).

Our model of the dorsal visual stream included the lateral geniculate nucleus, *V1, V5*, and the PP. Although connections between regions are generally reciprocal, for simplicity we modeled only unidirectional paths.

To assess effective connectivity in a condition-specific fashion, we used time series that comprised observations during the condition in question. Path coefficients for both conditions (*attention* and *no attention*) were estimated by using a maximum likelihood function. To test for the impact of changes in effective connectivity between *attention* and *no attention*, we defined a free model (allowing different path coefficients between *V1* and *V5* for *attention* and *no attention*) and a constrained model (constraining the *V1* → *V5* coefficients to be equal). Figure 29.4 shows the free-model and estimated path coefficients. The connectivity between *V1* and *V5* increases significantly during attention. Note also a significant difference in connectivity between *V5* and the PP.

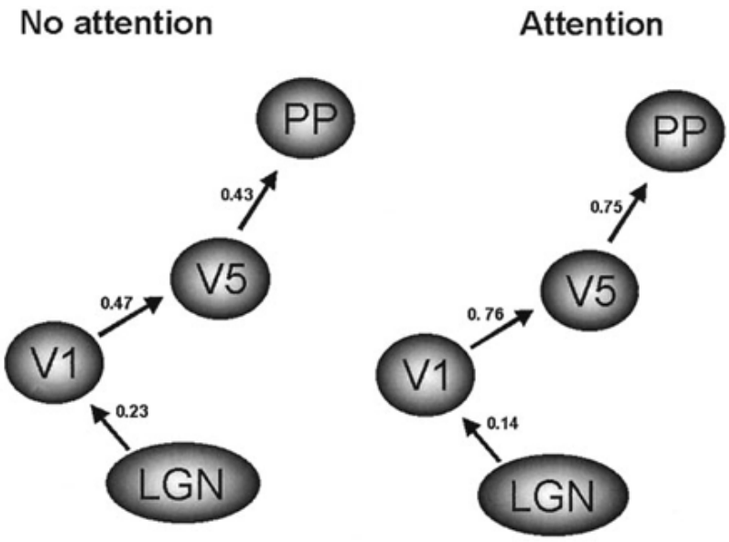


FIGURE 29.4. Structural equation model of the dorsal visual pathway, comparing *attention* and *no attention*. Connectivity between right primary visual cortex (*V1*) and motion-sensitive area (*V5*) is increased during *attention* relative to *no attention*. This is also shown for the connection between *V5* and the posterior parietal cortex. (From Büchel C, Friston KJ. Effective connectivity in functional brain imaging. *Neural Networks* 2000;13:871–882, with permission.)

The linear path model comparing *attention* and *no attention*

revealed increased effective connectivity in the dorsal visual pathway in relation to attention. The question that arises is, which part of the brain is capable of modulating this pathway? Based on lesion studies (26) and the system for directed attention described in ref. 27 , the PP is hypothesized to play such a modulatory role.

We extended our model accordingly to allow for nonlinear interactions, testing the hypothesis that the PP acts as a moderator of the connectivity between *V1* and *V5*. Assuming a nonlinear modulation of this connection, we constructed a new variable, *V1PP*, in our analysis. This variable, mediating the interaction, is simply the time series from region *V1* multiplied (element by element) by the time series of the right posterior parietal region.

The influence of this new variable on *V5* corresponds to the influence of the PP cortex on the connection between *V1* and *V5* (i.e., the influence of *V1* on *V5* is greater when activity in the PP is high). The model is shown in Fig. 29.5 . Because our nonlinear model could accommodate changes in connectivity between *attention* and *no attention*, the entire time series was analyzed (i.e., attention-specific changes are now explicitly modeled by the interaction term).

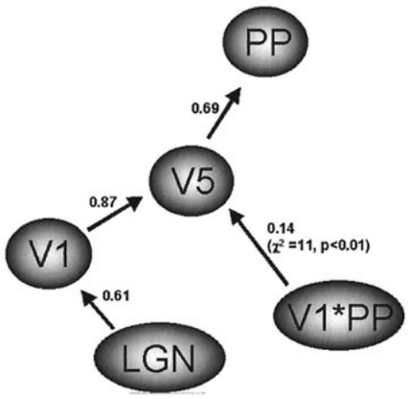


FIGURE 29.5. Structural equation model of the dorsal visual pathway incorporating the interaction effect of right posterior parietal cortex on the connection from right primary visual cortex (*V1*) to motion-sensitive area (*V5*). (From Büchel C, Friston KJ. Effective connectivity in functional brain imaging. *Neural Networks* 2000;13:871–882, with permission.)

As in the linear model, we tested for the significance of the interaction effect by comparing a restricted and a free model. In the restricted model, the interaction term (i.e., path from *V1PP* to *V5*) was set to zero. Omitting the interaction term led to a significantly reduced model fit ($p < .01$), which indicated the predictive value of the interaction term.

The presence of an interaction effect of the PP on the connection between *V1* and *V5* can also be illustrated by a simple regression analysis. If the PP shows a positive modulatory influence on the path between *V1* and *V5*, the influence of *V1* on *V5* should depend on the activity of the PP. This can be tested by splitting the observations into two sets, one containing observations in which the PP activity is high and another one in which the PP activity is low. It is now possible to perform separate regressions of *V5* on *V1* by using both sets. If the hypothesis of positive modulation is true, the slope of the regression of *V5* on *V1* should be steeper under high values of PP.

Variable Parameter Regression

As demonstrated in the previous sections, the basic linear model can be seen as a linear regression. The regression coefficient is then interpreted as a measure of the connectivity between areas. This interpretation of course implies that the influence is mediated by neural connections with an effective strength equal to the regression coefficient. Using this approach, one immediately makes the assumption that the effective connectivity does not change over observations because only a single regression coefficient for the whole time series is estimated. This is unsuitable for the assessment of effective connectivity in functional imaging because the goal in some experiments is to demonstrate changes in effective connectivity—for instance, as a function of different conditions (e.g., *attention* and *no attention*) or simply time itself. In the framework of regression analysis, there are three ways around this problem. Firstly, one can split the data in different groups according to the experimental condition (e.g., *attention* and *no attention*) and then test for the difference of the regression coefficients. However, we may not know *a priori* the time course of the changes that allow us to split the data in this way. A second, more general solution is to expand the explanatory variable in terms of a set of basis functions to account for changes in connectivity. Here, we present another alternative, variable parameter regression, that allows one to characterize the variation of the regression coefficient by using the framework of state–space models and the Kalman filter (28 ,29).

Mathematical Background

Consider the classic regression model

$$y = x\beta + u \qquad [2]$$

where *y* is the measured data vector, *x* is a vector of explanatory variables, and ß is the unknown parameter. Usually, ß is estimated as

$$\hat{\beta} = \text{pinv}(x)y \qquad [3]$$

However, ß can also be estimated recursively with the advantage that inversion of a smaller matrix is necessary. This approach is known as *recursive least squares* (30). This basic model is now extended to allow ß to evolve over time.

Variable parameter regression assumes *T*-ordered scalar observations (y_1,… y_T) generated by the following model:

$$y_t = x_t\beta_t + u_t,\ t = 1, \ldots, T, \qquad [4]$$

$$u_t \sim N(0, \sigma^2) \qquad [5]$$

where x_t is an *n*-dimensional row vector of known regressors and β_t is an *n*-dimensional column vector of unknown coefficients that corresponds to estimates of effective connectivity. u_t is drawn from a gaussian distribution. All observations are expressed as deviations from the mean.

A recursive algorithm known as the Kalman filter (29) can now be applied to estimate the state variable (β) at each point in time and also allows one to estimate the log-likelihood function of the model. A numeric optimization algorithm is then employed to maximize the likelihood function with respect to *P*. As the Kalman filter is a recursive procedure, the estimation of β_t is based on all observations up to time *t*. Therefore, the filtered estimates will be more accurate toward the end of the sample. This fact is corrected for with the Kalman smoothing algorithm, which is used *post hoc* and runs backward in time, taking account of the information made available after time *t*. Details of the Kalman filter and smoothing recursions can be found in standard textbooks of time series analysis and econometrics (31 ,32).

Example: Attention to Visual Motion

To illustrate variable parameter regression, we use the single-subject data set from the study of attention to visual motion. We concentrate on the effect of attention on the connection between the motion-sensitive area (*V5*) and the PP in the right hemisphere. Using structural equation modeling, we demonstrated that it is principally this connection, in the dorsal visual stream, that is modulated by attention (15). In the current analysis, we are interested in whether variable parameter regression is capable of reproducing these findings. We therefore have assessed the effective connectivity β_t by regressing PP on *V5*. An alternate direction search, numeric optimization, gave a x^2 statistic of 56.4. We therefore had to reject the null hypothesis of no variation at the 5% level. *P* was estimated to be 0.074 and σ^2 was 0.23. The ordinary regression coefficient β for the model $y = x\beta + u$ was estimated at 0.73. Figure 29.6 A,B shows the trajectories of the smoothed and filtered estimates β_t (T) together with the associated standard errors. It is clearly evident that β_t is higher during the *attention* conditions than during the *no attention* conditions. Figure 29.6D relates our technique to an ordinary regression. In this analysis, we constrained the variance term *P* to zero and reestimated β_t. The trajectory of β_t now converges to β, the ordinary regression coefficient of the model $y = x\beta + u$. As expected, the smoothed estimates are simply a constant (i.e., β = 0.73).

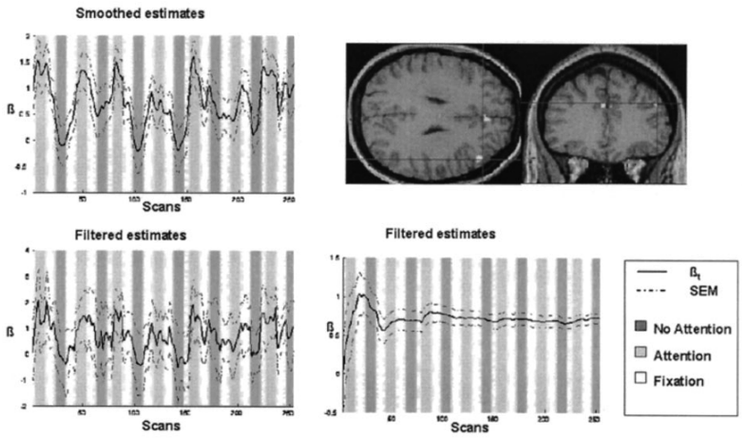


FIGURE 29.6. A,B: The trajectory of the smoothed and filtered estimates β_t (T) together with the associated standard errors for the variable parameter estimation of effective connectivity between motion-sensitive area (*V5*) and posterior parietal cortex (PP). It is evident that β_t (the dynamic regression coefficient) is higher during the *attention* conditions than during the *no attention* conditions. **C:** Areas that significantly covaried with the time-dependent measure of effective connectivity between *V5* and the PP [i.e., β_t (T)]. The output statistical image *SPM{Z{* thresholded at $p < .001$ (uncorrected) overlaid on coronal and axial slices of the subject's structural MRI. The maximum under the cross-hairs was at 45, 21, 39 mm, $Z = 4$. **D:** The relationship between our technique and an ordinary regression analysis. In this analysis, the variance term *P* was set to zero (i.e., fixed regression model). The trajectory of β_t now converges to β (= 0.73), the regression coefficient of the model $y = x\beta + u$. (From Büchel C, Friston KJ. Dynamic changes in effective connectivity characterized by variable parameter regression and Kalman filtering. *Hum Brain Mapping* 1998;6:403–408, with permission.)

We interpret β_t as an index of effective connectivity between area *V5* and the PP. In our example, the connection between *V5* and the PP resembles the *site* of attention modulation. This leads to an interesting extension, in which one might hypothesize that a third region is responsible for the observed variation in effective connectivity indicated by the trajectory of β_t (T). In other words, after specifying the *site* and nature of attentional modulation, we now want to know the location of the *source*. We addressed this by using β_t (T) as an explanatory variable in an ordinary regression analysis to identify voxels that covaried with this measure of effective connectivity. Figure 29.6C shows the result of this analysis. Among areas with statistically significant ($p < .001$, uncorrected) positive covariation were the dorsolateral prefrontal cortex and the anterior cingulate cortex. This result confirms the putative modulatory role of the dorsolateral prefrontal cortex in attention to visual motion, as suggested by previous analyses (15).

Effective Connectivity versus Categorical Comparisons

One obvious advantage of the assessment of effective connectivity is that it allows one to test hypotheses about the integration of cortical areas. For example, in the presence of modulation, the categoric comparison between *attention* and *no attention* might reveal prestriate, parietal, and frontal activations. However, the only statement possible is that these areas show higher cortical activity during the *attention* condition as opposed to the *no attention* condition. The analysis of effective connectivity revealed two additional results. Firstly, attention affects the pathway from *V1* to *V5* and from *V5* to PP. Secondly, the introduction of nonlinear interaction terms allowed us to test a hypothesis about how these modulations are mediated. The latter analysis suggested that the PP exerts a modulatory influence on area *V5*.

The measurements used in all examples in this chapter were *hemodynamic* in nature. This limits an interpretation at the level of *neuronal* interactions. However, the analogy between the form of the nonlinear interactions described above and voltage-dependent (i.e., modulatory) connections is a strong one. It is possible that the modulatory impact of PP on *V5* is mediated by predominantly voltage-dependent connections. We know of no direct electrophysiologic evidence to suggest that extrinsic backward PP to *V5* connections are voltage-dependent; however, our results are consistent with this. An alternative explanation for modulatory effects, which does not necessarily involve voltage-dependent connections, can be found in the work of Aertsen and Preissl (10). These authors show that effective connectivity varies strongly with, or is modulated by, background neuronal activity. The mechanism relates to the efficacy of subthreshold excitatory postsynaptic potentials in establishing dynamic interactions. This efficacy is a function of postsynaptic

depolarization, which in turn depends on the tonic background of activity.

CONCLUSIONS

Part of "29 - Interactions Among Neuronal Systems Assessed with Functional Neuroimaging "

This chapter has reviewed the basic concepts of effective connectivity in neuroimaging. We have introduced several methods to assess effective connectivity—multiple linear regression, covariance structural equation modeling, and variable parameter regression. In the first example, structural equation modeling was introduced as a device that allows one to combine observed changes in cortical activity and anatomic models. An application of this technique revealed changes in effective connectivity between the dorsal and the ventral stream over time in a paired-associates learning paradigm. The temporal pattern of these changes was highly correlated with individual learning performance, and therefore changes in effective connectivity predicted learning speed. The second example of structural equation modeling focused on backward modulatory influences of high-order areas on connections among lower-order areas. Both examples concentrated on changes in effective connectivity and allowed us to characterize the interacting areas of the network at a functional level. Variable parameter regression was then introduced as a flexible regression technique that allows the regression coefficient to vary smoothly over time. Again, we confirmed the backward modulatory effect of higher cortical areas on those areas situated lower in the cortical hierarchy. Although this field is less than mature, the approach to neuroimaging data and regional interactions discussed above is an exciting endeavor that is starting to attract more and more attention.

REFERENCES

1. Lashley KS. *Brain mechanisms and intelligence*. Chicago: University of Chicago Press, 1929.

2. Shallice T. *From neuropsychology to mental structure*. Cambridge: Cambridge University Press, 1988.

3. Warrington EK, Shallice T. The selective impairment of auditory short-term memory. *Brain* 1969;92:885–896.

4. Atkinson RC, Shiffrin RM. Human memory: a proposed system and its control processes. In: Spence KW, Spence JT, eds. *The psychology of learning and motivation: advances in research and theory* (vol 2). New York: Academic Press, 1968.

5. Petersen SE, Fox PT, Snyder AZ, et al. Activation of extrastriate and frontal cortical areas by words and word-like stimuli. *Science* 1990;249:1041–1044.

6. Posner MI, Petersen SE, Fox PT, et al. Localization of cognitive operations in the human brain. *Science* 1988;240:1627–1631.

7. Marshall JC, Newcombe F. Patterns of paralexia: a neurolinguistic approach. *J Psycholinguist* Res 1973;2:175–199.

8. Friston KJ, Frith CD, Liddle PF, et al. Functional connectivity: the principal component analysis of large (PET) data sets. *J Cereb Blood Flow Metab* 1993;13:5–14.

9. Friston KJ, Frith CD, Frackowiak RSJ. Time-dependent changes in effective connectivity measured with PET. *Hum Brain Mapping* 1993;1:69–80.

10. Aertsen A, Preissl H. Dynamics of activity and connectivity in physiological neuronal networks. New York: VCH Publishers, 1991.

11. Gerstein GL, Perkel DH. Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science* 1969;164:828–830.

12. Gerstein GL, Bedenbaugh P, Aertsen A. Neuronal assemblies. *IEEE Trans Biomed Eng* 1989;36:4–14.

13. Friston KJ, Ungerleider LG, Jezzard P, et al. Characterizing modulatory interactions between V1 and V2 in human cortex with fMRI. *Hum Brain Mapping* 1995;2:211–224.

14. McIntosh AR, Grady CL, Ungerleider LG, et al. Network analysis of cortical visual pathways mapped with PET. *J Neurosci* 1994;14:655–666.

15. Büchel C, Friston KJ. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modeling and fMRI. *Cereb Cortex* 1997;7:768–778.

16. McIntosh AR, Gonzalez-Lima F. Structural equation modeling and its application to network analysis in functional brain imaging. *Hum Brain Mapping* 1994;2:2–22.

17. Moseley ME, Cohen Y, Kucharczyk J, et al. Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system. *Radiology* 1990;176:439–445.

18. Kenny DA, Judd CM. Estimating nonlinear and interactive effects of latent variables. *Psychol Bull* 1984;96:201–210.

19. Büchel C, Coull JT, Friston KJ. The predictive value of changes in effective connectivity for human learning. *Science* 1999;283:1538–1541.

20. Milner B, Johnsrude I, Crane J. Right medial temporal-lobe contribution to object-location memory. *Philos Trans R Soc Lond B Biol Sci* 1997;352:1469–1474.

21. Desimone R. Neural mechanisms for visual memory and their role in attention. *Proc Natl Acad Sci USA* 1996;93:13494–13499.

22. Wiggs CL, Martin A. Properties and mechanisms of perceptual priming. *Curr Opin Neurobiol* 1998;8:227–233.

23. O'Craven KM, Savoy RL. Voluntary attention can modulate fMRI activity in human MT/MST. *Invest Ophthalmol Vis Sci* 1995; 36[Suppl]:S856.

24. Treue S, Maunsell HR. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* 1996;382:539–541.

25. Assad JA, Maunsell JH. Neuronal correlates of inferred motion in primate posterior parietal cortex. *Nature* 1995;373:518–521.

26. Lawler KA, Cowey A. On the role of posterior parietal and prefrontal cortex in visuo-spatial perception and attention. *Exp Brain Res* 1987;65:695–698.

27. Mesulam MM. Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Ann Neurol* 1990;28:597–613.

28. Büchel C, Friston KJ. Dynamic changes in effective connectivity characterized by variable parameter regression and Kalman filtering. *Hum Brain Mapping* 1998;6:403–408.

29. Kalman RE. A new approach to linear filtering and prediction problems. *Trans Am Soc Med Eng J Basic Eng* 1960;D82:35–45.

30. Harvey AC. *Time series models*. London: Harvester & Wheatsheaf, 1993.

31. Chow GC. *Econometrics*. New York: McGraw-Hill, 1983.

32. Harvey AC. *Forecasting, structural time series models and the Kalman filter*. Cambridge: Cambridge University Press, 1990.

*That is, signal input into the neural system as a result of external stimulation.

*The variance–covariance structure describes in detail the dependencies between the different variables (in this case, the measured regional responses to stimulation).

†The free parameters are estimated by minimizing a function of the observed and implied covariance matrix. To date, the most widely used objective function in structural equation modeling is the maximum likelihood (ML) function.

*All individual behavioral learning curves were well approximated by the function $1\text{-}e^{-kx}$, where $0 < k <$. 1 indexes learning speed. Small values of k indicate slower learning.

30

Measuring Brain Connectivity with Functional Imaging and Transcranial Magnetic Stimulation

Mark S. George

Daryl E. Bohning

Mark S. George: Departments of Psychiatry, Radiology, and Neurology, Medical University of South Carolina, Charleston, South Carolina.

Daryl E. Bohning: Department of Radiology, Medical University of South Carolina, Charleston, South Carolina.

THE PROBLEM OF ATTRIBUTING CAUSALITY WITH OBSERVATIONAL FUNCTIONAL BRAIN IMAGING

Part of "30 - Measuring Brain Connectivity with Functional Imaging and Transcranial Magnetic Stimulation "

Developments in functional imaging during the past two decades have allowed for significant advances in understanding how the brain functions at a systems, circuit, or organ level. Positron emission tomography (PET), single-photon emission computed tomography (SPECT), and blood oxygen level-dependent (BOLD) functional magnetic resonance imaging (fMRI) now allow researchers to image brain activity (usually related to oxygen or glucose use) with crisp spatial and temporal resolution. For example, fMRI can spatially resolve structures as small as 1 to 2 mm and view brain activity in time blocks as brief as 2 to 3 seconds. Although this time resolution is crude relative to the speed of neuronal activity and information flow between brain regions (on the order of milliseconds), these tools are nevertheless able to demonstrate the activity of clusters of brain cells through a sustained time domain in association with a behavior or task. Unfortunately, these slow time frames cannot image the directional flow of information through the brain, although exciting research in this area is under way. Thus, functional imaging tools alone have been limited in their ability to demonstrate how brain regions work in a coordinated and connected fashion to modulate information and regulate and produce behavior.

Therefore, *a fundamental problem with conventional functional imaging to date has been the inability to probe and understand the causal relationship between regional brain activity and behavior*. For example, if a brain region uses more glucose (fluorodeoxyglucose PET, or FDG PET) or oxygen (^{15}O PET or BOLD fMRI) while a subject performs a behavioral act, one can safely say that this regional activity *correlates with* the behavior. Most functional imaging researchers have correctly and appropriately used the term *correlate*, rather than *cause*, knowing well that the exact causal relationship of the regional activity to the behavior remains unclear after even the most fastidious study. For example, is the region producing the behavior? Or is the region trying to inhibit or modulate the behavior? Or is the region only incidentally activated as part of the neural network?

A recent advance in this field involves combining functional imaging with transcranial magnetic stimulation (TMS), a new technology that noninvasively stimulates the cortex. Used alone without brain imaging, TMS has been useful as a crude mapping tool for motor functions (1). Recently, by combining TMS with functional imaging, researchers have begun to test directly theories about how information flows within the brain (i.e., the functional connectivity of different brain regions). Thus, with this new combination of imaging and noninvasive stimulation, the field can now move a step closer to making causal statements of brain function. In this chapter, we introduce the technology of TMS and describe some of the important issues involved in integrating TMS with imaging to address brain connectivity. We conclude by reviewing the most recent studies in this new field in which researchers have combined noninvasive brain stimulation (TMS) with functional brain imaging.

TRANSCRANIAL MAGNETIC STIMULATION

Part of "30 - Measuring Brain Connectivity with Functional Imaging and Transcranial Magnetic Stimulation "

Transcranial magnetic stimulation is a new method for noninvasively stimulating the brain (2 ,3). With TMS, a

brief but powerful electric current is passed through a small coil of wires held against the scalp. This generates a powerful local magnetic field, which passes unimpeded through the skull and induces a weaker and somewhat less focal electric current in the brain (4 ,5 and 6). The highly localized TMS magnetic field typically has a strength of about 1 to 1.5 tesla (T) [about 30,000 times the earth's magnetic field, or about the same intensity as the static magnetic field used in clinical magnetic resonance imaging (MRI)] (7). Although different coil designs allow for more focal or more diffuse stimulation, current technology limits the depth of direct stimulation to just below the skull in superficial cortex. The magnetic field declines exponentially with distance from the coil. MRI techniques have enabled researchers actually to *image the magnetic field* of the TMS coil (8). Unfortunately, the actual important physiologic effects are likely a consequence of the *electric current density* and the *induced electric field in the area of cortex* (Appendix I). Current theories hold that the induced electric fields cause neuronal depolarization or changes in neuronal activity, which result in information flow and neurotransmitter release. Newer MRI sequences in development may someday soon allow us to image the electric current density directly and, by applying this technology to high-resolution structural imaging, actually image the induced electric field (D. LeBihan, *personal communication; May, 1999*).

Transcranial magnetic stimulation can be performed in outpatient laboratory settings in awake alert subjects (Fig. 30.1) and does not intentionally cause a seizure, nor does it require anesthesia (9). Subjects usually notice no adverse effects except for occasional mild headache and temporary discomfort at the site of the stimulation. Repeated rhythmic stimulation is called repetitive TMS (rTMS). Recent technologic advances have led to the development of magnetic stimulators that can repeatedly stimulate faster than once per second (1 Hz). By convention, stimulation faster than 1 Hz is called *fast rTMS*, and stimulation slower than 1 Hz is *slow rTMS*. This distinction is important because some evidence from work in animals (10) and humans (11) suggests that stimulation at different frequencies may have divergent and even antagonistic effects on neuronal activity (12 ,13). Importantly also, the risk for seizures in healthy adults is virtually nil with slow rTMS, and so in the United States, research with slow rTMS does not require an investigational device exemption from the Food and Drug Administration (14).

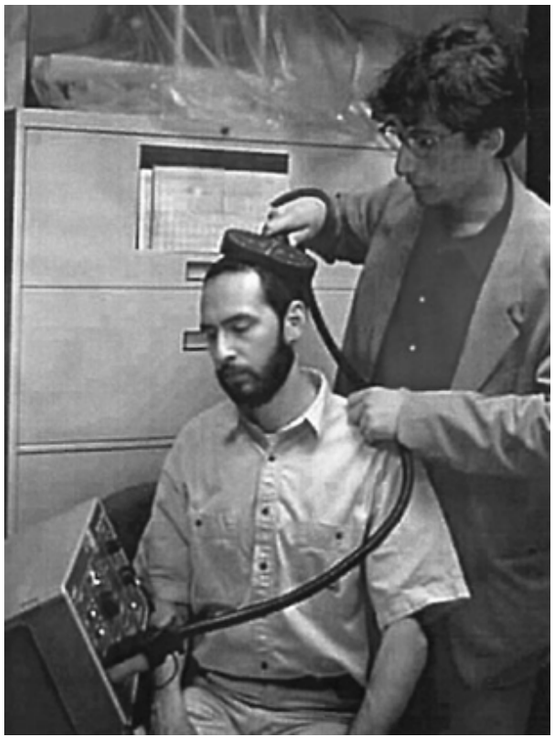


FIGURE 30.1. The chain of events by which transcranial magnetic stimulation produces changes in the brain and resulting behavior. Transcranial Magnetic Stimulation (TMS): Time-varying electrical current in a coil produces ⇒ Focal 2 Tesla magnetic field passes unimpeded through skull ⇒ Induces current in neurons ⇒ Behavioral change.

Over primary motor cortex, a TMS pulse of sufficient intensity causes movement in the opposite arm or leg (an intensity called the *motor threshold*). Similarly, a single pulse of TMS over visual cortex can produce a subjective flash of light (or phosphene). Precisely timed pulses can also interfere with, or augment, other complex tasks (see ref. 1 for review). Thus, TMS alone without imaging has been used as a relatively spatially crude mapping technique, largely over the motor cortex. rTMS at frequencies of 4 Hz or higher applied over Broca's area can cause temporary speech arrest (15). This ability to block function temporarily is frequency-dependent. It is unclear which neurons are stimulated with TMS, and whether and how this varies as a function of intensity or frequency. It is also not known how TMS causes speech arrest—whether through synaptic tetany or activation of local inhibitory interneurons.

In summary, the ability to stimulate the cortex noninvasively with TMS in an awake, alert human is an important new tool and scientific advance. At present, knowledge is limited about the physiologic and pharmacologic actions of TMS, especially as they may vary as a function of frequency, intensity, or length, in different brain regions, and in disease states versus health. Coupling TMS with imaging will likely produce new knowledge in two different areas. First, it will probably advance understanding of how TMS affects brain and *thereby refine the clinical applications* and therapeutic uses of TMS in neuropsychiatry. More importantly, from the perspective of cognitive neuroscience, combining TMS with functional imaging *will open up new avenues for the investigation of brain circuits, connectivity, and the causal chain in brain–behavior relationships* and is thus a powerful new research tool.

BACKGROUND OF THE CONCEPT OF BRAIN CONNECTIVITY AND CIRCUITS

Current Approaches to Functional Imaging Analysis and Connectivity

The first step in functional imaging is to find out which areas of the brain show activity (based on increased blood flow) when a subject performs some mental or physical task, or mentally responds to a stimulus. Conceptually, this is simple—one compares images of the brain acquired during periods when it is responding to some well-defined test stimulus with images acquired when it is performing some well-defined control task. The assumption is that the differences in the two sets of images represent the differences in brain activity during the test stimulus and during the control task. However, these are fairly subtle effects. A small change in the signal can confound the data in unknown ways. The fMRI signal is inherently noisy and often changes because of instability of the instrumentation and environmental influences on the subject. For this reason, the problem of determining the areas of the brain that are being activated by the test stimulus and assigning a probability to that determination has been extensively studied, and the field has developed commonly accepted methodologies for processing PET and fMRI data (16 ,17 ,18 ,19 ,20 ,21 and 22).

Determination of Regional Activation

The concept of constructing an interpolated spatial map of a statistical parameter, *significance probability mapping*, was developed in the analysis of multichannel electrophysiologic (EEG) data (23 ,24). In early functional imaging with PET, Fox and Mintun (16) introduced what they called *change distribution analysis*, which consisted of a subtraction of subject-averaged PET images. Present fMRI-processing methodology draws on both these ideas. In its simplest form, a pixel-by-pixel *t* test is performed by comparing the distribution of activation values for two different conditions during the course of an experiment. This gives a *t* map (i.e., an image in which each pixel represents the Student's *t* statistic for the comparison of the test condition relative to the reference condition at that location). By using the associated *p* values and the number of degrees of freedom, the *t* values can be converted to *z* values (gaussian distribution: mean 0, variance 1) to obtain *z* maps. This is the basis for statistical parametric mapping (25), formally described as the construction of spatially extended statistical processes, or maps, to test a hypothesis (usually about neurophysiology) directly. Generally based on a linear and parametric model, statistical parametric maps (SPMs) are image processes with voxel values that are, under the null hypothesis, distributed according to a known probability density function, usually gaussian. In the same way that a *t* value is interpreted by reference to Student's *t* distribution, an SPM is interpreted by referring to the probabilistic behavior of stationary gaussian fields (26) and can be used to make statistical inferences about regionally specific findings (e.g., the probability of finding an activation focus by chance). In general, SPMs characterize experimentally elicited changes in terms of (multiple) activation foci. Regions of the SPM with high or low values are interpreted as regional activations. Thus, it is possible to locate areas of the brain that are “active” during the execution of some task (i.e., the signal in those areas has a time-varying pattern that correlates with the pattern of the conditions in the experiment) (17).

Functional and Effective Brain Connectivity

Intuitively, it is common to think of *brain functional connectivity* as two or more separate anatomic areas of the brain that influence each other in the performance of some mental or physical task (e.g., recalling a name or moving a finger), or to produce a mental state (e.g., sadness). This action or state of the brain, in turn, affects the body through the somatosensory system, the sympathetic and parasympathetic nervous system, and of course the brain's neuropharmacologic/neuroendocrine hypothalamic–pituitary–adrenal system. Both in electrophysiology and functional neuroimaging, connectivity of different areas of the brain has been based on the correlation between regions. In the case of electrophysiology, this means the EEG signal (24 ,27 ,28 ,29 ,30 ,31 and 32), and in functional neuroimaging, this means the time course of regional blood flow or glucose use (18 ,21 ,33 ,34)).

Friston et al. (19) emphasize the distinction between *functional connectivity*, the temporal correlations between remote neurophysiologic events, and *effective connectivity*, the influence of one neural system on another (i.e., a functional as opposed to a causal relationship). Viewed in this way, functional connectivity is simply the observed covariance among different brain systems. It is an operational definition and says nothing about the causal relations of the observed correlations. To characterize distributed brain systems, the functional connectivity (covariance) matrix, obtained from a time series of neurophysiologic measurements, is subjected to principal component analysis (PCA) (20 ,35) (Appendix II). The resulting eigenimages (principal components or spatial modes) each identify a spatially distributed system, comprising regions of the brain that are jointly implicated by virtue of their functional interactions (connectivity). This analysis of neuroimaging time series is predicated on established techniques in electrophysiology (both EEG and multiunit recordings). For example, in the analysis of multichannel EEG data, the underlying spatial modes that best characterize the observed spatiotemporal

dynamics are identified with a Karhunen–Loeve expansion. Commonly, this expansion is in terms of the eigenvectors of the covariance matrix associated with the time series. The spatial modes are then identical to the principal components identified with a PCA.

Structural Equation Modeling

Principal component analysis and factor analysis approaches attempt to integrate the spatially distributed activations found in SPMs into functional systems characterized by the eigenimages or spatial modes. However, one would like to go further and explore the influence of one area on another (effective connectivity), not just the correlation (functional connectivity). Many anatomic connections are reciprocal, and simple pairwise correlations cannot resolve asymmetric influences. Structural equation modeling is an attempt to address this problem.

In describing their neural structural equation models, McIntosh and Gonzalez-Lima (21) use the terms *anatomic model* and *functional model* (36). The *anatomic model* simply represents the discrete anatomic brain regions and the neuroanatomic connections between them used in the structural equation models. These anatomic models have been derived from the observation of patients with brain lesions and from animal studies, or inferred from neuroimaging studies and the analysis of SPMs. The interregional correlations of activity are used to assign numeric weights to the connections in the anatomic model, which leads to the functional model. A *functional model*, therefore, represents the influences of regions within the model on each other through the anatomic connections, and both the magnitude and the sign of the path coefficients can be estimated. In some respects, the functional model is close to the notion of effective connectivity (20 ,37) because it depicts the influence of one region on another. The difference is that the influences in the functional model, unlike effective connections, are explicitly depicted as direct and indirect effects through the anatomic model. Effective connectivity, as defined by Aertsen et al. (38), resembles most closely direct effects in that an effective connection is the influence of one neural element on another irrespective of direct or indirect influences. In structural equation modeling, effective connections, or total effects, are further decomposed into direct and indirect effects by use of the anatomic model. A similar distinction can be made in covariance analysis, which is often characterized as exploratory (objective) or confirmatory (theoretical) analysis. PCA and factor analysis are essentially exploratory techniques because no constraints are placed on how the variance in the system is expressed. Structural equation modeling is typically thought of as a confirmatory approach (confirmatory factor analysis) because a causal model is usually being confirmed or disconfirmed (39).

Transcranial Magnetic Stimulation as a Probe to Alter Connectivity Networks

Although the PCA and structural equation modeling techniques are well grounded statistically and quite powerful, they cannot eliminate the possibility that apparent interactions between two regions may be a consequence of other factors. The activity of two theoretically linked regions may be modulated, either directly or via changes in neurotransmitter release, by neurologic activity in a third region that is outside of the field of view of the imaging experiment, or has not been included in the structural equation anatomic model. The two regions may also be responding to different aspects of the test stimulus, either inherently because of the nature of the task, or because of engagement of the subject in performing the task.

Since it was first developed, TMS has been used to test nerve connections, nerve excitability, and nerve conduction times. One might think of this as two anatomic areas with a single connection. Paus et al. (34) and our laboratories at the National Institute of Mental Health (40 ,41) and the Medical University of South Carolina (42 ,43) demonstrated that TMS might be combined with neuroimaging to explore the connectivity of more complex three-dimensional networks in the brain to allow the direct assessment of neural connectivity without requiring the subject to engage in any specific behavior.

Because TMS seems to have a disruptive effect in most areas of the brain, its most likely use will be to suppress the activity of a region of the brain or disrupt communication between areas. This may be done by simply applying the TMS pulse at the moment the task is performed or the stimulus is applied, and noting the changed response pattern. It may also turn out that it will be possible to apply TMS after a precisely timed delay to modulate responses (44) and so investigate brain communications at time resolutions far greater than that of the hemodynamic response, approaching that of EEG. Thus, TMS provides a noninvasive means of perturbing brain circuits both spatially and at high temporal resolution. Because it is a noncognitive stimulus, its effects are less dependent on subject engagement (“attention and performance”), and because it is a more direct and quantifiable stimulus, it more closely relates to basic neurophysiologic parameters such as nerve excitability and conduction times.

INTEGRATING TRANSCRANIAL MAGNETIC STIMULATION WITH FUNCTIONAL IMAGING: PROBLEMS AND CHALLENGES

Part of "30 - Measuring Brain Connectivity with Functional Imaging and Transcranial Magnetic Stimulation "

Stimulating the brain with TMS while simultaneously imaging brain activity presents a host of unique technical problems, including (a) physically placing the coil in the scanner

and over the appropriate brain regions, (b) determining whether the TMS coil interferes with the functional image, and (c) integrating the brief time domains of TMS with the slower temporal resolution of most modern imaging tools. We discuss several of these issues and recent attempts at dealing with them.

Placement of the Coil—Structural or Functional Guidance

One of the most obvious problems in combining imaging and stimulation revolves around how to position the TMS coil over the skull. Most researchers have used either a structural or functional guidance system.

Structural Guidance

The shape of the coil determines the magnetic field in the brain, and thus the pattern of induced electric current (7 ,45). For circular coils, the magnetic field is most intense near the windings. When a circular coil is placed flat against the scalp, it induces a toroidal ring of electric current in the underlying cortex that is of the same size as the coil itself but more diffuse. The electric current distributions are assumed to be broad and the effects distributed. In contrast, with figure 8 coils, a focus at the intersection of the two loops is roughly twice as intense as that obtained with a circular coil and the same current (Fig. 30.2). Although the distribution of induced current is still fairly broad, stimulation over motor cortex demonstrates that it is sufficiently focal to cause movement in one location; moving the coil less than a centimeter or even slightly changing the angle results in no movement, or movement in different muscle groups.

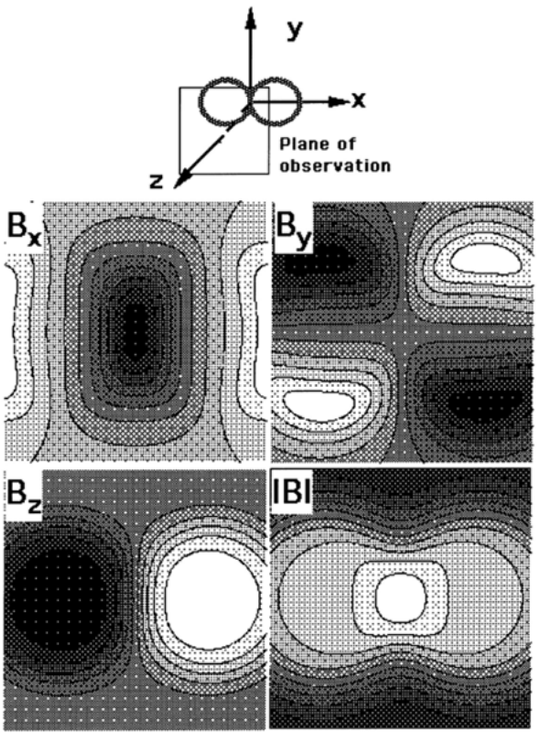


FIGURE 30.2. Magnetic field of single-turn figure 8 coil in plane parallel to and one-fourth diameter from the coil: *x, y, z* components and magnitude of field. (From Cohen LG, Roth BJ, Nilsson J, et al. Effects of coil design on delivery of focal magnetic stimulation. Technical considerations. *Electroencephalogr Clin Neurol* 1990;75:350–357, with permission.)

The coil can be positioned in several ways, based on the underlying brain structure. Perhaps the best method is to acquire a structural MRI scan of the head and then use image-guided systems to align the TMS coil precisely over a specific brain region. Several groups are exploring this option by using structural MRI and then integrating the TMS with PET (46 ,47). Performing the same mechanical alignment within an MRI scanner is more challenging because of the problems that arise when metal is used within a powerful magnetic field. An intermediate approach, employed by the McGill group (34), is to position the coil according to a probabilistic brain system keyed to landmarks on the subject being studied, rather than according to the subject's known anatomy as determined by MRI. This method is much easier and is adequate if one is planning on using only group statements for the statistical analysis (which requires spatially transforming the imaging data into a common brain atlas). Unfortunately, because the structural morphology of the brain and the functional location of behaviors vary greatly between individuals, the probabilistic method suffers from the problem that it is not certain *within an individual* whether the coil is positioned over the structure or region being studied.

The main difficulty with both image-based and structural localization is that unless one knows the exact relation of the induced currents relative to the position of the TMS coil, one really does not know where to stimulate most efficiently. The induced electric current is tissue-dependent, so it is not the same at different places on the scalp. Only if the currents had the same relation to the position of the coil, and one could use MR guidance to place the coil over the same sulcus or gyrus in a subject's brain, could one even be sure of stimulating the same structural area. However, brain conformation varies, so the induced currents might still impinge on the cerebral cortex at a different angle.

Paus and colleagues (34) approached this problem by obtaining an image volume with MR from each subject and

spatially transforming it into Talaraich space (a widely used common brain space) (48). After determining the Talairaich location shown by neuroimaging studies to correspond to a function, they performed the inverse transform back into the MR image space for each subject. Finally, using stereotactic guidance (49), they positioned the coil over this point, in effect ensuring that they were probably stimulating the functional location of the behavior in all subjects. Krings et al. (50) used a frameless stereotactic system to coregister TMS motor maps with fMRI data obtained during performance of a motor task. In two patients, they also performed direct cortical stimulation, finding good correspondence among all three methods. This probabilistic technique is more or less acceptable, depending on how consistently from one subject to the next a function is located within identifiable anatomic structures, and on how the shapes of those structures vary. Thus, this technique still entails the problem that function does not strictly map to the same location in different individuals.

Placement Based on Function

A different approach is to use TMS, electromyography, or functional imaging to determine the regions activated during a behavior and then position the coil directly over the functioning region. For behaviors like movement or phosphene production over visual cortex, one can bypass the imaging step, simply finding the scalp location with the desired behavioral effect and then keeping the TMS coil at this spot throughout an imaging study (*functional behavioral* approach to placement). Unfortunately, outside primary motor and vision areas, TMS does not produce easily viewed effects, so that this direct functional approach becomes impossible.

Despite its apparent simplicity, the *functional behavioral* approach of using elicited movement to guide coil placement to perform TMS over the motor cortex is associated with certain problems. The movement elicited by the TMS is a reassuring, if somewhat imprecise, way of being certain that one is in the correct area. However, even with comparable visible movement, one can be on one side or the other of the target area, and it is difficult to know how much or how little additional stimulation is occurring. Because this method reliably causes activation in large corticospinal circuits, we have used the functional behavioral approach for initial studies of interleaved TMS/fMRI effects over motor cortex (51 ,52).

Whichever method of locating the site of stimulation is used, it is important that the TMS coil be positioned accurately and repeatably, and then held securely in place so that its position relative to the brain is maintained throughout the stimulation. Each group seems to have developed its own mounting systems. We have developed a system for accurately and repeatably positioning the TMS coil within an MRI scanner (53). Both structural and functionally guided techniques have their place, and eventually systems will likely be developed for relating the two.

Interaction of the Transcranial Magnetic Stimulation Coil and Image Acquisition

Yet a different technologic problem in this new area revolves around whether and to what extent the TMS coil interferes with the functional image acquisition. Because of this concern, the early combinational studies used imaging techniques in which TMS was delivered in a location other than that where the actual brain imaging was performed (54 ,55). Both FDG PET and perfusion SPECT allow one to administer TMS and deliver the radiopharmaceutical agent away from the scanner. The tracer crosses the blood–brain barrier and then settles into active regions. The subject can then be transported to the scanner for image acquisition.

Because of radiation dose limits and slow time resolution, neither of these techniques (FDG PET and perfusion SPECT) is suited for thoroughly examining circuits and behavior with the combination of TMS and imaging. For this purpose, it is much better to have the TMS coil directly within the scanner. However, one then has to understand to what extent, if any, the TMS coil interferes with the acquisition of the functional images.

In PET, groups can perform an initial transmission scan before the functional image and then subtract the minor reduction in tracer counts caused by the TMS coil. Obviously, solid core coils are not suited for this type of combinational imaging. With fMRI, the TMS coil can produce both static and dynamic artifacts. Although it may be possible to correct for the static effects, as in PET, some groups are so concerned about static artifacts that they have developed systems that quickly lift the TMS coil 2 to 3 cm from the scalp during the actual MR image acquisition following a train of pulses (56). Both mechanical and pneumatic systems have been developed to do this. Although it minimizes the potential impact of the artifact, mechanically moving the TMS coil produces shimming and alignment issues on its own and does not allow for true interleaved imaging in real time. The dynamic artifacts produced by TMS within an fMRI scanner are both more complicated and more difficult to account for (see Fig. 30.3 and ref. 53 for a full discussion). Substantial progress has been made, so that this is not a major concern.

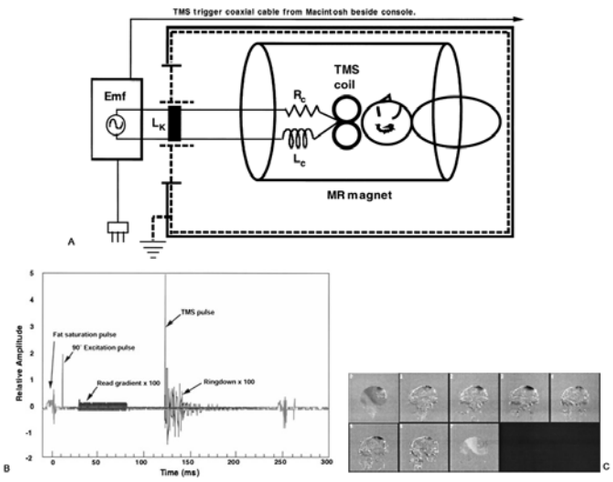


FIGURE 30.3. Researchers at the Medical University of South Carolina have recently developed the technique of performing transcranial magnetic stimulation (TMS) within the bore of a conventional 1.5-T magnetic resonance imaging scanner, the setup of which is depicted in **(A)**. This process produces dynamic TMS-induced eddy currents **(B)** and static TMS-induced eddy currents **(C)**. As seen in **(B)**, these dynamic eddy currents are approximately twice as strong as the read gradient for about 20 milliseconds, and then drop to approximately the same size as the read gradient for another 20 milliseconds. Although the major eddy currents have died out by 40 to 50 milliseconds after the TMS pulse, some longer, low-level currents are still present that cause significant image artifact **(C)**.

Integration of Temporal Domains of the Scanner and Transcranial Magnetic Stimulation

The final picture produced by each of the functional imaging tools represents summed brain activity over a measure

of time. The averaged time domain ranges from 20 to 30 minutes for FDG PET, to 40 to 60 seconds for ^{15}O PET and perfusion SPECT, to 2 to 3 seconds for BOLD fMRI, to milliseconds for EEG and electromyography. The actual TMS pulse is very brief, on the order of 300 microseconds. Thus, it is important in all combined imaging studies to understand the relationship between the TMS activity and the summed functional image. Moreover, because of the concerns of potentially causing a seizure with long trains of high-intensity, high-frequency TMS, only certain TMS parameters can be used constantly over the time domains of some forms of imaging. Unfortunately, some TMS effects, such as speech arrest, occur only with high-intensity and high-frequency stimulation.

Steady State

In this model, researchers perform TMS throughout an entire scan and then compare results with those of another scan in which all conditions are the same except for the TMS. Even with this design, and stimulation frequencies of approximately 1/s (1 Hz), most of the imaging is performed with the actual TMS machine off as a function of time.

Block Design

A different model is to scan in blocks in which periods of TMS are separated by periods of rest. An example of this is shown in Fig. 30.4 , which describes the interleaved TMS/fMRI

setup. Figure 30.5 portrays results of a recent study in which TMS was administered over motor cortex. This study showed that TMS at intensities slightly greater than motor threshold (110%) activates approximately the same number of pixels in the same region as does a volition movement (Fig. 30.5C). This study also revealed the relative magnitude of the TMS effect and the temporal relationship to changes in blood flow.

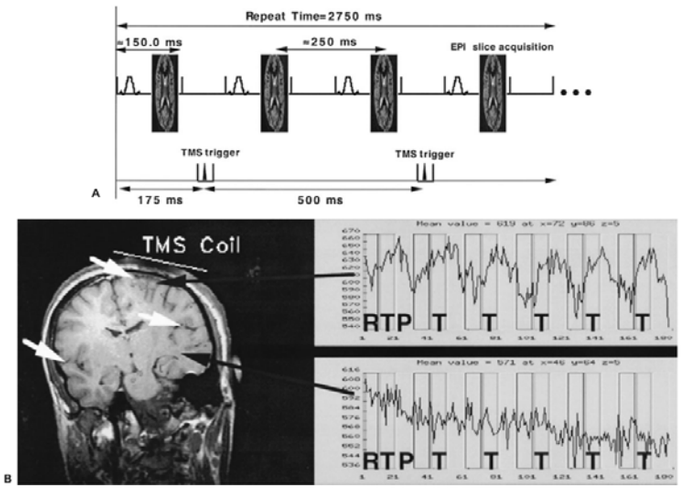


FIGURE 30.4. To perform interleaved transcranial magnetic stimulation/magnetic resonance imaging (TMS/fMRI), one must coordinate the TMS pulses with the MRI signal acquisition and interleave the two. **A:** An example of this process for a TMS rate of 1 to 2 Hz. **B:** An example of serial blood flow changes underneath the TMS coil (over left motor cortex) and in a control region. Note the increase from rest, *r*, in absolute blood oxygenation-dependent (BOLD) activity underneath the coil when it discharges at 1 Hz (task, *T*), and how it decreases afterward (post, *P*).

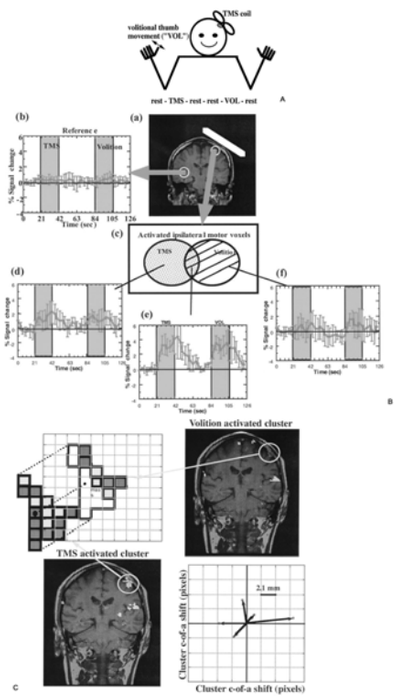


FIGURE 30.5. One of the first studies in which this interleaved technique was used attempted to detect differences between volitional and transcranial magnetic stimulation (TMS)-induced movement of the thumb. In **(A)**, the TMS device was placed over the left motor cortex of subjects, who alternately had TMS move their thumb *(TMS)* and then volitionally moved their thumb in response to a tone *(VOL)*. In **(B)** are averaged group time series of brain activity during TMS, volition, or a noise control region *(upper left)*. Note that for voxels that were activated in both tasks, the percentage rise in blood oxygenation-dependent (BOLD) activity does not differ from baseline. Thus, TMS produced BOLD changes that are dynamically similar to those of regular movement. In **(C)**, the center of mass of the BOLD signal is virtually the same for both TMS and volition, within the limit of resolution of the magnetic resonance imaging scanner (2 mm).

Single-Event fMRI

The method that is currently closest to the actual timing of TMS and brain events is single-event fMRI, or averaged-single-trials fMRI. With this method of scanning, images are steadily acquired at a rapid rate while the performance of a single event is rapidly interspersed. One can image the brain activity associated with a single TMS pulse by repeating the event many times and averaging the images acquired at similar times after the events, much as electrophysiologists have done with evoked responses (57). Although the BOLD response is relatively sluggish (on the order of 2 to 3 seconds), some groups are experimenting with the initial slope of the response to attempt to increase time resolution (Fig. 30.6). Applying TMS pulses to different brain regions with different interpulse interval times (milliseconds) may represent a unique way of improving the temporal resolution

of BOLD fMRI and of studying information flow within circuits. With further refinement, the combination of single-pulse and paired-pulse TMS and averaged-single-trials fMRI will probably be of considerable interest in *in vivo* neurophysiology.

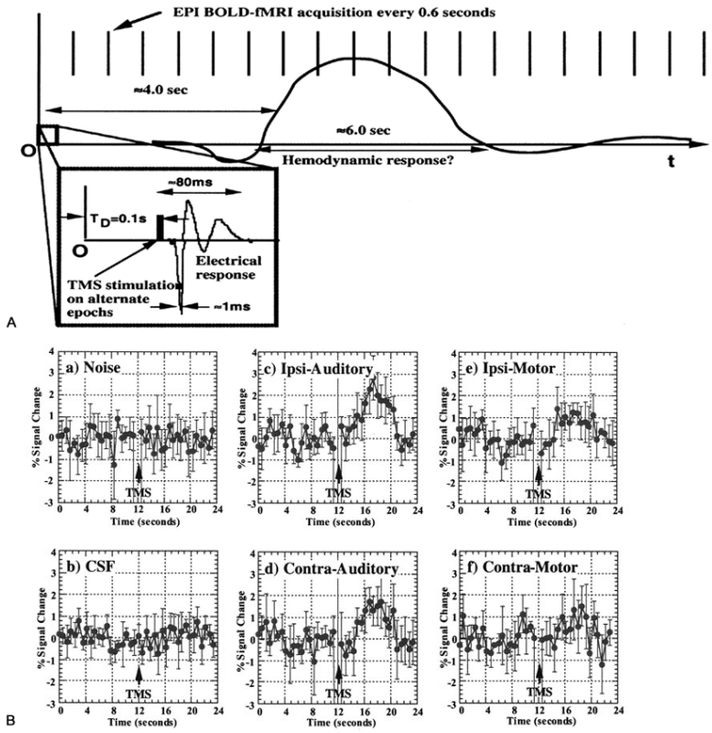


FIGURE 30.6. One can also use an averaged-single-trials approach of transcranial magnetic stimulation (TMS) and functional magnetic resonance imaging (fMRI)—that is, one can discharge a single TMS pulse and then measure the blood oxygenation-dependent (BOLD) response *(top)*. The time series above show the BOLD response in a control region, for the auditory cortex, and in motor cortex. Note that a single pulse of TMS over motor cortex sufficient to cause the opposite thumb to move produces more blood flow changes in auditory cortex (caused by noise) than it does in the motor cortex under the coil.

REVIEW OF TRANSCRANIAL MAGNETIC STIMULATION FUNCTIONAL IMAGING STUDIES TO DATE

Part of "30 - Measuring Brain Connectivity with Functional Imaging and Transcranial Magnetic Stimulation "

Transcranial Magnetic Stimulation Interleaved

Fluorodeoxyglucose PET

The first published combination of TMS and functional neuroimaging in real time was performed with FDG PET in a patient before and after rTMS treatment for refractory depression (54). At a separate time, these investigators also injected the glucose while the patient was intermittently stimulated at 20Hz over the left prefrontal cortex for 20 minutes. In comparison with her depressed scan at baseline, her total brain metabolism rose following weeks of TMS treatment. Also, the scan that was taken during left prefrontal TMS showed marked increases in activity, especially over the prefrontal cortex. Conclusions from this single case study are limited.

Complexity of the Issues as Demonstrated by Initial Simple Studies over Motor Cortex

A basic question for TMS and functional imaging is what happens to blood flow or activity in motor cortex while TMS is stimulating the thumb. A straightforward hypothesis would be that TMS increases blood flow in a manner similar to that produced by volitional movement. Confusion ensued when an early and still unpublished study of 1-Hz stimulation over the motor cortex for thumb showed decreased glucose uptake at the putative site of stimulation and in the contralateral motor cortex (40). Stimulation was performed at 1 Hz because FDG takes 20 minutes to settle into neurons and is thus a composite picture of brain activity over 20 minutes. This paradoxical decrease in localized brain activity both under the coil and at the mirror or contralateral site during TMS was surprising, but findings of decreased brain activity like this had been found in some electrophysiologic studies (12). The final image was a summed picture of 20 minutes of brain activity. It is likely that TMS has multiple different effects during that time—increased activity immediately with stimulation, decreases during the rest time between TMS pulses, and dynamic changes across the 20 minutes. Peter Fox and one of the chapter authors (MSG) (58) next sought to test this finding directly by using ^{15}O PET rather than FDG PET, with the TMS coil directly in the scanner. ^{15}O PET has a shorter time frame (approximately 1 minute for tracer uptake) than ^{18}FDG PET (20 to 30 minutes). Therefore, imaging with ^{15}O PET during stimulation requires that the TMS coil be placed inside the PET gantry. Using the exact same design as in the FDG study, but scanning every 10 minutes for 1 minute, we found that slow (1-Hz) rTMS over the motor cortex caused an *increase* in cerebral blood flow, although this was noted in only four subjects (58). Both of these studies used a functional behavioral placement

of the TMS coil over the optimal position for movement of the thumb and stimulated at or near the motor threshold with visual confirmation that the TMS was producing activity in the motor circuit. Thus, the results of these two initial studies were confusing and frankly contradictory.

To add even more confusion, Paus et al. (34) in the same year published a study combining ^{15}O PET and TMS. In this study, intermittent fast (10-Hz) rTMS over the frontal eye fields for 1 minute caused dose-dependent *increases* in blood flow at the stimulation site and in visual cortex. In other words, when they increased the number of 10-Hz trains within the minute, blood flow increased. Surprisingly, when the same investigators used the same rTMS parameters in the same subjects but shifted the coil to motor cortex, they found a *dose-dependent reduction in cerebral blood flow* (59). Importantly, they positioned the coil based on a probabilistic brain, and they also stimulated below motor threshold. No thumb movement occurred in these subjects.

Thus, the initial dream of using TMS and imaging to address connectivity problems in the brain has been hindered by a lack of consensus about basic imaging and TMS questions. Using yet a different technology, BOLD fMRI, our group in several studies consistently found that over much shorter time domains (7 to 30 seconds), TMS at motor threshold or above, positioned by a functional behavioral approach, consistently produced increases in blood flow at the stimulation site and in connected regions, such as the contralateral motor cortex and cerebellum (51 ,52). The issue now appears to be settled; the same National Institutes of Health group that found decreases with FDG PET has recently completed a more fastidious ^{15}O PET study. In this study, Speer and colleagues found dose-dependent *increases in blood flow* in motor cortex with 1-Hz TMS, as was noted in the study of Fox et al. (58) and confirmed with the BOLD fMRI technique by our laboratory (A. Speer, *personal communication; May, 2000*).

There is now a small consensus in the existing literature that blood flow increases under the motor cortex in a dose-dependent manner when the TMS coil is positioned by finding the appropriate spot for optimal thumb movement (functional behavioral technique) and stimulation is above motor threshold (and activates large excitatory neurons). When the TMS coil is positioned in this same region by a probabilistic approach, dose-dependent decreases have sometimes been found. Thus, some of the discrepancy in the literature can be explained not only by differing time domains of the imaging technologies, but also by potential differential effects caused by the method of coil placement. In this vein, using an identical study paradigm as their most recent TMS motor study, the National Institutes of Health group (Speer and colleagues) stimulated the same subjects over the prefrontal cortex, defined simply as a certain distance from the motor area (a very crude probabilistic approach, such as has been employed in many of the TMS challenge and clinical studies). Paradoxically, this group found in these same subjects *dose-dependent decreases* in blood flow over prefrontal cortex. In earlier work at 80% motor threshold, we found similar decreases in eight healthy adults when we used perfusion SPECT, with a tracer uptake time of 30 to 40 seconds, to image cerebral blood flow during fast (20-Hz) left dorsolateral prefrontal cortex rTMS (60). In comparison with a control scan with sham TMS, we found relative decreases under the coil site and in the anterior cingulate and orbitofrontal cortex. In contrast, a recent BOLD fMRI study over prefrontal cortex by our group found increases in blood flow at 120% motor threshold (61). With fMRI, one can examine individual differences, and a great deal of heterogeneity of response was noted across subjects. We are currently performing repeatability studies within subjects over time to address the inherent noise in this scanning system and the question of whether repeated TMS/fMRI studies yield consistent results.

The two most likely explanations for the opposite findings over motor and prefrontal cortex are that different brain regions react differently, or that the method of TMS coil placement matters, and that the effects of clear stimulation of large corticospinal neurons may be different from those of nonspecific stimulation of local inhibitory neurons with only probabilistic positioning. Obviously, a series of studies is needed to settle this most important issue. For example, an important next study would be to test directly the issue of blood flow as a function of functional behavioral versus probabilistic coil placement and see if functional positioning produces increases in blood flow and probabilistic placement decreases, presumably secondary to differential stimulation of excitatory versus inhibitory neurons.

BOLD fMRI

As mentioned above, the most promising, but also the most technically challenging, TMS imaging modality is a combination of TMS and fMRI. Bohning et al. (51) first demonstrated the capability of interleaving TMS and BOLD fMRI with good spatial and temporal resolution. This technique was initially thought impossible by many because of concerns about introducing a focal TMS magnetic field (1 to 2 T) inside a clinical MRI scanner of 1.5 T. Our group has found that this technique, with the right precautions, is both feasible and safe. Considerable progress has been made in devising a system for interleaving TMS with fMRI (53). Figure 30.4 shows one subject's brain with areas of TMS-induced activation superimposed in color. The time–activity curve shows the changes in BOLD signal over the course of the experiment as the TMS machine is alternately triggered at 1 Hz for 18 seconds and then turned off.

Work to date with this technique has shown that it is sensitive enough to detect subtle differences in brain blood flow response that result from minor changes in TMS intensity (52).

Additionally, direct comparisons of blood flow changes in motor cortex caused by TMS or volition show a surprising similarity between TMS-induced changes and those associated with normal movement (71). For example, the location of the peak blood flow change is the same for TMS and normal movement (within 2 mm). Also, stimulating at around 1 Hz and just at motor threshold activates roughly the same amount of brain tissue, and to the same degree. Thus, although many have the perception that TMS is causing supraphysiologic changes in the brain, these data imply that TMS at these parameters is acting remarkably like normal physiology.

Using Interleaved TMS/fMRI to Address Issues of Connectivity: An Initial Study

Several electrophysiologic studies have suggested that 1-Hz TMS over time domains of 3 minutes or more is inhibitory (12). To test whether this inhibitory effect occurs at time domains of several seconds, we performed TMS within an fMRI scanner and measured blood flow with the interleaved TMS/fMRI BOLD technique. Within a 1.5-T MRI scanner, five adults were stimulated by applying a figure 8 TMS coil over the left motor cortex at the optimal spot for producing movement in the contralateral (right) thumb (abductor pollicis brevis). In 21-second epochs, subjects alternated between rest and a sequential finger opposition task in their left (nondominant) hand. In alternating-movement trials, TMS was applied either at 120% motor threshold or 10% of stimulator output (below the threshold for movement) (Fig. 30.7). Coronal echo-planar BOLD images were acquired continuously throughout, interleaved so that TMS occurred 100 milliseconds after every fourth image acquisition.

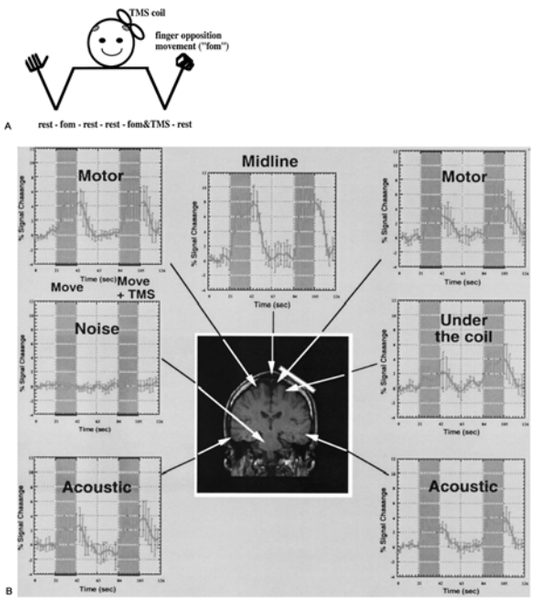


FIGURE 30.7. Several transcranial magnetic stimulation (TMS) electrophysiology studies have demonstrated that low-frequency TMS over one motor cortex can inhibit the opposite motor region. In this study, we applied TMS over the left motor cortex and had subjects perform a complex task with their nondominant (left) hand. TMS was applied on alternate epochs. We hypothesized that TMS would inhibit the blood oxygenation-dependent (BOLD) response in the right motor cortex. We did not see this. Interestingly, TMS produced an increase in BOLD response under the coil in an area of cortex that was already active. This simple study in which TMS was used to test connectivity highlights many of the issues raised in this chapter concerning how and where to apply TMS, and whether baseline activity in the underlying brain matters in terms of response.

With this technique and at these short time intervals, TMS did not inhibit the local or remote BOLD response during movement. In fact, TMS of the left hemisphere caused a local 1.5% increase in blood flow in addition to the 2.5% activation caused by the complex movement. We therefore concluded that the application of TMS over motor cortex for 21 seconds during a motor task-enhanced motor cortex activation does not inhibit the BOLD response. Actually imaging the remote inhibitory or modulatory effect of TMS at a different site remains for the future.

Averaged Single Trials

There is no doubt that a single TMS pulse applied to motor cortex is capable of causing a neuronal response because its consequences can be clearly seen in the form of an overt movement of the contralateral extremity. However, to date, the only functional neuroimaging technique in which the response to single-pulse TMS has been observed is EEG. Our laboratory recently sought to determine if interleaved TMS and fMRI could be used with an averaged-single-trials protocol to detect BOLD response to neuronal activation induced by a single TMS pulse, and to measure its time course.

The technique is important because it allows a comparison of different TMS events by means of their associated BOLD responses. For example, with a single-event technique. it might be possible to compare different TMS intensities, or coil orientations, or single versus paired stimulation (through one coil or possibly two different coils, one conditioning coil and one test coil). Such studies could provide a bridge between electrophysiology (variation of motor evoked potential amplitudes) and fMRI (variation of BOLD response). Moreover, combining TMS with precise timing relative to a behavior with the averaged-single-trials technique would likely make it possible to image the activity of brain circuits and their connections.

In an initial study in this area, five healthy volunteers were studied with interleaved TMS/fMRI and an averaged-single-trials protocol (57). The BOLD fMRI response to single TMS pulses over the motor cortex was detectable in both ipsilateral motor cortex under the TMS coil and contralateral motor cortex, and also bilaterally in auditory cortex. The associated BOLD signal increase showed the typical fMRI hemodynamic response time course. The response of the brain to a single TMS pulse over motor cortex at 120% of the level required to induce thumb movement (1.0% to 1.5% signal increase) was comparable in both level and duration with the auditory cortex response to the sound accompanying the TMS pulse (1.5% to 2.0% signal increase) (Fig. 30.5).

Thus, ultimately, TMS combined with fMRI may allow for more exact positioning of the TMS coil, with information obtained about the magnetic field produced and also about alterations in physiology and biochemistry. Refinements of the averaged-single-trials technique and precise timing of TMS offer the potential of increasing the temporal resolution of fMRI and promoting its evolution into a tool for assessing connectivity. Much background work is needed before this combined technique can achieve its potential.

Quantitative EEG

Ilmoniemi et al. (62) have combined high-resolution quantitative EEG (qEEG) with TMS and reported regional changes in spectral content that shifted over very brief episodes of time and corresponded with known regional connections with primary motor cortex. High-resolution EEG clearly has the best temporal window of all the techniques (in the millisecond range), although the spatial resolution is poor. Unfortunately, this group in Finland is the only one to date to be able to circumvent the technical problem of recording EEG immediately after TMS and so avoid the artifact produced by the TMS pulse. This area has not advanced as rapidly as expected in the last 3 years, perhaps because of the complexity of the technique.

Transcranial Magnetic Stimulation and Multimodal Integration

Lastly, several groups have now used TMS in a complementary way to address systems neuroscience questions. The TMS aspect of the study serves to confirm or validate a result from a purely functional imaging study. Two recent studies illustrate many of the important aspects of this type of work.

Kosslyn and colleagues (63) used TMS to investigate secondary visual cortex (area 17) and visual imagery. As a first part of this study, a traditional ^{15}O PET study was performed in subjects while they visually imagined a stimulus. As predicted from previous imaging studies in humans and animal studies, area 17 activated during this task. In a different cohort, with the use of probabilistic positioning, TMS over putative area 17 interrupted this visual imagery task. This study suffered in several respects, most notably in not knowing whether the TMS actually was applied over area 17. It nevertheless demonstrated the potential of using TMS as a convergent method of testing brain–behavior theories.

In a more elegant and rigorous application of this approach, Desmurget and colleagues (64) used TMS and imaging to test the role of the posterior parietal cortex in correcting the ongoing trajectory of movements (64). They scanned healthy subjects while they pointed to visual targets that either remained stationary or moved during saccadic eye movements. Then, using a functional image-based positioning system, they applied TMS over the left posterior parietal cortex during stimulus target presentation. The TMS disrupted the normal path corrections that occur in moving objects. Thus, in this study, TMS indicated the necessary and critical role of an area in the performance of a behavior and extended the traditional observational imaging approach.

CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

Part of "30 - Measuring Brain Connectivity with Functional Imaging and Transcranial Magnetic Stimulation "

Transcranial magnetic stimulation combined with functional imaging offers the promise of a better understanding of brain circuits and the causal relationship between behaviors and activity in distributed brain regions. Several studies with a variety of imaging modalities have begun to use this approach. These studies have largely demonstrated that before the combination of TMS and imaging can be used fully, much more work is needed to improve methods. Very basic questions remain largely unexplored. These include how best to position the coil (functional behavioral versus probabilistic), how to adjust the intensity for nonmotor areas of cortex, and whether to account for differences in depth into the brain (e.g., atrophy). Additionally, a true understanding of TMS effects on the brain are still lacking, and this incomplete knowledge contributes to the lack of understanding of how best to use TMS to address systems neuroscience questions. For example, do different frequencies of stimulation produce varying effects on local metabolism, and if so, how? More complete knowledge of the local pharmacologic effects of TMS as a function of the many parameters of use would greatly advance our ability to apply TMS/imaging in neuroscience research. It is obvious that a great deal of systematic work is needed to understand this interesting tool.

However, although a better knowledge of TMS brain effects would expand and improve its use as a neuroscience tool, the ability to combine noninvasive stimulation of the brain with real-time functional imaging is an important new technique that will no doubt add to our ability to understand brain connectivity.

APPENDIX I

Determining The Appropriate Model For Calculating The Induced Electric Current In The Brain

Although, typically, the spherical model has been used, this assumes that the brain is a sphere with uniform conductivity inside spherical shells with different conductivities, corresponding to the skull and scalp. One group has gone so far as to use tissue segmentations based on MR images and estimates of gray and white matter and cerebrospinal fluid from the literature and the theoretic field of the TMS coil to perform finite element computations of the electric currents induced in individual brains (69). Unfortunately, these computations of the electric currents were performed with special field computation software and a supercomputer (69). Although the assumptions of most computational models that the brain is spherical and the cortex is an isotropic, homogeneous volume conductor are simplistic, some important observations have been made. The charge accumulation on the tissue surface tends to cancel the perpendicular component of the induced electric field, shielding the brain. This forces the resultant electric field to lie predominantly parallel to the tissue surface and fall rapidly with depth (65 ,66 ,67 and 68). These observations are also expected to be valid for models that more faithfully represent the actual shape and composition of the brain by treating it as a summation of finite elements. To take into account the inhomogeneous conductivity of the brain, Cerri et al. (69) used MR images to segment the brain into white matter, gray matter, and cerebrospinal fluid, and a conductivity versus gray level interpolation function derived from tissue conductivity data in the literature to obtain a three-dimensional conductivity map. They then divided the brain into discrete resistive cells (quasistatic approximation) and used a supercomputer to determine the current distribution that would be induced in the three-dimensional resistive network by

an external magnetic field pulse However, such methods are not generally available and are still an approximation. A means of imaging the induced electric field is what is really needed.

APPENDIX II

Principal Component Analysis And Singular Value Decomposition

Principal Component Analysis

Principal component analysis is a mathematical device that uses the intrinsic spatial and temporal properties of a set of fMRI data to reduce its dimensionality. PCA does not refer to a specific statistical model, entails few assumptions, and makes no comment on the significance of the resulting spatial modes. By orthogonalizing the covariance matrix, PCA extracts its important features in terms of principal components, or eigenvectors. These vectors are the linear combinations that account for independent or orthogonal amounts of variance in the observed data. In terms of functional connectivity, a principal component represents a spatially distributed brain system, comprising a subset of brain region, within which many temporal intercorrelations exist. Because any one principal component is orthogonal to the remaining principal components, these systems are functionally unconnected from each other, even though any single area may be implicated in more than one system. To perform PCA, a mathematical technique called *singular value decomposition* is usually used (70).

Singular Value Decomposition

Given a set of M linear algebraic equations relating a set of N unknowns, x_j, j = 1,2,...

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1N}x_N &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2N}x_N &= b_2 \\ \cdots \qquad\qquad\qquad\qquad\qquad &\cdots \\ a_{M1}x_1 + a_{M2}x_2 + \cdots + a_{MN}x_N &= b_M \end{aligned}$$

or, in matrix form,

$$A \cdot x = b$$

where the as and bs are known. If $N = M$, there are as many equations as unknowns, and there is a good chance of finding a unique solution set of x_js.

If $M < N$ or $M = N$ but the equations are not all linearly independent, then there are effectively fewer equations than unknowns. In this case, either there is no solution, or else there is more than one solution vector x. In the latter event, the solution space consists of a particular solution x_p added to any linear combination of (typically) N-M vectors (which are said to be in the nullspace of the matrix A). The task of finding the solution space of A is called *singular value decomposition* of a matrix A.

Singular value decomposition explicitly constructs orthonormal bases for the nullspace (N-M dimensions) and the range (M dimensions) of the matrix, finding the least-squares best compromise solution.

REFERENCES

1. Grafman J. TMS as a primary brain mapping tool. In: George MS, Belmaker RH, eds. *Transcranial magnetic stimulation (TMS) in neuropsychiatry*. Washington, DC: American Psychiatric Press, 2000:115–140.

2. George MS, Belmaker RH. *Transcranial magnetic stimulation in neuropsychiatry*. Washington, DC: American Psychiatric Press, 2000.

3. George MS, Lisanby SH, Sackeim HA. Transcranial magnetic stimulation: applications in neuropsychiatry. *Arch Gen Psychiatry* 1999;56:300–311.

4. Barker AT, Jalinous R, Freeston IL. Non-invasive magnetic stimulation of the human motor cortex. *Lancet* 1985;1:1106–1107.

5. Saypol JM, Roth BJ, Cohen LG, et al. A theoretical comparison of electric and magnetic stimulation of the brain. *Ann Biomed Eng* 1991;19:317–328.

6. Roth BJ, Saypol JM, Hallett M, et al. A theoretical calculation of the electric field induced in the cortex during magnetic stimulation. *Electroencephalogr Clin Neurophysiol* 1991;81:47–56.

7. Bohning DE. Introduction and overview of TMS physics. In: George MS, Belmaker RH, eds. *Transcranial magnetic stimulation in neuropsychiatry*. Washington, DC: American Psychiatric Press, 2000:13–44.

8. Bohning DE, Pecheny AP, Epstein CM, et al. Mapping transcranial magnetic stimulation (TMS) fields *in vivo* with MRI. *Neuroreport* 1997;8:2535–2538.

9. Wassermann EM. Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the International Workshop in the Safety of Repetitive Transcranial Magnetic Stimulation, June 5–7, 1996. *Electroencephalogr Clin Neurophysiol* 1998;108:1–16.

10. Post RM, Kimbrell TA, Frye M, et al. Implications of kindling and quenching for the possible frequency dependence of rTMS. *CNS Spectrums: the International Journal of Neuropsychiatric Medicine* 1997;2:54–60.

11. Pascual-Leone A, Valls-Sole J, Wasserman EM, et al. Responses to rapid-rate transcranial magnetic stimulation of the human motor cortex. *Brain* 1994;117:847–858.

12. Wassermann EM, Wedegaertner FR, Ziemann U, et al. Crossed reduction of motor cortex excitability by 1 Hz transcranial magnetic stimulation. *Neurosci Lett* 1998;250:141–144.

13. Kimbrell TA, Little JT, Dunn RT, et al. Frequency dependence of antidepressant response to left prefrontal repetitive transcranial magnetic stimulation (rTMS) as a function of baseline cerebral glucose metabolism. *Biol Psychiatry* 1999;46:1603–1613.

14. Lorberbaum JP, Wassermann EM. Safety concerns of transcranial magnetic stimulation. In: George MS, Belmaker RH, eds. *Transcranial magnetic stimulation in neuropsychiatry*. Washington, DC: American Psychiatric Press, 2000:141–162.

15. Epstein CM, Lah JJ, Meador K, et al. Optimum stimulus parameters for lateralized suppression of speech with magnetic brain stimulation. *Neurology* 1996;47:1590–1593.

16. Fox PT, Mintun MA. Noninvasive functional brain mapping by change-distribution analysis of averaged PET images of $H_2{}^{15}O$ tissue activity. *J Nucl Med* 1989;30:141–149.

17. Bandettini PA, Jesmanowicz A, Wong EC, et al. Processing strategies for time-course data sets in functional MRI of the human brain. Magn Reson Med 1993;30:161–173.

18. Strother SC, Kanno I, Rottenberg DA. Principal component analysis, variance partitioning, and "functional connectivity." *J Cereb Blood Flow Metab* 1995;15:353–360.

19. Friston KJ, Frith CD, Liddle PF, et al. Functional connectivity: the principal-component analysis of large (PET) data sets. *J Cereb Blood Flow Metab* 1993;13:5–14.

20. Friston KJ. Functional and effective connectivity in neuroimaging: a synthesis. *Hum Brain Mapping* 1994;2:56–78.

21. McIntosh AR, Gonzalez-Lima F. Structural equation modeling and its application to network analysis in functional brain imaging. *Hum Brain Mapping* 1994;2:2–22.

22. Friston KJ, Worsley KJ, Frackowiak RSJ, et al. Assessing the significance of focal activations using their spatial extent. *Hum Brain Mapping* 1994;1:210–220.

23. Gerstein GL, Perkel DH. Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science* 1969;164:828–830.

24. Gerstein GL, Bedenbaugh P, Aertsen AM. Neuronal assemblies. *IEEE Trans Biomed Eng* 1989;36:4–14.

25. Friston KJ. Commentary and opinion: II. Statistical parametric mapping: ontology and current issues. *J Cereb Blood Flow Metab* 1995;15:361–370.

26. Adler RJ. *The geometry of random fields*. New York: John Wiley and Sons, 1981.

27. Gochin PM, Miller EK, Gross CG, et al. Functional interactions among neurons in inferior temporal cortex of the awake macaque. *Exp Brain Res* 1991;84:505–516.

28. Gevins A, Zeitlin GM, Yingling CD, et al. EEG patterns during cognitive tasks. I. Methodology and analysis of complex behaviors. *Electroencephalogr Clin Neurophysiol* 1979;47:693–703.

29. Gevins A, Doyle JC, Cuttilo BA, et al. Electrical potentials in human brain during cognition: a new method of reveals dynamic patterns of correlation. *Science* 1981;213:918–922.

30. Gevins A, Bressler SL, Morgan NH, et al. Event-related covariances during a bimanual visuomotor task. I. Methods and analysis of stimulus- and response-locked data. *Electroencephalogr Clin Neurophysiol* 1989;74:58–75.

31. Thatcher RW, Krause PI, Hrybyck M. Cortico-cortical associations and EEG coherence: a two-compartment model. *Electroencephalogr Clin Neurophysiol* 1986;64:123–143.

32. Tucker DM, Roth DL, Bair TB. Functional connections among cortical regions: topography of EEG coherence. *Electroencephalogr Clin Neurophysiol* 1986;63:242–250.

33. Friston KJ, Frith CD, Fletcher P, et al. Functional topography: multidimensional scaling and functional connectivity in the brain. *Cereb Cortex* 1996;6:156–164.

34. Paus T, Jech R, Thompson CJ, et al. Transcranial magnetic stimulation during positron emission tomography: a new method for studying connectivity of the human cerebral cortex. *J Neurosci* 1997;17:3178–3184.

35. Alexander GE, Moeller JR. Application of the scaled subprofile model to functional imaging in neuropsychiatric disorders: a principal component approach to modeling brain function in disease. *Hum Brain Mapping* 1994;2:79–94.

36. McIntosh AR, Grady CL, Ungerleider LG, et al. Network analysis of cortical visual pathways mapped with PET. *J Neurosci* 1994;14:655–666.

37. Friston KJ, Frith CD, Frackowiak RSJ. Time-dependent changes in effective connectivity measured with PET. *Hum Brain Mapping* 1993;1:69–79.

38. Aertsen AMH, Gerstein GL, Habib MK, et al. Dynamics of neuronal firing correlation: modulation of "effective connectivity." *J Neurophysiol* 1989;61:900–917.

39. Loehlin JC. *Latent variable models: an introduction to factor, path and structural analysis*. Hillside, NJ: Lawrence Erlbaum Associates, 1987.

40. Wassermann EM, Kimbrell TA, George MS, et al. Local and distant changes in cerebral glucose metabolism during repetitive transcranial magnetic stimulation (rTMS). *Neurology* 1997;48:A107-P02.049(abst).

41. Kimbrell TA, George MS, Danielson AL, et al. Changes in cerebral metabolism during transcranial magnetic stimulation. *Biol Psychiatry* 1997;41:108S-#374(abst).

42. George MS, Wassermann EM, Kimbrell T, et al. An overview of initial studies combining conventional functional imaging (PET, SPECT, fMRI) with transcranial magnetic stimulation (TMS) to actively probe brain-behavior relationships. *J Neuropsychiatry Clin Neurol* 1997;9:131.

43. Stallings LE, Speer AM, Spicer KM, et al. Combining SPECT and repetitive transcranial magnetic stimulation (rTMS)—left prefrontal stimulation decreases relative perfusion locally in a dose-dependent manner. *Neuroimage* 1997;5:S521(abst).

44. Chen W, Zhu X-H, Ogawa S, et al. Probing neural events by fMRI at the neural time scale of milliseconds. *Proceedings of the International Society of Magnetic Resonance in Medicine* 2000;8:#501(abst).

45. Cohen LG, Roth BJ, Nilsson J, et al. Effects of coil design on delivery of focal magnetic stimulation. Technical considerations. *Electroencephalogr Clin Neurol* 1990;75:350–357.

46. Wassermann EM, Wang B, Zeffiro TA, et al. Locating the motor cortex on the MRI with transcranial magnetic stimulation and PET. *Neuroimage* 1996;3:1–9.

47. Roberts DR, Vincent DJ, Speer AM, et al. Multi-modality mapping of motor cortex: comparing echoplanar BOLD fMRI and transcranial magnetic stimulation. *J Neural Transm* 1997;104:833–843.

48. Talairach J, Tournoux P. *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging*. New York: Thieme Medical Publishers, 1988.

49. Peters T, Davey B, Munger P, et al. Three-dimensional multi-modal image guidance for neurosurgery. *IEEE Trans Med Imaging* 1996;15:121–128.

50. Krings T, Buchbinder BR, Butler WE, et al. Transcranial magnetic stimulation and functional magnetic resonance imaging: complementary approaches in the evaluation of cortical motor function. *Neurology* 1997;48:1406–1416.

51. Bohning DE, Shastri A, Nahas Z, et al. Echoplanar BOLD fMRI of brain activation induced by concurrent transcranial magnetic stimulation (TMS). *Invest Radiol* 1998;33:336–340.

52. Bohning DE, Shastri A, McConnell K, et al. A combined TMS/fMRI study of intensity-dependent TMS over motor cortex. *Biol Psychiatry* 1999;45:385–394.

53. Shastri A, George MS, Bohning DE. Performance of a system for interleaving transcranial magnetic stimulation with steady state magnetic resonance imaging. *Electroencephalogr Clin Neurophysiol* 1999;51:55–64.

54. George MS, Wassermann EM, Williams WA, et al. Daily repetitive transcranial magnetic stimulation (rTMS) improves mood in depression. *Neuroreport* 1995;6:1853–1856.

55. George MS, Stallings LE, Speer AM, et al. Prefrontal repetitive transcranial magnetic stimulation (rTMS) changes relative perfusion locally and remotely. *Hum Psychopharmacol* 1999;14:161–170.

56. Josephs O, Athwal BS, Mackinnon C, et al. Transcranial magnetic stimulation with simultaneous undistorted functional magnetic resonance imaging. *Proceedings of the seventh annual meeting of the International Society of Magnetic Resonance in Medicine* 1999:1696.

57. Bohning DE, Shastri A, Wassermann EM, et al. BOLD-fMRI response to single-pulse transcranial magnetic stimulation (TMS). *J Magn Reson Imaging* 2000;11:569–574.

58. Fox P, Ingham R, George MS, et al. Imaging human intra-cerebral connectivity by PET during TMS. *Neuroreport* 1997;8:2787–2791.

59. Paus T, Jech R, Thompson CJ, et al. Dose-dependent reduction of cerebral blood flow during rapid-rate transcranial magnetic stimulation of the human sensorimotor cortex. *J Neurophysiol* 1997;79:1102–1107.

60. Nahas Z, Speer AM, Molloy M, et al. Preliminary results concerning the roles of frequency and intensity in the antidepressant effect of daily left prefrontal rTMS. *Biol Psychiatry* 1998;43:94s-#315(abst).

61. Nahas Z, Lomarev M, Roberts DR, et al. Unilateral left prefrontal transcranial magnetic stimulation produces intensity-dependent bilateral effects as measured with interleaved BOLD fMRI. *Biol Psychiatry (in press)*.

62. Ilmoniemi RJ, Virtanen J, Ruohonen J, et al. Neuronal response to magnetic stimulation reveals cortical reactivity and connectivity. *Neuroreport* 1997;8:3537–3540.

63. Kosslyn SM, Pascual-Leone A, Felician O, et al. The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 1999;284:167–170.

64. Desmurget M, Epstein CM, Turner RS, et al. Role of the posterior parietal cortex in updating reaching movements to a visual target. *Nat Neurosci* 1999;2:563–567.

65. Tofts PS. The distribution of induced currents in magnetic stimulation of the nervous system. *Phys Med Biol* 1990;35:1119–1128.

66. Branson NM, Tofts PS. Analysis of the distribution of current induced by a changing magnetic field in a volume conductor. *Phys Med Biol* 1991;36:161–168.

67. Cohen LG, Cuffin BN. Developing a more focal magnetic stimulator. Part I: some basic principles. *J Clin Neurophysiol* 1991;8:102–111.

68. Roth BJ, Saypol JM, Hallett M. A theoretical calculation of the electric field induced in the cortex during magnetic stimulation. *Electroencephalogr Clin Neurophysiol* 1991;81:47–56.

69. Cerri G, DeLeo R, Moglie F, et al. An accurate 3-D model for magnetic stimulation of the brain cortex. *J Med Eng Technol* 1995;19:7–16.

70. Press WH, Flannery BP, Teukolsky SA, et al. *Numerical recipes in C: the art of scientific computing*. New York: Cambridge University Press, 1986.

71. Bohning DE, Shastri A, McGavin L, et al. Brain activity for transcranial magnetic stimulation (TMS) induced and volitional movement are similar in location and level. *Invest Radiol* 2000;35:676–683.

31

In Vivo Molecular Imaging: Ligand Development and Research Applications

Masahiro Fumita

Robert B. Innis

Masahiro Fumita: Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.

Robert B. Innis: Departments of Psychiatry and Pharmacology, Yale University School of Medicine, New Haven, Connecticut.

In positron emission tomography (PET) and single-photon emission computed tomography (SPECT), tracers labeled with radioactive isotopes are used to measure protein molecules (e.g., receptors, transporters, and enzymes). A major advantage of these two radiotracer techniques is extraordinarily high sensitivity (~ 10^{-9} to 10^{-12} *M*), many orders of magnitude greater than the sensitivities available with magnetic resonance imaging (MRI) (~ 10^{-4} *M*) or magnetic resonance spectroscopy (MRS) (~ 10^{-3} to 10^{-5} *M*). For example, MRI detection of gadolinium occurs at concentrations of approximately 10^{-4} *M* (1), and MRS measures brain levels of γ-aminobutyric acid (GABA) and glutamine at concentrations of approximately 10^{-3} *M* (2 ,3). In contrast, PET studies with [^{11}C]NNC 756 in which a conventional bismuth germanate-based scintillator is used can measure extrastriatal dopamine D1 receptors present at a concentration of approximately 10^{-9} *M* (4). Because many molecules of relevance to neuropsychiatric disorders are present at concentrations of less than 10^{-8} *M*, radiotracer imaging is the only currently available *in vivo* method capable of quantifying these molecular targets.

PET and SPECT quantify the distribution of radioactivities in the brain, the direct *in vivo* correlates of *in vitro* autoradiographic film techniques such as receptor autoradiography, Western blots, and Northern blots. Thus, the future possibilities of radiotracer imaging are broad and exciting—and include targets of receptors, signal transduction, and gene expression. From this broader perspective, PET and SPECT methodologies are described as "*in vivo* molecular imaging."

Although *in vivo* molecular imaging is a promising technique, several barriers— physical, monetary, and chemical—to its successful application in neuropsychiatric disorders must be addressed. Physical barriers include limited anatomic resolution and the need for even higher sensitivity. However, recent developments with improved detector crystals (e.g., lutetium oxyorthosilicate) and three-dimensional image acquisition have markedly enhanced both sensitivity and resolution. (5). Commercially available PET devices provide resolution of 2 to 2.5 mm (6 ,7). Furthermore, the relatively high cost of imaging with SPECT, and especially PET, can be partially subsidized by clinical use of the devices. Recent approval of U.S. government (i.e., Medicare) reimbursement of selected PET studies for patients with tumors, epilepsy, and cardiac disease has significantly enhanced the sales of PET cameras and their availability for partial use in research studies. Thus, the major barriers for the expanded use of PET are not physical or monetary, but rather chemical in nature. Simply stated, the major barrier to radiotracer imaging of molecular targets may well be the difficulties associated with developing the radiotracers themselves. Labeling the appropriate precursor typically is not the major impediment. Almost all candidate ligands contain carbon and hydrogen, and the positron-emitting nuclides ^{11}C and ^{18}F can usually be incorporated as an isotopic variant (^{11}C for ^{12}C) or an atomic substitute (^{18}F for ^{1}H). As described in the next section, the most common obstacle to the development of *in vivo* tracers is the relatively small window of appropriate combinations of lipophilicity, molecular weight, and affinity. For a molecule to pass the blood–brain barrier, relatively small molecular weights, less than 400 to 600 daltons (d), and moderate lipophilicity are required. However, in brain, relatively low lipophilicity and high affinity are required to achieve high ratios of specific to nonspecific binding. In contrast, the requirements for *in vitro* ligands are much less stringent because the blood–brain barrier can be removed by homogenization or tissue sectioning and most nonspecific binding washed away. The special requirements of *in vivo* probes (low molecular weight, high affinity, and just the right amount of lipophilicity) are discussed in the next section.

REQUIRED PROPERTIES OF AN *IN VIVO* TRACER

In addition to a high degree of affinity and selectivity, several other properties are required for useful *in vivo* tracers.

Combination of Small Molecular Weight, Appropriate Lipophilicity, and High Affinity

Tissue uptake of a drug (whether radioactive or not) is often analyzed within the theoretic framework of a *compartment*, which is defined as a space in which the concentration of a drug is uniform. Within brain tissue, the time-dependent concentration of drug is described for at least three compartments—free [$C_{2f}(t)$], nonspecifically bound [$C_{2ns}(t)$], and specifically bound [$C_3(t)$] radiotracer (Fig. 31.1). Free and nonspecifically bound ligand, $C_{2f}(t)$ and $C_{2ns}(t)$, cannot be washed away, the way they are in *in vitro* studies. Therefore, a high ratio of specific to nondisplaceable uptake in brain (C_3/C_2) is required to obtain reliable data; to reduce the number of unknown variables, $C_{2f}(t)$ and $C_{2ns}(t)$ are often combined in a single compartment and defined as a compartment, $C_2(t)$) (Fig. 31.2).

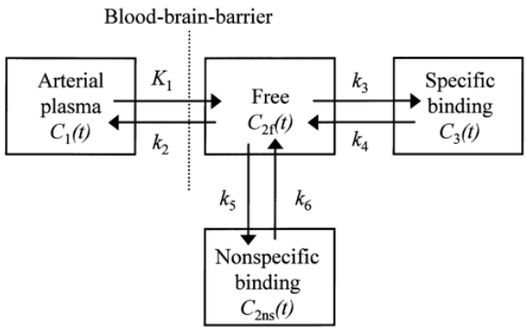


FIGURE 31.1. Compartmental description of radioactively labeled tracer

uptake in brain.

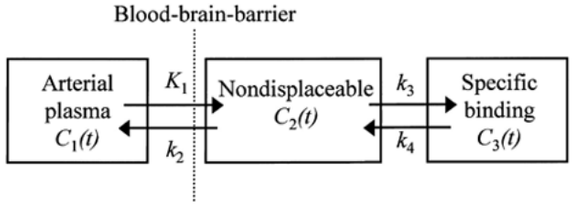


FIGURE 31.2. Three-compartment model (two tissues). To reduce the number of variables, $C_{2f}(t)$ and $C_{2ns}(t)$ in Fig. 31.1 are combined in a single compartment, $C_2(t)$.

Three major factors (i.e., lipophilicity, molecular weight, and affinity) determine the *in vivo* characteristics of a tracer. It is easy to understand that small molecular weight and a high degree of lipophilicity are required to pass the blood–brain barrier because it is composed of a lipid bilayer. However, lipophilicity has opposing effects on the brain uptake of a tracer. Increasing lipophilicity enhances the permeability of the compound, but it also tends to increase plasma protein binding, thereby decreasing the concentration of free drug available to cross the blood–brain barrier. From rat experiments in which 27 tracers of various chemical classes were used, Levin (8) obtained the following simple equation to derive a capillary permeability coefficient, P_c, from an octanol/water partition coefficient, P, and the molecular weight, M_r.

$$\text{Log } P_c = -4.605 + 0.4115 \bullet \log[P(M_r)^{-1/2}]$$

This equation was obtained for the tracers with molecular weights of less than 400 and relatively low lipophilicity (average log P, -0.34; standard deviation, 2.3). Because a higher lipophilicity is required for brain imaging tracers to be taken up adequately in brain, the equation only partially characterizes capillary permeability. High lipophilicity (higher value of log P) increases the binding to the plasma components (protein and cell membranes) (9) and reduces the capillary permeability coefficient expressed relative to the total plasma concentration of drug (10 ,11). Therefore, both low lipophilicity and high lipophilicity decrease the penetration of imaging agents across the blood–brain barrier, so that a parabolic curve is created (Fig. 31.3).

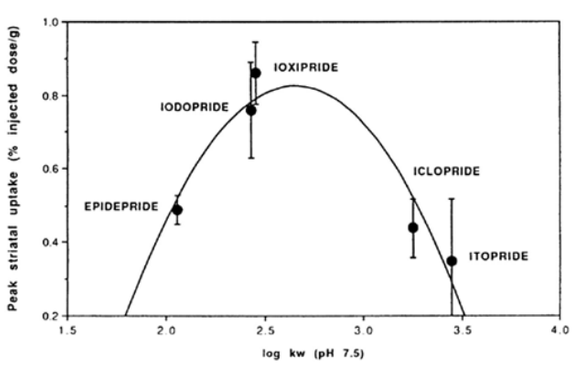


FIGURE 31.3. Relationship between apparent lipophilicity (log K_W) at pH 7.5 and peak striatal uptake (percentage of administered dose per gram of tissue in rat brain). Peak uptake occurred between log K_W of 2.4 and 2.8. Each point is the mean of four animals ± standard deviation. (From Kessler RM, Ansari MS, de Paulis T, et al. High-affinity dopamine D2 receptor radioligands. 1. Regional rat brain distribution of iodinated benzamides. *J Nucl Med* 1991;32:1593–1600, with permission.)

Although low lipophilicity decreases nonspecific binding in brain, it also decreases blood–brain barrier permeability (11). For any particular chemical series, optimal parameters

of affinity and lipophilicity generate a tracer with the best “signal-to-noise” ratio to measure the target molecule, as shown in Fig. 31.3 for a series of benzamide ligands for the dopamine D2 receptor. It should be noted that iodination for SPECT tracers usually increases lipophilicity in addition to molecular weight. Therefore, depending on the position in the optimization curve (Fig. 31.3), either an iodinated or a non-iodinated compound may show the most desirable *in vivo* properties.

Rapid Labeling of Precursor

The radionuclide must be incorporated quickly into appropriate precursor molecules because of the relatively short half-lives ($t_{1/2}$) of the isotopes (e.g., ^{11}C, 20.4 minutes; ^{18}F, 110 minutes; ^{123}I, 13.2 hours). Therefore, precursors must be available that allow quick labeling in one (but usually no more than two) synthetic steps just before the imaging study is performed.

Appropriate Clearance from Specific Binding Compartment

Following the bolus injection of radioactive tracer, the time–activity curve of an organ (e.g., brain) and its subcomponents (e.g., striatum) is characterized by uptake (rising) and washout (declining) phases. If the uptake phase is slow relative to the $t_{1/2}$ of the radionuclide, reasonably accurate data may be acquired only for the rising portion of the time–activity curve. Although parameters related to receptor density and affinity can be derived for selected targets with only the uptake portion, most quantitative methods calculate such parameters with both uptake and washout phases of the tissue time–activity curve. Thus, the tissue clearance of the tracer must typically be matched with the $t_{1/2}$ of the radionuclide. For example, ^{123}I can be used to quantify tracers with much slower tissue kinetics than ^{11}C. The rate of tissue clearance is in part determined by the affinity of the tracer, with ligands of higher affinity tending to “stick” longer to the target molecules. So, as with lipophilicity, the affinity of the candidate tracer should be high enough that significant tissue uptake occurs, but it should not be so high that washout is delayed beyond the usable measurement time of the radionuclide.

In summary, a low molecular weight is almost always mandatory, at least for tracers that cross the blood–brain barrier via passive diffusion. Lipophilicity should be high enough to allow adequate permeability of blood–brain barrier, but not so high as to cause unacceptable binding to plasma proteins or high levels of nonspecific binding in brain. Finally, the affinity of the tracer must balance the opposing goals of tight binding and fast washout from the brain. That is, a high affinity ligand is needed to provide high levels of tight binding of the ligand to the preceptor. However, if the binding is of such high affinity that the ligand shows negligible washout from the brain during the course of a typical study, then the washout rate cannot be determined and critical kinetic data are unavailable to calculate receptor levels in the brain. Such parameters are not required for therapeutic agents (in which fast uptake may not be even helpful) or for *in vitro* tracers (in which most nonspecific binding can be washed away). For example, the nondisplaceable uptake of [^{18}F]haloperidol (12) and [^{11}C]cocaine (13) is unacceptably high for optimal PET imaging of their molecular targets, dopamine D2 receptor and dopamine transporter, respectively, although they can provide valuable data about the disposition of the psychoactive drugs themselves. It should also be noted that desirable properties of radioactive tracers are usually different from those of therapeutic agents. For, example, slow clearance of haloperidol from brain (12 ,14) may maintain effective receptor occupancy for a long period of time. If the distribution in the nonspecific binding compartment [$C_{2ns}(t)$ in Fig. 31.1] is large, the compartment may act as a reservoir of the drug and provide stable levels in brain. The stringent requirements for an optimal radioactive tracer easily explain why only a tiny percentage of *in vitro* tracers and therapeutic agents are useful as *in vivo* imaging ligands.

Negligible Influence of Radioactively Labeled Lipophilic Metabolites

If the parent tracer generates lipophilic radioactive metabolites, they may enter the brain in significant concentration and confound the imaging study. If they do not bind to the molecular target (inactive metabolites), they will increase nonspecific binding [$C_{2ns}(t)$ in Fig. 31.1] and thereby decrease the signal-to-noise measurement of the target. On the other hand, if the radioactive metabolites are active (i.e., bind to the target), quantification is highly confounded because the measured signal represents undetermined proportions of parent tracer and metabolite, each of which may have a different affinity for the target. The problem of lipophilic radioactive metabolites may sometimes be avoided by appropriate selection of the labeling position. For example, a tracer for the serotonin 5-HT_{1A} receptor WAY 100635 can be labeled with ^{11}C at either an external (*O*-methyl) or internal (carbonyl) position. [*O*-methyl-^{11}C]WAY 100635 is rapidly metabolized in humans (15). The metabolic cleavage of WAY 100635 generates two moieties; the ^{11}C-containing methyl component is lipophilic and enters the brain, but the ^{11}C-containing carbonyl component is polar and does not enter the brain. The internal carbonyl labeling is more difficult than the external *O*-methyl labeling, but this clever radiochemistry has markedly improved the signal-to-noise measurement of 5-HT_{1A} receptors in brain (16).

Many tracers currently used for imaging studies produce at least somewhat lipophilic metabolites. However, the quantities produced or their kinetics in passing blood–brain barrier are such that they do not commonly confound the measurements. For example, if the uptake and washout of the parent tracer are fast relative to the production of radioactive

metabolites, then their component of the total measured activity may be negligible during the imaging study.

IN VIVO QUANTIFICATION OF TRACER UPTAKE

Part of "31 - Molecular Imaging: Ligand Development and Research Applications "

In vivo quantification of molecular targets with radiotracer imaging is much more complicated than *in vitro* measurements for several reasons: (a) For *in vivo* experiments, tracers are intravenously administered and not directly applied to the target tissue. Therefore, the delivery of a tracer to the target tissue is influenced by its peripheral clearance (i.e., metabolism and excretion). (b) Total tissue activity is measured, and the separate specifically bound, nonspecifically bound, and free components (Fig. 31.1) are usually estimated with relatively complicated mathematical analyses. (c) The spatial resolution of cameras is limited, and the activity in a region of interest is influenced by tracer uptake in adjacent areas. The details of *in vivo* quantification are beyond the scope of this chapter, and interested readers should refer to other sources (e.g., refs. 17 ,18). The following section provides a simplified overview of the typical parameters and methods of quantification used in radiotracer imaging.

Binding Potential

In addition to having appropriate chemical and physical characteristics, a useful ligand must provide imaging results that are "amenable to quantification" because analogue/visual images are likely to be of limited utility in neuropsychiatric research. The most commonly measured receptor parameter is the binding potential (*BP*), which equals the product of receptor density (B_{max}) and affinity ($1/K_d$, where K_d is the dissociation binding constant). Thus, increased uptake could reflect either an increased number of receptors or the same number of receptors, each of which has a higher affinity for the ligand. To measure both B_{max} and K_d separately, at least two experiments or two injections are necessary, in which formulations of the tracer with both high and low specific activity are used (19 ,20). Because the injection of a tracer with low specific activity (i.e., high mass dose) causes significant occupancy of the molecular target, the potential pharmacologic effects of the tracer must be both safe and acceptable within the experimental paradigm. If these studies are not performed, B_{max} and K_d cannot be measured separately, and only their ratio ($BP = B_{max}/K_d$) is used as the outcome measure (21).

Methods to Measure Binding Potential

In vivo quantification has followed the well-established Law of Mass Action applied to ligand–receptor interactions under equilibrium conditions:

$$\frac{B}{F} = \frac{B_{max}}{K_d + F} \qquad [1]$$

where *B* is the concentration of radiotracer bound to the receptor and *F* is the concentration of free radiotracer (i.e., not bound to proteins) in the vicinity of the receptor. Because radiotracer imaging typically involves the injection of a miniscule mass dose of ligand, the concentration of free radiotracer is quite low. That is, $F<<K_d$, with the result that

$$\frac{B}{F} = \frac{B_{max}}{K_d} \qquad [2]$$

Thus, *BP* can be simply estimated as the equilibrium ratio of bound (*B*) to free (*F*). With this fairly standard three-compartment model (i.e., two tissue compartments and one blood compartment; Fig. 31.2), *B* is denoted as C_3 and *F* as $f_1{*}C_1$. Thus, $B/F = C_3/(f_1{*}C_1)$. This equation makes the reasonable assumption for drugs that pass blood–brain barrier by passive diffusion that the concentration of free tracer in plasma ($f_1{*}C_1$) equals that in brain under equilibrium conditions. Note that Eqs. 1 and 2 refer to *equilibrium* binding conditions. Following the bolus injection of tracer, the ratio of receptor-bound tracer (*B*) and free tracer (*F*) changes dramatically and is not under equilibrium conditions. Figure 31.4 provides a schematic representation of radiotracer concentrations in brain and plasma following a bolus injection with Fig. 4A showing early time points and Fig. 4B showing the entire data. With use of the complete time–activity curves from brain (measured with a PET or a SPECT camera) and arterial plasma (directly sampled and measured in a gamma counter), the goal of compartmental modeling is to estimate the ratio of *B* and *F* under *equilibrium* conditions. In other words, if the free level could be maintained at a constant level, how many times higher than the free level (*F*) would the concentration of receptor-bound tracer (*B*) finally and stably be?

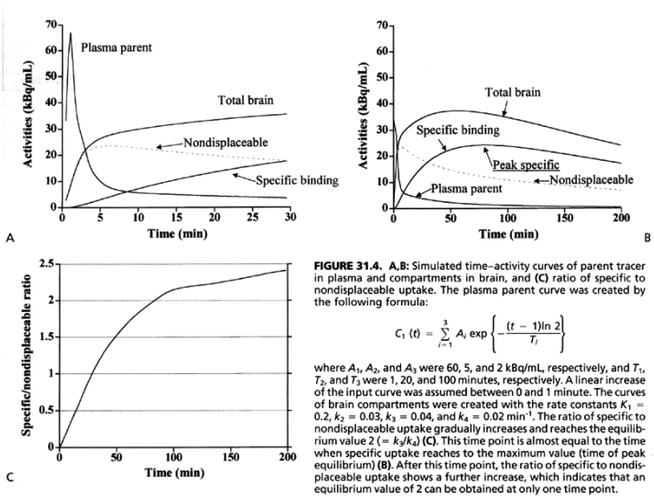


FIGURE 31.4. A,B: Simulated time–activity curves of parent tracer in plasma and compartments in brain, and **(C)** ratio of specific to nondisplaceable uptake. The plasma parent curve was created by the following formula:

$$C_1(t) = \sum_{i=1}^{3} A_i \exp\left\{-\frac{(t-1)\ln 2}{T_i}\right\}$$

where A_1, A_2, and A_3 were 60, 5, and 2 kBq/mL, respectively, and T_1, T_2, and T_3 were 1, 20, and 100 minutes, respectively. A linear increase of the input curve was assumed between 0 and 1 minute. The curves of brain compartments were created with the rate constants K_1 = 0.2, k_2 = 0.03, k_3 = 0.04, and k_4 = 0.02 min^{-1}. The ratio of specific to nondisplaceable uptake gradually increases and reaches the equilibrium value 2 (= k_3/k_4) **(C)**. This time point is almost equal to the time when specific uptake reaches to the maximum value (time of peak equilibrium) **(B)**. After this time point, the ratio of specific to nondisplaceable uptake shows a further increase, which indicates that an equilibrium value of 2 can be obtained at only one time point.

FIGURE 31.4. A,B: Simulated time–activity curves of parent tracer in plasma and compartments in brain, and **(C)** ratio of specific to nondisplaceable uptake. The plasma parent curve was created by the following formula:

$$C_1(t) = \sum_{i=1}^{3} A_i \exp\left\{-\frac{(t-1)\ln 2}{T_i}\right\}$$

where A_1, A_2, and A_3 were 60, 5, and 2 kBq/mL, respectively, and T_1, T_2, and T_3 were 1, 20, and 100 minutes, respectively. A linear increase of the input curve was assumed between 0 and 1 minute. The curves of brain compartments were created with the rate constants K_1 = 0.2, k_2 = 0.03, k_3 = 0.04, and k_4 = 0.02 min^{-1}. The ratio of specific to nondisplaceable uptake gradually increases and reaches the equilibrium value 2 (= k_3/k_4) **(C)**. This time point is almost equal to the time when specific uptake reaches to the maximum value (time of peak equilibrium) **(B)**. After this time point, the ratio of specific to nondisplaceable uptake shows a further increase, which indicates that an equilibrium value of 2 can be obtained at only one time point.

Several methods to estimate receptor parameters have been applied in radiotracer imaging and are briefly summarized below.

1. Compartmental modeling. Often viewed as the "gold standard," compartmental modeling typically requires concurrent and lengthy measurements of parent compound (separated from radioactively labeled metabolites) in plasma and of the brain time–activity curve. Kinetic parameters (K_1, k_2, k_3, k_4) are estimated from this so-called arterial input function (i.e., C_1) and the "impulse response function" of the brain (i.e., sum of C_2 and C_3). The goal of compartmental modeling is to determine the values of the rate constants between these compartments (Fig. 31.2), which when applied to the measured values of C_1 generate a brain time–activity curve similar to that actually measured with the PET or SPECT camera. The underlying concept is that the equilibrium ratio of *B* and *F* is equal to a ratio of kinetic rate constants.

$$BP = \frac{B_{max}}{K_d} = \frac{K_1 \bullet k_3}{k_2 \bullet k_4 \bullet f_1}$$

The major disadvantage of kinetic analysis of compartmental

modeling may often be logistic in nature because arterial sampling may be poorly tolerated, measurement of parent radiotracer in plasma may be technically difficult, and prolonged periods of data acquisition may be needed for both plasma and brain activities. The subsequent methods were developed in large part as “simpler” techniques to obtain receptor measurements that closely correlate with or directly equal those obtained in a complete compartmental analysis.

2. Peak equilibrium method. Based on good theoretic grounds, the “peak equilibrium method” [commonly attributed to Farde et al. (19) for radiotracer imaging] selects a unique equilibrium time as that when specific binding achieves its maximal level. Specific binding is operationally defined as activity in a target region (e.g., striatum) minus that in a background tissue region (e.g., cerebellum). From a practical perspective, a subject is continuously imaged for an hour or two following a bolus injection of tracer. Activities in target and background regions are plotted, and specific binding is calculated at each time point as the difference of the two curves (Fig. 31.4B). At the time of peak specific activity, a measure proportional to binding potential is calculated as specific/nondisplaceable = (target - background)/background, as in (striatum-cerebellum)/cerebellum for a D2 tracer.

One major advantage of the peak equilibrium method is that plasma measurements are not required. Its major limitations include the following:

(a) Background activity in brain is proportional but not equal to the free level of tracer (*F*). Thus, the outcome measure is not equal to *BP*; rather this ratio of specific to nondisplaceable uptake is proportional to *BP*.

(b) Nonspecific binding in the background region may not equal that in the target region, a situation that can be caused, for example, by different proportions of white and gray matter in the two regions. The assumption of equivalent nonspecific binding has been evaluated (usually in animals) with displacement of radiotracer binding by high doses of a nonradioactive drug that also binds to the site. The assumption would be supported by equivalent levels of residual activities in target and background regions with complete receptor blockade.

(c) The time curves of free levels in target and background regions would be predicted not to overlap exactly.

For example, during the uptake portion of the curve, the free level would be lower in the target region because both receptors and nonspecific sites bind the tracer. Thus, at the time of peak specific uptake, the free level in the target region would not be the same as that in the background region. Depending on the kinetics of specific and nonspecific binding, the resulting discrepancy might be significant.

3. Constant infusion methods. Stable levels of drugs, including radiotracers, can be achieved with constant (sometimes called *continuous*) intravenous infusion of the drug (Fig. 31.5A). At some point, typically referred to as *steady state*, the amount of drug entering the blood will equal that leaving the vascular compartment. The levels of both total and free drug in plasma will subsequently be stable. At a somewhat later time point, the amount of drug binding to a receptor in an organ will equal that coming off the receptor; the level of receptor-bound drug will subsequently be stable. In an analogy to *in vitro* receptor binding studies, this stable condition is a state of equilibrium receptor binding.

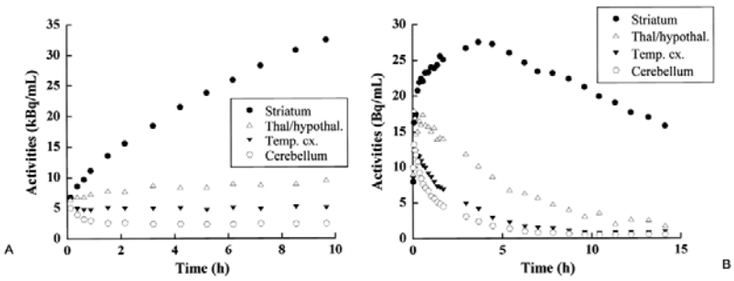


FIGURE 31.5. Brain time–activity curves in a bolus plus constant infusion/equilibrium **(A)** and a bolus/kinetic study **(B)** of [^{123}I]epidepride. **A:** [^{123}I]Epidepride was given as a bolus (145 MBq) followed by constant infusion with bolus/infusion ratio of 6.0 hours (the amount of the tracer given for 6 hours by constant infusion was equal to that of the bolus) in a 33-year-old man. Note that equilibrium was achieved only in low-density regions (thalamus/hypothalamus and temporal cortex), not in a high-density region (striatum), with this bolus/infusion ratio. To achieve equilibrium in all regions, including striatum, a higher bolus/infusion ratio and longer infusion are required. **B:** [^{123}I]Epidepride (371 MBq) was give as a bolus to a 24-year-old man. (Part **B** reprinted from Fujita M, et al. Kinetic and equilibrium analyses of [^{123}I]epidepride binding. *Synapse* 1999;34:290–304, with permission.)

The constant infusion (or so-called equilibrium) method is computationally much simpler than compartmental modeling and does not require extensive brain imaging or arterial plasma measurements. The concentration of receptor-bound tracer (*B*) can be estimated as target minus background. The level of free tracer in plasma (*F*) can be measured in either venous or arterial plasma because the body as a whole is in a condition of steady state. From a practical perspective, the subject is connected to an intravenous pump for several hours, and then a single image is acquired (e.g., 30-minute duration) and a single venous blood sample is obtained. The major disadvantage of this technique is that many hours of infusion may be required to achieve steady-state conditions in both plasma and brain. In addition, data are acquired during a relatively short interval (e.g., 30 minutes) of a long infusion period (e.g., 7 hours). Thus, activity measurements before and after the relatively brief acquisition are not used, and the resulting radiation exposure to the subject can be viewed as "wasteful." In contrast, "all" activity is measured and used in the analysis of a bolus/kinetic study (Fig. 31.5B). From a practical perspective, the total amount of injected activity is often fairly equivalent for a bolus/kinetic and a constant infusion/equilibrium study. However, the total activity in brain after many hours of decay may be quite low and cause statistical counting errors.

In summary, three basic methods can be used to estimate receptor binding potential: (a) compartmental analysis of a bolus/kinetic study, (b) peak equilibrium, and (c) equilibrium imaging following constant infusion of the tracer (with or without an initial bolus of tracer). For all three methods, the target parameter is typically B_{max}/K_d, which equals the equilibrium value of B/F under tracer occupancy conditions (i.e., < 10% of the receptor occupied by tracer). The "true" binding potential (B_{max}/K_d) can be calculated if the free concentration of tracer in plasma is measured with the assumption

that free concentration in plasma equals that in brain. In this case, the measurement of *BP* is the ratio of receptor-bound activity in brain to free plasma activity (*F*). Because a blood sample is not used in the peak equilibrium method, the true *BP* cannot be measured. An alternate outcome measure for each of these three methods uses the nondisplaceable activity in a background region of brain as a value proportional to free tracer concentration. This so-called poor man's binding potential is a ratio of activities in different regions of the brain (specific to nondisplaceable), and therefore plasma measurements are not required.

ESTIMATION OF ENDOGENOUS NEUROTRANSMITTER LEVELS

Part of "31 - Molecular Imaging: Ligand Development and Research Applications "

Molecular imaging probes can be used not only to measure a specific target but also, under appropriate conditions, to estimate concentrations of endogenous compounds that compete with the tracer for binding to the target. For example, a D2-receptor probe can be used not only to measure D2 receptors but also the extent of competition of this binding caused by endogenous dopamine. In fact, the most extensively studied indirect measurements have been the interaction of dopamine with D2-receptor ligands. These studies are briefly reviewed, and the special characteristic of a good ligand for such indirect measurements are discussed.

Tonic and Phasic Release of Dopamine

Dopamine transmission in striatum is thought to occur in two different modes, tonic and phasic (22 ,23). Tonic dopamine release represents the steady-state level of dopamine in the extracellular space, which is estimated to be in the nanomolar range. On the other hand, in phasic release, high extracellular concentrations of dopamine (millimolar range) are released within or near a synapse during an action potential. Close relationships have been proposed between abnormalities in phasic and tonic dopamine release and the symptoms of schizophrenia. Namely, excessive phasic release causes psychosis, and decreased tonic release causes cognitive deficits and negative symptoms (24).

Phasic release has typically been initiated by intravenous administration of a stimulant such as amphetamine or methylphenidate. These agents elevate synaptic dopamine concentrations either by releasing dopamine in a reverse manner via a dopamine transporter (amphetamine) or by blocking dopamine transporter-mediated reuptake of dopamine (methylphenidate). In an imaging study, the elevation of synaptic dopamine levels is estimated by the decrease in D2 radiotracer binding following stimulant administration in comparison with control conditions. Just as careful quantification is required for direct radiotracer binding to a molecular target, a similar if not even more rigorous approach is required for these indirect methods to ensure that measurements of both the tracer and the competing displacer represent “equilibrium” values and not just transient pharmacokinetic “artifacts.” To support the validity of these measurements, microdialysis studies have been performed in conjunction with D2-receptor imaging and a stimulant challenge (25 ,26). Although D2-ligand displacement correlated with the increase in extracellular dopamine measured with microdialysis, the relative increase in extracellular dopamine (1,000% to 4,000%) was much greater than the percentage of displacement of the ligand binding (5% to 15%). Thus, D2-receptor displacement is a “low-gain” monitor (i.e., underestimation) of the increase in extracellular dopamine. The reasons that the changes in binding are so much lower (although still, it is hoped, linear) relative to the increase in extracellular dopamine are unclear. This “low-gain” monitoring may reflect the fact that the imaging measurements cannot temporally track and therefore lag behind the chemical changes they are designed to measure. In other words, the displacement of radiotracer from the brain region occurs over a much slower time course than the relatively rapid changes in extracellular dopamine. Nevertheless, these stimulant-induced displacement studies appear to provide some reflection of changes in synaptic dopamine levels because they are relatively well correlated and because depletion of tissues levels of dopamine can block the effects of amphetamine (27).

The tonic release of dopamine has been estimated by the increase in D2-receptor binding induced by the depletion of endogenous dopamine. The removal of dopamine “unmasks” receptors, which then become available for radiotracer binding. The percentage of unmasking reflects the percentage of D2 receptors occupied by dopamine under basal or tonic conditions. Dopamine depletion has been induced in both animals and humans, with a resulting increase in D2 radiotracer binding (28 ,29). One limitation of these studies, especially in humans, is the difficulty of knowing whether depletion is essentially complete, so that the full extent of dopamine occupancy of the receptor has been measured. For example, if differences in unmasking are found in two subjects, does that reflect different levels of endogenous dopamine—or just different levels of dopamine depletion? A second limitation of this depletion paradigm is that increased receptor binding may not reflect “unmasking” but rather an up-regulation of D2-receptor levels, similar to that often found in denervation supersensitivity. One mechanism to minimize this potential confound is to perform the measurements as soon after dopamine depletion as possible. However, one clear advantage of the depletion paradigm in comparison with the stimulant-induced increase is that the depleted levels can typically be stably maintained during the scan. Thus, the relative slowness of the imaging measurements does not present a pharmacokinetic confound, as it does in studies with stimulant-induced release of dopamine.

Both stimulant and depletion studies have been performed in patients with schizophrenia. In general agreement

with the hypothesis of Grace, the “phasic” increase in synaptic dopamine (assessed following amphetamine administration) has been found to be elevated in drug-free schizophrenic patients, and the decrease in D2 radiotracer binding correlates with a transient increase in psychotic symptoms (30 ,31 and 32). In addition, stimulant depletion studies have found greater unmasking of striatal D2 receptors in patients with schizophrenia, which suggests that basal/tonic synaptic dopamine levels are higher in this disorder (33).

Affinity of Radiotracer

The relationship between affinity of the radiotracer and the sensitivity of its binding to endogenous dopamine is a source of confusion. Under *in vitro*equilibrium conditions and at tracer levels of radioactively labeled ligand, both the Michaelis–Menten and Cheng–Prusoff equations predict that the percentage of receptor occupancy by a competitive inhibitor (e.g., dopamine or other neurotransmitters) depends on the affinity of the inhibitor for the receptor and is independent of the affinity of the radioactively labeled ligand. However, such equilibrium binding conditions are achieved for neither the tracer nor the displacer if each is injected as a bolus. Even under these conditions, the sensitivity of radioligand binding to endogenous dopamine levels is theoretically (at least based on the *in vitro* theories) independent of the affinity of the radioactively labeled ligand when both the tracer and the displacer have achieved equilibrium binding conditions. However, if either the radiotracer (as in the bolus injection paradigm) or endogenous dopamine (as in stimulant-induced release) changes dynamically over time, the equilibrium condition is not achieved, and the apparent sensitivity of the radioligand to endogenous dopamine levels is determined by the kinetic properties of the radioligand (34 ,35). Equilibrium conditions can be achieved for both tracer and displacer in the dopamine depletion paradigm. For example, equilibrium can be achieved with bolus plus constant infusion of the radiotracer, and stable dopamine depletion with AMPT (α-methyl-*p*-tyrosine). The high-affinity D2 radioligand [^{123}I]epidepride provides an instructive example of the differences seen in kinetic and equilibrium studies. The kinetics of its uptake in brain are slow and do not show displacement by transiently increased dopamine levels induced with amphetamine (36). However, stable low levels of dopamine induced with AMPT show unmasking of D2 receptors (37).

In Vivo Confounding Factors

Although the displacement of radioligand binding by neurotransmitter can be simply described with *in vitro* tissue homogenates, several factors complicate the interpretation of *in vivo* experimental results. (a) The affinity states of some receptors for agonists, but not antagonists, is regulated by the receptor to guanyl nucleotide-binding proteins (38). Therefore, *in vivo* measurements are influenced by the affinity states and agonist or antagonist properties of the radiolabeled tracer, and results obtained with agonist and antagonist tracers may be different. (b) Agonist binding typically causes receptor internalization of G protein-coupled receptors (39 ,40). Thus, the endogenous agonist dopamine presumably facilitates the intracellular trafficking of D2 receptors (41), and radiotracers may differ in their affinity for the internalized versus membrane-bound receptor. (c) A significant proportion of D2 receptors are located extrasynaptically (42 ,43), where the concentration of dopamine is lower than in the synapse. Thus, neurotransmitters may occupy a smaller percentage of extrasynaptic receptors than of receptors within the synapse, and the *in vivo* measurement may not truly reflect synaptic neurotransmitter levels.

INTERACTION AMONG NEUROTRANSMISSION SYSTEMS

Part of "31 - Molecular Imaging: Ligand Development and Research Applications "

Abnormalities in psychiatric disorders likely represent the complex interaction of several neurotransmitter systems in the brain. PET imaging has recently been used to examine aspects of neurotransmitter interactions. For example, Dewy and colleagues (44 ,45 and 46) have pioneered studies on interactions among dopamine, GABA, and acetylcholine (ACh) systems in striatum. GABA neurons in the striatum have inhibitory effects on nigral dopamine neurons, nigral dopamine neurons have inhibitory effects on striatal ACh neurons, and striatal ACh neurons have facilitating effects on striatal GABA neurons. By estimating dopamine levels in striatum as described above, Dewey and collaborators showed in human or anesthetized nonhuman primates that the blockade of cholinergic transmission by benztropine (44) or scopolamine (45) decreased [^{11}C]raclopride binding (increase in dopamine levels) and that stimulation of GABAergic transmission by γ-vinyl-GABA (a suicide inhibitor of GABA transaminase) and lorazepam (a benzodiazepine agonist) increased [^{11}C]raclopride binding (decrease in dopamine levels) (46). In addition, they showed that a dopamine antagonist, *N*-methylspiroperidol, induced a decrease in [*N*-^{11}C-methyl]benztropine binding, indicating an increase in ACh levels (44).

Other interactions have also been studied with PET. In two human studies, an *N*-methyl-D-aspartate (NMDA) antagonist, ketamine, decreased [^{11}C]raclopride binding in striatum (47 ,48). In two human studies with similar techniques, the binding of [^{11}C]raclopride was decreased by stimulation of 5-hydroxytryptamine (5-HT) transmission with fenfluramine (a 5-HT releaser) (49) or psilocybin (a mixed 5-HT_{2A} and 5-HT_{1A} agonist) (50). However, these results are discordant with those of previous studies in baboon, in which citalopram (a 5-HT uptake inhibitor) increased [^{11}C]raclopride binding (51). Key aspects of the interaction between dopamine and 5-HT neurotransmitter

systems may well be mediated by glutamate. One significant limitation of these studies is that no useful glutamatergic PET probes have been developed to examine this important mediating neurotransmitter system. Furthermore, the linkage of pharmacologic challenges can be difficult to interpret. For example, if a disorder is associated with an abnormal dopamine outcome measured with PET in response to a 5-HT challenge, is the abnormal response caused by altered sensitivity of the dopamine or 5-HT system?

The results of these studies of interactions among neurotransmission systems have been interpreted under a simple assumption that the binding of [^{11}C]raclopride and other tracers is affected by synaptic neurotransmitter levels. This simple assumption has been questioned by elaborate studies by Tsukada et al. (52). They measured dopamine synthesis, dopamine transporter, and dopamine D2 receptor with L-[β-^{11}C]methyldopa (L-[β-^{11}C]DOPA), [^{11}C]β-CIT, and [^{11}C]raclopride, respectively, in combination with microdialysis in conscious rhesus monkeys. Scopolamine did not change extracellular dopamine levels in the striatum but increased [^{11}C]raclopride binding by decreasing its affinity at the dopamine D2 receptor (52) Furthermore, ketamine decreased [^{11}C]raclopride binding in the striatum without increasing extracellular dopamine levels, and it increased both dopamine synthesis and dopamine availability (53).

Although interpretation may be difficult, and although the pharmacokinetics of either the tracer or displacer and changes in the synthesis and reuptake of neurotransmitters and affinity of receptor binding may complicate the experiment, the authors feel that challenges linked with radiotracer imaging are likely to provide useful information to allow a better understanding of the pathophysiology of neuropsychiatric disorders.

USE OF RADIOTRACER IMAGING IN THERAPEUTIC DRUG DEVELOPMENT

Part of "31 - Molecular Imaging: Ligand Development and Research Applications "

Radiotracer imaging can provide useful information about molecules that are either the direct target of or indirect markers for the effects of therapeutic drugs. For example, if both the tracer and therapeutic drug competitively bind to the same target, then imaging can provide direct information on the extent and kinetics of receptor occupancy in the relevant tissue. In addition, the molecular target measured by the tracer (e.g., amyloid) can be used as a “surrogate” biological marker to assess the efficacy of the therapeutic agent (e.g., one designed to decrease amyloid deposition).

Measurement of Receptor Occupancy

Molecular imaging can provide useful guidance for two aspects of drug administration: dose and dosing interval. The dose is most easily chosen with a known target occupancy and accepted range. For example, typical antipsychotic agents occupy 50% to 80% of striatal D2 receptors (54). Thus, a new typical antipsychotic agent should show a reasonably acceptable side effect profile when given at doses that occupy this range of D2 receptors. With regard to dosing interval, the drug may be retained in tissue much longer than in plasma, and, therefore dosing intervals based on plasma pharmacokinetics may be too frequent. Such a situation has clearly been shown for antipsychotic agents, with which a high rate of receptor occupancy extends well beyond low plasma levels (55 ,56).

The measurement of receptor occupancy and pharmacokinetics in brain combined with the evaluation of adverse reactions in a small number of healthy subjects may provide valuable information for “go/no go” decisions at early stages of clinical drug development. If targeted receptor occupancy is achieved without causing adverse reactions, studies in patients are justified. If not, further studies may not be indicated. Even in the absence of a target level for receptor occupancy, it is reasonably safe to assume that doses associated with greater than 95% occupancy are unnecessarily high, and that those with less than 10% occupancy are unlikely to be efficacious.

The 5-HT$_{2A}$ antagonist M100907 may provide the best example of the use of PET receptor occupancy to provide useful information on dose and dosing interval. By measuring receptor occupancy with [^{11}C]*N*-methylspiperone in healthy human subjects, initially an appropriate amount of a single dose (57) and then an appropriate dose and dosing interval were determined (56). Further, a similar level of receptor occupancy was recently confirmed with [^{11}C]M100907 in a small number of patients with schizophrenia (58).

Dopamine Transporter Imaging as a Biological Marker in Parkinson Disease

Imaging of a biological marker may provide information that is useful either for diagnosis or as a monitor of disease progression. In Parkinson disease, the two most successful imaging targets used as biological markers are measures of dopamine synthesis with [^{18}F]FDOPA and of dopamine terminal innervation with ligands for dopamine transporter.

The symptoms of Parkinson disease are caused by the progressive loss of dopamine neurons in the nigrostriatal pathway. Therefore, a reasonable method of diagnosis would be to “count” the number of dopamine neurons noninvasively. Such measurements are possible with *in vivo* imaging tracers: [^{18}F]FDOPA and various SPECT/PET tracers for dopamine transporter. However, these tracers do not detect the same biological process. Whereas [^{18}F]FDOPA detects metabolic activities at dopamine nerve terminals, the tracers for dopamine transporter simply measure the density of the target protein. Because of this difference, the sensitivity to detect the decrease in dopamine neurons

may be different. In fact, both human and animal studies have indicated that the imaging of dopamine transporter is more sensitive to dopamine neuronal loss (59). Dopamine transporter can be quantified with SPECT tracers, which can be used with lower cost than can PET studies. Therefore, imaging of the dopamine transporter in movement disorders is widely performed in many developed countries.

It is clinically important but sometimes difficult to differentiate essential tremor and Parkinson disease. Dopamine transporter imaging clearly distinguishes these two groups, with only a small overlap (60 ,61). However, dopamine transporter imaging is not able to distinguish idiopathic Parkinson disease from other parkinsonian syndromes, such as multiple system atrophy (62). Further, these techniques have shown bilateral loss of dopamine transporter in hemi-Parkinson disease (i.e., the earliest stage of the disorder), which suggests that even preclinical disease may be detected (63).

Restoring dopamine levels with L-DOPA is still the core medication treatment of Parkinson disease. However, long-term treatment with L-DOPA frequently results in fading of the therapeutic effect and the development of serious motor and psychiatric side effects. Although palliative treatment with L-DOPA is clearly of significant clinical benefit, current drug development is oriented toward neuroprotective treatments designed to slow the loss of dopamine neurons and the consequent progression of symptoms. As a biological marker of dopamine terminal innervation of the striatum, dopamine transporter imaging may well be useful to monitor whether such new therapies actually slow the loss of dopamine neurons. In fact, SPECT imaging is sensitive to and can quantify dopamine transporter loss in longitudinal studies of patients with Parkinson patients treated with conventional therapies (64 ,65 and 66). Therefore, with appropriate sample sizes, this imaging technique can quantify a relevant biological marker as a surrogate measure of the efficacy of putative neuroprotective therapies.

IMAGING POST-RECEPTOR SIGNAL TRANSDUCTION

Part of "31 - Molecular Imaging: Ligand Development and Research Applications "

The majority of the imaging studies performed to date have focused on the synapse: transporters as presynaptic targets, receptors as postsynaptic targets, and indirect measurements of the transmitter as a type of “intrasynaptic” target. However, measurements of membrane receptors merely “scratch the surface” of the cell and ignore the multitude of important intracellular mechanisms required for biological effects. Therefore, two promising areas for expansion in the near future are post-receptor signal transduction and subsequent changes in gene expression. Abnormalities in second messenger systems have been postulated to play important pathophysiologic roles in many psychiatric illnesses, including mood disorders (67) and drug addiction (68). However, only a small number of PET and SPECT studies have been performed on these systems. Although tracers have been developed to image three major biochemical cascades [i.e., cyclic adenosine monophosphate (cAMP), phosphoinositide (PI), and arachidonate pathways], all existing ligands have moderate to significant technical limitations, and better ones are clearly needed. Some of these tracers are lipids, and their nonspecific binding is too high. In addition, because most of these tracers do not bind to a single type of protein but are metabolized by several enzymes, the interpretation of the results is not clear.

cAMP Cascade

Imaging of this signal transduction system was initially attempted by labeling its activator, forskolin, and a related analogue, 1-acetyl-7-deacetylforskolin, with ^{11}C (69 ,70). The binding of [^{11}C]forskolin may be correlated with the activation of adenylate cyclase (71). However, brain uptake of these tracers is very low (69).

Recently, imaging of this cascade was tried with an inhibitor of phosphodiesterase IV (PDE-IV), [^{11}C]rolipram (72). PDE-IV is the major subtype in brain of PDE, which hydrolyzes cAMP into 5′-AMP. PDE-IV is composed of four major enzyme subtypes, PDE-IV A, B, C, and D, and all four subtypes are found in human brain (73). Rolipram binds to and inhibits all four PDE-IV subtypes with high affinity. In one report of a rat *ex vivo* study, [^{11}C]rolipram was a promising tracer exhibiting high specific brain uptake (72). Further evaluation of the binding selectivity and kinetics must be performed in nonhuman primates and humans subjects.

Phosphoinositide System

Imahori and colleagues have studied the PI system *in vivo* using ^{11}C-labeled (74) and ^{123}I-labeled (75) 1,2-diacylglycerol. They showed specific incorporation of the tracer in the chemical components of the rat PI system, such as phosphatidic acid, phosphatidylinositol, phosphatidylinositol 4-phosphate, and phosphatidylinositol 4,5-*bis*-phosphate (74). Further, the tracer uptake was increased by an agonist at the muscarinic ACh receptor, which is known to be coupled with the PI system (74). However, absolute quantification of the tracer uptake is probably difficult because of the high lipophilicity. As explained above (“Required Properties of an *In Vivo* Tracer”), high lipophilicity causes a high level of binding to plasma components (protein and cell membranes), which reduces delivery of the tracer into brain. Under these circumstances, intersubject variability in the amount of tracer delivered to brain may be primarily determined by its binding to plasma components and not by neuronal activity of the PI system. If intersubject variability in f_1 (the free fraction of tracer in plasma) is noted, tracer

uptake cannot be compared among different subjects. Tracers with high lipophilicity and high binding to plasma components enter the brain slowly, and [^{11}C]diacylglycerol did indeed show such behavior (76). Further, whatever amount of a highly lipophilic tracer actually crosses the blood–brain barrier tends to exhibit a high rate of nonspecific binding. In summary, because of the difficulty in absolute quantification, the utility of this tracer as a quantitative measure may be significantly limited.

Arachidonate

The utility of [^{11}C]arachidonate to detect *in vivo* activity of phospholipase A_2 has been studied rigorously. After intravenous injection, [^{11}C]arachidonate is readily taken up (“pulse labeling”) and incorporated into cerebral phospholipids and other stable brain compartments (77). By stimulating receptors that are linked to phospholipase A_2, probably via the PI pathway, the labeled phospholipids are catalyzed to generate arachidonate, and the regionally localized enhancement of phospholipid turnover increases the uptake of [^{11}C]arachidonate. The cholinomimetic arecoline has been shown to increase [^{14}C]arachidonate levels, and this effect can be blocked with a muscarinic ACh-receptor antagonist (78). For the quantification of the activity of this signal transduction system, the measurement should not be affected by cerebral blood flow. Chang et al. (79) showed that the uptake of [^{11}C]arachidonate is relatively flow-independent in monkeys. As with radioactively labeled diacylglycerol, high lipophilicity may be a significant limitation in absolute quantification. The potential dependence of brain uptake on the plasma free fraction may preclude between-subject studies, although within-subject experimental designs may still be valid (e.g., before and after pharmacologic challenge).

IMAGING REGULATION OF GENE EXPRESSION

Part of "31 - Molecular Imaging: Ligand Development and Research Applications "

Signal transduction initiated with presynaptic firing does not terminate with the interaction of a transmitter with its receptors and consequent second messenger generation. Instead, signal transduction may extend to the nucleus with alterations in gene expression. A well-known example is the induction of the protooncogene c-*fos* by receptor–ligand interactions (80). In fact, oncogenes encoding growth factors, membrane receptors, cytoplasmic and membrane-associated protein kinases, guanosine triphosphate-binding proteins (GTP), and transcription factors play important roles in signal transduction and altered gene expression. These genes and their cognate proteins will be important future targets for brain imaging. Much of what we know about these proteins in the central nervous system is derived from studies of cancer biology. Thus, imaging oncogenes and their protein products is also an exciting target for nuclear oncologists, in part because they are pharmacophores for drug development. Several modalities are applied in cancer imaging, including (a) imaging oncogene products with radioactively labeled antibodies, (b) imaging messenger RNAs with labeled antisense oligonucleotides (81), (c) imaging reporter gene products with labeled reporter probes (82 ,83), and (d) applying conventional techniques with labeled small molecules that bind to particular oncogene products.

Because the blood–brain barrier presents a special obstacle in neuroimaging, most techniques successfully used in cancer imaging are difficult to apply in brain imaging. For this reason, the first approach with antibodies is not possible in brain imaging unless the integrity of the blood–brain barrier is disrupted, as in the case of brain tumors. The second approach, in which antisense oligonucleotides are used, is also difficult because these multiply charged compounds do not cross the blood–brain barrier in any appreciable amount. To make radioactively labeled oligonucleotide probes pass the blood–brain barrier, complicated techniques are required, including the utilization of receptor-mediated transport (e.g., insulin and insulin-like growth factor) (84).

The third technique, the use of reporter probes, is being rigorously pursued, with possible applications in gene therapies. The best imaging example is the use of labeled thymidine analogues (e.g., 5-iodo-2′-fluoro-2′-deoxy-1-ß-D-arabino-furanosyl-uracil) in cells expressing herpes simplex thymidine kinase. The probe is phosphorylated by the viral but not by mammalian thymidine kinases and is thereby trapped within the cell, as in the brain uptake of 2-deoxyglucose analogues (85). The basic idea is that gene expression can be monitored by radiolabeled substrate (“reporter probe”), which is metabolized by a transfected gene (“reporter gene”) product and trapped in cells but not metabolized by endogenous enzymes of the host (82 ,83). A reporter gene can be different from a therapeutic gene as long as parallel levels of expression are expected by sharing a common promoter.

At the moment, imaging of reporter probes is used to detect expression only of exogenously introduced genes. Endogenous gene expression, which is interesting in psychiatric research, could be studied by using transgenes containing endogenous promoters fused to a reporter gene (83). A limitation of these techniques in brain imaging is that the widely used reporter probes, the radiolabeled substrates of herpes simplex type 1 thymidine kinase, do not show good permeability of the blood–brain barrier. This limitation can be overcome by using dopamine D2 receptor as a reporter gene and D2 ligands as reporter probes (86 ,87). Because the expression of functional D2 receptors may cause unwanted effects, further studies are being performed on the use of D2-receptor mutants, which are not coupled with intracellular signaling but still maintain binding affinity for D2 ligands.

Parkinson disease provides a useful example for gene therapy of a neuropsychiatric disorder (88). The concept of gene therapy for Parkinson disease has grown directly from the promising results obtained by grafting fetal dopamine-producing neurons. However, the limited availability and ethically controversial nature of the tissue source have restricted the utility of fetal grafts in this disorder. As an alternative, a relatively unlimited supply of homogenous, well-characterized viral vectors could theoretically be produced to deliver tyrosine hydroxylase, the rate-limiting enzyme in dopamine synthesis. Attempts have also been made in animal models to deliver neuroprotective or neurotrophic factors, such as superoxide dismutase and glial cell line-derived neurotrophic factor, to prevent continued degeneration of dopamine neurons. Similar techniques of gene therapy have been investigated in motor neuron degenerative diseases and Alzheimer disease. Reporter genes whose probes can cross the blood–brain barrier, such as D2 receptors, can monitor the expression of these transfected genes. On the other hand, dopamine release from the grafts could be monitored by a conventional technique utilizing competition of radioligand binding to D2 receptors, as described in the section on estimation of endogenous neurotransmitter levels. In fact, this technique of receptor displacement has been used to detect dopamine release in patients with embryonic nigral transplants (89).

The fourth approach, the relatively conventional one of imaging with small probes for relevant gene or oncogene protein products, is hampered by the development of useful and selective probes. As described in the section on the required properties of an *in vivo* tracer, it is difficult to fulfill all the requirements for a successful brain-imaging agent. However, once good imaging agents are developed, these targets can in general be imaged without the complicated techniques required in the other three modalities, described above. Many new anticancer agents are being developed, and a significant number of these agents target signal transduction systems, which may also play pathophysiologic roles in psychiatric disorders (90 ,91). For example, Ras farnesyltransferase is a target for cancer chemotherapy and potentially also for brain imaging. After post-translational modifications, including farnesylation, Ras binds to the cell membrane and transmits signals. Many inhibitors of Ras farnesyltransferase have been developed as anticancer medications (92). Among them, a recently developed agent has a high affinity (93) and may be used as a template from which brain imaging agents can be developed. In addition, a small molecule ligand for epidermal growth factor receptor has been labeled with ^{11}C and has shown brain uptake (94).

Rapid developments in molecular biology and the advent of gene-targeting techniques have enabled the study of individual genes in mice by means of *in vitro* experimental techniques. Recently developed animal-dedicated PET devices (e.g., “rat PET” and “microPET”) achieve high resolution of about 2 mm and can now image these animals*in vivo* (6). These imaging studies may make it possible to apply new findings in molecular biology to the study of patients with neuropsychiatric disorders in exciting ways.

CONCLUSIONS

Part of "31 - Molecular Imaging: Ligand Development and Research Applications "

Progress in molecular neurobiology has dramatically changed our understanding of psychiatric disorders. A significant proportion of these findings have been obtained from animal experiments and postmortem human studies. A major challenge for neuroimaging in future years will be to extended the application of these *in vitro* probes to *in vivo* imaging of patients. Radiotracer imaging with PET, and to a lesser extent with SPECT, is ideally suited for such *in vivo* applications because of its extraordinarily high sensitivity and improving anatomic resolution (now about 2 mm). This chapter has reviewed what is arguably the most difficult barrier to accomplishing *in vivo* molecular imaging—the development of useful and quantifiable tracers. The blood–brain barrier is a challenge to both the delivery of radiolabeled tracers and the quantification of brain uptake of the tracers. However, many successful tracers have been developed to date. These probes have largely been synthesized as analogues of agents active at synaptic transmission and have provided measures of presynaptic, postsynaptic, and even “intrasynaptic” targets. Relatively little progress has been made in measuring intracellular signal transduction or gene expression. These two areas are clearly important targets for future ligand development. By bridging new findings in molecular neuroscience and clinical studies, *in vivo* molecular imaging will likely contribute to a greater understanding of psychiatric pathophysiology and, it is hoped, enhance the development of improved pharmacotherapies.

REFERENCES

1. Nunn AD, Linder KE, Tweedle MF. Can receptors be imaged with MRI agents? *Q J Nucl Med* 1997;41:155–162.

2. Rothman DL, Petroff OA, Behar KL, et al. Localized 1H NMR measurements of gamma-aminobutyric acid in human brain *in vivo. Proc Natl Acad Sci U S A* 1993;90:5662–5666.

3. Michaelis T, Merboldt KD, Bruhn H, et al. Absolute concentrations of metabolites in the adult human brain *in vivo*: quantification of localized proton MR spectra. *Radiology* 1993;187:219–227.

4. Abi-Dargham A, Simpson N, Kegeles L, et al. PET studies of binding competition between endogenous dopamine and the D1 radiotracer [^{11}C]NNC 756. *Synapse* 1999;32:93–109.

5. Bendriem B, Townsend DW. Volume imaging tomographs. In: Bendriem B, Townsend DW, eds. *The theory and practice of 3D PET*. Dordrecht: Kluwer Academic, 1998;111–132.

6. Chatziioannou AF, Cherry SR, Shao Y, et al. Performance evaluation of microPET: a high-resolution lutetium oxyorthosilicate PET scanner for animal imaging. *J Nucl Med* 1999;40:1164–1175.

7. Schmand M, Eriksson L, Casey M, et al. Performance results of a new DOI detector block for a high resolution PET–LSO research tomograph HRRT. *IEEE Trans Nucl Sci* 1998;45:3000–3006.

8. Levin VA. Relationship of octanol/water partition coefficient and molecular weight to rat brain capillary permeability. *J Med Chem* 1980;23:682–684.

9. Rowley M, Kulagowski JJ, Watt AP, et al. Effect of plasma protein binding on *in vivo* activity and brain penetration of glycine/NMDA receptor antagonists. *J Med Chem* 1997;40:4053–4068.

10. Dishino DD, Welch MJ, Kilbourn MR, et al. Relationship between lipophilicity and brain extraction of C-11-labeled radiopharmaceuticals. *J Nucl Med* 1983;24:1030–1038.

11. Kessler RM, Ansari MS, de Paulis T, et al. High-affinity dopamine D2 receptor radioligands. 1. Regional rat brain distribution of iodinated benzamides. *J Nucl Med* 1991;32:1593–1600.

12. Schlyer DJ, Volkow ND, Fowler JS, et al. Regional distribution and kinetics of haloperidol binding in human brain: a PET study with [^{18}F]haloperidol. *Synapse* 1992;11:10–19.

13. Fowler JS, Volkow ND, Wolf AP, et al. Mapping cocaine binding sites in human and baboon brain *in vivo. Synapse*1989;4:371–377.

14. Kornhuber J, Schultz A, Wiltfang J, et al. Persistence of haloperidol in human brain tissue. *Am J Psychiatry* 1999;156:885–890.

15. Osman S, Lundkvist C, Pike VW, et al. Radioactive metabolites of the 5-HT$_{1A}$ receptor radioligand, [*O*-methyl-^{11}C]WAY-100635, in rat, monkey and humans plus evaluation of the brain uptake of the metabolite, [*O*-methyl-^{11}C]WAY-100634, in monkey. *J Label Compound Radiopharm* 1995;37:283.

16. Pike VW, McCarron JA, Lammerstma AA, et al. Exquisite delineation of 5-HT$_{1A}$ receptors in human brain with PET and [carbonyl-^{11}C]WAY-100635. *Eur J Pharmacol* 1996;301:R5–R7.

17. Huang S-C, Phelps ME. Principles of tracer kinetic modeling in positron emission tomography and autoradiography. In: Phelps M, Mazziotta J, Schelbert H, eds. *Positron emission tomography and autoradiography: principles and applications for the brain and heart*. New York: Raven Press, 1986:287–346.

18. Carson RE. Parameter estimation in positron emission tomography. In: Phelps ME, Mazziotta JC, Schelbert HR, eds. *Positron emission tomography and autoradiography: principles and applications for the brain and heart*. New York: Raven Press, 1986:347–390.

19. Farde L, Eriksson L, Blomquist G, et al. Kinetic analysis of central [^{11}C]-raclopride binding to D2 dopamine receptors studied with PET—a comparison to the equilibrium analysis. *J Cereb Blood Flow Metab* 1989;9:696–708.

20. Delforge J, Pappata S, Millet P, et al. Quantification of benzodiazepine receptors in human brain using PET, [^{11}C]flumazenil, and a single-experiment protocol. *J Cereb Blood Flow Metab* 1995;15:284–300.

21. Mintun MA, Raichle ME, Kilbourn MR, et al. A quantitative model for the *in vivo* assessment of drug binding sites with positron emission tomography. *Ann Neurol* 1984;15: 217–227.

22. Grace AA. Phasic versus tonic dopamine release and modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neurosci* 1991;41:1–24.

23. Grace AA. Cortical regulation of subcortical systems and its possible relevance to schizophrenia. *J Neural Transm [Gen Sect]* 1993;91:111–134.

24. Moore H, West AR, Grace AA. The regulation of forebrain dopamine transmission: relevance to the pathophysiology and psychopathology of schizophrenia. *Biol Psychiatry* 1999;46:40–55.

25. Endres CJ, Kolachana BS, Saunders RC, et al. Kinetic modeling of [^{11}C]raclopride: combined PET–microdialysis studies. *J Cereb Blood Flow Metab* 1997;17:932–942.

26. Laruelle M, Iyer RN, al-Tikriti MS, et al. Microdialysis and SPECT measurements of amphetamine-induced dopamine release in nonhuman primates. *Synapse* 1997;25:1–14.

27. Innis RB, Malison RT, al-Tikriti M, et al. Amphetamine-stimulated dopamine release competes *in vivo* for [^{123}I]IBZM binding to the D2 receptor in nonhuman primates. *Synapse* 1992;0:177–184.

28. Ross SB, Jackson DM. Kinetic properties of the accumulation of ^{3}H-raclopride in the mouse brain *in vivo. Naunyn Schmiedebergs Arch Pharmacol*1989;340:6–12.

29. Laruelle M, D'Souza C, Baldwin RM, et al. Imaging D2 receptor occupancy by endogenous dopamine in humans. *Neuropsychopharmacology* 1997;17:162–174.

30. Laruelle M, Abi-Dargham A, van Dyck CH, et al. SPECT imaging of amphetamine-induced dopamine release in drug-free schizophrenic subjects. *Proc Natl Acad Sci U S A* 1996;93:9235–9240.

31. Breier A, Su TP, Saunders R, et al. Schizophrenia is associated with elevated amphetamine-induced synaptic dopamine concentrations: evidence from a novel positron emission tomography method. *Proc Natl Acad Sci U S A* 1997;94:2569–2574.

32. Abi-Dargham A, Gil R, Krystal J, et al. Increased striatal dopamine transmission in schizophrenia: confirmation in a second cohort. *Am J Psychiatry* 1998;155:761–767.

33. Abi-Dargham A, Kegeles L, Zea-Ponce Y, et al. Removal of endogenous dopamine reveals elevation of D2 receptors in schizophrenia. *J Nucl Med* 1999;40:30P(abst).

34. Morris ED, Fisher RE, Alpert NM, et al. *In vivo* imaging of neuromodulation using positron emission tomography: optimal ligand characteristics and task length for detection of activation. *Hum Brain Mapping* 1995;3:35–55.

35. Endres CJ, Carson RE. Assessment of dynamic neurotransmitter changes with bolus or infusion delivery on neuroreceptor ligands. *J Cereb Blood Flow Metab* 1998;18:1196–1210.

36. Al-Tikriti MS, Baldwin RM, Zea-Ponce Y, et al. Comparison of three high-affinity SPECT radiotracers for the dopamine D2 receptor. *Nucl Med Biol* 1994;21:179–188.

37. Fujita M, Verhoeff NPLG, Varrone A, et al. Imaging extrastriatal dopamine D2 receptor occupancy by endogenous dopamine in healthy humans. *Eur J Pharmacol* 2000;387:179–188.

38. Wreggett KA, Seemn P. Agonist high- and low-affinity states of the D2 dopamine receptor in calf brain. *Mol Pharmacol* 1983;25:10–17.

39. Dumartin B, Caillé I, Gonon F, et al. Internalization of D1 dopamine receptor in striatal neurons *in vivo* as evidence of activation by dopamine agonists. *J Neurosci* 1998;18:1650–1661.

40. Ito K, Haga T, Lameh J, et al. Sequestration of dopamine D2 receptors depends on coexpression of G protein-coupled receptor kinases 2 or 5. *Eur J Biochem* 1999;260:112–119.

41. Chugani DC, Ackermann RF, Phelps ME. *In vivo* [^{3}H]spiperone binding: evidence for accumulation in corpus striatum by agonist-mediated receptor internalization. *J Cereb Blood Flow Metab* 1988;8:291–303.

42. Yung KK, Bolam JP, Smith AD, et al. Immunocytochemical localization of D1 and D2 dopamine receptors in the basal ganglia of the rat: light and electron microscopy. *Neuroscience* 1995;65:709–730.

43. Khan ZU, Gutierrez A, Martin R, et al. Differential regional and cellular distribution of dopamine D2-like receptors: an immunocytochemical study of subtype-specific antibodies in rat and human brain. *J Comp Neurol* 1998;402:353–371.

44. Dewey SL, Brodie JD, Fowler JS, et al. Positron emission tomography (PET) studies of dopaminergic/cholinergic interactions in the baboon brain. *Synapse* 1990;6:321–327.

45. Dewey SL, Smith GS, Logan J, et al. Effects of central cholinergic blockade on striatal dopamine release measured with positron emission tomography in normal human subjects. *Proc Natl Acad Sci USA* 1993;90:11816–11820.

46. Dewey SL, Smith GS, Logan J, et al. GABAergic inhibition of endogenous dopamine release measured *in vivo* with ^{11}C-raclopride and positron emission tomography. *J Neurosci* 1992;12:3773–3780.

47. Breier A, Adler CM, Weisenfeld N, et al. Effects of NMDA antagonism on striatal dopamine release in healthy subjects: application of a novel PET approach. *Synapse* 1998;29:142–147.

48. Smith GS, Schloesser R, Brodie JD, et al. Glutamate modulation of dopamine measured *in vivo* with positron emission tomography (PET) and ^{11}C-raclopride in normal human subjects. *Neuropsychopharmacology* 1998;18:18–25.

49. Smith GS, Dewey SL, Brodie JD, et al. Serotoninergic modulation of dopamine measured with [^{11}C]raclopride and PET in normal human subjects. *Am J Psychiatry* 1997;154:490–496.

50. Vollenweider FX, Vontobel P, Hell D, et al. 5-HT modulation of dopamine release in basal ganglia in psilocybin-induced psychosis in man—a PET study with [^{11}C]raclopride. *Neuropsychopharmacology* 1999;20:424–433.

51. Dewey SL, Smith GS, Logan J, et al. Serotoninergic modulation of striatal dopamine measured with positron emission tomography (PET) and *in vivo* microdialysis. *J Neurosci* 1995;15:821–829.

52. Tsukada H, Harada N, Nishiyama S, et al. Cholinergic neuronal modulation alters dopamine D2 receptor availability *in vivo* by regulating receptor affinity induced by facilitated synaptic dopamine turnover: positron emission tomography studies with microdialysis in the conscious monkey brain. *J Neurosci* 2000;20:7067–7073.

53. Tsukada H, Harada N, Nishiyama S, et al. Ketamine decreased striatal [^{11}C]raclopride binding with no alterations in static dopamine concentrations in the striatal extracellular fluid in the monkey brain: multiparametric PET studies combined with microdialysis analysis. *Synapse* 2000;37:95–103.

54. Farde L, Nordstrom A-L, Wiesel F-A, et al. PET analysis of central D1 and D2 dopamine receptor occupancy in patients treated with classic neuroleptics and clozapine—relationship to extrapyramidal side effects. *Arch Gen Psychiatry* 1992;49:538–544.

55. Nyberg S, Farde L, Halldin C. Delayed normalization of central D2 dopamine receptor availability after discontinuation of haloperidol decanoate. Preliminary findings. *Arch Gen Psychiatry* 1997;54:953–958.

56. Grunder G, Yokoi F, Offord SJ, et al. Time course of 5-HT_{2A} receptor occupancy in the human brain after a single oral dose of the putative antipsychotic drug MDL 100,907 measured by positron emission tomography [published erratum appears in *Neuropsychopharmacology* 1998;19:161]. *Neuropsychopharmacology* 1997;17:175–185.

57. Andree B, Nyberg S, Ito H, et al. Positron emission tomographic analysis of dose-dependent MDL 100,907 binding to 5-hydroxytryptamine-2A receptors in the human brain. *J Clin Psychopharmacol* 1998;18:317–323.

58. Talvik-Lotfi M, Nyberg S, Nordström A-L, et al. High $5HT_{2A}$ receptor occupancy in M100907-treated schizophrenic patients. *Psychopharmacology* 2000;148:400–403.

59. Stoof JC, Winogrodzka A, van Muiswinkel FL, et al. Leads for the development of neuroprotective treatment in Parkinson's disease and brain imaging methods for estimating treatment efficacy. *Eur J Pharmacol* 1999;375:75–86.

60. Asenbaum S, Pirker W, Angelberger P, et al. [^{123}I]beta-CIT and SPECT in essential tremor and Parkinson's disease. *J Neural Transm* 1998;105:1213–1228.

61. Lee MS, Kim YD, Im JH, et al. ^{123}I-IPT brain SPECT study in essential tremor and Parkinson's disease. *Neurology* 1999;52:1422–1426.

62. Varrone A, Marek KL, Jennings D, et al. [^{123}I]β-CIT SPECT imaging demonstrates reduced density of striatal dopamine transporters in Parkinson's disease and multiple system atrophy. *Mov Disord (in press)*.

63. Booij J, Tissingh G, Winogrodzka A, et al. Imaging of the dopaminergic neurotransmission system using single-photon emission tomography and positron emission tomography in patients with parkinsonism. *Eur J Nucl Med* 1999;26:171–182.

64. Booij J, Winogrodzka A, Bergmans P, et al. [I-123]-beta-CIT and [I-123]FP-CIT SPECT are useful methods to monitor progression of dopaminergic degeneration in early-stage Parkinson's disease. *J Nucl Med* 1999;40:28P(abst).

65. Seibyl JP, Innis RB, Early ML, et al. Baseline striatal dopamine transporter uptake measured with [I-123]β-CIT SPECT may predict the rate of disease progression in idiopathic Parkinson's disease. *J Nucl Med* 1999;40:27P(abst).

66. Marek K, Innis R, van Dyck C, et al. [^{123}I]β-CIT SPECT imaging assessment of the rate of Parkinson's disease progression. *Neurology (in press)*.

67. Duman RS, Heninger GR, Nestler EJ. A molecular and cellular theory of depression. *Arch Gen Psychiatry* 1997;54:597–606.

68. Nestler EJ. Molecular neurobiology of drug addiction. *Neuropsychopharmacology* 1994;11:77–87.

69. Sasaki T, Enta A, Nozaki T, et al. Carbon-11-forskolin: a ligand for visualization of the adenylate cyclase-related second messenger system. *J Nucl Med* 1993;34:1944–1948.

70. Sasaki T, Furukata K, Iimori T, et al. 1-Acetyl-7-deacetylforskolin: a potential non-specific inactive analog of forskolin for estimation of its specific high-affinity binding and adenylyl cyclase stimulation *in vitro*. *Life Sci* 1995;57:1367–1373.

71. Nelson CA, Seamon KB. Binding of [^{3}H]forskolin to solubilized preparations of adenylate cyclase. *Life Sci* 1988;42:1375–1383.

72. Lourenco CM, DaSilva JN, Warsh JJ, et al. Imaging of cAMP-specific phosphodiesterase-IV: comparison of [^{11}C]rolipram and [^{11}C]Ro 20-1724 in rats. *Synapse* 1999;31:41–50.

73. Engels P, Fichtel K, Lubbert H. Expression and regulation of human and rat phosphodiesterase type IV isogenes. *FEBS Lett* 1994;350:291–295.

74. Imahori Y, Fujii R, Ueda S, et al. Membrane trapping of carbon-11-labeled 1,2-diacylglycerols as a basic concept for assessing phosphatidylinositol turnover in neurotransmission process. *J Nucl Med* 1992;33: 413–422.

75. Ohmori Y, Imahori Y, Ueda S, et al. Radioiodinated diacylglycerol analogue: a potential imaging agent for single-photon emission tomographic investigations of cerebral ischaemia. *Eur J Nucl Med* 1996;23:280–289.

76. Matsumoto K, Imahori Y, Fujii R, et al. Evaluation of phosphoinositide turnover on ischemic human brain with 1-[1-11C]-butyryl-2-palmitoyl-rac-glycerol using PET. *J Nucl Med* 1999;40:1590–1594.

77. Robinson PJ, Noronha J, DeGeorge JJ, et al. A quantitative method for measuring regional *in vivo* fatty-acid incorporation into and turnover within brain phospholipids: review and critical analysis. *Brain Res Rev* 1992;17:187–214.

78. DeGeorge JJ, Nariai T, Yamazaki S, et al. Arecoline-stimulated brain incorporation of intravenously administered fatty acids in unanesthetized rats. *J Neurochem* 1991;56:352–355.

79. Chang MC, Arai T, Freed LM, et al. Brain incorporation of [1-^{11}C]arachidonate in normocapnic and hypercapnic monkeys, measured with positron emission tomography. *Brain Res* 1997;755:74–83.

80. Morgan JI, Curran T. Role of ion flux in the control of c-*fos* expression. *Nature* 1986;322:552–555.

81. Hnatowich DJ. Antisense and nuclear medicine. *J Nucl Med* 1999;40:693–703.

82. Blasberg RG, Tjuvajev JG. Herpes simplex virus thymidine kinase as a marker/reporter gene for PET imaging of gene therapy. *Q J Nucl Med* 1999;43:163–169.

83. Gambhir SS, Barrio JR, Herschman HR, et al. Assays for noninvasive imaging of reporter gene expression. *Nucl Med Biol* 1999;26:481–490.

84. Pardridge WM. Drug delivery to the brain. *J Cereb Blood Flow Metab* 1997;17:713–731.

85. MacLaren DC, Toyokuni T, Cherry SR, et al. PET imaging of gene expression. *Synapse* 1999;34:290–304.

86. MacLaren DC, Gambhir SS, Satyamurthy N, et al. Repetitive, non-invasive imaging of the dopamine D2 receptor as a reporter gene in living animals. *Gene Ther* 1999;6:785–791.

87. Ogawa O, Umegaki H, Ishiwata K, et al. *In vivo* imaging of adenovirus-mediated over-expression of dopamine D2 receptors in rat striatum by positron emission tomography. *Neuroreport* 2000;11:743–748.

88. Barkats M, Bilang-Bleuel A, Buc-Caron MH, et al. Adenovirus in the brain: recent advances of gene therapy for neurodegenerative diseases. *Prog Neurobiol* 1998;55:333–341.

89. Piccini P, Brooks DJ, Bjorklund A, et al. Dopamine release from nigral transplants visualized *in vivo* in a Parkinson's patient. *Nat Neurosci*1999;2:1137–1140.

90. Buolamwini JK. Novel anticancer drug discovery. *Curr Opin Chem Biol* 1999;3:500–509.

91. Seymour L. Novel anti-cancer agents in development: exciting prospects and new challenges. *Cancer Treat Rev* 1999;25:301–312.

92. Leonard DM. *Ras* farnesyltransferase: a new therapeutic target. *J Med Chem* 1997;40:2971–2990.

93. Njoroge FG, Taveras AG, Kelly J, et al. (+)-4-[2-[4-(8-chloro-3,10-dibromo-6,11-dihydro-5H-benzo[5, 6]cyclohepta[1,2-ß]-pyridin-11(*R*)-yl)-1-piperidinyl]-2-oxo-ethyl]-1-piperidinecarboxamide (SCH-66336): a very potent farnesyl protein transferase inhibitor as a novel antitumor agent. *J Med Chem* 1998;41:4890–4902.

94. Fredriksson A, Johnstrom P, Thorell JO, et al. *In vivo* evaluation of the biodistribution of ^{11}C-labeled PD153035 in rats without and with neuroblastoma implants. *Life Sci* 1999;65:165–174.

32

Event-Related Potentials and Magnetic Fields in the Human Brain

Steven A. Hillyard

Marta Kutas

Steven A. Hillyard: Department of Neurosciences, University of California, San Diego, La Jolla, California.

Marta Kutas: Department of Cognitive Science, University of California, San Diego, La Jolla, California.

To uncover the neural bases of a cognitive process it is important both to identify the participating brain regions and determine the precise time course of information transmission within and among those regions. Although neuroimaging techniques based on cerebral blood flow or metabolism (e.g., positron emission tomography [PET] and functional magnetic resonance imaging [fMRI]) are providing increasingly detailed pictures of the anatomic regions activated during cognitive activity, these methods lack the temporal resolution to reveal the rapid-fire patterning of neuronal communication. Noninvasive recordings of the electrical and magnetic fields generated by active neuronal populations, however, can reveal the timing of brain activity related to cognition with a very high, msec-level resolution. This chapter gives an overview of how these temporally precise recording techniques have been used to analyze perceptual and cognitive mechanisms in the human brain.

The changes in field potentials that are time-locked to sensory, motor, or cognitive events are known as event-related potentials (ERPs) and the corresponding magnetic field changes are termed event-related fields (ERFs). Both ERPs and ERFs consist of precisely timed sequences of waves of varying field strength and polarity (Fig. 32.1). These observed peaks and troughs in the waveform are often referred to as “components.” Some authors, however, prefer to use the term component to refer to portions of the waveform that originate from particular neural structures, whereas others consider ERP/ERF components to be those waveform features that are associated with a particular cognitive process or manipulation (2). Both ERPs and ERFs are generated primarily by the flow of ionic currents in elongated nerve cells during synaptic activity. Whereas synaptic currents flowing across nerve cell membranes into the extracellular fluid produce ERPs, the flow of synaptic current through neuronal processes produce ERFs, thereby giving rise to concentric magnetic fields surrounding the cell. When a sufficient number of neurons having a similar anatomic configuration are synchronously active, their summated fields may be strong enough to be detectable at the scalp.

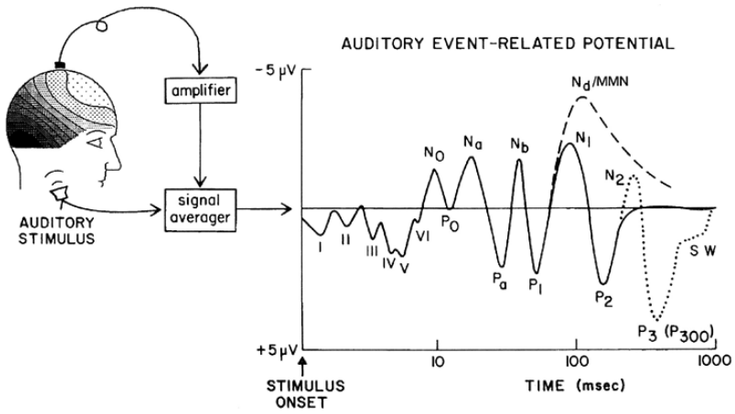


FIGURE 32.1. The characteristic waveform of the auditory event-related potential following a brief stimulus such as a click or tone. The individual components (peaks and troughs) are evoked with specific time delays (latencies) after stimulus onset. Note the logarithmic time base, which makes it possible to visualize the earliest waves (I–VI) generated in the auditory brainstem pathways. Longer latency negative (N) and positive (P) components are generated in different cortical areas. Dashed line shows increased negativity elicited by attended sounds (negative difference) or by deviant sounds (mismatch negativity), and dotted line shows N2 and P3 components to task-relevant target stimuli. Adapted from Münte TF, Schiltz K, Kutas M. When temporal terms belie conceptual order. *Nature* 1998;395:71–73.

When ERPs or ERFs are recorded from the surface of the head, the locations of the active neural generators can only be estimated rather than visualized directly. The calculation of generator locations from surface field distributions is known as the inverse problem, which typically has no unique solution. However, the validity of inverse source estimations can be considerably improved by using algorithms and models that take into account the geometry of the cortical surface, biophysical properties of the intervening tissues, constraints from neuroimaging data, and statistical likelihood of alternative source locations (3 ,4). In general, the localization of active neural populations is more straightforward with surface recordings of ERFs than with ERPs, because magnetic fields, unlike electrical fields, are minimally distorted by the physical properties of the intervening tissues.

ERP and ERF recordings have been used extensively to investigate the spatiotemporal patterns of brain activity associated with a variety of perceptual, cognitive, and linguistic processes. The general research strategy has been to discover the mapping between the components of the waveform and specific cognitive processes that are engaged by a particular task. When an ERP/ERF component can be shown to be a valid index of the neural activity underlying a cognitive operation, that component can yield valuable information about the presence or absence of that operation and its timing with respect to other cognitive events. In many cases, such data have been related to psychological models of the underlying processing operations and used to test alternative theoretical positions. In addition, by localizing the neural generators of such components, inferences can be made

about the participating anatomic circuits that can be interfaced with neuroimaging and neuropsychological data. This chapter describes recent advances made using this approach for analyzing the neural and cognitive mechanisms of preattentive sensory processing, selective attention, mental chronometry, memory storage and retrieval, and language comprehension and production. The use of ERP/ERF recordings to evaluate cognitive disorders associated with neurobehavioral and psychopathologic syndromes also is reviewed.

PREATTENTIVE SENSORY PROCESSING

Part of "32 - Event-Related Potentials and Magnetic Fields in the Human Brain "

Much of the early ERP waveform, and some later components as well, represent sensory-evoked neural activity in modality-specific cortical areas. These evoked components vary with the physical parameters of the stimuli and in many cases are associated with the preattentive encoding of stimulus features. In the visual modality, for example, the early C1 component (onset latency 50 to 60 msec) originates in retinotopically organized visual cortex (5) and varies in amplitude according to the spatial frequency of the stimulus (6). Similarly, the early auditory cortical components P50 and N100 (and their magnetic counterparts, M50 and M100) arise in part from generators in tonotopically organized supratemporal auditory cortex and reflect the encoding of perceived pitch (7).

In general, ERP amplitudes decrease when the time between successive stimulus presentations is made shorter than the refractory or recovery period of the component under study. Although the neural processes underlying ERP refractory effects are not well established, some candidate mechanisms include synaptic fatigue, active inhibition, and the persistence of sensory memory for the preceding stimulus. In line with this latter idea, the refractory period of the auditory M100 has been found to have a similar time course to that of sensory or “echoic” memory for stimulus loudness (8).

The P50 and Sensory Gating

The refractory properties of the auditory P50 (P1) component have been studied extensively over the past 20 years as a possible marker of abnormal sensory input control in schizophrenia (9 ,10 ,11 and 12). In the standard paradigm, pairs of auditory stimuli are presented with an ISI of 0.5 sec, and the amplitude ratio of the P50 evoked by the second stimulus (S2) relative to the first stimulus (S1) is calculated. In general, schizophrenic subjects do not show as large a reduction in the P50 amplitude to S2 relative to S1 as do normal controls. This refractory reduction of P50 amplitude to S2 has been interpreted as a sign of preattentive sensory gating, which occurs because the initial S1 automatically activates an inhibitory system that suppresses responsiveness to S2 (9 ,10). This inhibitory system presumably prevents irrelevant information from ascending to higher levels of cortical processing. The abnormally large S2/S1 amplitude ratio for P50 seen in schizophrenics was thus considered evidence for impaired sensory gating, which was suggested to be the principal sensory deficit of the disease process.

This pattern of more rapid P50 recovery in schizophrenia has been widely reported, but there have been some notable exceptions that raise questions about the exact conditions needed to produce the effect (13 ,14 and 15). A more serious question, however, is whether existing studies have, in fact, demonstrated a reliably abnormal S2/S1 ratio of the auditory P50 in schizophrenics. This concern stems from the way the

P50 has typically been measured—as the maximal positive amplitude within a time window (e.g., 40 to 80 msec) that encompasses the P50 peak. Such peak measures may be artificially inflated by increased levels of background noise in the EEG recordings, originating from either intracranial or extracranial sources. Thus, if a patient group has higher EEG noise levels, then the measured P50 amplitudes tend to be greater because the noise peaks are at times mistaken for the actual peak of the P50. This type of error is more pronounced when measuring the P50 to S2, because its amplitude is diminished relative to the noise owing to refractory effects. Reports of increased variability and lower reliability of P50 measures in schizophrenics (12 ,16) suggest that background noise levels are indeed higher in the patient groups. Several studies, however, have reported that the P50 evoked by S1 is smaller in schizophrenics (11 ,12 ,16), which suggests that overall response amplitudes may be lower and/or response latencies more variable in these patients. Further studies are needed to determine whether the actual S2/S1 amplitude ratio is reliably higher in schizophrenics, or whether this reported effect is a product of noise-sensitive measurement procedures.

Even if the S2/S1 ratio for P50 were determined to be greater in schizophrenia, there would still be some question about its functional interpretation. There is little evidence that refractory reductions of ERP amplitudes to stimuli repeated at 0.5 sec ISIs are associated with any reduction in the perceptual information reaching higher centers. Nor does it appear that the amplitude of P50 to S2 is reduced only when S2 is irrelevant (17), calling into question the hypothesis that the P50 refractory effect reflects the selective gating of irrelevant versus relevant sensory inputs. In addition, it has been reported that schizophrenic patients showing the most severe perceptual anomalies did not differ from normals in their P50 refractory effects (15). Thus, there seems to be scant evidence that reduced P50 refractoriness in schizophrenia, if such exists, is related to the selective gating or filtering of irrelevant sensory information in the auditory cortical pathways.

Auditory Feature Encoding

The preattentive coding of auditory features is indexed with considerable precision by the mismatch negativity (MMN) component and its magnetic counterpart, the MMNm. The MMN is a scalp-negative component with a latency of 120 to 250 msec that is specifically elicited by a deviant sound in a repetitive auditory sequence (18 ,19 and 20). The MMN can be triggered by any discernible change in the ongoing sounds, such as deviations in frequency, intensity, duration, rise-time, timing, and spatial location. MMNs also have been recorded to changes in more complex sound properties such as the phonetic structure of speech sounds and the patterning of tone sequences (20 ,21). Näätänen (22) has proposed that the MMN is generated by an automatic comparator process that contrasts current auditory input against a multidimensional trace of the previous repeating sound's features held in preperceptual sensory memory. This mismatch detection process may represent an early stage in the alerting and orienting of the organism toward novel and potentially important changes in the acoustic environment.

The MMN provides a window on auditory sensory or “echoic memory” because it is initiated by a mismatch with the memory traces of the preceding sounds (23). Indeed, the maximal interstimulus interval (ISI) at which the MMN can be maintained is of the order of 10 sec, corresponding well with behavioral estimates of the duration of echoic memory (19 ,20). The MMN also can be used to study more permanent auditory memory traces, such as those for the phonemic characteristics of one's native language (19) as they emerge during the first year of life (21).

It has been proposed that a supratemporal component of the MMN originating in auditory cortex reflects the preattentive sensory store and automatic change detection process, whereas a frontal component indexes the involuntary orienting of attention to the deviant event (24 ,25). For speech sounds, however, the MMN/MMNm appears to arise from sources in auditory cortex of the left hemisphere, in accordance with proposals that the left posterior temporal cortex is the locus of language-specific auditory traces (19).

Given that the MMN may be elicited with minimal cooperation from the subject, it has found wide applicability for the diagnosis and evaluation of a variety of neurobehavioral and psychiatric disorders (24 ,26). Schizophrenic patients have reduced or prolonged MMNs to pitch or duration deviants, with the degree of abnormality depending on the specific parameters of the stimulus deviance (24 ,27 ,28). These findings provide clear evidence for a deficit in preattentive auditory processing in schizophrenia, although there is some debate about whether the impairment is primarily in temporal processing (28) or auditory encoding and trace formation (27). A different pattern of abnormality has been observed in Alzheimer's disease, with MMN amplitude reductions becoming more prominent at longer ISIs (29), suggesting a more rapid decay of auditory sensory memory traces. MMN abnormalities indicative of auditory processing deficits have also been reported in cases of learning disorders, language and speech impairments, depression, autism, parkinsonism, and HIV infection. (See refs. 19 , 21 , 24 , and 29 for reviews.) Drugs that interfere with NMDA-receptor mediated neurotransmission also disrupt the MMN, which is consistent with models of schizophrenia that posit a disturbance in glutaminergic/NMDA functioning (27).

SELECTIVE ATTENTION

Part of "32 - Event-Related Potentials and Magnetic Fields in the Human Brain "

The brain's attentional systems include a central control network with projections to the sensory pathways of the different modalities that enable the selective modulation of

incoming information. A good deal of research on attention over the past few decades has been aimed at determining whether incoming sensory information is selected at “early” or “late” levels of processing—that is, before or after stimulus properties are fully analyzed (30). Current evidence from both behavioral and physiologic studies indicates that attention can select stimuli at different levels of the sensory pathways, depending on the features being attended and the task requirements.

Auditory Attention

In the auditory modality, ERPs have demonstrated that attentional selection occurs at early levels of cortical processing, but not in the brainstem pathways (31). In dichotic listening tasks with rapidly presented tones to the left and right ears, the earliest ERP component that is reliably influenced by paying attention selectively to one ear is a small positive wave with a latency of 20 to 50 msec (termed the P20–50), which has been localized using magnetoencephalography (MEG) to sources in or near primary auditory cortex. This short-latency modulation provides evidence for an attentional mechanism of sensory gain control at the earliest levels of cortical processing. A much stronger attentional modulation of auditory input takes place at 50 to 70 msec after stimulus onset in the form of a negative difference (Nd) potential that augments the amplitude of the evoked N1 wave to attended-channel sounds (Fig. 32.1). This N1/Nd attention effect also has been localized to auditory cortex by MEG recordings and is considered to be an index of further processing of attended sound information, or alternatively, of the closeness of match between incoming stimuli and the acoustic features that define the attended channel (30). These negative ERP modulations indicate the precise timing with which different auditory features are attended or rejected (32) and provide strong evidence for early selection theories of attention. Schizophrenic subjects reportedly show abnormally reduced Nd amplitudes when attending to multiple sound features, suggesting a deficit in their control functions for allocating attentional resources during selective listening (33).

In recent studies, auditory ERPs have been used to study how attention is allocated in a noisy environment with multiple, competing sound sources (34). When subjects listened selectively to sounds coming from one loudspeaker in a free-field array, the spatial focusing of auditory attention took place in two distinct stages: an early, broadly tuned input selection occurring over the first 80 to 180 msec (indexed by N1/Nd) was followed by a more sharply focused selection of target sounds that began at around 250 msec (indexed by the late positive P3 component). These findings indicate that auditory spatial attention is deployed as a sharply tuned gradient around an attended sound source in a noisy environment. Under these conditions congenitally blind persons were found to have sound localization capabilities superior to those of sighted control subjects (35). ERP data indicated that this enhanced capability of blind persons is mediated at least in part by an attentional tuning mechanism that operates within the first 100 msec after sound onset.

Visual Attention

Covertly directing attention to a specific location in the visual fields typically results in faster and more accurate detections or discriminations of stimuli at that location. Recordings of brain activity in both humans and animals have identified a number of sites along the visual pathways where afferent information is modulated under the influence of visual-spatial attention. Neurophysiologic studies in monkeys demonstrated strong influences of spatial attention on neural activity in extrastriate cortical regions, including retinotopic areas V2, V3A, and V4 and higher areas of both the ventral (inferior temporal lobe) and dorsal (area MT, posterior parietal lobe) processing streams (36). These findings are congruent with human ERP studies showing that stimuli at attended locations elicit enlarged P1 (70 to 130 msec) and N1 (150 to 190 msec) components (Fig. 32.2), which have been localized to generators in extrastriate visual cortex (37). This amplitude enhancement of the P1 and N1 waves occurs with little or no change of the component latencies, suggesting that spatial attention exerts a gain control or selective amplification of attended inputs within the visual-cortical pathways in the interval between 70 and 200 msec after stimulus onset (38).

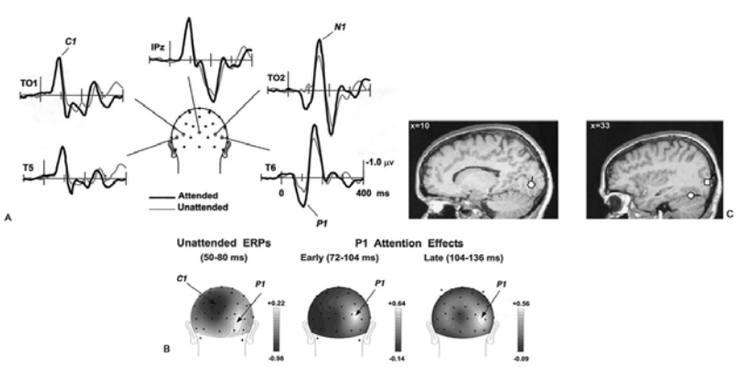


FIGURE 32.2. A: Visual event-related potential waveforms from several scalp sites in response to stimuli in left visual field in study by Martínez and colleagues (41). Subjects were required to attend to one field at a time while randomized stimulus sequences were presented concurrently to left and right fields. Note increased amplitude of P1 and N1 components when stimuli were attended and lack of change in C1 component. **B:** Voltage topographies of C1 component and of the P1 attention effect (increased positivity with attention) in two different time ranges. **C:** Locations of dipolar sources calculated for the C1 component (in primary cortex of calcarine fissure, *left*) and the early and late P1 attention effects (dorsal and ventral extrastriate cortex, respectively, *right*).

In experiments that combined PET neuroimaging and ERP recordings, the calculated dipolar sources of the P1 attention effect were found to correspond closely with regions of increased blood flow in retinotopically organized extrastriate areas, including areas V3/V4 and the posterior fusiform gyrus (37). Significantly, however, the earlier C1 component (onset at 50 msec), which appears to originate from generators in primary visual cortex (area V1), was found to remain invariant with changes in the direction of spatial attention. These findings suggest that spatial attention modulates the flow of visual information at a higher level than the primary cortex.

Recent studies in monkeys, however, reported that stimulus-evoked activity in area V1 may be affected by spatial attention when competing stimuli are present in the visual field (39). The participation of V1 in spatial attention has also been inferred from recent fMRI studies in humans (40). In a study combining fMRI with ERP recordings, Martínez and colleagues (41) observed focal increases in blood flow (BOLD signal) in area V1 as well as areas V2, V3/VP, and V4 at sites corresponding to the retinotopically mapped position of the attended stimulus; however, the amplitude of the C1 component again remained invariant (Fig. 32.2). The earliest influence of attention was on the P1 component (75 to 130 msec), which was localized through dipole modeling to dorsal and ventral extrastriate sources. It was suggested

that the attentional modulation of V1 activity revealed by fMRI may take place at a latency longer than the initial geniculo-striate response represented by the C1 and is consequent on delayed feedback of enhanced visual signals back to V1 from higher extrastriate areas. Such long-latency modulations in V1 have been observed in animals and may enhance figure/ground contrast in attended regions of the visual field (39).

The spatial allocation of attention has also been studied with the steady-state visual evoked potential (SSVEP), which is the oscillatory response of the visual cortex evoked by regularly repeating stimuli such as a light that flickers at a rate of 8 Hz or more (42). The amplitude of the SSVEP elicited by such a flickering stimulus is substantially increased in amplitude when attention is directed to its location, thereby indexing amplification of visual inputs within the spotlight of attention. The continuous nature of the SSVEP as a measure of cortical facilitation makes it suitable for measuring the time course of attention shifts among stimuli in the visual fields.

In contrast with spatial attention, when stimuli are selected on the basis of nonspatial features such as color, shape, or spatial frequency, a different pattern of ERP components is observed. Stimuli having the attended feature elicit a prominent “selection negativity” (SN) over the posterior scalp that begins at 120 to 220 msec and may extend for several hundreds of msec (37). The onset of the SN provides a precise measure of the time point at which a particular feature is discriminated and selectively processed, and localization of its neural generators points to the brain regions involved in the selection. When stimuli are selected on the basis of two or more features concurrently, recordings of the SN can indicate whether the features are selected independently or in an interactive, contingent manner (43).

MENTAL CHRONOMETRY

Part of "32 - Event-Related Potentials and Magnetic Fields in the Human Brain "

Motor preparation, execution, and evaluation are indexed by a series of electric (and magnetic) components both preceding and following movement onset. Prime among the ERPs indexing preparation is the readiness potential (RP), which is a slowly ramping negativity that starts about 1 s before the onset of a voluntary or self-paced movement and peaks around movement onset (44). The initial, bilaterally symmetric portion of the RP is generated in the supplementary

motor area (SMA). Approximately 200 to 500 msec before movement onset the RP becomes asymmetric, being maximal over central scalp contralateral to the active musculature. This lateralized portion of the RP has a somatotopic distribution over the motor cortex and has been localized to the primary motor cortex (45). This asymmetric portion of the RP can be seen in stimulus-locked averages (46) and serves as the basis for an index of differential motor preparation termed the lateralized readiness potential (LRP). Subtracting for each hand separately the activity over the ipsilateral from that over the contralateral central site, and then averaging the activity for the two hands together derive the LRP.

The LRP has proven especially useful in studies of mental chronometry aimed at answering questions about the dynamics of information processing. For example, LRP data have shown that whether a person responds quickly and accurately is in large part a function of whether he or she is prepared to do so before the stimulus appears. Appropriate preparation leads to fast and correct responses, whereas inappropriate preparation leads to fast but wrong responses, and no preparation at all leads to slow but accurate responses. More important, LRP data revealed that people develop biases that influence how they prepare to respond. (See refs. 47 and 48 for review.)

In a number of sensory-motor discrimination tasks, LRP recordings have provided strong support for “continuous flow” models that specify that transmission of perceptual information to the response system occurs continuously rather than in discrete, all-or-none stages. Such studies have revealed partial transmission of perceptual information (e.g., about letter identity) to the response system and have provided a means of tracking the time course of the extraction of various types of information. The LRP also has been used to determine the timing of “the point of no return;” that is, the time in the course of response preparation beyond which response execution cannot be stopped (47).

Readiness potentials have been examined in a number of patient populations. They are abnormal in a large majority of individuals with Parkinson's disease (PD), presumably because of abnormal activation of the SMA by the basal ganglia (49 ,50). Because the early part of the RP is sensitive to attention, it has been suggested that motor performance in PD patients might be improved by having them attend to movements that they might otherwise try to execute automatically (51). Individuals with tardive dyskinesia (TD) also show abnormal RPs that are larger in amplitude than those of normal controls and schizophrenic patients without TD (52). Unlike voluntary leg movements, involuntary myoclonic leg movements in patients with restless leg syndrome do not elicit an RP, suggesting that these involuntary movements have a subcortical or spinal origin (53).

Purposeful movements are generally monitored and evaluated so that remedial action may be taken if an error in committed. Performance monitoring of this sort is indexed by an ERP known as the error-related negativity (ERN), which peaks about 100 msec after the onset of the electromyographic activity associated with an erroneous response. ERN amplitude covaries with the perceived inaccuracy of a response (54) and is influenced by how similar the given response is to the correct one. An ERN is also elicited by a feedback stimulus that lets the subject know an error was made. ERN generators have been localized to the anterior cingulate gyrus and are modulated by lateral prefrontal cortex (55). Patients with lateral prefrontal cortical damage are impaired in correcting their behaviors and produce equal-sized ERNs for correct and incorrect responses. ERN amplitude is sensitive to mood and personality variables (56), especially when correct responses are rewarded and/or incorrect responses are punished (57).

MEMORY

Part of "32 - Event-Related Potentials and Magnetic Fields in the Human Brain "

Working Memory

Unexpected events typically require us to revise or update our current working mental model of the environment. Donchin and colleagues (58) proposed that this updating of the working memory (temporary, limited capacity) system is indexed by the P3 (also known as the P300 or P3b) component. (See refs. 59 and 60 for alternative views.) The P3b is a positive, broadly distributed component with a centro-parietal maximum and peak latency between 300 and 800 ms. It is elicited by infrequent target events in a sequence of higher probability nontarget events that are being attended, although irrelevant stimuli that draw attention may also trigger a P3. In general P3 amplitude grows with the relevance, salience, and utility of the target or “oddball” stimulus. The P3 can be elicited by many different simple and complex events, including the occasional absence of a predicted stimulus. (See ref. 61 for review.) The differences in the distribution and timing of P3s in various modalities are consistent with the proposal that there are multiple working memory stores.

P3 amplitude is inversely related to the overall probability of the target events, and varies with the fine structure of event sequences. The more difficult the categorization of the target events, the longer the P3 latency. P3 latency is not correlated with the variance in reaction time that is caused by response execution, thereby making it a rather pure measure of stimulus evaluation/categorization time. The combined sensitivity of the P3 to attention and stimulus evaluation makes it a good index of the availability and allocation of capacity-limited perceptual resources. Measurements of P3 latency and RT together can be used to pinpoint the processing locus of individual differences in performance, as was done, for example, to analyze cognitive slowing with normal aging (62). Similarly, P3 data have demonstrated that the prolongation of response time for the second of two decisions made in rapid succession (“psychological

refractory period") is owing to interference at a stage that follows perceptual categorization, presumably that of response selection (63).

Updating working memory has consequences for an individual's subsequent performance. For example, the relative amplitude of the P3 on a trial when a subject commits an error is predictive of the performance (accuracy and response speed) on the next trial; moreover, the larger the P3 to an item, the greater the likelihood that it will be remembered subsequently. (See ref. 64 for review.) Intracranial recordings in individuals with epilepsy have revealed P3-like potentials in the hippocampal region of the medial temporal lobe (65). The scalp-recorded P3, however, primarily reflects activity in a number of neocortical (frontal, central, parietal, temporal-parietal junction) and perhaps subcortical generators and is mediated by several neurotransmitter systems (66). Not surprisingly, therefore, many patient populations show abnormally small or delayed P3bs including schizophrenic patients, individuals at risk for alcoholism, patients with probable Alzheimer's dementia, and individuals with attention deficit and hyperactivity disorder, among others (61 ,66).

The storage of information in working memory may be modulated by attention and appears to be strongly suppressed during the "attentional blink" that follows detection of the first of two target stimuli presented in a rapid sequence. Luck and colleagues showed that the P3 was absent in response to targets that went undetected during the attentional blink, suggesting that no updating of working memory occurred. (See ref. 67 for review.) However, undetected target words did elicit late negative ERPs, indicating that they had been identified at the level of lexical/semantic processing. Thus, the ERP data provided strong evidence that the attentional blink acts at the postperceptual stage of working memory storage.

A frontally distributed late positivity (P3a) is elicited by rare and unexpected stimuli for which there is no memory template readily available. It appears to reflect an orienting response to stimulus novelty and is reduced in patients with prefrontal cortical injury (68). Maintenance of information in working memory is also reflected in sustained ERPs lasting on the order of seconds. These potentials vary in their scalp distribution as a function of the information being held in working memory, consistent with proposals of independent short-term buffers for verbal and nonverbal information, among others (69).

Long(er)-Term Memory

Encoding

Encoding processes (transforming sensory input into a lasting representation) are associated with an increased positivity between 200 and 800 msec that spans several components. Items that call for preferential processing because they stand out, for example, are better recalled and elicit larger P3 components (70). Likewise, the more deeply (semantically) an item is analyzed, the more likely it is to be remembered, and this is reflected in greater late positivity (71). Even among items that are all deeply processed, those that will in fact be remembered later elicit a larger positivity during encoding than those that will be forgotten (Fig. 32.3). These late components produced during encoding that are predictive of subsequent memory performance are collectively termed Dm effects (71).

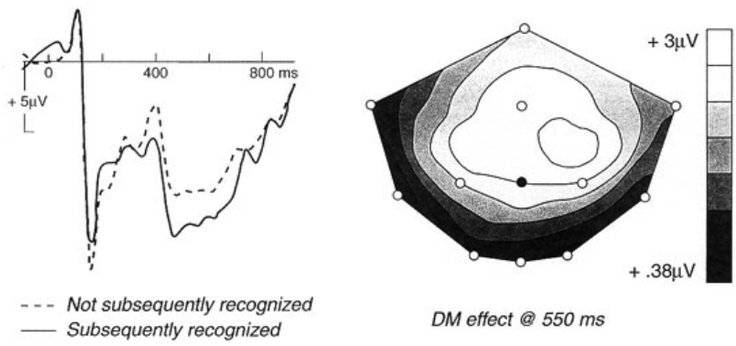


FIGURE 32.3. Averaged event-related potentials (ERPs) from midline parietal site *(filled in circle in map on the right)* sorted as a function of subsequent memory in a cued recall test. The responses to words subsequently recalled *(solid line)* are overlapped with those subsequently not recalled *(dashed line)*. Participants were presented the first three letters of a word and asked to use this stem as a clue for verbally recalling the words they had just studied. The voltage map of this difference related to memory (Dm) effect at 550 ms was computed by subtracting the ERPs to words subsequently not recalled to ERPs from those subsequently recalled. **A:** Semantically anomalous word. **B:** Unexpected word. Adapted from Paller KA. Recall and stem-completion priming have different electrophysiological correlates and are modified differentially by directed forgetting. *J Exp Psychol Learn Mem Cogn* 1990;16:1021–1032.

Dm effects are larger in semantic than in nonsemantic tasks and are not seen for items that have no preexisting representation in long-term memory. Van Petten and colleagues (73) suggested that this positivity indexes the richness of associative elaboration engendered by the to-beremembered event. Consistent with this proposal, the Dm effect varies with the encoding task and information retrieved from long-term memory and shows substantial variability in onset latency, duration, and scalp topography (74).

Retrieval

Retrieval processes are indexed by several ERP effects that vary with whether or not the rememberer is in a retrieval mode, whether memory is queried directly or indirectly, what aspect of the memory is being queried, and whether or not the retrieval attempt is successful (75 ,76). Retrieval itself is indexed by slow potentials sustained over several seconds with an amplitude determined by the difficulty of the retrieval and a scalp topography determined by the nature

of the information retrieved (77). These results fit with the notion that the brain areas involved in explicit memory are the same as those carrying out the initial encoding and perception and argue against the concept of a single, amodal memory store.

In a typical retrieval paradigm items are presented twice, and ERPs to the first and second (i.e., repetition) presentations are compared. When subjects are asked to recognize and detect the repeated items, the task is considered to probe memory directly or explicitly. By contrast, when the old or new distinction is irrelevant, as in tasks involving lexical decision, semantic judgment, or identification of visually degraded words, the stimuli may only tap memory indirectly or implicitly and may not produce actual recollection. In both implicit and explicit memory tasks, stimulus repetition produces large and reliable ERP effects. The first is a reduction in the amplitude of negativity between 250 and 500 msec (N400) that is associated with semantic processing (76). The N400 is reduced by repetition, whether or not the task explicitly calls for detection of repeated items, even in amnesic individuals with damage to the medial temporal lobes (78). Some authors have linked a frontal subcomponent of the N400 to repetition independent of recognition (79).

Another ERP consequence of word repetition is a change in the amplitude of a late positive component (LPC), which typically begins around 400 to 500 msec, and and is somewhat larger over the left than right scalp. There is mounting evidence that this LPC reflects conscious recollection. Factors that influence perceptual priming do not modulate LPC amplitude (80), whereas factors that influence recognition memory do. There is an LPC repetition effect whether memory is tested implicitly or explicitly. When participants are asked to indicate whether an item is old or new, correctly recognized old items elicit larger LPCs than do unrecognized old items or correctly recognized new items, although its distribution across the scalp varies somewhat with the materials (81). The LPC to correctly recognized old items is larger for confident than less confident decisions, and for items that participants actually “remember” (82).

When a subject attempts to remember some aspect of the context in which an item was studied or some attribute of the item that it shared with others in the study task, a large, late, frontally distributed (sometimes right lateralized) positivity is elicited (83). This large positivity over prefrontal sites occurs in addition to the standard LPC effect. The prefrontal locus of this ERP source retrieval effect fits with the known impairments that patients with frontal lobe damage have in retrieving source information about items that they recognize (84).

LANGUAGE

Part of "32 - Event-Related Potentials and Magnetic Fields in the Human Brain "

Semantic Analysis

The semantic analysis of verbal and nonverbal stimuli is indexed by the N400 component (85). The N400 is a broadly distributed component, with a negative-going peak over centroparietal sites often with a slightly right predominance; in young adults, it has an onset around 200 msec and a peak around 400 msec. The largest N400s are elicited by unexpected, semantically anomalous words in a sentence (Fig. 32.4). However, all *potentially* meaningful items (e.g., words and pseudowords, environmental sounds, pictures,

faces) can elicit some N400 activity with an amplitude that is determined by a variety of factors. With little or no contextual constraint, N400 amplitudes are inversely related to the frequency of the eliciting word in the language (86).

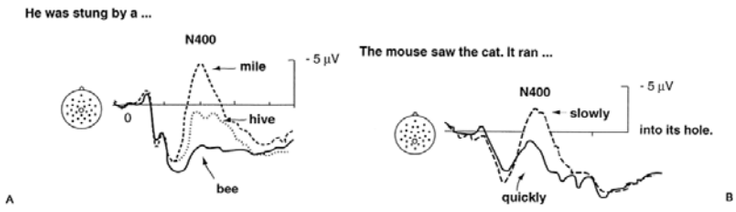


FIGURE 32.4. A: A prototypical N400 recorded at the vertex in response to a semantically anomalous word *(dashed line)* at the end of a sentence, compared with the ERP to the expected ending *(solid line)*, and an anomalous ending that is semantically related to the expected ending *(dotted line)*. Note reduced N400 for nonsensical ending ("hive") that is semantically related to expected ending. **B:** The N400 recorded at a midline parietal site elicited by a word that fits with the ongoing discourse context *(solid line)* versus that to a word that is less expected and does not fit as well with the ongoing discourse context *(dashed line)*. Data taken from van Berkum JJA, Hagoort P, Brown CM. Semantic integration in sentences and discourse: evidence from the N400. *J Cog Neurosci* 1999;11:657–671.

The N400 is typically considered an ERP index of semantic processing or contextual integration because its amplitude is modulated by its relation and fit to the ongoing context, be it a single word, a sentence, or a multisentence discourse. (See refs. 87 and 88 for review.) N400 amplitudes are enlarged to a word in unrelated word pair or in an incongruous or weak context relative to the response to the same word in a related pair or strong congruous sentence. The N400 in these cases is almost identical in timing and distribution over the head, indicating that by 400 msec at the latest, lexical, sentential, and discourse processes all converge to influence language comprehension in a similar manner. Visual half-field studies of the N400 show that the left hemisphere, in particular, uses the organization of semantic memory tapped by context words to aid in its online predictions, whereas the right hemisphere waits and integrates (89).

As would be expected of an index of semantic processing and contextual integration, N400 amplitude is greatly attenuated and its latency delayed in aphasic patients with moderate to severe comprehension deficits (90) N400 latency is also prolonged with normal aging and various dementias. Although ERP evidence for a differential organization of semantic memory in schizophrenia is equivocal, a delay in N400 latency has been reported. (See ref. 87 for review.) Intracranial recordings from patients with epilepsy show potentials functionally similar to the scalp N400 in the anterior fusiform gyrus (91).

Syntactic Analysis

The processing of language at a syntactic level is indexed by a several ERP components, both negative and positive. Many, although not all, syntactic violations elicit a late positivity variously called the P600 or the syntactic positive shift (SPS) (92 ,93 and 94). The P600 is typically elicited when some aspect of sentence structure violates the rules of the language—for example, if the subject of the sentence does not agree with its verb in number or if a word in a phrase is out of order. The P600 also may be elicited when processing difficulties arise at a structural level (87). Some researchers have proposed that the P600 belongs to the family of P3 waves (95). In addition to the P600, many syntactic violations also elicit a left anterior negativity (LAN), which some researchers have interpreted as an index of working memory usage (96 ,97).

The fact that an N400 or P600 is elicited shortly after a semantically anomalous or grammatically incorrect word, respectively, regardless of its ordinal position in a sentence, is most consistent with those models of sentence processing that emphasize the immediate and online nature of comprehension (98). That is, the language processing system seems to use all information as it becomes available, often to predict what words or ideas are likely to come next (89 ,99). That processing at a semantic and syntactic level yields different patterns of electrophysiologic activity suggests that the processes differ, if not the representations. Further, the presence of different ERP patterns to various syntactic violations indicates that syntax is not a unitary phenomenon mediated by a single neural generator. Many aspects of sentence processing at semantic, syntactic, referential, thematic, prosodic, and discourse levels are indexed by transient ERP effects and/or slow potentials that encompass the entire sentence (100 ,101). In short, the reported patterns of ERP effects are inconsistent with a view of language comprehension that gives syntactic analysis precedence over semantic analysis or a system wherein syntactic processes are isolated from all other processes. Instead, ERP data provide considerable evidence for parallel processing, interaction, and top-down effects during language processing. The brains of readers and listeners work very much online using all information as it becomes available to anticipate upcoming items, concepts, and schemas to achieve the aim of an efficient and error-free understanding of the incoming language (even if at times these predictions may lead to misunderstanding).

Language Production

As in language comprehension, many of the controversies in language production revolve around the issue of the relative timing of the different levels of processing that are engaged. Although there is a general consensus that producing a coherent utterance involves information at the levels of meaning, syntax, and phonetics there is no agreement as to whether meaning comes first and then phonologic form (i.e., a serial model), these processes overlap somewhat in time (i.e., a cascade model), or they unfold in parallel (102). Two ERP measures—the LRP and N200—can be used to track the time course of information availability as people prepare to speak, even if they never actually utter a word. In studies using a two choice go/no-go paradigm, subjects were shown a picture of an item on each trial about which they were asked to make two decisions (Fig. 32.5). Across experiments, decisions were based on semantic, syntactic, and phonologic aspects of the pictured item and its name. The timing of the N200 and LRP on no-go trials indicated that semantic information becomes available before syntactic information (by about 80 msec), which is in turn available before phonologic information (by about 40 msec) (103 ,104). Electrophysiologic data from the scalp thus support a serial model of speech production, indicating that people first figure out what they want to say and then choose exactly how to say it.

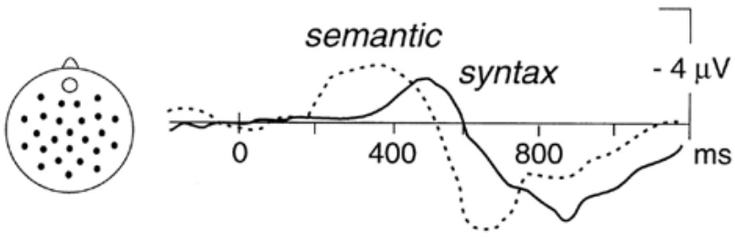


FIGURE 32.5. Overlapped are the N200 difference waves (no-go minus go event-related potentials) recorded at a midline prefrontal site *(as marked on head icon)* when the decision to respond or not (go/no-go) was contingent on a semantic *(dashed line)* versus a syntactic *(solid line)* attribute of the pictured object. Note that the N200 effect contingent on the semantic analysis peaks around 380 ms, whereas that contingent on syntax peaks around 500 ms. Data from Schmitt B, Münte TF, Kutas M. Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology* 2000 Jul;37(4):473–484.

CONCLUSION

Specific components of ERPs and ERFs recorded from the surface of the head are sensitive to a wide range of sensory, perceptual, motor, mnemonic, and linguistic processes. It appears that many cognitive acts engender synchronous neural activity patterns that produce electrical and magnetic fields precisely time-locked to informational transactions in the brain. Recordings of ERPs/ERFs thus provide critical information about the timing and neural substrates of the processing stages that underlie cognitive activity. These physiologic data are being used increasingly to test alternative functional models and to constrain psychological theories (2 ,64 ,105).

Considerable progress has been made in demonstrating reliable associations between ERP/ERF components and a wide range of psychopathologic syndromes. In no case, however, is a single ERP/ERF component absent or abnormal in such a way as to be diagnostic. Rather, a given syndrome (e.g., schizophrenia) usually manifests abnormalities in one or more parameters of several different ERP components, and a given component (e.g., the P3) appears abnormal across a range of neurobehavioral syndromes. This is to be expected, except in the unlikely (and perhaps nonexistent) case in which the psychopathology would only affect a single, isolated cognitive subprocess that had a unique ERP/ERF marker. Thus, instead of seeking a single ERP marker, it seems more likely that various patient populations will be distinguished by different profiles of ERP effects across a number of different tasks (much like the approach taken in neuropsychological testing). Many of the same interpretational issues that are of concern with neuropsychological testing may become relevant for testing with an ERP battery, together with some that are specific to these physiologic measures. For example, the considerable synaptic plasticity of the neocortex suggests that even normal individuals' component amplitudes and latencies are likely to show considerable variability, depending on their life experiences. Such variability within the normal population clearly exacerbates the difficulty of uniquely identifying ERP/ERF markers of specific clinical syndromes. Further progress in achieving diagnostic specificity and sensitivity may require comparing ERPs/ERFs across multiple tasks in each patient to reveal reliable abnormalities that are related to specific cognitive manipulations. Such ERP/ERF abnormalities will become increasingly informative about the specific processing mechanisms that are dysfunctional in patient groups as the cognitive specificity of the distinctive components is sharpened through studies in normals and as better methods are developed for measuring and isolating those components. These developments should make it possible to incorporate ERP/ERF data into multimeasure diagnostic batteries to aid in classifying and subtyping psychopathologic syndromes.

Recent technical advances have made it possible to obtain more accurate information about the neural bases of ERPs/ERFs and their relationships with cognitive and behavioral variables. The neural generators of surface recorded ERP/ERF activity can be localized with increased precision using algorithms that exploit more accurate bioelectric models of the head and constrain the generators to lie within the cortical mantle as reconstructed from MRI scans. (See ref. 105 for review.) Source localizations can be further improved by incorporating functional imaging data (e.g., from fMRI) into the inverse calculations, thereby providing a more veridical picture of the spatiotemporal patterning of cognitive-related brain activity (3 ,4 ,106). New approaches also have been developed for decomposing these complex patterns of brain activity arising from multiple, concurrently active generators into functionally meaningful subcomponents. Among these, the technique of Independent Component Analysis (107) has shown considerable promise for decomposing ERP data sets from multiple task conditions into temporally independent and spatially localizable components that may be related to cognitive operations on the one hand and to fMRI activation patterns on the other. Newer spatiotemporal filtering procedures (e.g., wavelet filtering) have improved our ability to extract the ERP/ERF signal from ongoing brain activity and other background noise. (See ref. 105 for review.) These methods ultimately may allow reliable detection of event-related signals on a single-trial basis without relying on the usual computer averaging procedure. Single-trial analyses are important not only for achieving a closer correspondence between brain activity and behavioral performance but also for ascertaining the degree of trial-to-trial variability that may characterize different clinical syndromes. All of these techniques will substantially increase the utility of ERP/ERF recordings for analyzing the neural bases of both normal and disordered cognition.

ACKNOWLEDGMENTS

Part of "32 - Event-Related Potentials and Magnetic Fields in the Human Brain "

This work was supported by NIH grants MH-25594, HD-22614, and AG-08313. We thank Carole Montejano, Matt Marlow, and Tom Urbach for assistance with manuscript preparation.

REFERENCES

1. Hillyard SA. Electrical and magnetic brain recordings: contributions to cognitive neuroscience. *Curr Opin Neurobiol* 1993;3:217–224.

2. Kutas M, Dale A. Electrical and magnetic readings of mental functions. In: Rugg MD, ed. *Cognitive neuroscience*. Cambridge, MA: MIT Press, 1997:197–242.

3. Dale AM, Liu AK, Fischl BR, et al. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 2000;26:55–67.

4. Schmidt DM, George JS, Wood CC. Bayesian inference applied to the electromagnetic inverse problem. *Hum Brain Map* 1999;7:195–212.

5. Clark VP, Fan S, Hillyard SA. Identification of early visually evoked potential generators by retinotopic and topographic analyses. *Hum Brain Map* 1995;2:170–187.

6. Kenemans JL, Baas JMP, Mangun GR, et al. On the processing of spatial frequencies as revealed by evoked-potential source modeling. *Clin Neurophys* 2000;111:1113–1123.

7. Pantev C, Elbert T, Ross B, et al. Binaural fusion and the representation of virtual pitch in the human auditory cortex. *Hear Res* 1996;100:164–170.

8. Lu ZL, Williamson SJ, Kaufman L. Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science* 1993;258:1668–1670.

9. Adler LE, Pachtman E, Franks RD, et al. Neurophysiological evidence for a defect in neural mechanisms involved in sensory gating in schizophrenia. *Biol Psychiatry* 1982;17:639–654.

10. Freedman R, Adler LE, Gerhardt GA, et al. Neurobiological studies of sensory gating in schizophrenia. *Schizophr Bull* 1987;13:669–678.

11. Clementz BA, Geyer MA, Braff DL. P50 suppression among schizophrenia and normal comparison subjects: a methodological analysis. *Biol Psychiatry* 1997;41:1035–1044.

12. Patterson JV, Gierczak M, Hetrick WP, et al. Effects of temporal variability on P50 and the gating ratio in schizophrenia. *Arch Gen Psychiatry* 2000;57:57–64.

13. Kathmann N, Engel RR. Sensory gating in normals and schizophrenics: a failure to find strong P50 suppression in normals. *Biol Psychiatry* 1990;27:1216–1226.

14. Guterman Y, Josiassen RC. Sensory gating deviance in schizophrenia in the context of task related effects. *Int J Psychophysiol* 1994;18:1–12.

15. Jin Y WE, Bunney J, Sandman CA, et al. Is P50 suppression a measure of sensory gating in schizophrenia. *Biol Psychiatry* 1998;43:873–878.

16. Jin Y, Potkin SG, Patterson JV, et al. Effects of P50 temporal variability on sensory gating in schizophrenia. *Psychiatry Res* 1997;70:71–81.

17. Jerger K, Biggins C, Fein G. P50 suppression is not affected by attentional manipulations. *Biol Psychiatry* 1992;31:365–377.

18. Näätänen R. The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear Hear* 1995;16:6–18.

19. Näätänen R, Escera C. Mismatch negativity: clinical and other applications. *Audiology Neuro-Otology* 2000;5:105–110.

20. Picton TW, Alain C, Otten L, et al. Mismatch negativity: different water in the same river. *Audiology Neuro-Otology* 2000;5:111–139.

21. Kraus N, Cheour M. Speech sound representation in the brain. *Audiology Neuro-Otology* 2000;5:140–150.

22. Näätänen R, Alho K. Mismatch negativity: the measure for central sound representation accuracy. *Audiology Neuro-Otology* 1997;2:341–352.

23. Ritter W, Deacon D, Gomes H, et al. The mismatch negativity of event-related potentials as a probe of transient auditory memory: a review. *Ear Hear* 1995;16:52–67.

24. Gené-Cos N, Ring HA, Pottinger RC, et al. Possible roles for mismatch negativity in neuropsychiatry. *Neuropsychiatry Neuropsychol Behav Neurol* 1999;12:17–27.

25. Escera C, Alho K, Schröger E, et al. Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology Neuro-Otology* 2000;5:151–166.

26. Csépe V, Molnar M. Towards the possible clinical application of the mismatch negativity component of event-related potentials. *Audiology Neuro-Otology* 1997;2:354–369.

27. Javitt DC. Intracortical mechanisms of mismatch negativity dysfunction in schizophrenia. *Audiology Neuro-Otology* 2000;5:207–215.

28. Michie PT, Budd TW, Todd J, et al. Duration and frequency mismatch negativity in schizophrenia. *Clin Neurophys* 2000;111:1054–1065.

29. Pekkonen E. Mismatch negativity in aging and in Alzheimer's and Parkinson's diseases. *Audiology Neuro-Otology* 2000;5:216–224.

30. Näätänen R. *Attention and brain function*. Hillsdale, New Jersey: LEA, 1992.

31. Hillyard SA, Mangun GR, Woldorff MG, et al. Neural systems mediating selective attention. In: Gazzaniga MS, ed. *The cognitive neurosciences*. Cambridge, MA: MIT Press, 1995:665–681.

32. Woods DL, Alho K, Algazi A. Stages of auditory feature conjunctions: an event-related brain potential study. *J Exp Psychol Hum Percept Perform* 1994;20:81–94.

33. Michie PT, Fox AM, Ward PB, et al. Event-related potential indices of selective attention and cortical lateralization in schizophrenia. *Psychophysiology* 1990;27:209–228.

34. Teder-Sälejärvi WA, Hillyard SA, Röder B, et al. Spatial attention to central and peripheral auditory stimuli as indexed by event-related potentials (ERPs). *Brain Res Cogn Brain Res* 1999;8:213–227.

35. Röder B, Teder-Sälejärvi W, Sterr A, et al. Improved auditory spatial tuning in blind humans. *Nature* 1999;400:162–166.

36. Desimone R. Visual attention mediated by biased competition in extrastriate visual cortex. *Philos Trans R Soc Lond B Biol Sci* 1998;353:1245–1255.

37. Hillyard SA, Anllo-Vento L. Event-related brain potentials in the study of visual selective attention. *Proc Natl Acad Sci USA* 1998;95:781–787.

38. Hillyard SA, Vogel EK, Luck SJ. Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos Trans R Soc Lond B Biol Sci* 1998;353:1257–1270.

39. Posner MI, Gilbert CD. Attention and primary visual cortex. *Proc Natl Acad Sci USA* 1999;96:2585–2587.

40. Tootell RBH, Hadjikhani N, Hall EK, et al. The retinotomy of visual spatial attention. *Neuron* 1998;21:1409–1422.

41. Martínez A, Anllo-Vento L, Sereno MI, et al. Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nature Neurosci* 1999:364–369.

42. Müller MM, Teder-Sälejärvi WA, Hillyard SA. The time course of cortical facilitation during cued shifts of spatial attention. *Nature Neurosci* 1998;1:631–634.

43. Hillyard SA, Teder-Sälejärvi WA, Münte TF. Temporal dynamics of early perceptual processing. *Curr Opin Neurobiol* 1998;8:202–210.

44. Deecke L, Scheid P, Kornhuber HH. Distribution of readiness potential, pre-motion positivity, and motor potential of the human cerebral cortex preceding voluntary finger movements. *Exp Brain Res* 1969;7:158–168.

45. Böcker KB, Brunia CH, Cluitmans PJ. A spatio-temporal dipole model of the readiness potential in humans. II. Foot movement. *Electroencephalogr Clin Neurophysiol* 1994;91:286–294.

46. Kutas M, Donchin E. Preparation to respond as manifested by movement-related brain potentials. *Brain Res* 1980;202:95–115.

47. Coles MGH, Smid HGOM, Scheffers MK, et al. Mental chronometry and the study of human information processing. In: Rugg MD, Coles MGH, eds. *Electrophysiology of mind: event-related brain potentials and cognition*. New York: Oxford University Press, 1995:86–131.

48. Fabiani M, Gratton G, Coles MGH. Event-related brain potentials methods, theory and applications. In: Cacioppo J, Tassinary L, Bernston G, eds. *Handbook of psychophysiology*. Cambridge: Cambridge University Press, 2000:53–84.

49. Oishi M, Mochizuki Y, Du C, et al. Contingent negative variation and movement-related cortical potentials in Parkinsonism. *Electroencephalogr Clin Neurophysiol* 1995;95:346–349.

50. Dick JP, Rothwell JC, Day BL, et al. The Bereitschaftspotential is abnormal in Parkinson's disease. *Brain* 1989;112:233–244.

51. Cunnington R, Iansek R, Bradshaw JL. Movement-related potentials in Parkinson's disease: external cues and attentional strategies. *Mov Disord* 1999;14:63–68.

52. Adler LE, Pecevich M, Nagamoto H. Bereitschaftspotential in tardive dyskinesia. *Mov Disord* 1989;4:105–112.

53. Trenkwalder C, Bucher SF, Oertel WH, et al. Bereitschaftspotential in idiopathic and symptomatic restless legs syndrome. *Electroencephalogr Clin Neurophysiol* 1993;89:95–103.

54. Scheffers MK, Coles MG. Performance monitoring in a confusing world: error-related brain activity, judgements of response accuracy, and types of errors. *J Exp Psychol Hum Percept Perform* 2000;26:141–151.

55. Gehring WJ, Knight RT. Prefrontal-cingulate interactions in action monitoring. *Nature Neurosci* 2000;3:516–520.

56. Luu P, Collins P, Tucker DM. Mood, personality, and self-monitoring: negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *J Exp Psychol Gen* 2000;129:43–60.

57. Dikman ZV, Allen JJ. Error monitoring during reward and avoidance learning in high- and low-socialized individuals. *Psychophysiology* 2000;37:43–54.

58. Donchin E, Coles MGH. Is the P300 component a manifestation of context updating. *Behav Brain Science* 1988;11:357–374.

59. Verleger R. Event-related potentials and cognition: a critique of the context updating hypothesis and an alternative interpretation of P3. *Behav Brain Science* 1988;11:343–427.

60. Picton TW. The P300 wave of the human event-related potential. *J Clin Neurophysiol* 1992;9:456–479.

61. Polich J, Kok A. Cognitive and biological determinants of P300: an integrative review. *Biol Psychol* 1995;41:103–146.

62. Bashore TR, Ridderinkhof KR, van der Molen MW. The decline of cognitive processing speed in old age. *Curr Direct Psychol Science* 1997;6:163–169.

63. Luck SJ. Sources of dual-task interference: evidence from human electrophysiology. *Psychol Sci* 1998;9:223–227.

64. Rugg MD, Coles MGH, eds. *Electrophysiology of mind*. Oxford: Oxford University Press, 1995.

65. McCarthy G, Wood CC, Williamson PD, et al. Task-dependent field potentials in human hippocampal formation. *J Neurosci* 1989;9:4253–4268.

66. Frodl-Bauch T, Bottlender R, Hegerl U. Neurochemical substrates and neuroanatomic generators of the event-related P300. *Neuropsychobiology* 1999;40:86–94.

67. Luck SJ, Hillyard SA. The operation of selective attention at multiple stages of processing: evidence from human and monkey electrophysiology. In: Gazzaniga MS, ed. *The new cognitive neurosciences*. Cambridge, MA: MIT Press, 2000:687–700.

68. Knight RT, Nakada T. Cortico-limbic circuits and novelty: a review of EEG and blood flow data. *Rev Neurosci* 1998;9:57–70.

69. Ruchkin DS, Berndt RS, Johnson RJ, et al. Modality-specific processing streams in verbal working memory: evidence from spatio-temporal patterns of brain activity. *Brain Res Cogn Brain Res* 1997;6:95–113.

70. Karis D, Fabiani M, Donchin E. "P300" and memory: individual differences in the von Restorff effect. *Cog Psychol* 1984;16:177–216.

71. Paller KA, Kutas M, Mayes AR. Neural correlates of encoding in an incidental learning paradigm. *Electroencephalogr Clin Neurophysiol* 1987;67:360–371.

72. Paller KA. Recall and stem-completion priming have different electrophysiological correlates and are modified differentially by directed forgetting. *J Exp Psychol Learn Mem Cogn* 1990;16:1021–1032.

73. Van Petten C, Senkfor AJ. Memory for words and novel visual patterns: repetition, recognition and encoding effects in the event-related brain potential. *Psychophysiology* 1996;33:491–506.

74. Wagner AD, Koutstaal W, Schacter DL. When encoding yields remembering: insights from event-related neuroimaging. *Philos Trans R Soc Lond B Biol Sci* 1999;354:1307–1324.

75. Düzel E, Cabeza R, Picton TW, et al. Task-related and item-related brain processes of memory retrieval. *Proc Natl Acad Sci USA* 1999;96:1794–1799.

76. Rugg MD, Allan K. Event-related potential studies of memory. In: Tulving E, Craik FIM, eds. *The Oxford handbook of memory*. New York: Oxford University Press, 2000:521–537.

77. Rösler F, Heil M, Henninghausen E. Slow potentials during long-term memory retrieval. In: Heinze HJ, Münte TF, Mangun GR, eds. *Cognitive electrophysiology*. Boston: Birkhauser, 1994:149–168.

78. Olichney J, Van Petten C, Paller KA, et al. Word repetition in amnesia: electrophysiological measures of impaired and spared memory. *Brain* 2000;123:1948–1963.

79. Rugg MD, Walla P, Schloerscheidt AM, et al. Neural correlates of depth of processing effects on recollection: evidence from brain potentials and positron emission tomography. *Exp Brain Res* 1998;123:18–23.

80. Paller KA, Gross M. Brain potentials associated with perceptual priming vs explicit remembering during the repetition of visual word-form. *Neuropsychologia* 1998;36:559–571.

81. Senkfor AJ, Van Petten C. Who said what? An event-related potential investigation of source and item memory. *J Exp Psychol Learn Mem Cogn* 1998;24:1005–1025.

82. Rubin SR, Petten CV, Glisky EL, et al. Memory conjunction errors in younger and older adults: event-related potential and neuropsychological data. *Cog Neuropsychol* 1999;16:459–488.

83. Wilding EL, Rugg MD. An event-related potential study of recognition memory with and without retrieval of source. *Brain* 1996;119:889–905.

84. Janowsky JS, Shimamura AP, Squire LR. Memory and metamemory: comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology* 1989;17:3–11.

85. Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 1980;207:203–205.

86. Van Petten C, Kutas M. Influences of semantic and syntactic context on open and closed class words. *Mem Cogn* 1991;19:95-112.

87. Kutas M, Federmeier K, Coulson S, et al. Language. In: Cacioppo J, Tassinary L, Bernston G, eds. *Handbook of psychophysiology*. Cambridge: Cambridge University Press, 2000:576–601.

88. van Berkum JJA, Hagoort P, Brown CM. Semantic integration in sentences and discourse: evidence from the N400. *J Cog Neurosci* 1999;11:657–671.

89. Federmeier KD, Kutas M. Right words and left words: electrophysiological evidence for hemispheric differences in language processing. *Brain Res Cogn Brain Res* 1999;8:373–392.

90. Swaab TY, Brown CM, Hagoort P. Spoken sentence comprehension in aphasia: event-related potential evidence for a lexical integration deficit. *J Cog Neurosci* 1997;9:39–66.

91. Nobre AC, Allison T, McCarthy G. Word recognition in the human inferior temporal lobe. *Nature* 1994;372:260–263.

92. Osterhout L, Holcomb PJ. Event-related brain potentials elicited by syntactic anomaly. *J Mem Lang* 1992;3:785–806.

93. Hagoort P, Brown CM, Groothusen J. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Lang Cogn Proc* 1993;8:439–483.

94. Münte TF, Matzke M, Johannes S. Brain activity associated with syntactic incongruencies in words and pseudo-words. *J Cog Neurosci* 1997;9:318–329.

95. Coulson S, King JW, Kutas M. Expect the unexpected: event-related brain response to morphosyntactic violations. *Lang Cogn Proc* 1998;13:21–58.

96. Hahne A, Friederici AD. Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *J Cog Neurosci* 1999;11:194–205.

97. Kluender R, Kutas M. Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. *J Cog Neurosci* 1993;5:196–214.

98. Just MA, Carpenter PA. A theory of reading: from eye fixations to comprehension. *Psychol Rev* 1980;87:329–354.

99. Van Petten C, Coulson S, Rubin S, et al. Time course of word identification and semantic integration in spoken language. *J Exp Psychol Learn Mem Cogn* 1999;25:394–417.

100. Steinhauer K, Alter K, Friederici AD. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neurosci* 1999;2:191–196.

101. Münte TF, Schiltz K, Kutas M. When temporal terms belie conceptual order. *Nature* 1998;395:71–73.

102. Levelt WJM. Speaking: from intention to articulation. Cambridge: MIT Press, 1989.

103. Schmitt B, Münte TF, Kutas M. Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology* 2000 Jul;37:473–484.

104. van Turennout M, Hagoort P, Brown CM. Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *J Exp Psychol Learn Mem Cogn* 1997;23:787–806.

105. Münte TF, Urbach TP, Duzel E, et al. Event-related brain potentials in the study of human cognition and neuropsychology. In: Boller F, Grafman J, Rizzolatti G, eds. *Handbook of neuropsychology*. Amsterdam: Elsevier Science, 2000:1–97.

106. Ahlfors SP, Simpson GV, Dale AM, et al. Spatio-temporal activity of a cortical network for processing visual motion revealed by MEG and fMRI. *J Neurophys* 1999;82:2545–2555.

107. Makeig S, Westerfield M, Jung T-P, et al. Functionally independent components of the late positive event-related potential during visual spatial attention. *J Neurosci* 1999;19:2665–2680.

Section IV
Drug Discovery and Evaluation

Herbert Y. Meltzer

Drug Discovery and Evaluation - Introduction

It is fitting to include in *Neuropsychopharmacology: The Fifth Generation of Progress* a group of chapters concerned with many of the important issues related to: (a) the development of new drugs (e.g., the role of preclinical models to develop and test targets for new drugs; (b) the role of biological markers and imaging studies to provide indices of drug action in humans or predict clinical response; (c) advances in clinical trial design; (d) the optimal utilization of existing drugs (e.g., pharmacokinetics); (e) pharmacogenomics in relation to drug metabolism, discovery, and development; (f) the ethical issues concerning clinical trials and neuropsychiatric research in general; (g) governmental (at least United States) regulation of the process of developing and utilizing new drugs; and (h) the evaluation of the impact of drug treatment on outcome of neuropsychiatric disorders from the perspectives of economics, clinical endpoints, and humanistic considerations. Although the majority of these issues are the same in many parts of the world where this book will be read, there are clear differences in ethical attitudes toward clinical research and governmental regulation of drug development and utilization around the world (despite the Declaration of Helsinki). It is beyond the scope of this volume to consider the worldwide variations in these issues. All the chapters that have preceded this section bear on the subject matter addressed here because new drug development depends so heavily on our understanding of the function of the brain in general, neurotransmitter and modulator receptors and elimination mechanisms, theories of the etiology of the major neuropsychiatric disorders, efficacy and side effects of existing treatments, and an understanding of the mechanism of action of existing drugs.

The process of new drug development has changed greatly in the few years since the last volume in this series and is likely to change even more rapidly in the immediate future. A great advance was the development of combinatorial chemistry and rapid, robotic characterization of the pharmacologic profile of the vast libraries of compounds produced by this shotgun approach, which, when successful, leads to the elegant and expensive custom syntheses of candidate compounds by sophisticated organic chemical procedures that are now often computer-derived or by methods involving cell and molecular biology to produce peptides or other organic substances. Together with greatly improved methods for analysis of structure-activity relationships, it has been possible to develop putative pharmaceuticals with the desired pharmacologic profile. The old cliché, beware what you desire because you may get it, is relevant here, because it much easier now to come up with the desired pharmacologic profile than it is to be certain that what is sought is what should be sought. There is not yet sufficient understanding of what is needed in the way of an optimal antidepressant, anxiolytic, antipsychotic, mood stabilizer, antidementia, or other type of drug, especially when a truly novel compound is sought.

The chapter by Geyer and Markou on the role of preclinical models in the development of psychotropic drugs mainly focuses on animal models for the major psychiatric disorders. These authors point out that this approach may seem somewhat old-fashioned compared to approaches such as high throughput screening and utilization of molecular biological techniques to develop targets based on gene expression and identification methods; however, they correctly state that preclinical models are *required* (emphasis added) to provide initial assessment of the functional effects of novel compounds in the integrated organism. We are not yet to the point where new chemical entities go directly into patients or even normal volunteers without some evidence that clinically relevant effects might be present. We

can expect major advances in the development of preclinical models as our knowledge of disease processes and our ability to alter the genome in laboratory animals increase. Knockout and knockin mouse models will increasingly guide drug discovery and testing. The importance of research designed to identify new drug targets based on the Human Genome Project and the ensuing effort to characterize the genes involved in neuropsychiatric disorders and the action of drugs used to treat neuropsychiatric illness is discussed in various chapters throughout this volume rather than in a single chapter in this section.

The use of biomarkers (i.e., natural history markers), biological activity marker, and surrogate markers is thoroughly explored by Wong and colleagues, who note that the importance of biomarkers as a means to reduce the cost of drug development, improve the ability to predict outcome, and expedite the identification of desired endpoints (e.g., no more than 80% occupancy of striatal $_2$ receptors in order to minimize the development of extrapyramidal side effects), is increasing all the time. The extraordinary development of a variety of brain imaging methods, including magnetic resonance imaging, functional magnetic resonance imaging, single photon emission computed tomography, and positron emission tomography, appears to be particularly suited for this purpose. Given the cost associated with a failed clinical trial, someday it may be possible to bring brain imaging into routine clinical practice to guide drug dosage and choice. However, more classical methods such as neuroendocrine testing or examining the effect of treatments on peripheral processes such as changes in saliva, serum, and blood cells still can be valuable at various stages of drug development.

Clinical trials abound in psychiatry. Good clinical trials are much more rare. Problems in trial design, identification of appropriate patients, recruitment, retention, and ethical issues surrounding the use of placebos are very much with us and show signs of becoming more rather than less intractable in the near future. The cost of clinical trials in Western countries has grown enormously, leading to fewer and smaller trials that are often market-driven rather than designed to answer the most important research questions. Kane describes a number of efforts that have been made to improve clinical trial design and to cope with the increasing limitations that current ethical viewpoints have placed on this process. It is a sign of progress that broader outcome measures other than global psychopathology are increasingly the focus of clinical trials. For example, the recognition that cognition may be a more important endpoint than the reduction of positive or negative symptoms in the evaluation of a new drug for schizophrenia is an enormous advance because it refocuses the goal of new drug development and allows for distinguishing between new and existing drugs on a much more meaningful basis.

Pharmacokinetics, pharmacodynamics, and drug disposition in relation to new drug development and their subsequent utilization have never been better summarized than they are in the chapter by Greenblatt and colleagues. Advances in this area yield the information needed to use drugs wisely. This area of research has matured to the point where the fundamental principles are well understood and can be readily incorporated into the processes of drug development and utilization. Information about drug interactions that affect efficacy, elimination, and toxicity are ever more essential in the current area of polypharmacy.

Özdemir and colleagues discuss pharmacogenetics, the field that explores individual differences in drug responses that depend on genetic factors and genetic-environmental interactions. A key part of this refers to genetic variations in the liver enzymes that metabolize drugs. Greenblatt and colleagues consider this as well as the genetic factors that determine pharmacodynamics, and thus directly impact on efficacy and side effects. It is clear that this is critically important to psychopharmacology and will become even more so with the completion of the Human Genome Project. They also consider how genomic research will play an increasingly important role in drug discovery and development, including the design of safer and more efficient clinical trials. Their term "personalized therapeutics" is provocative. Have doctors not tried to do this since time immemorial? Genetic information will aid the process to be more science than art (rather than the reverse).

Is medicine in need of new ethical compasses in clinical research? This would seem to be the view of the authors of the United States-based National Bioethics Advisory Commission (NBAC) or the newest version of the Declaration of Helsinki. NBAC, in particular, has singled out research on the mentally ill for more stringent regulation and unique standards. The chapter on ethical aspects of neuropsychiatric research by Pinals and Appelbaum thoughtfully analyzes these recommendations as well as the fundamental principles that should guide policy in this area. Professional societies such as the ACNP have developed guidelines for investigators in an effort to show that there is an awareness of the obligation to protect subjects who agree to research to the greatest possible extent with minimal diminution of the information to be gained from the study. A "safe" study from which we can learn little or nothing of use because of design flaws may be less ethical than others where the absolute risk may be greater, but still within acceptable limits, whereas the potential gain in knowledge is far higher. Institutional review boards (IRB) are becoming increasingly restrictive around clinical research with the mentally ill, based in part on a poor understanding of the intactness of their decisional capacities. It is imperative that methods to assure competence to give consent that meets the legitimate concerns of IRBs are employed in all trials.

Paul Leber, formerly head of the Division of Psychopharmacology of the United States Food and Drug Administration (FDA), explicates the policies of his former employer with the oratorical flourish he is renowned for now transmuted into the written word with equal elegance. This chapter

explains what was (and probably still is) guiding the policies of the regulators of the FDA, which has worldwide influence directly and indirectly. This chapter can serve as a primer on getting a new drug application approved by that agency.

Mahmoud and colleagues succinctly and clearly describe the process of evaluating treatment outcomes utilizing the Economic, Clinical, Humanistic outcomes (ECHO) model. These three perspectives on outcome are intimately intertwined. Trouble arises when any one of the aspects is overemphasized to the neglect of the others. There has been increasing awareness of the need for societal consensus on the importance of outcomes as the cost of achieving the best outcomes now possible has risen greatly as a percentage of gross national product in both developing and developed countries. Advances in medical research will surely suffer if there is insufficient attention to demonstrating that new therapies are more cost-effective, not just more effective, than previous methods. The critical issue of distinguishing between efficacy and effectiveness research and the need for both are thoroughly discussed.

In conclusion, this section on new drug development and clinical research issues should be of great importance to any reader who is interested in the broader picture of alleviating neuropsychiatric disease burden through psychopharmacology.

33

The Role of Preclinical Models in the Development of Psychotropic Drugs

Mark A. Geyer

Athina Markou

Mark A. Geyer: Department of Psychiatry, School of Medicine, University of California, San Diego, La Jolla, California.

Athina Markou: Department of Neuropharmacology, The Scripps Research Institute, La Jolla, California.

This chapter critically discusses how preclinical models, primarily animal models, can be used in neurobiological research to promote the development of psychotropic drugs as therapeutics for psychiatric disorders. The authors' previous chapter in *Neuropsychopharmacology: The Fourth Generation of Progress* (1) extensively discussed the process of developing, validating, and working with animal models relevant to psychiatric disorders. Various approaches to model development and validation criteria for animal models were defined and evaluated. These basic principles of model development and validation were further elaborated in the context of reviewing animal models of depression and schizophrenia. The present chapter is intended as a continuation and addition to the previous chapter. Thus, assuming the fundamental principles of model development and validation established in the previous chapter and briefly reviewed here, the present chapter focuses on additional aspects of model development, validation, and use that need to be taken into consideration when using models as aids to the development of *therapeutic approaches* for psychiatric disorders. These principles are clarified further by discussing a few exemplary issues relating to animal models used in the study of depression, schizophrenia, and anxiety. Although the development of pharmacologic treatments for psychiatric disorders is typically the major focus, the same basic principles of model use also can be applied in the development of nonpharmacologic therapeutics for these disorders.

DEFINITION OF A PRECLINICAL MODEL

Part of "33 - The Role of Preclinical Models in the Development of Psychotropic Drugs "

A model is defined as any experimental preparation developed for the purpose of studying a condition in the same or different species. A model is comprised of both the independent variable (i.e., inducing manipulation) and the dependent variable (i.e., the measure[s] used to assess the effects of the manipulation). The choice of the inducing manipulation is usually based on hypotheses about the etiology of the disorder of interest or nontheoretic exploratory attempts to induce the abnormality (as reflected in the dependent measures) that is considered relevant to the psychiatric disorder of interest. Pathologies having homology with those in humans can be induced in animals more readily if the etiology of the disease is known. Unfortunately, the etiologies of psychiatric diseases are largely unknown, making the choice of the independent variable particularly difficult. The choice of the dependent measures is usually based on operational definitions of abnormalities believed to be pathognomonic, or at least symptomatic, of the disorder of interest. As with the inducing manipulations, the selection of diagnostic criteria and determination of the core features of a psychiatric disorder are also debatable. Thus, the selection of both the inducing manipulations and dependent measures that comprise a model of a psychiatric disorder are based largely on theoretic arguments regarding both the etiology and core aspects of the disorder. The choice of dependent variables is somewhat easier than the choice of the inducing manipulation because it can be based on operational definitions of observable aspects of the disease, even if the chosen measure is not a core symptom of the disorder.

Preclinical models could involve either human or nonhuman experimental preparations. Typically, models are nonhuman animal preparations that attempt to mimic a human condition, including human psychopathology. Nevertheless, as implied in the definition of a model provided above, preclinical models could also be human experimental preparations. Whether a human or a nonhuman model should be used depends largely on the purpose of the model and the experimental question of interest (see the following). The vast majority of preclinical models in use are nonhuman because such models provide two distinct advantages over

human preclinical models. First, nonhuman models enable the investigation of the neurobiology of the phenomena of interest using invasive techniques that cannot be used in humans. Second, if used properly, nonhuman animal models can significantly reduce the cost of drug development by increasing (or decreasing) the degree of confidence in a particular pharmacologic approach before undertaking expensive and time-consuming clinical trials in the psychiatric population of interest. Nevertheless, it should be clarified that human preclinical models can also contribute importantly to this latter goal, if used properly (see the following).

PURPOSES OF A PRECLINICAL MODEL

Part of "33 - The Role of Preclinical Models in the Development of Psychotropic Drugs "

In developing and assessing an animal model, it is imperative to consider the *explicit* purpose intended for the model (2), because the intended purpose determines the criteria that the model must satisfy to establish its validity and utility. For example, is the purpose of the experimental system to model specific signs and symptoms or to model the entire diagnostic syndrome? Is the purpose of the model to promote our basic understanding of the neurobiological, genetic, environmental, and other factors that contribute to a mental disorder or the development of therapeutic agents for this disorder? Is the purpose of the model to rapidly and efficiently screen compounds to identify drugs that may have similar therapeutic properties to an existing class of compounds, or the identification of therapeutic targets that may lead to the development of compounds having novel mechanisms of action? The preceding are just a few general examples of the various purposes that a model may be intended to fulfill. Such purposes and uses explicitly guide the development and validation process for a particular model. Following, the necessary and sufficient criteria for evaluating preclinical models are reviewed briefly. (See refs. 1 and 3 ,4 ,5 ,6 and 7 for more extensive discussions.) Then, some general issues about preclinical models are discussed that also relate to the premise that the intended purpose of a model determines the validation criteria that the model must satisfy.

NECESSARY AND SUFFICIENT CRITERIA FOR EVALUATING PRECLINICAL MODELS

Part of "33 - The Role of Preclinical Models in the Development of Psychotropic Drugs "

The validity of a model refers to the extent to which a model is useful for a given purpose. Thus, depending on the desired purpose of the test that one wishes to validate, different types of validity are relevant. Further, in considering the validity of a model, both the independent and dependent variables need to be evaluated. The reliability and predictive validity of the model system are relevant to both the independent variable and dependent measures and are the most important criteria to satisfy. (See refs. 1 and 8 ,9 and 10 for definitions of the various types of validity.) The additional criteria relevant to the independent variable (i.e., inducing manipulation) include etiologic, construct, and face validity, with etiologic validity being the most relevant. The criteria relevant to the dependent variable include construct, convergent, discriminant, and face validity. Undoubtedly, the more types of validity a model satisfies, the greater its value, utility, and relevance to the human condition. Nevertheless, it could be considered circular logic if a model was required to satisfy all types of validity before being considered useful. Hence, it has been argued previously that predictive validity and reliability are the only *necessary and sufficient* criteria for the *initial* evaluation of any animal model (1).

Predictive validity of a model is broadly defined as the ability to make accurate predictions about the human phenomenon of interest based on the performance of the model (1 ,9). In reference to animal models of human psychopathology, the term predictive validity is often used in a narrow sense to refer to the model's ability to identify drugs with potential therapeutic value in humans (i.e., pharmacologic isomorphism) (2). Although correct, this use of the term is limited because it ignores other important ways in which a model can lead to successful predictions (7). For example, the identification of any variables that have similar influences in both the experimental preparation and modeled phenomenon can demonstrate predictive value of the experimental preparation and enhance one's understanding of the phenomenon.

Reliability refers to the consistency and stability with which the variables of interest are observed, and is relevant to both the independent and dependent variables (1 ,5).

APPROACHES AND ISSUES RELATED TO MODEL DEVELOPMENT

Part of "33 - The Role of Preclinical Models in the Development of Psychotropic Drugs "

Modeling Specific Signs or Symptoms Versus Modeling Diagnostic Syndromes

As discussed previously (1), early attempts at model development focused on reproducing in animals a psychiatric syndrome in its entirety. Such an approach, although useful in advancing the field at the time, has been largely abandoned because of increasing awareness that such an approach is typically impractical, unrealistic, and fruitless for the following reasons. (a) The defining symptoms of psychiatric disorders and even the diagnostic categories themselves are being revised and redefined continuously. (b) One would not expect homology on *all* aspects of a disorder between two species (e.g., one would not expect a rodent to exhibit a complete schizophrenia syndrome). (c) Modeling a syndrome in its entirety would require the validation of multiple endpoints. (d) There is considerable heterogeneity within each of the major diagnostic categories of psychiatric disorders. (e) The validation process for such a model needs to be extensive, thorough, and all-inclusive, and thus not different from the scientific process aimed at elucidating

the neurobiological and behavioral mechanisms mediating a psychiatric disorder.

Most recent approaches to the development of animal models rely on mimicking only specific signs or symptoms associated with psychopathologic conditions, rather than mimicking an entire syndrome. These specific signs or symptoms may be: (a) observables that have been identified in psychiatric populations that may or may not be pathognomonic for or even diagnostic symptoms of the disorder, but can be defined objectively and measured reliably; or (b) more theoretically based measures of psychological constructs that are believed to be relevant to the psychiatric disorder under investigation (2 ,7). The latter approach involves the definition of a hypothetical construct and subsequent establishment of operational definitions suitable for the experimental testing of the validity of the construct in both human and nonhuman animals. The narrow focus of this approach generally leads to pragmatic advantages in the conduct of mechanistic studies addressing the neurobiological substrates of the specific behavior under study. Furthermore, the study of putatively homologous behaviors in both human and nonhuman subjects effectively addresses and bypasses the nonconstructive criticism that complex mental disorders cannot possibly be modeled in nonhuman animals. (See ref. 1 for a more extensive discussion comparing these two approaches to modeling.)

An illustrative example of this approach is provided by some of the models now being used to identify antipsychotic drugs, based on the hypothesis that schizophrenia involves deficits in attentional filtering or gating (i.e., the psychological construct). Theoretically, schizophrenia patients suffer from impairments in filtering or gating of sensory stimuli that lead to an inundation of information and consequent cognitive fragmentation. The hypothetical construct of attentional filtering has been defined operationally and explored in multiple paradigms and in both human and animal studies. For example, numerous studies of schizophrenia patients have demonstrated deficits in behavioral habituation, which is a prerequisite to selective attention, prepulse inhibition (PPI) of startle, a preattentional sensorimotor gating phenomenon, and the gating of auditory P50 event-related potentials (ERPs) (11 ,12 and 13) (see Chapter 51). Each of these operational measures is potentially relevant to the construct of deficient filtering of incoming information, hypothesized to be a common element in the schizophrenia disorders (14 ,15). Each of these operational measures is also amenable to cross-species studies of analogous or homologous behaviors. (See the following for a discussion of these terms.) The fact that schizophrenia patients exhibit deficits in all three measures provides converging support for the hypothesis that schizophrenia involves disturbances in the filtering of sensory and cognitive information (i.e., construct validity). Nevertheless, a recent study explicitly testing the convergent validity of this hypothetical construct has prompted some further refinements in our thinking. Specifically, in a group of normal subjects, P50 gating was strongly correlated with the amount of startle habituation and only weakly with PPI (16), despite the fact that P50 gating appears to be more similar phenomenologically to PPI than habituation. Similar findings have been reported in the parallel animal paradigms using the same operational measures (17). This situation illustrates how phenomenologic similarity (i.e., face validity) can sometimes lead to erroneous conclusions until further detailed behavioral investigations of the construct(s) are undertaken. Furthermore, as reviewed elsewhere (see Chapter 50), habituation, PPI, and P50 gating exhibit some differences as well as similarities in their sensitivity to pharmacologic manipulations used to mimic schizophrenia-like changes in animals. Of critical importance is the relationship, if any, between these experimental measures of filtering deficits and clinical complaints of sensory overload or signs of thought disorder that prompted the original hypothetical construct (i.e., extrapolation from animals to humans). Surprisingly, within a cohort of schizophrenia patients (19), those with deficient P50 sensory gating reported fewer complaints of sensory overload than did patients with normal P50 gating (i.e., the opposite of the predicted relationship). With regard to the PPI sensorimotor gating measure, however, significant correlations have been observed between deficient PPI, and both distractibility (20) and measures of thought disorder based on an abstract problem-solving task (21) in schizophrenia patients. Hence, it appears that the three main operational measures of deficient attentional filtering do not all measure the same hypothetical construct. Thus, in parallel with the heterogeneous group of schizophrenia-like disorders, the construct of deficient filtering may not be a unitary construct, although it could still represent a phenomenologically common outcome of differing etiologies in different forms of schizophrenia. It is important to recognize that each of these measures is demonstrably affected in (presumably heterogeneous) groups of schizophrenia patients and each has engendered animal models that have varying degrees of predictive validity for the identification of antipsychotic treatments. It remains to be seen whether different subgroups of schizophrenia patients will exhibit only one or another of these deficits. If so, the parsing of the original hypothetical construct may lead to empirical distinctions among patient subgroups that could have important implications for the application of specific treatment approaches.

Discovery of Novel Versus "Me-Too" Treatments

Another extreme approach to model development and use relates to the limited purpose of systematically and efficiently screening and identifying potential therapeutic treatments without explicitly assessing the mode of action that leads to the therapeutic effect. In such a case, the model may or may not mimic the actual psychiatric disorder.

Rather, the model is only intended to reflect the efficacy of known therapeutic agents, and consequently lead to the discovery of new pharmacotherapies. Thus, the principle guiding this approach has been termed “pharmacologic isomorphism” (2). As discussed elsewhere (2 ,7), the fact that such models are developed and validated by reference to the effects of known therapeutic drugs frequently limits their ability to identify new drugs having novel mechanisms of action. Accordingly, an inherent limitation of this approach is that it is not designed to identify new therapeutics that may treat either the symptoms of the disorder that are refractory to current treatments, or patient populations that are resistant to current treatments. An example of such a limitation is found in the use of drug-discrimination paradigms used to identify new treatment compounds. In these paradigms, the animal is trained to recognize the drug state induced by a prototypical drug. Typically, the animal is required to press either the right or left of two levers, depending on whether it had been treated with the vehicle or training drug. Potential new therapeutics are then identified by their ability to substitute for the prototypical drug on which the animal was trained. Because these paradigms rely only on the subjective drug-induced cue to which each animal responds and not on an endpoint that can be validated by reference to other behaviors in animals or humans, such procedures can only identify drugs having a similar effect on some unknowable dimension. If the screening involves several paradigms, the profile of the drug can be compared qualitatively and quantitatively to the profiles of known compounds. Such profiles, when combined with “a special kind of flair for the problem” (22), may lead to reasonable predictions about the potential of the compound in the clinic. The ability to rapidly and efficiently identify treatments that may be shown clinically to have some advantages over the older treatments is an advantage of this approach. Nevertheless, these screening paradigms do not provide ways to predict whether the “me-too” drug will have any clinical advantages (e.g., fewer side effects, treatment of refractory symptoms or patient populations) over the “prototypes,” other than in relation to potency.

Modeling Specific Aspects of Treatment Effects: Chronic Versus Acute Drug Treatments

Because both the etiologies and core features of psychiatric disorders are still poorly understood, much research addressing the neurobiology of these disorders has focused on the study of the mode of action of known therapeutics. The targets of clinically effective therapeutics have provided excellent starting points in the investigation of the neurobiology of psychiatric disorders. When taking this approach, it is important that specific aspects of the treatment effects are taken into consideration and incorporated into the paradigms used. For example, it is recognized that chronic treatment with antidepressants is required before a therapeutic effect is observed. This therapeutic delay is not only a severe limitation of current antidepressant treatments but also a hurdle in determining the mechanisms through which antidepressants produce their beneficial effects. Because of this delay in the emergence of the therapeutic effects, it is assumed that these effects are mediated by neuroadaptations that develop as a result of the chronic drug administration. Much research has focused on these neuroadaptations in order to understand the neurobiology of depression; because the therapeutic effect may be produced through “normalization” of the specific abnormalities characterizing depression. It is possible, however, that the therapeutic effect may be produced by separate systems or mechanisms that counteract the abnormalities that are etiologic in depression.

The preceding discussion is relevant, not only to approaches that may be taken in studying psychiatric disorders, but also to the question of whether animal paradigms that demonstrate positive results after acute administration of an established antidepressant are indeed valid models of depression rather than just screening paradigms. It could be argued that with acute drug administration the mechanisms leading to the reversal of the behavioral deficit are not the same as the ones leading to the clinical therapeutic effect. Such arguments certainly have merit. An animal paradigm that not only indicates therapeutic efficacy but also the time-course of such effects is a powerful tool for both neurobiological investigations and drug discovery. The vast majority of animal models of depression do not readily satisfy this criterion despite extensive efforts over decades. Thus, the question is how to best design and interpret data from paradigms that appear to reveal primarily acute therapeutic effects. In many of the animal studies, the acute drug doses are much higher than doses that would be tolerated by humans, especially on the first drug administration. Higher doses may be more likely to produce an acute effect. This argument is supported by data with the forced swim model where it was shown that either high doses of antidepressant drugs or chronic treatment with low doses of antidepressants, ineffective when administered acutely, reversed immobility in the swim test (23). Further, it has been argued that antidepressants may produce immediate improvement of *some* symptoms in humans, but this acute effect may be hard to detect statistically because the initial improvement may be small and seen only in some, but not all, symptoms (23 ,24). Thus, it is possible that reversal of a specific behavioral deficit in a model after acute treatment may indeed be consistent with the clinical reality about a specific symptom. This experimental question is an example of a case in which preclinical animal data could guide the design of clinical investigations that would help assess and improve animal models. Finally, animal models that can only detect acute effects, when guided by good working hypotheses, can be used for target identification by investigating the mechanisms

that lead to reversal or exacerbation of the deficit of interest.

Issues Regarding the Use of "Normal" (Healthy) Versus Perturbed Animals

Although most animal models rely on an explicit inducing manipulation, some models test the effects of putative therapeutic compounds under baseline conditions, that is, without first inducing a deficit in the subjects. Even if such an animal model has predictive validity, it may not be useful in furthering our understanding of the pathophysiology underlying the disorder if the effects of therapeutic treatments depend on an interaction with the underlying pathophysiology. Specifically, the mechanisms through which drugs produce their effects in "normal" versus perturbed animals may differ, even if the primary neurochemical effect may be the same (25 ,26). Indeed, antidepressant drugs have been shown in some studies to have no effects in "normal" human or nonhuman subjects, whereas inducing "therapeutic" effects in patients or perturbed subjects. For example, fluoxetine, a selective serotonin reuptake inhibitor (SSRI) with antidepressant properties, does not induce euphoria or elevate mood in nondepressed healthy individuals even after chronic treatment (27 ,28). Similarly, treatment with tricyclic antidepressants or SSRIs has no effect on behaviors assessing reward function in unperturbed animals. More specifically, treatments with desmethylimipramine, a tricyclic antidepressant, or low doses of fluoxetine typically have no effect on intracranial self-stimulation reward thresholds (29 ,30 ,31 ,32 ,33 ,34 ,35 ,36 and 37), whereas reversing reward deficits observed after drug withdrawal or during chronic mild stress (31 ,32 ,38). Moreover, recent findings indicated that the coadministration of fluoxetine together with a relatively selective serotonin 1A receptor antagonist had *opposite* effects in "normal" rats (i.e., decreased reward), whereas reversing reward deficits (i.e., increased reward) in perturbed animals in which a reward deficit had been induced (34). Recent clinical and preclinical findings have suggested that the coadministration of a SSRI together with a serotonin 1A receptor antagonist leads to accelerated or augmented antidepressant effects compared to those seen after treatment with the SSRI alone, presumably by enhancing serotonergic neurotransmission to levels above those seen with the SSRI alone (39 ,40 and 41). In conclusion, the study of animals that exhibit a deficit that is pathognomonic of depression, rather than "normal" healthy animals, may be critical to the study of both the underlying pathophysiology and its treatment.

In other situations, however, it may be advantageous to utilize animal models that examine baseline behaviors. For example, known antipsychotic drugs can be identified with reasonable predictive power using the conditioned avoidance response paradigm (42). The conditioned avoidance response paradigm has then been applied to the testing of potentially novel mechanisms that may have efficacy in the treatment of psychosis (43). Latent inhibition is another paradigm in which antipsychotics produce changes in baseline behavior that are relevant to schizophrenia. Acutely ill schizophrenia patients exhibit deficits in latent inhibition that are reduced by antipsychotic treatment (44). Similarly, when the appropriate testing parameters are used in the analogous animal paradigm, both typical and atypical antipsychotics improve measures of latent inhibition under baseline conditions (45). In contrast to pharmacologically induced models (see the following), models such as the conditioned avoidance response and latent inhibition, in which known antipsychotics influence behaviors under baseline conditions, may be more effective in identifying new therapeutic targets for antipsychotic effects. As discussed elsewhere (see Chapter 50), most of the schizophrenia animal models used historically to identify antipsychotic agents have relied on the induction of abnormal behaviors by the administration of a dopaminergic agonist and then define an antipsychotic as a drug that reverses the agonist effect. Hence, most such models are primarily and perhaps exclusively sensitive to drugs that block dopamine receptors and may not detect novel mechanisms that could have efficacy without involving dopamine receptor blockade. Ultimately, only further research with each class of psychiatric treatments will determine the relative utility of models that use an explicit inducing condition versus models that rely on changes in baseline behavior.

Inducing Conditions: Drug-Induced Versus Nonpharmacologic and Genetic Models

With the exception of paradigms assessing treatment effects in nonperturbed animals, the inducing manipulations constitute a critical aspect of a model. The selection of inducing conditions for animal models of psychiatric disorders is difficult because the etiologies of psychiatric disorders are generally unknown and are likely to be heterogeneous within each diagnostic category. Inducing conditions could be environmental manipulations, drug manipulations, lesions, genetic manipulations, or combinations of the preceding. Further, all of the preceding manipulations could be implemented during development or combined with developmental manipulations or factors. An inducing condition may be selected: (a) based on theoretic arguments about the environmental and/or neurobiological factors that lead to the disorder; (b) because it induces a deficit that is considered pathognomonic of the disorder of interest, even though no theoretic arguments are made about the etiology of the disorder; or (c) based on purely practical considerations about the predictive value of the model without theoretic arguments about either the etiology of the disorder or the relevance of the dependent measure to aspects of the symptomatology characterizing the disorder. The selection of

each inducing condition has advantages and disadvantages and is often based on the following considerations.

Acute or chronic drug manipulations have the advantage of probing the function of a specific receptor or neurotransmitter system that either is implicated in the etiology of the disorder or produces the desired deficit. The main disadvantage of an acute drug manipulation, and often of chronic drug manipulations, is that it readily leads to "receptor" or "neurotransmitter tautology." For example, a deficit induced by a specific receptor agonist is very likely to be reversed most effectively by a receptor antagonist at the same receptor, as in the case of dopamine agonist–antagonist interactions in most animal models of antipsychotic action. The same applies to neurotransmitter systems. Nevertheless, reversal of a drug-induced deficit by a compound acting on a different neurotransmitter system is a powerful indication of system interactions that may be relevant to the pathophysiology and/or treatment of the disorder. Chronic drug manipulations offer additional advantages and disadvantages. Chronic drug administration is likely to lead to compensatory adaptations to the acute effects of the drug that are likely to be longer lasting than the effects of a single drug administration and to involve additional systems that are not involved in the acute drug effects. Nevertheless, the resulting neuroadaptations may be irrelevant to the disorder unless there is a relationship between the deficits induced by the drug and etiology of the psychiatric symptoms.

An example of a pharmacologic model is the use of the reward deficits seen during withdrawal from a variety of drugs of abuse as a model of the core symptom of "diminished interest or pleasure" in rewarding stimuli that characterizes depression (46). In rats, converging evidence indicates that withdrawal from psychostimulant drugs is associated with reward deficits expressed as elevations in brain reward thresholds (33 ,47), decreased breaking-points under a progressive ratio schedule for a sucrose reinforcer (48), and decrements in motivation for sexual reinforcement (49). The advantage of this model is the induction of deficits in reward and motivational processes that are hypothesized to be, not only pathognomonic of depression, but also deficits expressed as negative symptoms of schizophrenia. Thus, these paradigms focus on the study of a hypothetical construct that may have relevance to core symptoms seen in multiple diagnostic categories. These deficits are most likely homologous to similar deficits seen in people abusing these drugs because the etiology of the deficit is the same in both the animal model and humans. Nevertheless, it is not known whether pharmacologically induced deficits in reward and motivational processes are homologous, or just analogous, to similar deficits seen in nondrug abusing psychiatric populations. That treatments with clinically effective antidepressants reverse the drug-induced reward deficits in both rats and humans suggests that the deficits may be homologous across species (32 ,34 ,38 ,50).

Environmental manipulations often induce only short-lasting deficits in healthy subjects because a healthy system is able to "bounce back" readily once the inducing conditions have been removed. Potential interactions, however, between the environmental manipulation and a genetic predisposition may lead to long-lasting behavioral or neurobiological changes having relevance to the disorder of interest. Finally, environmental manipulations are important to use and incorporate into animal models because it appears that psychiatric disorders often result from interactions between "nature" and "nurture" to a larger extent than most nonpsychiatric diseases. Another advantage of environmental manipulations is that such manipulations are likely to affect integrated brain functions rather than a single component of a system.

Lesion manipulations offer different advantages and disadvantages compared to environmental and drug manipulations. An advantage of lesion manipulations over chronic drug manipulations is that lesions may lead to deficits and/or neuroadaptations in a variety of brain systems rather than just the one or few affected by a drug. A disadvantage of traditional lesion manipulations is that the initial lesion manipulation in most cases is a rather large insult to a specific brain site. Thus, the circuitry affected is dependent on the interconnections of this specific brain site. Nevertheless, recent advances in genetic techniques are allowing very precise "lesions" (knockouts) or increased expression (knockins) of specific proteins in selected brain sites in adult animals. Such technological advances, when combined with more traditional behavioral and pharmacologic aspects of well-developed models, are likely to advance our understanding of psychiatric disorders.

Developmental manipulations are gaining in popularity primarily because there is increased awareness that many psychiatric disorders develop gradually through childhood and adolescence and are lifelong. In some cases, investigators combine one of the previously discussed inducing manipulations with a developmental manipulation (e.g., applying the inducing manipulations during development or in a genetically altered animal). For example, decreases in PPI of startle, an operational measure of sensorimotor gating deficits that are evident in patients with schizophrenia, have been demonstrated to result from socially isolating rats from weaning until after puberty (51). Social isolation of rats in early stages of development has been used to produce a variety of behavioral abnormalities that have been related to both schizophrenia and depression. Recent studies have shown that 6 to 8 weeks of social isolation during development, but not during adulthood (52), produces deficits in PPI that are at least partially reversible by the administration of neuroleptic dopamine antagonists (51) or by clinically effective atypical antipsychotics having antagonist activity at multiple receptors (53 ,54). Furthermore, postweaning isolation rearing of rats also results in deficits in the gating of the N40 event-related potential, that are analogous to the deficits in P50 gating observed in schizophrenia (55).

Because schizophrenia commonly emerges in early adulthood, developmental factors have provided the basis for some etiologic hypotheses (56 ,57). Hence, further study of the gating deficits produced by isolation rearing of rats may establish a nonpharmacologic and developmentally relevant animal model of the gating deficits observed in patients with schizophrenia. Potentially, in contrast to the drug-induced models of gating deficits, such a model might have etiologic validity and might be sensitive to antipsychotic drugs having novel mechanisms of action.

Genetic manipulations are popular because of the recent surge of interest in genetic contributions to psychiatric disorders. Such interest promises to enable the development of a class of animal models based on hypothesized etiologic validity. As specific genes and gene products become linked to specific disorders, molecular biologists will be able to generate mutant or transgenic animals having genetic abnormalities that are potentially homologous to those seen in humans. Behavioral and pharmacologic studies of these genetically engineered animals will then be important in identifying the phenotypic changes associated with the mutation, testing hypotheses about the etiology of the disease, and exploring potential therapeutic treatments. The combination of genetic and molecular biological approaches with behavioral and pharmacologic approaches may well revitalize interest in etiologically based models of psychiatric disorders. It is important to recognize that genetic manipulations necessarily begin with the fetus and often lead to compensatory adaptations throughout the course of development. Hence, developmental factors must be taken into account and studied when working with such an early genetic alteration. The latter is an example of a case where a technological limitation can lead to new creative ways of studying the function of a system and how it may contribute to our understanding of the processes mediating a disease.

The increased use of strain differences and genetically engineered mutants in drug discovery programs will necessitate both practical and conceptual modifications to the development and validation of animal models. Among the most fundamental differences between these genetic models and most previous models involves the distinction between trait and state measures. Most of the traditional models used to explore psychiatric treatments have relied on relatively short-term changes in the state of the animal, as modified by inducing manipulations such as stressors or drugs. In contrast, by definition, genetically based models rely on traits rather than states. For example, it has become commonplace to use approach/avoidance conflict tests to examine the possibility that gene knockout mice exhibit alterations in what is called “anxiety.” Approach/avoidance conflict tests, such as the elevated plus-maze or the light/dark box, have been widely used in rodent studies of anxiolytic drugs. Anxiolytic drugs increase approach behavior in such paradigms, presumably because they reduce the anxiety that competes with the animal's tendency to explore novel stimuli and environments. Such an observation with an unknown drug could as readily be interpreted as an increase in novelty seeking (i.e., approach) rather than a decrease in anxiety (i.e., avoidance). The fact that known anxiolytic drugs increase approach behavior has provided substantive validation of approach/avoidance conflict tests for the identification of changes in state anxiety. Accordingly, such conflict tests are now being used widely in the characterization of mutant mice in attempts to identify changes in trait anxiety. It should be recognized, however, that the validation of a measure as predictive of a change in state may or may not validate the measure as reflective of a change in the conceptually related trait. That is, the observation of a shift in approach/avoidance behavior in a mutant mouse that is similar to that produced by an anxiolytic drug cannot readily support the conclusion that the mutant mouse exhibits low levels of trait anxiety rather than high levels of approach behavior, as in the trait of high novelty seeking. Only by examining approach/avoidance behavior across a range of contexts can one determine which pole of the approach/avoidance conflict is altered in the mutant animal (58).

Dependent Measures: Value of Analogous and Homologous Measures Across Species

As with the choice of inducing manipulations, the choice of dependent measures is not simple when developing animal models of psychiatric disorders, primarily because the major and core features of human psychopathology are still poorly understood and still debatable. Thus, what should be considered an adequate or appropriate endpoint for a model in psychiatry is not always clear. Whenever possible, it is preferable to work with homologous rather than analogous endpoints. The terms analogy and homology originated in comparative anatomy and refer to the morphology and function of a structure. Structures or behaviors across species that are similar in origin (i.e., neurosubstrates), form, and function are termed homologous, whereas structures or behaviors that have different origins or neurosubstrates, superficially similar form, and have similar function are termed analogous (59). Another term that has been used to refer to analogous endpoints is isomorphism (60). Thus, in some sense, the terms homology and analogy refer to both the symptomatology and the underlying substrates that relate to the etiology. Although homologous measures are preferable, they are rare. Fortunately, analogous measures can also be valuable. It is because of the assumption of homology, or at least analogy, among the physiological and behavioral characteristics of various species that extrapolations can be made from nonhuman animals to humans (61). The establishment of multiple forms of validation for a particular model provides convergent evidence in support of the postulate of cross-species homology.

When developing an animal model related to a psychiatric

disorder, it is important to determine what features of the disorder(s) the experimental preparation is intended to model. Investigators often begin by attempting to identify the core features of the particular disorder of interest. Nevertheless, it is clear that appreciable diversity of both etiologies and symptom profiles exists within each of the major psychiatric diagnostic categories. Furthermore, very few specific symptoms are unique to any specific diagnosis, but occur in multiple diagnostic categories. Hence, it is most productive to focus on specific features observed in patients as endpoints for use in the development of animal models, rather than on clusters of symptoms. A related implication of this reasoning is that multiple different animal models, in terms of the endpoint used, may all be useful in parallel. Thus, it is advantageous to utilize an array of models rather than rely too heavily on any one model. The endpoints used in such models could be *in vivo* behaviors, biological markers, or *in vitro* behaviors of biological systems or preparations. Operational definitions, especially of *in vivo* behavioral measures, assist in determining the theoretic relationship between the observable and the construct of interest (62 ,63). Finally, the observables should be measured objectively and reliably.

Human Preclinical Models: Relationship between Animal and Human Phenomena

Human preclinical models can also contribute significantly to drug development. Unfortunately, such human models appear to be underutilized, and relatively little effort is focused on the development of such human models. An advantage of using human preclinical models is that one would not have to be concerned about cross-species generalizations. Nevertheless, even with human models, questions regarding the etiology of the disorder or the relationship between the dependent measure and the symptoms still need to be addressed using the same principles as when extrapolating across species. Relative to animal models, human preclinical models are necessarily more constrained by the additional ethical considerations regarding the use of humans in research.

The most typical example of a human preclinical model is when a drug-induced state is used in healthy volunteers to mimic some aspects of a disorder of interest. For example, the glutamate antagonist ketamine is used to induce a state that mimics some aspects of acute schizophrenia in healthy volunteers (64 ,65). Then, using brain imaging, psychological assessments, and pharmacologic interactions, the neurobiology of this drug-induced state can be studied to gain insight into the possible substrates underlying the psychotic state in schizophrenia patients. Such a human preclinical model can also play an important role in assessing novel treatments. For example, it has been found that the atypical antipsychotic clozapine reduces the exacerbation of symptoms in schizophrenic patients given ketamine (66). In contrast, typical antipsychotics such as haloperidol are ineffective in treating psychotic episodes induced by drugs such as ketamine or phencyclidine (PCP). Hence, studies of ketamine effects in either human or animal preclinical models may aid in the identification of additional atypical antipsychotics having efficacy in the treatment of patients who are nonresponsive to typical antipsychotics. Indeed, in the PPI models of schizophrenia, the disruptive effects of glutamate antagonists on PPI of startle are reversed by atypical, but not by most typical, antipsychotics (67). Interestingly, this effect of clozapine-like antipsychotics is mimicked by the putative antipsychotic M100907, a selective serotonin-2A antagonist (68). In general, one goal of translational research is to utilize the knowledge gained from human preclinical and clinical studies to guide invasive neurobiological studies in animals, which in turn can be translated back to the human clinical studies. In the present example, this strategy would suggest that further studies could now determine whether M100907 reverses disruptions in PPI produced by ketamine in healthy human volunteers. Furthermore, clinical trials of the efficacy of M100907 in schizophrenia could be designed to test the hypothesis that only schizophrenic patients whose deficits in PPI are reversed acutely by M100907 would respond clinically to prolonged treatment with M100907. This admittedly speculative example illustrates some of the potential advantages derived from the use of homologous, or at least analogous, measures in animal and human preclinical models as well as in clinical trials. Such translational research is needed in the field of psychiatric disorders in order to guide both the refinement of the animal models and the development of new drugs.

DRUG DISCOVERY AND DEVELOPMENT: PRECLINICAL MODELS AND CLINICAL TRIALS

Part of "33 - The Role of Preclinical Models in the Development of Psychotropic Drugs "

An emerging belief is that animal preclinical models represent a bottleneck in psychotropic drug discovery (69). Compared to high-speed chemical synthesis, high-throughput screening of libraries of compounds, and rapid gene seeking and sequencing techniques, the use of preclinical models as screening techniques appears slow. Nevertheless, such preclinical models of human psychopathology are required to provide initial assessments of the functional effects of novel compounds in the integrated organism. Only such *in vivo* functional measures can confirm predictions about the potential effects of psychotropic drugs in patients. It is unrealistic to attempt to go from the “test tube” to the clinic when attempting to treat complex mental, cognitive, and emotional disturbances that do not yet have clearly defined neurobiological substrates, or even correlates. The “relative paucity of preclinical behavioral models predictive of clinical efficacy” (69) reflects the paucity of our quantitative measures of the human phenomena related to psychiatric

disorders, as well as the limited investment in the development of animal behavioral models over the past few decades. Despite the excitement in the field of neuroscience about the recent progress made in understanding brain function (70), there is also an appreciation of how little is known about the neurobiology of psychiatric disorders compared to advances in other fields of medicine (71). Given the rapidity of techniques available to target discovery and drug screening efforts relative to the limited state of our knowledge about psychiatric disorders, the role of *in vivo* preclinical models as the intermediary between these extremes needs to be considered carefully. Some preclinical models, such as the tail suspension or swim tests for antidepressants and the prepulse inhibition test for antipsychotics, are amenable to relatively rapid screening without knowledge or understanding of the compounds' mechanism of action. Paradigms that are far more laborious can be used in the identification of new targets through the investigation of the interacting systems that contribute to the disorder's symptomatology or the therapeutic effects of established drug treatments. After identification of such novel targets, drug development efforts can be focused in identifying a compound with the desired mechanism of action and other desired properties, such as no toxicity and limited actions at systems that would produce side effects. Converging evidence from other basic research efforts would be crucial in such an undertaking. Even though the previously described process is time consuming and requires well-integrated multidisciplinary research efforts, this process may lead to the breakthroughs in psychiatric drug development that have been long awaited.

After a candidate drug has been identified through the use of both animal and human preclinical models and safety issues have been addressed, then the therapeutic efficacy of the compound is tested in the clinical population. Unfortunately, such clinical trials often do not have sufficient power, in the statistical and experimental design meaning of the term, to detect potentially beneficial effects of novel candidate compounds. Because of the high cost of drug development, pharmaceutical companies are interested in pursuing drugs that have the potential to be used in a large market that is often a broadly defined diagnostic category. This situation is aggravated by the fact that diagnostic categories in psychiatry are still rather crudely defined by rating scales rather than by objective and quantitative measures. For example, it is often assumed, at least implicitly, that there is diagnostic homogeneity within a particular patient population. It is also assumed that the boundaries of psychiatric categories as currently defined are rather absolute. In fact, most psychiatric disorders do not have clear pathological or biological markers and are defined as constellations of symptoms that are on a continuum with normality (71). Even though such diagnostic issues are constantly discussed and debated among psychiatrists, such issues are often put aside in clinical trials. Typically, the main focus in clinical trials is on the global measures of remission that are acceptable to regulatory agencies. Unfortunately, the reliance on rating scales in clinical trials provides little specific information that is useful in guiding either human or animal preclinical studies. Although understandable in view of economic forces, the infrequent use of a selected battery of scientifically established objective measures even in the early phases of clinical trials limits the further development of translational research involving cross-species comparisons and model validation. Hence, clinical trials do not benefit sufficiently from the scientific information provided by academic research and seldom provide the kind of empiric measures that are needed to adequately validate related animal models. Emphasis on multiple biological or psychological measures of disease progression with or without treatment with the drug of interest could potentially provide valuable information about the mechanisms that underlie various aspects or symptoms of the disease and thus lead to pragmatic advances in our understanding of the neurobiology of psychiatric disorders. Communication from the clinic back to the preclinical behavioral laboratory will enable the refinement of established models and the creation of new ones. In turn, understanding of the disorder could benefit if clinical trials included measures suggested to be relevant from preclinical research in either human or animal models. Another situation that limits progress in drug development is a recent movement to discourage clinical trials that include a placebo control group. Instead, the new compounds are expected to show greater efficacy compared to established therapeutics for the particular disorder. Overall, this state of affairs significantly limits the potential to identify new drugs that may: (a) be more beneficial than established treatments to a subpopulation of patients; or (b) produce global improvement through amelioration of symptoms that are not adequately assessed by the established measures of efficacy. Thus, it is difficult to make real advances in the development of new drugs because the current system encourages a circular logic and approach.

Another limitation of clinical trials that contributes to this circular approach is the absence of use of well validated, objective, and reliable measures of psychopathology in addition to the available clinical measures. As with animal models, clinical trials also need to incorporate measures that objectively and reliably assess specific psychological constructs or processes that appear to be altered in the population of interest. The validation of any animal model can be only as sound as the information available in the relevant preclinical human literature and the clinical literature (7). It is very fruitful when conceptually related experiments are undertaken in both the relevant patient population and the putative preclinical human and animal models. That is, studies of appropriate patients are needed to establish the operational definitions of the hypothetical construct, and the construct's relevance to the particular disorder. In concert, parallel studies of the theoretically homologous construct,

process, or dimension are required to determine the similarity of the animal model to the human phenomenon. Development of animal models requires parallel development of clinical measures that allow meaningful comparisons. Clinical studies need to be informed by results from animal studies as much as the reverse is true. An important and advantageous aspect of the approach described herein is that the validation of the hypothetical construct and its cross-species homology can be established by studies of normal humans and animals, in addition to psychiatrically disordered patients or experimentally manipulated animals. Thus, this approach adds to and benefits from the psychological and neurobiological literature relevant to the hypothetical construct upon which the model is based. In a sense, this approach explicitly recognizes that the experimental study of the disorder in humans involves as much of a modeling process as does the study of the disorder in an animal model. Thus, more translational science is needed to relate animal findings to humans and vice versa.

ACKNOWLEDGMENTS

Part of "33 - The Role of Preclinical Models in the Development of Psychotropic Drugs "

This work was supported by National Institute of Health (NIH) grants DA02925 and MH52885, and the VISN 22 Mental Illness Research, Education, and Clinical Center to MAG; and NIH grant DA11946, Tobacco-Related Disease Research Program Grant from the State of California 7RT-0004, and a Novartis research grant to AM. The authors thank Dr. Daniel Hoyer for his comments and input, and Mike Arends for computer and library searches and editorial assistance. This is publication 13601-NP of The Scripps Research Institute.

REFERENCES

1. Geyer MA, Markou A. Animal models of psychiatric disorders. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:787–798.

2. Matthysse S. Animal models in psychiatric research. *Progress Brain Res* 1986;65:259–270.

3. Geyer MA, Segal DS. Behavioral psychopharmacology. In: Judd LL, Groves PM, eds. *Psychobiological foundations of clinical psychiatry*. Philadelphia: JB Lippincott, 1991:1–15.

4. Hinde RA. The use of differences and similarities in comparative psychopathology. In: Serban G, Kling A, eds. *Animal models in psychobiology*. New York: Plenum, 1976:187–202.

5. Markou A, Weiss F, Gold LH, et al. Animal models of drug craving. *Psychopharmacology* 1993;112:163–182.

6. McKinney WT, Bunney WE. Animal models of depression. *Arch Gen Psychiatry* 1969;21:240–248.

7. Segal DS, Geyer MA. Chapter 45: Animal models of psychopathology. In: Judd LL, Groves PM, eds. *Psychobiological foundations of clinical psychiatry*. Philadelphia: JB Lippincott, 1985:1–14.

8. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 1959;56:81–105

9. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull* 1955;52:281–302.

10. Mosier CI. A critical examination of the concepts of face validity. *Educ Psychol Measure* 1947;7:191–205.

11. Braff DL, Geyer MA. Sensorimotor gating and schizophrenia: human and animal model studies. *Arch Gen Psychiatry* 1990;47:181–188.

12. Deleted in press.

13. Geyer MA, Braff DL. Startle habituation and sensorimotor gating in schizophrenia and related animal models. *Schizophrenia Bull* 1987;13:643–668.

14. McGhie A, Chapman J. Disorders of attention and perception in early schizophrenia. *Br J Med Psychol* 1961;34:103–116.

15. Venables PH. Input dysfunction in schizophrenia. In: Maher BA, ed. *Progress in experimental personality research*. New York: Academic Press, 1964:1–47.

16. Schwarzkopf SB, Lamberti JS, Smith DA. Concurrent assessment of acoustic startle and auditory P50 evoked potential measures of sensory inhibition. *Biol Psychiatry* 1993;33:806–814.

17. Ellenbroek BA, van Luijtelaar G, Frenken M, et al. Sensory gating in rats: lack of correlation between auditory evoked potential gating and prepulse inhibition. *Schizophrenia Bull* 1999;25:777–788.

18. Deleted in press.

19. Jin Y, Bunney WE, Sandman CA, et al. Is P50 suppression a measure of sensory gating in schizophrenia? *Biol Psychiatry* 1998;43:873–878.

20. Karper LP, Freeman GK, Grillon C, et al. Preliminary evidence of an association between sensorimotor gating and distractibility in psychosis. *J Neuropsychiatry Clin Neurosci* 1996,8:60–66.

21. Perry W, Geyer MA, Braff DL. Sensorimotor gating and thought disturbance measured in close temporal proximity in schizophrenic patients. *Arch Gen Psychiatry* 1999;56:277–281.

22. Janssen PAJ. Screening tests and prediction from animals to man. In: Steinberg H, ed. *Animal behaviour and drug action*. Boston: Little, Brown, 1964:264–268.

23. Detke MJ, Johnson J, Lucki I. Acute and chronic antidepressant drug treatment in the rat forced swimming test model of depression. *Exp Clin Psychopharmacol* 1997;5:107–112.

24. Frazer A. Antidepressant drugs. *Depression* 1994;2:1–19.

25. Murphy DL. Animal models for human psychopathology: observations from the vantage point of clinical psychopharmacology. In: Serban G, Kling A, eds. *Animal models of human psychobiology*. New York: Plenum, 1976:265–271.

26. Newman ME. The need for cautiously extrapolating results obtained with normal animals (healthy individuals) to depressed ones. *J Neurochem* 1998;70:2641–2642.

27. Barr LC, Heninger GR, Goodman W, et al. Effects of fluoxetine administration on mood response to tryptophan depletion in healthy subjects. *Biol Psychiatry* 1997;41:949–954.

28. Gelfin Y, Gorfine M, Lerer B. Effect of clinical doses of fluoxetine on psychological variables in healthy volunteers. *Am J Psychiatry* 1998;155:290–292.

29. Hall FS, Stellar JR, Kelley AE. Acute and chronic desipramine treatment effects on rewarding electrical stimulation of the lateral hypothalamus. *Pharmacol Biochem Behav* 1990;37:277–281.

30. Fibiger HC, Phillips AG. Increased intracranial self-stimulation in rats after long-term administration of desipramine. *Science* 1981;214:683–685.

31. Moreau J-L, Jenck F, Martin JR, et al. Antidepressant treatment prevents chronic unpredictable mild stress-induced anhedonia as assessed by ventral tegmentum self-stimulation in rats. *Eur J Pharmacol* 1992;2:43–49.

32. Markou A, Hauger RL, Koob GF. Desmethylimipramine attenuates cocaine withdrawal in rats. *Psychopharmacology* 1992;109:305–314.

33. Lin D, Koob GF, Markou A. Differential effects of withdrawal from chronic amphetamine or fluoxetine administration on brain stimulation reward: interactions between the two drugs. *Psychopharmacology*, 1999;145:283–294.

34. Harrison AA, Liem YTB, Markou A. Fluoxetine combined with a serotonin-1A receptor antagonist reversed reward deficits observed during nicotine and amphetamine withdrawal in rats. *Neuropsychopharmacology* 2001;25:55–71.

35. McCarter BD, Kokkinidis L. The effects of long-term administration of antidepressant drugs on intracranial self-stimulation responding in rats. *Pharmacol Biochem Behav* 1988;31:243–247.

36. Valentino DA, Riccitelli AJ, Dufresne RL. Chronic DMI reduces thresholds for brain stimulation reward in the rat. *Pharmacol Biochem Behav* 1991;39:1–4.

37. Lee K, Kornetsky C. Acute and chronic fluoxetine treatment decreases the sensitivity of rats to rewarding brain stimulation. *Pharmacol Biochem Behav* 1998;60:539–544.

38. Kokkinidis L, Zacharko RM, Predy PA. Post-amphetamine depression of self-stimulation responding from the substantia nigra: reversal by tricyclic antidepressants. *Pharmacol Biochem Behav* 1980;13:379–383.

39. McAskill R, Mir S, Taylor D. Pindolol augmentation of antidepressant therapy. *Br J Psychiatry* 1998;173:203–208.

40. Artigas F, Romero L, de Montigny C. Acceleration of the effect of selected antidepressant drugs in major depression by 5-HT_{1A} antagonists. *Trends Neurosci* 1996;19:378–383.

41. Blier P, de Montigny C. Current advances and trends in the treatment of depression. *Trends Pharmacol Sci* 1994;5:220–226.

42. Wadenberg ML, Hicks PB. The conditioned avoidance response test re-evaluated: is it a sensitive test for the detection of potentially atypical antipsychotics? *Neurosci Biobehav Rev* 1999;23:851–862.

43. Hertel P, Fagerquist MV, Svensson TH. Enhanced cortical dopamine output and antipsychotic-like effects of raclopride by alpha2 adrenoceptor blockade. *Science* 1999;286:105–107.

44. Gray JA. Integrating schizophrenia. *Schizophrenia Bull* 1998;4:249–266.

45. Ruob C, Weiner I, Feldon J. Haloperidol-induced potentiation of latent inhibition: interaction with parameters of conditioning. *Behav Pharmacol* 1998;9:245–253.

46. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington, DC: American Psychiatric Press, 1994.

47. Markou A, Koob GF. Postcocaine anhedonia: an animal model of cocaine withdrawal. *Neuropsychopharmacology* 1991;4:17–26.

48. Barr AM, Phillips AG. Withdrawal following repeated exposure to d-amphetamine decreases responding for a sucrose solution as measured by a progressive ratio schedule of reinforcement. *Psychopharmacology* 1999;141:99–106.

49. Barr AM, Fiorino DF, Phillips AG. Effects of withdrawal from an escalating dose schedule of d-amphetamine on sexual behavior in the male rat. *Pharmacol Biochem Behav* 1999;64:597–604.

50. Nunes EV, Quitkin FM, Donovan SJ, et al. Imipramine treatment of opiate-dependent patients with depressive disorders: a placebo-controlled trial. *Arch Gen Psychiatry* 1998;55:153–160.

51. Geyer MA, Wilkinson LS, Humby T, et al. Isolation rearing of rats produces a deficit in prepulse inhibition of acoustic startle similar to that in schizophrenia. *Biol Psychiatry* 1993;34:361–372.

52. Wilkinson LS, Killcross, AS, Humby T, et al. Social isolation produces developmentally specific deficits in prepulse inhibition of the acoustic startle response but does not disrupt latent inhibition. *Neuropsychopharmacology*, 1994;10:61–72.

53. Bakshi VP, Swerdlow NR, Braff DL, et al. Reversal of isolation-rearing induced deficits in prepulse inhibition by Seroquel and olanzapine. *Biol Psychiatry* 1998;43:436–445.

54. Varty GB, Higgins GA. Examination of drug-induced and isolation-induced disruptions of prepulse inhibition as models to screen antipsychotic drugs. *Psychopharmacology* 1995;122:15–26.

55. Stevens KE, Freedman R, Collins AC, et al. Genetic correlation of hippocampal auditory evoked response and alpha-bungarotoxin binding in inbred mouse strains. *Neuropsychopharmacology* 1996;15:152–162.

56. Murray RM, Lewis SW. Is schizophrenia a neurodevelopmental disorder? *Br Med J* 1987;295:681–682.

57. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 1987;44:660–669.

58. Dulawa SC, Grandy DK, Low MJ, et al. Dopamine D4 receptor-knockout mice exhibit reduced exploration of novel stimuli. *J Neurosci* 1999;19:9550–9556.

59. Russell RW. Extrapolation from animals to man. In: Steinberg H, ed. *Animal behavior and drug action*. Boston: Little, Brown, 1964:410–418.

60. Kornetsky C. Animal models: promises and pitfalls. In: Hanin I, Usdin E, eds. *Animal models in psychiatry and neurology*. Oxford: Pergamon, 1977:1–7.

61. Russell RW. *The comparative study of behavior*. London: UK, 1951.

62. Boring EG. The use of operational definitions in science. *Psychol Rev* 1945;52:243–245.

63. Bridgman PW. Some general principles of operational analysis. *Psychol Rev* 1945;52:246–249.

64. Krystal JH, Karper LP, Seibyl JP, et al. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry* 1994;51:199–214.

65. Malhotra AK, Pinals DA, Weingartner H, et al. NMDA receptor function and human cognition: the effects of ketamine in healthy volunteers. *Neuropsychopharmacology* 1996;14:301–307.

66. Malhotra AK, Adler CM, Kennison SD, et al. Clozapine blunts *N*-methyl-D-aspartate antagonist-induced psychosis: a study with ketamine. *Biol Psychiatry* 1997;42:664–668.

67. Geyer MA, Braff DL, Swerdlow NR. Startle-response measures of information processing in animals: relevance to schizophrenia. In: Haug M, Whalen RE, eds. *Animal models of human emotion and cognition*. Washington, DC: APA Books, 1999:103–116.

68. Varty GB, Bakshi VP, Geyer MA. M100907, a serotonin 5-HT_{2A} receptor antagonist and putative antipsychotic, blocks dizocilpine-induced prepulse inhibition deficits in Sprague-Dawley and Wistar rats. *Neuropsychopharmacology* 1999;20:311–321.

69. Tallman JF. Neuropsychopharmacology at the new millennium: new industry directions. *Neuropsychopharmacology* 1999;20:99–105.

70. Duman RS. Neuropharmacology in the new millennium: promise for breakthrough discoveries. *Neuropsychopharmacology* 1999;20:97–98.

71. Andreasen NC. Linking mind and brain in the study of mental illnesses: a project for a scientific psychopathology. *Science* 1997;275:1586–1593.

34

Proof of Concept: Functional Models for Drug Development in Humans

Dean F. Wong

William Z. Potter

James R. Brasic

Dean F. Wong and James R. Brasic: Department of Radiology, Johns Hopkins Medical Institution, Baltimore, Maryland.

William Z. Potter: Lilly Research Laboratories, Indianapolis, Indiana.

A drug developed for human use classically goes through a number of steps, including discovery, extensive preclinical studies for safety in experimental animals, and then human safety and efficacy studies. The Food and Drug Administration (FDA) requires successful completion of all the above tasks before approval for therapeutic use in human beings. The exciting advances in human genomics and combinatorial chemistry promise substantial applications to new drugs for human diseases; however, the time and expense associated with the processes necessary to bring a new drug to market are rising exponentially. Thus, there is a great need for functional models of disease progression, including animal models of human disease and biomarkers in human clinical trials.

BIOMARKERS AND SURROGATE MARKERS

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

The monitoring of biologic disease processes increasingly employs biomarkers (Table 34.1). At a recent National Institutes of Health (NIH) and FDA conference (1) biomarkers for clinical efficacy were divided into several groups including natural history markers, biological activity markers, and surrogate markers.

TABLE 34.1. ABBREVIATIONS FREQUENTLY ENCOUNTERED WITH BIOMARKERS AND SURROGATE MARKERS

5HIAA	5-Hydroxyindoleacetic acid	MAO	Monoamine oxidase
5HT	Serotonin	MAOI	Monoamine oxidase inhibitor
5HT2A	Serotonin type 2A	MCCP	Meta-chlorophenylpiperazine
5HTP	5-Hydroxy-L-tryptophan	MCV	Mean red cell volume
ACTH	Adrenocorticotropic hormone	MDL	MDL 100,907
AD	Alzheimer's disease	MHPG	3-Methoxy-4-hydroxyphenylglyco
Ag	Agonist	MOI	Medical optical imaging
AIDS	Acquired immunodeficiency virus	MOS	Medical optical spectroscopy
ALS	Amyotropic lateral sclerosis	MRI	Magnetic resonance imaging
AMP	Amphetamine	MRM	Magnetic resonance microscopy
An	Antagonist	MRS	Magnetic resonance spectroscopy
APOE	Apolipoprotein E	MTD	Maximum tolerated dose
CNS	Central nervous system	NE	Norepinephrine
CSF	Cerebrospinal fluid	NERI	Norepinephrine reuptake inhibitor
D_2	Dopamine type 2	NIH	National Institutes of Health
D2R	Dopamine type 2 receptor	NIOS	Near-infrared optical spectoscopy
DA	Dopamine	NIRI	Near-infrared imaging
DAT	Dopamine transporter	NMR	Nuclear magnetic resonance
ECD	Ethylenediylbis-L-cystein diethylester	NMSP	N-methyl-spiperone
EEG	Electroencephalogram	OCD	Obsessive-compulsive disorder
EPR	Electron paramagnetic resonance	OCT	Optical coherence tomography
FDA	Food and Drug Administration	P-EEG	Pharmaco-electroencephalogram
FDG	Fluorodeoxyglucose	PET	Positron emission tomography
fMRI	Functional magnetic resonance imaging	PK	Pharmacokinetic(s)
GABA	Gamma aminobutyric acid	PRE AMP	Before amphetamine
GBR	GBR 12909	PRL	Prolactin
GGT	Gamma-glutamyl transferase	rCBF	Regional cerebral blood flow
GH	Growth hormone	SPECT	Single photon emission computed tomography
HCFA	Health Care Financing Administration	SSRI	Selective serotonin reuptake inhibitor
HIV	Human immunodeficiency virus	SWS	Slow wave sleep
HMPAO	Hexamethylpropyleneamine oxime	V_h	Volume of distribution
HVA	Homovanillic acid		
IV	Intravenous		

A relevant issue in clinical trials is the selection of an appropriate endpoint. Endpoints that are less deleterious than death or onset of a disease are desirable. Surrogate markers or endpoints are events of a more intermediate nature. These typically replace the final endpoints such as mortality. A surrogate marker is defined statistically as a "response variable for which a test of the null hypothesis of no relationship to the treatment groups under comparison is also a valid test of the corresponding null hypothesis based on the true endpoint" (2). A good biomarker, in contrast, provides information on the possible mechanisms of action of medications or the pathophysiology involved in a disease process. Surrogate markers that are helpful in early drug development (e.g., FDA Phase I and II trials) provide a prognostic indication of a clinical endpoint relevant to FDA Phase III trials.

Accelerated approval, which has become more common in FDA clinical trials, may occur based on surrogate marker effects although completion of longer-term clinical outcome and clinical endpoints may eventually be required. Prentice (1989) suggested that a surrogate marker must be both prognostic of disease progression and affected by treatment (2); the effect of the treatment on the surrogate marker should mediate the effect of the treatment on the clinical outcome or true outcome measure. There are not yet any surrogates for neuropsychiatric disorders that fully meet these criteria.

ROLE OF SURROGATES FOR DRUG DEVELOPMENT

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

The value of surrogates and biomarkers for drug development is recognized by the pharmaceutical industry. The needs to reduce development costs, improve the practical attainment of predictive outcomes, and expedite these endpoints are motivations for growing interest in this area. We present in the following an overview of the current state-of-the art regarding utilization of various surrogate markers, in a much broader sense of the term, which includes biomarkers, in drug discovery and development. We start with the most basic steps pertinent to initial studies in humans.

The integration of preclinical science and assessment of therapeutic potential in humans is done primarily through quantification of drug and active metabolites in accessible specimens (blood, urine, and cerebrospinal fluid). This is an advance on the maximum tolerated dose (MTD) approach, which necessarily prevailed before sufficiently sensitive assay

methodology was available. In current drug development, upper doses for early human studies are based on toxicokinetic studies in at least two animal species (most often rat and dog) whereby one does not exceed some predetermined ratio between the exposure (i.e., plasma concentration) for the dose in humans and the exposure associated with toxicity in the most sensitive species (3). Thus, depending on the ratio selected, exposures in humans may produce few side effects and be below those achieved in the older MTD approach.

In practice, one often embarks on clinical studies with a constricted range of exposures and an inability to address the question of whether therapeutic (or biochemical) effects are greater at the MTD until much later in development (4). If a compound does not produce significant side effects within this original range, doses much higher than those used in initial Phase I studies can be tested depending on the nature of the toxicity observed in animals at the limiting exposure. Even if a limiting dose is identified in volunteer subjects, a greater dose may be tolerated in patients (5). Sramek and colleagues (1997) (6) have defined this transition from dosing in healthy volunteers to patients to be “bridging.” For instance, the difference in tolerated dose for schizophrenic patients compared with normal volunteers is often greater than 25-fold (5). Interestingly, before pharmacokinetic (PK) data were routinely used to limit exposure, one explored an MTD relatively early and concluded that if response was not achieved by that dose, one did not have a viable clinical candidate.

Most marketed and late in development psychopharmacologic agents do show limiting side effects above the recommended dose range, especially if such doses are given initially. Starting with three or more times the lowest standard dose of a selective serotonin reuptake inhibitor (SSRI) will produce far more marked nausea (and even vomiting) than after building up to the same dose over several weeks, as is done for the treatment of obsessive–compulsive disorder (OCD). This may seem simply common sense but the perceived market advantage of the starting and ultimate therapeutic dose being the same discourages exploration of higher doses that require some sort of titration over time in order to be well tolerated. For instance, even though it was appreciated early on that an antidepressant, venlafaxine, might be more efficacious at higher doses that were associated with

more marked side effects (7), it was marketed at target lower doses (37.5 mg b.i.d.), which only produce effects consistent with serotonin reuptake inhibitory activity. In the over 225 mg per day range (which requires some degree of titration to avoid unacceptable side effects), venlafaxine produces effects consistent with norepinephrine (NE) as well as serotonin (5HT) uptake inhibition (8 ,9) and is reported to be superior to SSRIs (10 ,11). Obviously, it would be beneficial if the relevant dose ranges of such compounds could be better understood earlier rather than later in the life of a drug.

In this particular example, if valid markers of 5HT and NE uptake inhibition in the brain were investigated as early in development as possible, then the doses necessary to achieve both effects and those achieving only one could have been systematically pursued from the outset. Such markers would be “surrogate” markers for clinical doses, allowing one to test the widely held hypothesis that in a big enough population of patients with depression, one will find those who respond better to combined NE/5HT uptake inhibitors than to SSRIs alone. Table 34.2 shows examples of biochemical markers that may serve as surrogates.

	Neurotransmitters and Metabolites[b]			Plasma Hormones			Physiologic	
Drug Class	**NE and MHPG**	**5HT, 5HIAA**	**HVA**	**PRL**	**GH**	**ACTH/Cortisol**	**Temperature**	**SWS**
Serotonergic								
SSRIs		↓5HT, ↓5HIAA		±↑				
Indirect-Ag				↑	↑			
$5HT_{1A}$-Ag				±↑	↑	↑	↓	
$5HT_{1B/D}$-Ag				↓	↑			
$5HT_{2C}$-Ag				↑		↑	↑	↑↓[c]
Noradrenergic								
NERI	↑NE, ↓MHPG				↑			
α2-Ag	↓NE				↑			
α2-An	↑NE							
Dopaminergic								
D2-An			↑↓[c]					
DA-Ag			↓					
Mixed								
Tricyclic antidepressants	↓MHPG	↓5HIAA						
MAOIs	↓MHPG	↓5HIAA	↓					
Atypical antipsychotics	±↑NE			=[d]				

5HIAA, 5-hydroxy indoleacetic acid; 5HT, serotonin; ACTH, adrenocorticotropic hormone; Ag, agonist; An, antagonist; GH, growth hormone; HVA, homovanillic acid; MAOI, monoamine oxidase inhibitor; MHPG, 3-methoxy-4-hydroxyphenylglycol; NE, norepinephrine; NERI, norepinephrine reuptake inhibitor; PRL, prolactin; SSRI, selective serotonin reuptake inhibitor; SWS, slow wave sleep.
[a]This table indicates the utilization of neurotransmitters and their metabolites, plasma hormones, and physiologic phenomena as surrogate markers for the biochemical effects produced by drugs on the indicated classes. Although the specified drug classes may affect other measures than those indicated (e.g. SSRIs also decrease MHPG), only those hypothesized to reflect a primary action are shown.
[b]In plasma, platelets, urine, and/or cerebrospinal fluid.
[c]Acute increase followed by decrease in responders.
[d]Equal sign (=) indicates no change.

TABLE 34.2. MEASURES THAT SERVE AS SURROGATES FOR PREDICTED BIOCHEMICAL EFFECT OF DRUGS TARGETED TO MONOAMINE NEUROTRANSMITTERS, TRANSPORTERS, RECEPTORS, AND DEGRADATIVE ENZYMES[a]

In the section that follows we briefly review how traditional surrogate markers (in the broadest sense), which antedate the application of brain imaging technology, are used.

PHARMACOKINETICS

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

Analytical technology, especially that provided by much simpler to operate equipment, allows for precise quantification of very low concentrations in blood so that rapid and accurate determination of a new compound’s pharmacokinetics (PK) is possible after much lower doses than required formerly. If a half-life is too short or long or if nonlinear pharmacokinetics are marked (e.g., half-life increases with doses in the expected therapeutic range), then development may go no further. Similarly, if a compound shows unacceptably wide variations in clearance because its metabolism is dependent on a highly polymorphic isozyme of cytochrome P-450 (e.g., CYP2D6), then it will be seen as of lower potential commercial value. Usually, *in vitro* tests using human hepatic microsomal preparations or P-450 isozyme specific cloned systems are used to screen for this possibility prior to human use, but these are not always

fully predictive. Thus, concentrations measured over time in Phase I studies serve, at this most basic level, as a surrogate of whether a compound has the potential to become a successful drug (12).

Drug interactions sometimes can constitute a serious safety risk (e.g., inhibition of terfenadine metabolism by CYP3A4 inhibitors leading to fatal QTc prolongation in genetically susceptible individuals) and, almost always, a marketing disadvantage. The same preclinical approaches referred to in the preceding are used to screen out compounds that depend for their metabolism on a problematic pathway. *In vivo* PK (and sometimes drug interaction data) are also obtained in animals, but more to set dose and frequency of administration to guarantee adequate concentrations for expensive chronic toxicology studies than to predict what the PK characterization will be in humans.

Low bioavailability (i.e., concentration after oral versus intravenous doses) is a special problem at this stage of preclinical development. First, for most compounds an oral form is seen as the ideal (if not only) option and low bioavailability almost always means very high variability in exposure to parent compound for a given dose. For instance, 2% to 10% bioavailability means an automatic fivefold range before individual variations seen in metabolic clearance after intravenous (IV) administration are taken into account versus 60% to 80% bioavailability, which only entails a 11/3-fold variation. High variability in exposure per unit dose makes clinical studies more difficult. Second, to reach the exposures required to define the concentration at which chronic toxicity is observed, low bioavailability means that much more compound will be necessary. This is actually a far more significant issue than is generally appreciated because amounts of material are limited early in development. It can be very costly to develop efficient synthetic schemes; one does not want to invest too much in syntheses until one is fairly sure that a compound is safe at pharmacologically active concentrations.

TOXICOKINETICS

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

This introduces the critical role of toxicokinetics, the goal of which is to establish in animals (usually rats and dogs) the ratio between the concentration that produces either unacceptable physiologic effects (e.g., QTc prolongation, seizures) or organ damage (e.g., liver, bone marrow) and the one that produces the targeted biochemical or behavioral effect, a ratio of at least 2 to 1. As the ratio approaches 1.0 it is increasingly unlikely that a compound will be taken into humans unless the observed unwanted effects are readily reversible and one is dealing with a condition for which existing therapies are poor or nonexistent (Alzheimer's disease [AD] or amyotrophic lateral sclerosis [ALS]). There are instances in which humans ultimately prove tolerant to drugs much more than would be predicted by these ratios, but this is only established after exposure to large numbers of subjects.

Although the ratios established in preclinical studies are interpreted as providing a reasonable estimate of any ultimate therapeutic index in humans, they are specific for both species and mechanism of action. In other words, the concentration necessary to achieve a behavioral or biochemical effect and produce toxicity can vary more than an order of magnitude across species. The relationship among *in vitro* biochemical effects, activity in some *in vivo* model, and undesired effects is poorly understood even at the preclinical level, particularly when one is investigating novel targets. Because efficacious *in vivo* concentrations in animals are based either on activity in an animal model (e.g., anticonflict activity to identify a potential anxiolytic) or biochemical changes (e.g., changes in brain serotonin) that are not measured in humans, we do not even have validation for existing drugs relating preclinical and clinical data (13). Not surprisingly, given the weakness of the link between an efficacious dose in rats and a therapeutic dose in humans, the clinical development of central nervous system (CNS) drugs is notoriously inefficient to identify therapeutic concentrations.

Measures of receptor occupancy or specific degree of enzyme inhibition do not drive the selection of an efficacious concentration. In other words, the basic preclinical work that underlies any test of a specific hypothesis such as "X% of inhibition produces effect Y" is rarely, if ever, done. Therefore, in toxicokinetic studies one hopes for very high ratios that provide a great deal of flexibility to study a extremely wide range of concentrations in humans, especially with novel compounds.

Phase I

Select starting doses for healthy volunteers based on findings from the preceding toxicokinetic studies, which is often one-tenth of the dose that produces a no-adverse-event level in the most sensitive species studied (4). Most marketed central nervous system (CNS) drugs produce some sufficiently unpleasant (but not dangerous) side effect that sets the dose range for subsequent studies, the so-called maximum tolerated dose (MTD). Multiple dose studies, administering various fixed doses over a 2- to 4-week period are also part of the Phase I protocols to get a general idea of safety and tolerability under likely clinical conditions (14). The observed steady-state concentration during repeated dosing, broadly speaking, is the first surrogate for the desired effect, the obvious but often incorrect assumption being that a certain value assures a certain biochemical effect in the brain based on extrapolations from preclinical data.

Once the initial safety studies are completed, there may be additional studies in healthy volunteers to assess efficacy. In the case of sedatives and hypnotics, healthy volunteers can be used reliably with a combination of subjective reports, objective measures of performance (looking for decrements), and the electroencephalogram (EEG) (15 ,16 ,17 ,18 and 19). These measures, however, are direct ones, of a desired or undesired effect, not surrogates. On the other hand, the so-called

pharmaco-EEG (P-EEG) is sometimes introduced as a surrogate measure of the therapeutic action a compound is likely to produce (20 ,21), but is probably more usefully viewed as a direct measure that a compound has produced some change in surface brain electrical activity. The P-EEG provides neither evidence of a direct effect in the brain nor proof of a hypothesized causative role of specific pharmacologic activity of the molecule (22).

In the majority of CNS Phase I studies, aside from measures of drug concentration, the only usual objective measures are targeted to the heart and other vital organs. Side effects are often simply recorded as spontaneous reports on a grid of prespecified subjective descriptors; sometimes a formal checklist is used. Thus, quantitative surrogate measures of drug effect are sparse.

Phase 1B/2A (Proof of Concept) Studies

There is wide variation in practice as to what is done, both with healthy volunteers and patient populations, before going on to full Phase II studies. These latter studies are powered to detect a clinically significant improvement, usually over placebo, often utilizing multiple dose arms (each with at least 50 or more patients) and multiple clinical sites. In the case of a compound acting on a novel target, the period before full Phase II trials is argued to be the phase during which exploratory work should be done (23). For instance, if one had an ampakine such as CX-516 that improved learning in rats, then one might show enhancement of some aspects of cognition in healthy volunteers after safe doses were established in Phase I studies (24). The improved performance in healthy subjects would then be a surrogate for improvement in AD to identify the appropriate dose.

In the more usual case of a variation on an existing mechanism (e.g., SSRIs), this is the period to verify which dose produces the desired biochemical effect so as to move as quickly as possible into Phase II trials with doses that one is confident will work (25 ,26). In this latter instance, the direct accessible biochemical measure becomes a surrogate for an effect in brain that is hypothesized to produce the clinical benefit.

We briefly review a range of specific measures that have been used in enough instances over the last two decades to qualify as surrogate markers. There is no available literature as to the extent to which these measures actually drove decisions about what doses to use in Phase II trials, but one can assume that they must have been influential in some cases. What is certain is that at least in the United States, outside of polysomnography (for sleep-related disorders) and EEG (for seizures), the measures have had no acknowledged regulatory impact. Assessments of decrements in performance (especially ability to concentrate, recall, and carry out motor tasks) and measures of cognitive enhancement all seem to be better classified as direct measures of an intended or unintended drug effect rather than surrogates. This leaves mainly biochemical measures, which can be broadly classified as “neuroendocrine.”

Cerebrospinal Fluid

The monoamine hypotheses of depression and schizophrenia, which are, in turn, based on the pharmacology of antidepressants and antipsychotics, generated an interest in measures of norepinephrine (NE), serotonin (5HT), and dopamine (DA) in humans. In most instances these were initially studied in an attempt to distinguish patients from controls but came to be applied to validating predicted effects of psychotropic agents (27).

Thus, NE reuptake inhibitors (NERIs) decrease NE turnover as measured by NE and its metabolites in urine or cerebrospinal fluid (CSF) (28 ,29 ,30 and 31) and by increasing plasma NE (32) or, more precisely, its spillover rate (33 ,34). Another marker of NE reuptake inhibition is based on the need for exogenously infused tyramine to be taken up into NE nerve endings to exert its effects, the so-called tyramine pressor test (35) or, more recently, the dorsal hand vein constrictor test (36). These latter two procedures have been recently employed to establish the dose at which venlafaxine produces NE reuptake inhibition (8 ,9 ,36). Because NE metabolism varies according to whether it occurs inside or outside neurons (37), a more definitive and perhaps more sensitive measure would be the ratio of extraneuronal to intraneuronal metabolites before and after treatment in an integrated pool (e.g., a 24-hour urine) (38) as shown for desipramine (30). However, all of these measures reflect functional changes in the peripheral sympathetic nervous system; therefore, they are at best surrogates for effects in the brain.

Some of these same measures have been applied to study other compounds, which are predicted to affect NE function either following acute or chronic administration. These include bupropion, clozapine, alpha 2 antagonists, and monoamine oxidase inhibitors (MAOIs) (39 ,40 ,41 ,42 ,43 ,44 ,45 and 46). Interestingly, such studies are not yet available to support or refute claims that the relatively new antidepressant mirtazapine produces alpha 2 antagonism in humans.

Most, if not all, drugs classified as SSRIs have been shown to inhibit platelet uptake of 5HT, most simply determined by investigating 5HT depletion in platelets following chronic administration (25 ,47 ,48 ,49 and 50). Some of these studies were clearly done prior to Phase II trials so they can be inferred to have influenced the selection of dose. Doses of SSRIs known to inhibit platelet uptake in humans have also been shown to reduce the turnover of 5HT in the CNS as reflected by reductions of its major metabolite 5-hydroxyindoleacetic acid (5HIAA) in CSF (44 ,51). Similarly, MAOIs reduce both platelet MAO and 5-HIAA in CSF (51 ,52). There are not, however, systematic dose response studies to compare the sensitivity of the peripheral and central measures. Moreover, 5HIAA in CSF obtained by a lumbar puncture reflects a complicated process of all sources of formation and clearance of the metabolite.

DA has been studied almost exclusively in terms of its metabolite, homovanillic acid (HVA) in blood and CSF. There is a complex relationship among administration of a

dopamine type 2 (D_2) antagonist, duration of treatment, clinical state, and HVA in either compartment (53 ,54 ,55 ,56 ,57 and 58). Although changes usually can be explained as consistent with altered turnover as a function of receptor antagonism, time, and presence or lack of clinical response, measures of HVA are not really useful as a surrogate measure of D_2 antagonism. Decreased HVA in CSF can be expected after mixed Types A and B MAO inhibition in the brain (52) and might be relevant to assessing DA uptake inhibition. There is, however, no available positive control for the latter. Bupropion, the one compound studied in regard to possible DA uptake inhibition, did not decrease HVA in the CSF (40 ,44).

Prolactin

Prolactin has the potential to be used as a marker for drugs affecting these systems, because either DA or 5HT can affect this circulating hormone. In the simplest and most widespread instance, it has been used as an index of D_2 antagonism. Unlike typical neuroleptics, atypical neuroleptics produce no elevations of prolactin at therapeutic dosages (59 ,60). By extension, absence of prolactin elevation after antipsychotics could be considered a surrogate of low D_2 receptor occupancy in the striatum and, hence, imply low to absent motor side effects (61 ,62).

Prolactin is also elevated following various pharmacologic challenges such as: (a) those that are predicted to increase extracellular 5HT in the brain including fenfluramine, clomipramine, l-tryptophan, and 5hydroxytrytophan; and (b) those that stimulate various types of 5HT receptors, including meta-chlorophenylpiperazine (mCPP) and some, but not all, putatively selective $5HT_{1A}$ agonists (63 ,64 ,65 and 66). In all of these instances, prolactin increase becomes a surrogate of increased serotonergic transmission in one or more regions of the brain. Blockade of this effect by appropriate 5HT receptor specific antagonists could serve, in turn, as a surrogate of a compound's ability to functionally antagonize the specified receptor in human brain (67 ,68 ,69 and 70).

Growth Hormone

There was a period in which plasma growth hormone (GH) was used as a surrogate for increased noradrenergic transmission in human brain (presumably at the level of the hypothalamus) after, for instance, the α2 agonist, clonidine, or an NE uptake inhibitor (71 ,72). One could, in turn, infer α2 antagonism by a test compound if it blocked the effect.

More recently, stimulation of plasma GH has been considered evidence of activation of both $5HT_{1A}$ and $5HT_{1B/1D}$ receptors (see Table 34.2 on page 459) (63). Sumatriptan increases GH, apparently through activation of the $5HT_{1B/1D}$ receptors (73 ,74), with the most recent evidence using the more brain penetrant, zolmitriptan, implicating $5HT_{1D}$ postsynaptic receptors (75). It has been suggested that stimulation of the 1B/1D receptors inhibits somatostatin release (76); however, an increasing volume of experimental research indicates that 5HT can act directly on the adrenal gland and possibly on the anterior pituitary as well (77). This provides another example of the same measure serving as a surrogate for very different CNS effects, the interpretation of which depends on knowing the potential target for a compound in advance. It is doubtful that any substantial decisions concerning doses of compounds affecting either noradrenergic or GABAergic or serotonergic systems have ever been made based on GH release stimulation or inhibition.

ACTH/Cortisol

Stimulation of the hypothalamic–pituitary–adrenal axis as reflected by increases of ACTH and/or cortisol has also been used as a marker of drug action in the CNS. This approach has been most extensively applied to evaluation of putative 5HT agonists, sometimes coupled with pretreatment with whatever antagonists were available (e.g., ritanserin and pindolol) (63). Such studies can generate evidence of apparent selectivity in postsynaptic receptor responses to indirect and direct agonists. ACTH release induced by 5-hydroxy-L-tryptophan (5HTP) is argued to occur through indirect activation of 5HT2 receptors because it is antagonized by ritanserin (78), whereas the cortisol response is not affected by doses of pindolol expected to produce $5HT_{1A}$ antagonism (67). Pindolol does, however, antagonize ACTH responses to a variety of agents classified as partial to full $5HT_{1A}$ agonists (64 ,79 ,80). Again, given the complexities of the ACTH/cortical response and the imperfect selectivity of the pharmacologic agents, quantitative conclusions as to degree of specific receptor activation or antagonism are not possible.

Other Physiologic Measures

Sleep EEGs have already been referred to and may have utility as surrogates of, for instance, activation or antagonism of 5HT2C receptors as reflected by decreases or increases of slow wave sleep, respectively (81 ,82).

Temperature decreases are consistently observed following $5HT_{1A}$ agonists (63 ,83), and hence can serve as a surrogate marker of $5HT_{1A}$ agonist effects in the CNS. Evaluating the ability of a novel compound to antagonize the hypothermia produced by a $5HT_{1A}$ agonist may be the easiest way to see if it antagonizes $5HT_{1A}$ receptors in the human brain.

LIMITATIONS OF CURRENT SURROGATE MARKERS

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

As described, there is no validated link between the concentration of drug in blood (or even CSF) and a specific biochemical effect in human brain (not to mention a specific brain region). Thus, even in the case of SSRIs, the most

widely prescribed class of drugs in neuropsychopharmacology, we do not know how closely platelet 5HT depletion reflects 5HT uptake inhibition in brain. Although known therapeutic doses of SSRIs invariably have been shown to deplete platelet 5HT, the converse is not clear; that is, any dose that is effective in platelets will be therapeutic. Matters are even less certain when it comes to using available surrogates of NE uptake inhibition to establish the dose of a drug. And, as already noted, primary dosing decisions are not made on the basis of whether compounds affect prolactin, growth hormone, or ACTH/cortisol responses.

One could use the CSF concentration of a drug as a reasonable estimate of the free drug concentration in brain under true steady-state conditions (84 ,85 and 86), and infer from one's preclinical *in vitro* and *in vivo* studies that this will produce a specific effect (87). The problem is that even in the most refined *in vitro* system, as represented by cloned human receptors expressed in some vector, the relationship among receptor occupancy, drug effect, and free drug concentration may be extremely different from that observed *in vivo* in humans thanks to multiple uncontrollable differences among these systems (13). Furthermore, *in vivo* animal studies raise questions about species differences; therefore, how does one select the dose for clinical studies when testing a new compound that is well tolerated with a wide range of safe concentrations that are predicted to be pharmacologically active by one or more preclinical models? How does one know that the target in question has been blocked or stimulated so as be sure that one is testing the hypothesis that such an effect produces therapeutic benefit? The answer is, one does not know with the surrogates (discussed in the preceding). This brings us to a discussion of the promise of direct measures of drug effect in the brains of living humans.

IMAGING STUDIES

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

Measures of Substrate Metabolism

Regional cerebral blood flow (rCBF) is a well-established surrogate marker widely used for both clinical diagnostic procedures and new drug development. rCBF is commonly assessed utilizing estimates of cerebral perfusion, such as single photon emission computed tomography (SPECT) with radioligands labeled with ^{123}I or ^{99m}Tc; for example, [^{123}I]iodoamphetamine, [^{99m}Tc]ethylenediylbis-L-cystein diethylester (ECD), and [^{99m}Tc]hexamethylpropyleneamine oxime (HMPAO). Although repeated assessments with multiple radiotracer injections are usually impractical because of the long half-lives of both ^{123}I and ^{99m}Tc, linearization techniques have been developed to estimate sequential measurements (88). Furthermore, ^{99m}Tc is relatively available in practically every hospital with a radiology or nuclear medicine department worldwide; therefore, SPECT procedures can be performed widely on a clinical basis.

By contrast, positron emission tomography (PET), another procedure to estimate rCBF, requires an onsite cyclotron, so it is not available in many areas. Recent third-party camera reimbursement of some PET procedures (primarily [^{18}F]fluorodeoxyglucose (FDG)) are making PET cameras more available, but cyclotrons are still relatively scarce. The two most widely used radioisotopes used in PET are ^{11}C with a half-life of 20 minutes and ^{18}F with a half-life of 110 minutes. Additionally, ^{15}O, with a half-life of 2 minutes is primarily employed in brain perfusion studies. The most available PET radioligand is FDG, a tool to measure glucose metabolism. Its advantages include ease of use, availability from commercial cyclotrons throughout much of the world, and sensitivity to a number of studies to facilitate drug development and assess mechanism of action (89). Its disadvantage is lack of specificity.

An example of the value of PET studies to obtain potential surrogate markers for development of drugs for substance abuse is the dysfunction of regional glucose metabolism in the limbic system and areas of working memory in cocaine euphoria and cocaine craving (90). [^{15}O] PET studies demonstrate increases in rCBF in the limbic system and decreases in the basal ganglia as manifestations of craving for cocaine following exposure to videotapes suggesting cocaine use (91).

ESTIMATION OF DOPAMINE RECEPTOR OCCUPANCY BY ANTIPSYCHOTIC DRUGS

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

It has been hypothesized that positive symptoms of schizophrenia, such as hallucinations and delusions, result from increased stimulation of postsynaptic dopamine (DA) receptors by DA (92). This DA hypothesis of schizophrenia is substantiated by the observation that positive symptoms of schizophrenia abate when DA receptor blocking drugs, such as neuroleptics, occupy the postsynaptic DA receptors. In 1988, Farde and colleagues demonstrated the concept that effective neuroleptic dosages for schizophrenia correspond to 80% to 90% occupancy of DA type-2 receptors (D_2Rs) by the drug (95). Thus, occupancy of 80% to 90% of D_2Rs may constitute a surrogate marker for the dosage of a D_2 antagonist producing maximal beneficial effects with minimal adverse and toxic effects. Furthermore, clinically equivalent doses of neuroleptics are estimated by comparing proportions of receptors occupied by various psychotherapeutic agents (93 ,94). Several examples of this approach are given in the following.

Receptor Occupancy by Typical Neuroleptics

The occupancy of D_2Rs by neuroleptics has been extensively evaluated by PET and SPECT. Historically, this began with the study of typical neuroleptics, including haloperidol and

fluphenazine, in healthy normal control subjects and patients with schizophrenia utilizing [^{11}C]raclopride (95 ,96), spiperone derivatives, including [^{11}C]N-methyl-spiperone (NMSP) (97), [^{18}F], or ^{76}Br (98). This research has established that a therapeutic response corresponds to occupancy of 65% to 90% of the D_2Rs by the drug; however, occupancy of greater than 90% of the D_2Rs is not associated with a greater therapeutic effect. Therefore, occupancy of 65% to 90% of D_2R_2 is correlated with a therapeutic dose of typical neuroleptics as well as other clinical manifestations of pharmacologic efficacy. For example, patients with acute extrapyramidal side effects were found to have higher DA receptor occupancies (82%) than those without (74%) (99 ,100). Wolkin and colleagues (1989) found a comparable occupancy of haloperidol in responders versus nonresponders indicating that treatment nonresponse is not a function of insufficient CNS binding of the antipsychotic (100).

Kapur and colleagues recently confirmed that the D_2 occupancy is an important mediator of beneficial and adverse effects (101) in a study of first episode schizophrenia and haloperidol. Although the patients showed a wide range of D_2R occupancy (38% to 87%), the degree of D_2R occupancy predicted clinical improvement, hyperprolactinemia, and extrapyramidal side effects. Also, Daskalakis and associates (1998) (102) found a relationship between D_2R occupancy and hyperprolactinemia.

Receptor Occupancy by Atypical Neuroleptics

D_2R occupancy has been determined for humans treated with atypical neuroleptics. For example, receptors of people treated with clozapine exhibit low (20% to 65%) D_2R occupancy and high serotonin type 2A ($5HT_{2A}$) receptor occupancy (94 ,99 ,103). In studies directly comparing clozapine (450 mg/day) and haloperidol (5 mg/day), there was a reversal of receptor occupancy producing high D_2 blockade with haloperidol and high $5HT_{2A}$ blockade with clozapine. Indeed PET imaging has been an important approach for screening a number of new antipsychotics with an atypical profile. Risperidone has been shown to block 40% to 60% of D_2 receptors at only 1 mg (104) with even higher $5HT_{2A}$ receptor occupancy. Indeed, in patients with schizophrenia, only 4 mg/day is needed for 70% to 80% occupancy with minimal risk of extrapyramidal side effects (105). After an oral dose of 6 mg risperidone, 75% to 80% of D_2Rs and 78% to 88% of $5HT_{2A}$ receptors are occupied (106). Plasma levels of risperidone correlate with D_2R occupancy (107). The proportion of D_2R occupancy has been estimated for other atypical neuroleptics; for example, olanzapine, has been shown to have D_2 occupancy similar to risperidone and greater than clozapine (108). The usual clinical doses of olanzapine (10 to 20 mg/kg) produce 71% to 80% D_2R occupancy. These doses usually do not result in the adverse effects of dyskinesias and prolactin elevation. On the other hand, doses of greater than 30 mg/day olanzapine are associated with greater than 80% D_2R occupancy as well as dyskinesias and prolactin elevation. Olanzapine exhibits 59% to 63% D_2Rs occupancy after single 10-mg dosing. Furthermore, olanzapine shows a greater (68% to 84%) occupancy of D_2Rs than clozapine (20% to 67%) after 10 to 20 mg daily dosing in patients (60). These characteristics suggest that olanzapine differs from clozapine. Additionally, quetiapine is characterized by abatement of psychotic symptoms in association with a transient increase in D_2R occupancy (62).

Receptor Occupancy as a Surrogate Marker of Clinical Efficacy?

One of the most important questions about D_2R and $5HT_{2A}$ receptor occupancy is the prediction of doses yielding clinical efficacy. D_2R occupancy predicted the clinical improvement of 22 first episode schizophrenic patients randomly assigned to 1 or 2.5 mg/day of haloperidol for 2 weeks (101). Thus, D_2R occupancy is related to the clinical response to antipsychotics (99 ,109 ,110).

PET studies have helped demonstrate that high levels of D_2R occupancy occur at very low haloperidol doses (111). Kapur (1996) showed D_2R occupancy of 53% to 74% at only 2 mg haloperidol in first-episode patients (114). Kapur and colleagues (1997) also showed D_2 occupancies between of 53% to 88% for haloperidol doses of 1 to 5 mg. Thus, the conventional therapeutic practice of haloperidol doses of greater than 10 mg/day is too high for many schizophrenic patients because there is no increase in beneficial effect, whereas the risk of adverse effects increases in proportion to the dose (112). Nevertheless, Volavka and colleagues (113) (1995) showed that antipsychotic efficacy of haloperidol increases with plasma levels up to 12 ng/mL, plasma levels that would predict almost completely saturated D_2 receptors according to the Kapur and associates (1997) data (112). Similarly, Wolkin found D_2 receptor occupancy increasing with haloperidol plasma levels up to 15 ng/mL and almost complete D_2R occupancy saturation with haloperidol plasma levels above 15 to 20 ng/mL (100). Thus, Wolkin's (100) and Volavka's (113) findings confirm each other and contradict Kapur's (112 ,114) studies. These differences could be explained partly by the differences in radioligands and patient populations. Further research by other investigators with various populations is needed to resolve the controversy.

Another important implication of receptor occupancy with neuroleptics is the prediction of extrapyramidal side effects. The probability of acute dyskinesias is directly proportional to the proportion of D_2Rs occupied by the drug (99 ,110 ,115). Furthermore, receptor imaging demonstrates

a lower degree of D_2 receptor occupancy during treatment with atypical neuroleptics. For example, PET has been used to obtain the minimal effective dose of risperidone. The high D_2R occupancy associated with 6 mg or more risperidone daily suggests a high risk of acute dyskinesias. On the other hand, 4 mg risperidone daily results in 70% to 80% D_2R occupancy and a lesser risk of acute dyskinesias (116). An additional role for occupancy studies in drug development takes the form of the interpretation of the time course of receptor occupancy following a single drug dose. Such studies help determine the appropriate dosing regimen for future trials, such as once or twice a day dosing. For example, 70% to 90% of $5HT_{2A}$ receptors are occupied during the 24 hours after a single oral 20-mg dose of MDL100,907 (MDL), a selective serotonergic agent, whereas only 20% are occupied 24 hours after a 10-mg dose (117). These results suggest that a 20-mg dose may be administered once daily, whereas a 10-mg dose requires administration twice a day. Thus, occupancy studies constitute surrogate markers for the outcome variable and frequency of drug administration.

Another use of occupancy studies is the correlation of D_2R occupancy with plasma levels of neuroleptics. This approach has been successfully applied to estimate D_2R occupancy by haloperidol in patients with low doses of haloperidol (118). These preliminary results can be refined in future research with larger sample sizes.

In summary, receptor occupancies have helped establish the optimal dosage range of antipsychotic medications. These imaging methods also have a role to determine *in vivo* occupancy of new neuroleptics with multiple sites for D_2 and $5HT_{2A}$ binding. The studies probably have their greatest role in giving approximate dosage estimates for future clinical trials.

Other Roles for Neuroreceptor Imaging in Drug Development

Four major areas in which neuroreceptor imaging can assist in drug development are listed in Table 34.3 . The first and most well developed area is in helping target rational drug dosing as in the neuroleptic studies described in the preceding. The second is the elucidation of the biodistribution of the drug by radiolabeling the drug or a derivative of the drug. Examples of this application include the characterization of MDL, which is the first selective $5HT_{2A}$ antagonist developed primarily for schizophrenia. MDL has been radiolabeled with [^{11}C] in an isotopic form such that a stable carbon atom (atomic number ^{12}C) is substituted with a radioactive atom (^{11}C) without a change in the pharmacology or chemistry. This procedure facilitates characterization of the biodistribution and washout characteristics of the agent (117).

Rational drug dosing
Biodistribution of drug bound to radiolabels
^{11}C and ^{18}F for PET
^{123}I and ^{99m}Tc for SPECT
Therapeutic rationale for drug utilization
Mechanism of drug action

PET, positron emission tomography; SPECT, single photon emission computed tomography.

TABLE 34.3. COMPONENTS OF THE DRUG DEVELOPMENT PROCESS ACCOMPLISHED BY NEURORECEPTOR IMAGING

The third application of neuroreceptor imaging to drug development is to better understand the mechanism of action of drugs. One example of this is in the development of drugs for cocaine abuse. Unlike neuroleptics and antidepressants, drugs developed to inhibit the action of cocaine have failed clinical trials. Although cocaine has been shown to affect multiple neurotransmitter systems, current research efforts to develop effective treatment for cocaine dependence focus on the dopamine system. Cocaine is hypothesized to produce euphoria by increasing the intrasynaptic concentration of dopamine. Cocaine has high affinity for the dopamine transporter (DAT); therefore, contemporary research to treat cocaine dependence includes the development of noncompetitive inhibitors of cocaine at this site without affecting dopamine transport. One example of potential treatments for cocaine dependence is the development of GBR12909 (GBR), a potent DAT inhibitor. This pharmaceutical, originally developed in Europe as an antidepressant, has found a potential new application as a prototypical drug for cocaine abuse. Prior studies have shown that IV infusion of GBR to Rhesus monkeys selectively reduced (1 mg/kg) and eliminated (3 mg/kg) cocaine self-administration (119). Villemagne and colleagues (120) tested the hypothesis that doses of GBR, which reduce self-administration, also produce significant occupation of DAT. Doses of 1, 3, and 10 mg/kg demonstrated occupancy of 26%, 53%, and 72%, respectively, in *Papio anubis* baboons (Fig. 34.1 and Fig. 34.2). These data suggest that doses that suppress cocaine administration also provide high occupancy of the DAT. Preclinical research supports the hypothesis that elevations of mesolimbic DA mediate the addictive and reinforcing effects of methamphetamine and amphetamine. *In vivo* rodent microdialysis has demonstrated that GBR attenuates cocaine and amphetamine induced increases in mesolimbic DA. Utilizing PET scans of a continuous infusion of [^{11}C]raclopride in baboons, Villemagne and colleagues (120a) also showed that GBR attenuates amphetamine induced striatal DA release by 74% (Fig. 34.3 and Fig. 34.4). Thus, GBR is a potentially effective agent to treat cocaine and methamphetamine dependence. This experimental

model is also being utilized to evaluate other forms of GBR such as the long-acting decanoate derivative.

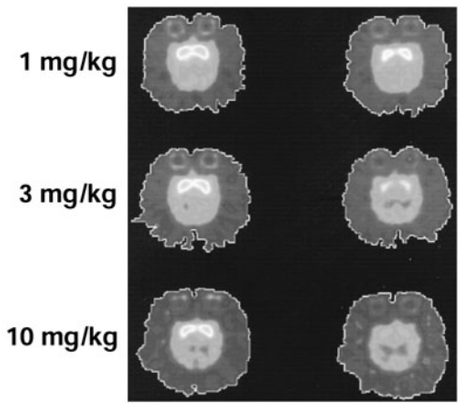


FIGURE 34.1. Reductions in dopamine transporter occupancy are shown in transaxial [^{11}C]WIN35,428 images in *Papio anubis* baboons before *(left)* and after *(right)* administration of three different doses of GBR. Each dose is given 90 minutes before [^{11}C]WIN35, 428 injection. The illustrations represent average PET images at midstriatal level between 70 and 90 minutes after the injection of the radiotracer normalized to the injected radioactivity. Modified from Villemagne V, Rothman RB, Yokoi F, et al. Doses of GBR12909 that suppress cocaine self-administration in non-human primates substantially occupy dopamine transporters as measured by [^{11}C]WIN35, 428 PET scans. *Synapse* 1999;32:44–50. Copyright © 1999, Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc. See color version of figure.

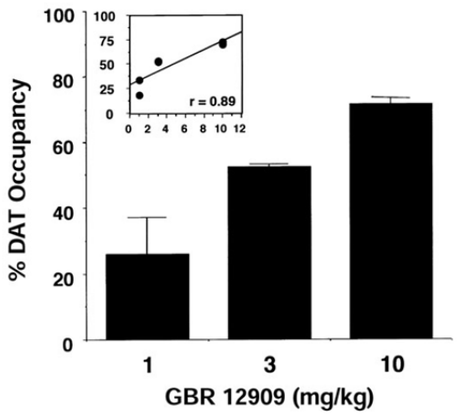


FIGURE 34.2. This histogram illustrates the percentage dopamine transporter (DAT) occupancy by GBR 12909 (GBR) as measured by positron emission tomography imaging with [^{11}C]WIN35, 428. DAT occupancy is represented as the percentage mean ± standard error of the mean differences between binding potentials at baseline and after GBR administration. Percentage occupancy is calculated by the formula as follows: [(Baseline binding potential - GBR binding potential)/Baseline binding potential] × 100. *Inset:* Relation between DAT occupancy and GBR dose (120). Modified from Villemagne V, Rothman RB, Yokoi F, Rice KC, Matecka D, Dannals RF, Wong DF. Doses of GBR12909 that suppress cocaine self-administration in nonhuman primates substantially occupy dopamine transporters as measured by [^{11}C]WIN35, 428 PET scans. *Synapse* 1999;32:44–50. Copyright © 1999, Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.

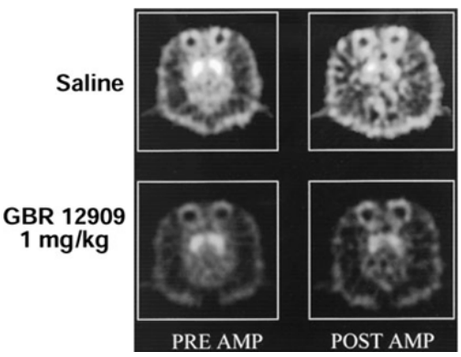


FIGURE 34.3. These images illustrate the binding of [^{11}C]raclopride to the basal ganglia of *Papio anubis* baboons treated with saline *(top row)* and GBR (1 mg/kg) *(bottom row)* after the administration of saline (3 mL/kg) (PRE AMP) *(left column)* or amphetamine (1 mg/kg) (POST AMP) *(right column)*. After the administration of saline (3 mg/kg) *(top row)* there is prominent binding of [^{11}C]raclopride to the basal ganglia at baseline (PRE AMP) *(upper left)* and significant reduction after the administration of amphetamine (1 mg/kg) (POST AMP) *(upper right)*. After the administration of GBR (1 mg/kg) *(bottom row)* there is reduced binding of [^{11}C]raclopride to the basal ganglia at baseline (PRE AMP) *(lower left)* and minimal reduction after the administration of amphetamine (1 mg/kg) (POST AMP) *(lower right)*. Modified from Villemagne VL, Wong DF, Yokoi F, Stephane M, Rice KC, Matecka D, Clough DJ, Dannals RF, Rothman RB. GBR12909 attenuates amphetamine-induced striatal dopamine release as measured by continuous infusion PET scans. *Synapse 1999;33:268–273*. Copyright © 1999, Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc. See color version of figure.

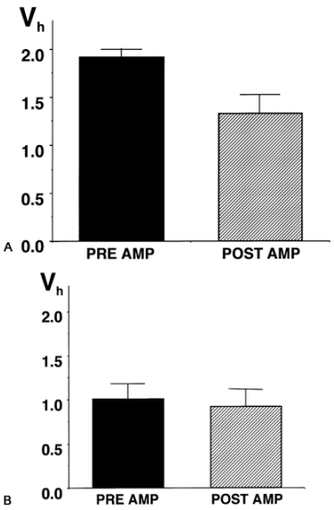


FIGURE 34.4. These histograms contrast the release of DA after administration of amphetamine (AMP) in *Papio anubis* baboons treated with saline *(left)* and GBR (1 mg/kg) *(right)*. For each histogram, the abscissa indicates the intravenous administration of saline (3 mL/kg) (PRE AMP) or amphetamine (1 mg/kg) (POST AMP) and the ordinate is the average volume of distribution (V_h) and the standard error of the mean. The left histogram illustrates a significant reduction in V_h consistent with the release of DA induced by amphetamine. The right histogram demonstrates a reduced baseline (PRE AMP) V_h after administration of GBR consistent with an increase in baseline extracellular intrasynaptic DA concentration (PRE AMP) and the absence of a significant change in V_h after the administration of amphetamine (POST AMP). Modified from Villemagne VL, Wong DF, Yokoi F, Stephane M, Rice KC, Matecka D, Clough DJ, Dannals RF, Rothman RB. GBR12909 attenuates amphetamine-induced striatal dopamine release as measured by [^{11}C]raclopride continuous infusion PET scans. *Synapse* 1999;33:268–273. Copyright © 1999, Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.

The fourth method in which neuroreceptor imaging can assist in drug development is the empirical evaluation of theories of disease, such as the DA hypothesis for schizophrenia. For example, Grace (1991) (121) proposed that schizophrenia is characterized by intrasynaptic concentrations of DA that are abnormally low in the basal tonic state and abnormally high in the simulated phasic state. This has been supported by numerous findings of elevated dopa decarboxylase measurements using [^{18}F]fluorodopa (122), elevated amphetamine induced dopamine release (123 ,124 and 125), and elevated D_2Rs (97 ,126). There is also some potential evidence that elevated intrasynaptic dopamine release is also found at baseline (126 ,127 and 128). In this example of the DA system, the combined strength of measuring presynaptic, postsynaptic, and intrasynaptic DA, for example, provides converging evidence to test this hypothesis. Development of additional ligands such as those for glutamate, glycine, and second messengers, will further expand the potential to evaluate the complex pathophysiology of schizophrenia.

MAGNETIC RESONANCE SPECTROSCOPY

Magnetic resonance spectroscopy (MRS) provides surrogate markers to determine clinical endpoints to evaluate new therapeutic interventions for specific diseases. For example, choline metabolites estimated through photon MRS function as surrogate markers of cognitive motor symptom severity in HIV-1 (129). Additionally, [^{19}F]MRS has been explicitly employed to examine the pharmacokinetics of psychotropic medications containing a fluorine group such as fluvoxamine (130). This allows direct comparison of brain and plasma concentrations, and brain elimination half-lives (131), although the sensitivity allows only micromolar measures in contrast to the nanomolar measures attained with PET. These methods have also been employed in animal models using [^{19}F]nuclear magnetic resonance (NMR) chemical shift imaging to study the cerebral distribution of general anesthetics *in vivo* (132).

MAGNETIC RESONANCE IMAGING

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

Magnetic resonance imaging (MRI) provides surrogate markers for disease progression to facilitate drug development. For example, T2-weighted cerebral MRI functions as a surrogate marker in early stages of demyelinating disease to predict disease progression and disability over the subsequent 10 years (132a and 132e).

Another example of the use of MRI as a surrogate marker for drug development is in the diagnosis and treatment of subjects with AD. Atrophy of the hippocampal formation has been correlated with memory and cognitive impairments. Reductions in the volume of the hippocampus have been predictive for the individuals who later develop memory impairments consistent with AD (133). Another marker of the vulnerability to develop AD is the measure of the apolipoprotein E (APOE) genotype. The relative risk for AD is increased for people with the gene (134). The APOE gene is associated with loss of hippocampal volume (135). Cross-sectional studies in 116 healthy volunteers, 59 to 85 years old, demonstrated significantly larger ventricular volumes and smaller gray and white matter volumes in older compared to younger individuals and in men compared to women over a period of only 1 year. An increase of over 1,500 mm^3 in ventricular volume was demonstrated during this time but no detectable change in the total or regional brain volumes. This suggests that determination of the pattern and rate of these changes longitudinally could be a future predictor of cognitive declines and dementia (136).

Functional MRI (fMRI), a technique in which subjects are asked to perform particular mental or physical tasks while the MRI is obtained, may provide biomarkers useful for drug development. For example, the rCBF of cocaine-dependent subjects administered intravenous cocaine exhibited increases in the nucleus accumbens, subcallosal cortex, and hippocampus and decreases in rCBF in the amygdala, temporal pole, and medial frontal cortex (137). Future studies utilizing this paradigm could additionally assess whether another compound administered before cocaine can antagonize or amplify its affects.

BIOMARKERS OF ALCOHOL ABUSE

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

The synergistic combination of carbohydrate-deficient transferrin (CDT), gamma-glutamyl transferase (GGT),

and mean red cell volume (MCV) functions as a possible biomarker for alcohol abuse (138).

NEWER MARKERS FOR DRUG DEVELOPMENT

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

Several techniques recently have been developed to provide the means to assess the efficacy of newly developed potential drugs. These procedures assist clinicians in both the (a) diagnosis of neuropsychiatric disorders, and (b) monitoring of the course of disease progression and the response to drug treatments and other therapeutic interventions. The application of the following novel procedures to drug development is the result of a collaborative effort jointly of the National Institutes of Health (NIH), the United States Food and Drug Administration (FDA), the Health Care Financing Administration (HCFA), and private industry (139). The protocols described in this section are likely to be instrumental in drug development for neuropsychiatric disorders in the future.

External Imaging of Internal Bioluminescent and Fluorescent Signals

Biological processes, such as the propagation of cancer cells, are studied in animal models by the detection of bioluminescent (140) and fluorescent signals recorded externally by means of sensitive photon detection systems (141). For example, optical imaging systems visualize near-infrared fluorescent molecular targets triggered by enzymes released by cancer in experimental animals (142). Other examples include the evaluation of potential agents to treat human cancer and infection by assessing the behavior of cells from patients in animals utilizing bioluminescent markers. The technique utilizes bioluminescent compounds bound to specific enzymes that signal the proliferation of cancers and bacteria. Human subjects with cancer or infection are treated with agents to alter the expression of genes that produce the enzymes needed for the growth of the cancer or bacteria. Then animals are injected with (a) cells from the previously cancerous or infected tissue of treated humans, and (b) bioluminescent markers that respond to the presence of the enzyme that was hopefully blocked in the humans. Detection of the bioluminescent markers in the animals indicates the proliferation of the enzyme to be blocked in the treated human and, therefore, the failure of the therapeutic attempt. This technique, which provides a surrogate for the clinical effects, is being applied to a variety of human diseases including cancer and infection (http://www.xenogen.com/). By analogy, this procedure could be applied to assess drugs developed to treat malignancy and infection of the CNS. For example, novel agents to alter the expression of the genes that code for the control of the production of enzymes required by cancer and infection of the CNS could be administered as therapeutic agents to suitable patients with malignancy and infection of the CNS. CSF from treated patients and bioluminescent compounds responsive to the presence of the enzyme that should be absent could then be injected into experimental animals. Detection of the bioluminescence in the experimental animal would then indicate failure of the new drug to prevent the growth of the tumor or infection. Thus, this technique offers the potential of surrogate markers to monitor the presence of CNS tumors and infections.

Simultaneous Optical and Magnetic Resonance Microscopy

Simultaneous optical and magnetic resonance imaging (MRI) is being developed in experimental animals. MRI contrast agents (143), such as fluorescently detectable magnetic resonance imaging agents, are utilized to permit light and magnetic resonance imaging microscopy at the same time (144). These compounds can be detected deep in tissue, not merely at the surface as with simple optical detection systems (e.g., for fluorescent dyes) (145). Although light microscopy can penetrate only 100 to 300 µm beneath the surface of organisms, MRI microscopy can penetrate 1 to 6 mm into an organism (144). For example, agents that can be simultaneously recognized by MRI microscopy and by fluorescent optical microscopy permit visualization of structures 1 to 6 mm below the surface of an organism approaching cellular resolution (i.e., 10 µm) (143 ,144).

Diffusion-Based Optical Imaging

Procedures to measure light emitted into opaque structures have been termed medical optical imaging (MOI), medical optical spectroscopy (MOS), near-infrared imaging (NIRI), and near-infrared optical spectroscopy (NIOS) (146). For example, three-dimensional optical coherence tomography (OCT) is a technique to image nontransparent biologic tissue by recording and analyzing light emitted into scattering media. OCT has been employed to visualize nerve fascicles in experimental animals for the microsurgical anastomoses of vessels and nerves (147). An example of diffusion-based optical imaging is the use of optical tomography to detect intraventricular hemorrhage in premature infants by external transmission of light emitted through the skull (146). These techniques may be utilized to develop treatments for human disease, including infections and malignancies (146). Hopefully, in the future they will be applied to CNS tumors and malignancies.

Magnetic Resonance Microscopy

Microscopic visualization of magnetic resonance images has detected transgene expression in experimental animals.

Identification of pathologic processes, including the proliferation of tumor cells in clinical settings (148), may be facilitated by this procedure. This technique offers the means to both detect the occurrence of malignancies and to monitor their growth (148). The application of this procedure to human CNS malignancy is a goal to be attained in the future.

Electron Paramagnetic Resonance

Electron paramagnetic resonance (EPR) (149) imaging and spectroscopy are procedures to spatially map parameters of physiologic importance by incorporating paramagnetic spin labels into the system of interest (150). This technique has been utilized to visualize oxygen concentration in the tissues of experimental animals (150) and to measure oxygen free radical generation in human endothelial cells exposed to anoxia and reoxygenation (151). EPR has been employed to estimate the production of nitric oxide in biological systems (152), a process vital to the measurement of the progression of pathologic processes (153), including cerebral ischemia (154) and malignancies.

CONCLUSION

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

Biomarkers and surrogate markers are tools currently utilized to develop new drugs. They provide evidence of the proof of concept required for successful Phase 1B/2A studies submitted to the FDA. Currently drugs are being developed by the use of neuroendocrine markers including CSF, prolactin, GH, ACTH, and cortisol. Imaging studies provide the means to estimate therapeutic dosages of new drugs. Surrogate markers include a variety of neuroimaging techniques including MRI, MRS, PET, and SPECT. Newer techniques for drug development are likely to include external imaging of internal bioluminescent and fluorescent signals, simultaneous optical and MRM, diffusion-based optical imaging, and EPR.

ACKNOWLEDGMENT

Part of "34 - Proof of Concept: Functional Models for Drug Development in Humans "

We acknowledge grant support from United States Public Health Service Grants K24DA00412 (Wong), and from DA09482, DA11080, NS38927, MH42821 (Wong/Brasic), the Rett Syndrome Research Foundation (RSRF) (Wong/Brasic), and from NARSAD and The Essel Foundation (Brasic).

REFERENCES

1. Mildvan D, Kagan J. Biomarkers as surrogate endpoints in clinical trials for HIV/AIDS: a model for other diseases? In: National Institutes of Health and Food and Drug Administration, eds. *Biomarkers and surrogate endpoints: advancing clinical research and applications program book*. Bethesda, Maryland: National Institutes of Health, 1999;15–16.

2. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med* 1989;8:431–440.

3. Swenberg JA. Bioassay design and MTD setting: old methods and new approaches. *Regul Toxicol Pharmacol* 1995;21:44–51; discussion 81–86.

4. Goldstein DJ, Potter WZ. Determination of appropriate dosage in clinical development guidance. In: Balant LP, Benitez J, Dahl SG, et al., eds. *Clinical pharmacology in psychiatry: finding the right dose of psychotropic drugs*. Brussels: European Commission, 1998:307–322.

5. Sramek JJ, Cutler NR, Hurley DJ, et al. The utility of salivary amylase as an evaluation of M3 muscarinic agonist activity in Alzheimer's disease. *Prog Neuro-Psychopharmacol Biol Psychiatr* 1995;19:85–91.

6. Sramek JJ, Cutler NR, Kurtz NM, et al. Phase I clinical development and finding the dose: the role of the bridging study. In: Sramek JJ, Cutler NR, Kurtz NM, et al., eds. *Optimizing the development of antipsychotic drugs*. New York: John Wiley and Sons, 1997:101–126.

7. Preskorn SH. Recent dose-effect studies regarding antidepressants. In: Balant LP, Benitez J, Dahl SG, et al., eds. *Clinical pharmacology in psychiatry: finding the right dose of psychotropic drugs*. Brussels: European Commission, 1998:45–61.

8. Preskorn SH. Debate resolved: there are differential effects of serotonin selective reuptake inhibitors on cytochrome P450 enzymes. *J Psychopharmacol* 1998;12:S89–S97.

9. Debonnel G, Blier P, Saint-Andre E, et al. Comparison of the effects of low and high doses of venlafaxine on serotonin and norepinephrine reuptake processes in patients with major depression and healthy volunteers. *Int J Neuropsychopharmacol* 1998;1:S17.

10. Nierenberg AN, Feighner JP, Rudolph R, et al. Venlafaxine for treatment-resistant unipolar depression. *J Clin Psychopharmacol* 1994;14:419–423.

11. Derivan A, Entsuah AR, Kikta D. Venlafaxine: measuring the onset of antidepressant action. *Psychopharmacol Bull* 1995;31:439–447.

12. Peck CC, Barr WH, Benet LZ, et al. Opportunities for integration of pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development. *J Clin Pharmacol* 1994;34:111–119.

13. Briley M. Clinical dose prediction on the basis of preclinical data. In: Balant LP, Benitez J, Dahl SG, et al., eds. *Clinical pharmacology in psychiatry: finding the right dose of psychotropic drugs*. Brussels: European Commission, 1998:15–19.

14. Bergstrom RF, Lemberger L. First time in man studies: an industrial perspective. *J Clin Pharmacol* 1990;30:212–217.

15. Ellinwood EH, Nikaido AM. Perceptual-neuromotor pharmacodynamics of psychotropic drugs. In: Meltzer HY, ed. *Psychopharmacology: the third generation of progress*. New York: Raven Press, 1987:1457–1466.

16. Hindmarch I. The behavioural toxicity of the selective serotonin reuptake inhibitors. *Int Clin Psychopharmacol* 1995;9:13–17.

17. Hindmarch I, Alford C, Barwell F, et al. Measuring the side effects of psychotropics: the behavioural toxicity of antidepressants. *Psychopharmacology* 1992;6:198–203.

18. Patat A, Perault MC, Vandel B, et al. Assessment of the interaction between a partial agonist and a full agonist of benzodiazepine receptors, based on psychomotor performance and memory, in healthy volunteers. *Psychopharmacology* 1995;9:91–101.

19. Greenblatt DJ. The benzodiazepines: kinetic-dynamic relationships. In: Herdman JRE, Delva NJ, Hockney RA, et al., eds. *Integration of pharmacokinetics, pharmacodynamics and toxicokinetics in rational drug development*. New York: Plenum, 1993:217–224.

20. Saletu B, Grunberg J. Drug profiling by computed electroencephalography and brain maps, with special consideration of sertraline and its psychometric effects. *J Clin Psychiatry* 1988;49:59–71.

21. Itil TM. The discovery of psychotropic drugs by computer-analyzed cerebral bioelectric potentials (CEEG). *Drug Dev Res* 1981;1:373–407.

22. Dago KT, Luthringer R, Lengelle R, et al. Statistical decision tree: a tool for studying pharmaco-EEG effects of CNS-active drugs. *Neuropsychobiology* 1994;29:91–96.

23. Sheiner LB. Learning versus confirming in clinical drug development. *Clin Pharmacol Ther* 1997;61:275–291.

24. Ingvar M, Ambros-Ingerson J, Davis M, et al. Enhancement by an ampakine of memory encoding in humans. *Exp Neurol* 1997;146:553–559.

25. Kenny M, Lenehan TJ, Lambe R, et al. The effect of PK5078, a new serotonin uptake inhibitor, on serotonin levels and uptake in human platelets, following administration to healthy volunteers. *Eur J Clin Pharmacol* 1983;25:23–28.

26. Lemberger L, Rowe H, Carmichael R, et al. Fluoxetine, a selective serotonin uptake inhibitor. *Clin Pharmacol Ther* 1978;23:421–429.

27. Potter WZ. Psychotherapeutic drugs and biogenic amines: current concepts and therapeutic implications. *Drugs* 1984;28:127–143.

28. Asberg M, Bertilsson L, Tuck D, et al. Indoleamine metabolites in the cerebrospinal fluid of depressed patients before and during treatment with nortriptyline. *Clin Pharmacol Ther* 1973;14:277–286.

29. Bertilsson L, Asberg M, Thoren P. Differential effect of chlorimipramine and nortriptyline on cerebrospinal fluid metabolites of serotonin and noradrenaline in depression. *Eur J Clin Pharmacol* 1974;7:365–368.

30. Linnoila M, Karoum F, Calil HM, et al. Alteration of norepinephrine metabolism with desipramine and zimelidine in depressed patients. *Arch Gen Psychiatry* 1982;39:1025–1028.

31. Bowden CL, Koslow SH, Hanin I, et al. Effects of amitriptyline and imipramine on brain amine neurotransmitter metabolites in cerebrospinal fluid. *Clin Pharmacol Ther* 1985;37:316–324.

32. Ross RJ, Zavadil AP, Calil HM, et al. The effects of desmethylimipramine on plasma norepinephrine, pulse and blood pressure in volunteers. *Clin Pharmacol Ther* 1983;33:429–437.

33. Veith C, Lewis N, Linares OA, et al. Sympathetic nervous system activity in major depression: basal and desipramine-induced alterations in plasma norepinephrine kinetics. *Arch Gen Psychiatry* 1994;51:411–422.

34. Esler M, Jackman G, Bobik A, et al. Determination of norepinephrine apparent release rate and clearance in humans. *Life Sci* 1979;25:1461–1470.

35. Ghose K. Tyramine pressor test: implications and limitations. *Methods and Find Exp Clin Pharmacol* 1984;6:455–464.

36. Abdelmawla AH, Langley RW, Szabadi E, et al. Comparison of the effects of venlafaxine, desipramine, and paroxetine on noradrenaline- and methoxamine-evoked constriction of the dorsal hand vein. *Br J Clin Pharmacol* 1999;48:345–354.

37. Kopin IJ. Catecholamine metabolism: basic aspects and clinical significance. *Pharmacol Rev* 1985;37:333–364.

38. Linnoila M, Guthrie S, Lane EA, et al. Clinical studies on norepinephrine metabolism: how to interpret the numbers. *Psychiatry Res* 1986;17:229–239.

39. Coupland NJ, Bailey JE, Wilson SJ, et al. A pharmacodynamic study of the α2-adrenergic receptor antagonist ethoxyidazoxan in healthy volunteers. *Clin Pharmacol Ther* 1994;56:420–429.

40. Golden RN, Rudorfer MV, Sherer MA, et al. Bupropion in depression. I. Biochemical effects and clinical response. *Arch Gen Psychiatry* 1988;45:139–143.

41. Glue P, Wilson S, Lawson C, et al. Acute and chronic idazoxan in normal volunteers: biochemical, physiological and psychological effects. *J Psychopharmacol* 1991;5:394–401.

42. Pickar D, Owen RR, Litman RE, et al. Clinical and biologic response to clozapine in patients with schizophrenia: crossover comparison with fluphenazine. *Arch Gen Psychiatry* 1992;49:345–353.

43. Sunderland T, Cohen RM, Molchan S, et al. High-dose selegiline in treatment-resistant older depressive patients. *Arch Gen Psychiatry* 1994;51:607–615.

44. Little JT, Ketter TA, Mathe AA, et al. Venlafaxine but not bupropion decreases cerebrospinal fluid 5-hydroxyindoleacetic acid in unipolar depression. *Biol Psychiatry* 1999;45:285–289.

45. Major LF, Murphy DL, Lipper S, et al. Effects of clorgyline and pargyline on deaminated metabolites of norepinephrine, dopamine and serotonin in human cerebrospinal fluid. *J Neurochem* 1979;32:229–231.

46. Potter WZ, Scheinin M, Golden RN, et al. Selective antidepressants and cerebrospinal fluid: lack of specificity on norepinephrine and serotonin metabolites. *Arch Gen Psychiatry* 1985;42:1171–1177.

47. Bjerkenstedt L, Flyckt L, Overo KF, et al. Relationship between clinical effects, serum drug concentration and serotonin uptake inhibition in depressed patients treated with citalopram. A double-blind comparison of three dose levels. *Eur J Clin Pharmacol* 1985;28:553–557.

48. Marsden CA, Tyrer P, Casey P, et al. Changes in human whole blood 5-hydroxytryptamine (5-HT) and platelet 5-HT uptake during treatment with paroxetine, a selective 5-HT uptake inhibitor. *J Psychopharmacol* 1987;1:244–250.

49. Ishigooka J, Nagata E, Takahashi A, et al. Simultaneous monitoring of inhibition of serotonin uptake by platelets and plasma drug concentrations following administration of duloxetine, a new antidepressant candidate, to healthy volunteers. *Current Ther Res* 1997;10:679–692.

50. Lemberger L, Rowe H, Carmichael R, et al. Pharmacologic effects in man of a specific serotonin-reuptake inhibitor. *Science* 1978;199:436–437.

51. Potter WZ, Rudorfer MV, Lesieur P, et al. Biochemical effects of selective 5HT-reuptake inhibitors in man. In: Gastpar M, ed. *Selective 5-HT reuptake inhibitors: novel or common place agents*. Basel: S Karger, 1988:18–30.

52. Murphy DL, Lipper S, Campell IC, et al. Comparative studies of MAO-A and MAO-B inhibitors in man. In: Singer TP, Von Korff RW, Murphy DL, eds. *Monoamine oxidase: structure, function, and altered functions*. New York: Academic Press, 1979:457–475.

53. Lambert GW, Eisenhofer G, Jennings G, et al. Regional homovanillic acid production in humans. *Life Sci* 1993;53:63–75.

54. Alfredsson G, Bjerkenstedt L, Edman G, et al. Relationships between drug concentrations in serum and CSF, clinical effects and monoaminergic variables in schizophrenic patients treated with sulpiride or chlorpromazine. *Acta Psychiatr Scand Suppl* 1984;311:49–74.

55. Davidson M, Losonczy MF, Mohs RC, et al. Effects of debrisoquin and haloperidol on plasma homovanillic acid concentration in schizophrenic patients. *Neuropsychopharmacology* 1987;1:17–23.

56. Pickar D, Labarca R, Linnoila M, et al. Neuroleptic-induced decrease in plasma homovanillic acid and antipsychotic activity in schizophrenic patients. *Science* 1984;225:954–957.

57. Davila R, Manero E, Zumarraga M, et al. Plasma homovanillic acid as predictor of response to neuroleptics. *Arch Gen Psychiatry* 1988;45:564–567.

58. Chang WH, Chen TY, Lin SK, et al. Plasma catecholamine metabolites in schizophrenics: evidence for the two-subtype concept. *Biol Psych* 1990;27:510–518.

59. Meltzer HY. Long-term effects of neuroleptic drugs on the neuroendocrine system. *Adv Biochem Psychopharmacol* 1985;40:59–68.

60. Nordstrom AL, Farde L. Plasma prolactin and central D2 receptor occupancy in antipsychotic drug-treated patients. *J Clin Psychopharmacol* 1998;18:305–310.

61. Avrantis LA, Miller BG. Multiple fixed doses of "Seroquel" (quetiapine) in patients with acute exacerbation of schizophrenia: a comparison with haloperidol and placebo: the Seroquel Trial 13 Study Group. *Biol Psychiatry* 1997;42:233–246.

62. Kapur S, Zipursky R, Jones C, et al. A positron emission tomography study of quetiapine in schizophrenia: a preliminary finding of an antipsychotic effect with only transiently high dopamine D_2 receptor occupancy. *Arch Gen Psychiatry* 2000;57:553–559.

63. Cowen PJ. Use of surrogate markers in clinical psychopharmacology. In: Balant LP, Benitez J, Dahl SG, et al., eds. *Clinical pharmacology in psychiatry: finding the right dose of psychotropic drugs*. Brussels: European Commission, 1998:151–160.

64. Seletti B, Benkelfat C, Blier P, et al. Serotonin$_{1A}$ receptor activation by flesinoxan in humans: body temperature and neuroendocrine responses. *Neuropsychopharmacology* 1995;13:93–104.

65. Murphy DL, Aulakh CS, Mazzola-Pomietto P, et al. Neuroendocrine responses to serotonergic agonists as indices of the functional status of central serotonin neurotransmission in humans: a preliminary comparative analysis of neuroendocrine endpoints versus other endpoint measures. *Behav Brain Res* 1996;73:209–214.

66. Golden RN, Hsiao J, Lane E, et al. The effects of intravenous clomipramine on neurohormones in normal subjects. *J Clin Endocrinol Metab* 1989;68:632.

67. Meltzer HY, Maes M. Effect of pindolol on the L-5-HTP-induced increase in plasma prolactin and cortisol concentrations in man. *Psychopharmacology (Berl)* 1994;114:635–643.

68. Seibyl JP, Krystal JH, Price LH, et al. Effects of ritanserin on the behavioural, neuroendocrine, and cardiovascular responses to meta-chlorophenylpiperazine in healthy human subjects. *Psychiatry Res* 1991;38:227–236.

69. Silverstone PH, Cowen PJ. The 5-HT$_3$ antagonist BRL 46470 does not attenuate m-chlorophenylpiperazine (mCPP)-induced changes in human volunteers. *Biol Psychiatry* 1994;36:309–316.

70. Goddard AW, Wood SW, Sholomskas DE, et al. Effects of the serotonin reuptake inhibitor fluvoxamine on yohimbine-induced anxiety in panic disorder. *Psychiatry Res* 1993;48:119–133.

71. Siever LJ. Role of noradrenergic mechanisms in the etiology of the affective disorders. In: Meltzer HY, ed. *Psychopharmacology: the third generation of progress*. New York: Raven Press, 1987:493–504.

72. Laakman G, Hinz A, Voderholzer U, et al. The influence of psychotropic drugs and releasing hormones on anterior pituitary hormone secretion in healthy subjects and depressed patients. *Pharmacopsychiatry* 1990;23:18–26.

73. Facchinetti F, Nappi RE, Sances G, et al. The neuroendocrine effects of sumatriptan, a specific ligand for 5-HT$_1$-like receptors. *Clin Endocrinol* 1994;40:211–214.

74. Herdman JRE, Delva NJ, Hockney RA, et al. Neuroendocrine effects of sumatriptan. *Psychopharmacology* 1994;113:561–564.

75. Whale R, Bhagwagar Z, Cowen PJ. Zolmitriptan-induced growth hormone release in humans: mediation by 5-HT$_{1D}$ receptors? *Psychopharmacology* 1999;145:223–226.

76. Mota A, Bento A, Penalva A, et al. Role of the serotonin receptor subtype 5-HT1D on basal and stimulated growth hormone secretion. *J Clin Endocrinol Metab* 1995;80:1973–1977.

77. Dinan TG. Serotonin and the regulation of hypothalamic-pituitary-adrenal axis function. *Life Sci* 1996;58:1683–1694.

78. Gartside SE, Cowen PJ. Mediation of ACTH and prolactin responses to 5-HTP by 5-HT$_2$ receptors. *Eur J Pharmacol* 1990;179:103–109.

79. Cowen PJ. Serotonin receptor subtypes in depression: evidence from studies in neuroendocrine regulation. *Clin Neuropharmacol* 1993;16:S6–S18.

80. Meltzer HY, Gudelsky GA, Lowy MT, et al. Neuroendocrine effects of buspirone: mediation by dopaminergic and serotonergic mechanisms. In: Tunnicliff G, Eison AS, Taylor OP, eds. *Buspirone: mechanisms and clinical aspects*. San Diego: Academic Press, 1991:177–192.

81. Sharpley AL, Elliott JM, Attenburrow MJ, et al. Slow wave sleep in humans: role of 5-HT$_{2A}$ and 5-HT$_{2C}$ receptors. *Neuropharmacology* 1994;33:467–471.

82. Katsuda Y, Walsh AES, Ware CJ, et al. Meta-Chlorophenylpiperazine decreases slow-wave sleep in humans. *Biol Psychiatry* 1993;33:49–51.

83. Murphy DL, Wichems C, Li Q, et al. Molecular manipulations as tools for enhancing our understanding of 5-HT neurotransmission. *Trends Pharmacol Sci* 1999;20:246–252.

84. Potter WZ, Muscettola G, Goodwin FK. Binding of imipramine to plasma proteins and to brain tissue: relationship to CSF tricyclic levels in man. *Psychopharmacology* 1979;63:187–192.

85. Neil-Dwyer G, Bartlett J, McAinsh J, et al. Beta-adrenoceptor blockers and the blood-brain barrier. *Br J Clin Pharmacol* 1981;11:549–553.

86. Rapeport WG, Mendelow AD, French G, et al. Plasma protein-binding and CSF concentrations of valproic acid in man following acute oral dosing. *Br J Clin Pharmacol* 1983;16:365–369.

87. Kornhuber J, Quack G. Cerebrospinal fluid and serum concentrations of the N-methyl-D-aspartate (NMDA) receptor antagonist meantime in man. *Neurosci Lett* 1995;195:137–139.

88. Yonekura Y, Nishizawa S, Mukai T, et al. SPECT with [99mTc]-d,l-hexamethyl-propylene amine oxime (HM-PAO) compared with regional cerebral blood flow measured by PET: effects of linearization. *J Cereb Blood Flow Metab* 1988;8:S82–S89.

89. Cooper M, Metz J, de Wit H, et al. Interclass drug effects and changes in regional brain glucose metabolism. *Psychopharmacol Bull* 1998;34:229–232.

90. Grant S, London ED, Newlin DB, et al. Activation of memory circuits during cue-elicited cocaine craving. *Proc Natl Acad Sci USA* 1996;93:12040–12045.

91. Childress AR, Mozley PD, McElgin W, et al. Limbic activation during cue-induced cocaine craving. *Am J Psychiatry* 1999;156:11–18.

92. Carlsson A. The current status of the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology* 1988;1:179–186.

93. Wong DF, Wagner HN Jr, Coyle J. Assessment of dopamine receptor blockage of neuroleptic drugs in the living human brain. *J Nucl Med* 1985;26:52.

94. Wong DF, Singer HS, Brandt J, et al. D2-like dopamine receptor density in Tourette syndrome measured by PET. *J Nucl Med* 1997;38:1243–1247.

95. Farde L, Grind M, Nilsson MI, et al. Remoxipride—a new potential antipsychotic drug. Pharmacological effects and pharmacokinetics following repeated oral administration in male volunteers. *Psychopharmacology (Berl)* 1988;95:157–161.

96. Farde L, Hall H. Positron emission tomography—examination of chemical transmission in the living human brain. Development of radioligands. *Arzneimittelforschung* 1992;42:260–264.

97. Wong DF, Wagner HN Jr, Tune LE, et al. Positron emission tomography reveals elevated D2 dopamine receptors in drug-naive schizophrenics. *Science* 1986;234:1558–1563. [Published erratum appears in *Science* 1987;235:623.]

98. Baron JC, Martinot JL, Cambon H. Striatal dopamine receptor occupancy during and following withdrawal from neuroleptic treatment: correlative evaluation by positron emission tomography and plasma prolactin levels. *Psychopharmacology* 1989;99:463–472.

99. Farde L. Selective D1- and D2-dopamine receptor blockade both induces akathisia in humans—a PET study with [^{11}C]SCH 23390 and [^{11}C]raclopride. *Psychopharmacology (Berl)* 1992;107:23–29.

100. Wolkin A, Brodie JD, Barouche F, et al. Dopamine receptor occupancy and plasma haloperidol levels. [Letter] *Arch Gen Psychiatry* 1989;46:482–484.

101. Kapur S, Zipursky R, Jones C, et al. Relationship between dopamine D(2) occupancy, clinical response, and side effects: a double-blind PET study of first-episode schizophrenia. *Am J Psychiatry* 2000;157:514–520.

102. Daskalakis Z, Christensen B, Zipursky R, et al. Relationship between D2 occupancy and prolactin levels in first episode psychosis. *Biol Psychiatry* 1998;43:113S–113S.

103. Farde L, Halldin C, Muller L, et al. PET study of [^{11}C]beta-CIT binding to monoamine transporters in the monkey and human brain. *Synapse* 1994;16:93–103.

104. Nyberg S, Farde L, Eriksson L, et al. 5-HT2 and D2 dopamine receptor occupancy in the living human brain. A PET study with risperidone. *Psychopharmacology* 1993;110:265–272.

105. Nyberg S, Eriksson B, Oxenstierna G, et al. Suggested minimal effective dose of risperidone based on PET-measured D2 and 5-HT2A receptor occupancy in schizophrenic patients. *Am J Psychiatry* 1999;156:869–875.

106. Farde L, Hall H, Pauli S, et al. Variability in D2-dopamine receptor density and affinity: a PET study with [^{11}C]raclopride in man. *Synapse* 1995;20:200–208.

107. Remington G, Kapur S, Zipursky R. The relationship between risperidone plasma levels and dopamine D2 occupancy: a positron emission tomographic study. *J Clin Psychopharmacol* 1998;18:82–83.

108. Kapur S, Zipursky RB, Remington G, et al. 5-HT2 and D2 receptor occupancy of olanzapine in schizophrenia: a PET investigation. *Am J Psychiatry* 1998;155:921–928.

109. Pilowsky LS, Costa DC, Ell PJ, et al. Antipsychotic medication, D2 dopamine receptor blockade and clinical response: a 123I IBZM SPET (single photon emission tomography) study. *Psychol Med* 1993;23:791–797.

110. Nordstrom AL, Farde L, Halldin C. High 5-HT2 receptor occupancy in clozapine treated patients demonstrated by PET. *Psychopharmacology (Berl)* 1993;110:365–367.

111. Kapur S, Zipursky RB, Jones C, et al. The D2 receptor occupancy profile of loxapine determined using PET. *Neuropsychopharmacology* 1996;15:562–566.

112. Kapur S, Zipursky R, Roy P, et al. The relationship between D2 receptor occupancy and plasma levels on low dose oral haloperidol: a PET study. *Psychopharmacology* 1997;131:148–152.

113. Volavka J, Cooper TB, Czobor P, et al. Plasma haloperidol levels and clinical effects in schizophrenia and schizoaffective disorder. *Arch Gen Psychiatry* 1995;52:837–845.

114. Kapur S, Remington G, Jones C, et al. High levels of dopamine D2 receptor occupancy with low-dose haloperidol treatment: a PET study [see comments]. *Am J Psychiatry* 1996;153:948–950.

115. Kapur S, Remington G, Zipursky RB, et al. The D2 dopamine receptor occupancy of risperidone and its relationship to extrapyramidal symptoms: a PET study. *Life Sci* 1995;57:L103–L107.

116. Nyberg S, Eriksson B, Oxenstierna G, et al. Suggested minimal effective dose of risperidone based on PET-measured D2 and 5-HT2A receptor occupancy in schizophrenic patients. *Am J Psychiatry* 1999;156:869–875.

117. Grunder G, Yokoi F, Offord SJ, et al. Time course of 5-HT2A receptor occupancy in the human brain after a single oral dose of the putative antipsychotic drug MDL 100,907 measured by positron emission tomography. *Neuropsychopharmacology* 1997;17:175–185. [Published erratum appears in *Neuropsychopharmacology* 1998;19:161.]

118. Fitzgerald PB, Kapur S, Remington G, et al. Predicting haloperidol occupancy of central dopamine D2 receptors from plasma levels. *Psychopharmacology (Berl)* 2000;149:1–5.

119. Glowa JR, Fantegrossi WE, Lewis DB, et al. Sustained decrease in cocaine-maintained responding in rhesus monkeys with 1-[2-[bis-(4-fluorophenyl)methoxy]ethyl]-4-(3-hydroxy-3-phenylpropyl)piperazinyl decanoate, a long-acting ester derivative of GBR 12909. *J Med Chem* 1996;39:4689–4691.

120. Villemagne V, Rothman RB, Yokoi F, et al. Doses of GBR12909 that suppress cocaine self-administration in non-human primates substantially occupy dopamine transporters as measured by [^{11}C]WIN35,428 PET scans. *Synapse* 1999;32:44–50.

120a. Villemagne VL, Wong DF, Yokoi F, et al. GBR12909 attenuates amphetamine-induced striatal dopamine release as measured by [^{11}C]raclopride continuous infusion PET scans. *Synapse* 1999;268–273.

121. Grace AA. Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis for the etiology of schizophrenia. *Neuroscience* 1991;41:1–24.

122. Reith J, Benkelfat C, Sherwin A, et al. Elevated dopa decarboxylase activity in living brain of patients with psychosis. *Proc Natl Acad Sci USA* 1994;91:11651–11654.

123. Laruelle M, Abi-Dargham A, Van Dyck CH, et al. Single photon emission computerized tomography imaging of amphetamine-induced dopamine release in drug-free schizophrenic subjects. *Proc Natl Acad Sci USA* 1996;93:9235–9240.

124. Abi-Dargham A, Gil R, Krystal J, et al. Increased striatal dopamine transmission in schizophrenia: confirmation in a second cohort. *Am J Psychiatry* 1998;155:761–767.

125. Breier A, Su TP, Saunders R, et al. Schizophrenia is associated with elevated amphetamine-induced synaptic dopamine concentrations—evidence from a novel positron emission tomography method. *Proc Natl Acad Sci USA* 1997;94:2569–2574.

126. Abi-Dargham A, Rodenhiser J, Printz D, et al. Increased baseline occupancy of D_2 receptors by dopamine in schizophrenia. *Proc Natl Acad Sci USA* 2000;97:8104–8109.

127. Wong DF, Gjedde A, Wagner HN, Jr. Quantification of neuroreceptors in the living human brain. I. Irreversible binding of ligands. *J Cereb Blood Flow Metab* 1986;6:137–146.

128. Gjedde A, Wong DF. Psychotic propensity associated with fourfold elevated dopamine binding to D2-like receptors in caudate nucleus. *Neuroimage* 1997;5:A44.

129. Chang L, Ernst T, Leonido-Yee M. Cerebral metabolite abnormalities correlate with clinical severity of HIV-1 cognitive motor complex. *Neurology* 1999;52:100–108.

130. Strauss WL, Layton ME, Dager SR. Brain elimination half-life of fluvoxamine measured by ^{19}F magnetic resonance spectroscopy. *Am J Psychiatry* 1998;155:380–384.

131. Bolo NR, Hode Y, Nedelec J-F, et al. Brain pharmacokinetics and tissue distribution in vivo of fluvoxamine and fluoxetine by fluorine magnetic resonance spectroscopy. *Neuropsychopharmacology* 2000;23:428–438.

132. Venkatasubramanian PN, Shen YJ, Wyrwicz AM. Characterization of the cerebral distribution of general anesthetics in vivo by two-dimensional ^{19}F chemical shift imaging. *Magn Reson Med* 1996;35:626–630.

132a. Fox NC, Jenkins R, Leary SM, et al. Progressive cerebral atrophy in MS. A serial study using registered, volumetric MRI. *Neurology* 2000;54:807–812.

132b. Jagust WJ, Noseworthy JH. Brain atrophy as a surrogate marker in MS. Faster, simpler, better? [Editorial]. *Neurology* 2000;54:782–783.

132c. Khan OA, Rothman MI. Use of magnetic resonance imaging techniques in multiple sclerosis: clinical applications and clues to pathogenesis. *The Neurologist* 1998;4:259–268.

132d. Nusbaum AO, Tang CY, Wei T-C, et al. Whole-brain diffusion MR histograms differ between MS subtypes. *Neurology* 2000;54:1421–1426.

132e. Weiner HL, Guttmann CRG, Khoury SJ, et al. Serial magnetic resonance imaging in multiple sclerosis: correlation with attacks, disability, and disease stage. *Neuroimmunol* 2000;104:164–173.

133. De Leon MJ, George AE, Golomb J, et al. Frequency of hippocampal formation atrophy in normal aging and Alzheimer's disease. *Neurobiol Aging* 1997;18:1–11.

134. Corder EH, Saunders AM, Strittmatter WJ, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families [see comments]. *Science* 1993;261:921–923.

135. Moffat SD, Szekely CA, Zonderman AB, et al. Longitudinal change in hippocampal volume as a function of apolipoprotein E genotype. *Neurology* 2000;55:134–136.

136. Resnick SM, Goldszal AF, Davatzkos C, et al. One-year age changes in MRI brain volumes in older adults. *Cereb Cortex* 2000;10:464–472.

137. Breiter HS, Gollub RL, Weiskoff RM. Acute effects of cocaine on human brain activity and emotion. *Neuron* 1997;19:591–611.

138. Sillanaukee P. Laboratory markers of alcohol abuse. *Alcohol Alcoholism* 1996;31:613–616.

139. Sullivan DC. Biomedical imaging symposium: visualizing the future of biology and medicine. *Radiology* 2000;215:638.

140. Contag CH, Spilman SD, Contag PR, et al. Visualizing gene expression in living mammals using a bioluminescent reporter. *Photochem Photobiol* 1997;66:523–531.

141. Contag CH, Jenkins D, Contag PR, et al. Use of reporter genes for optical measurements of neoplastic disease *in vivo*. *Neoplasia* 2000;2:41–52.

142. Mahmood U, Tung C-H, Bogdanov A Jr, et al. Near-infrared optical imaging of protease activity for tumor detection. *Radiology* 1999;213:866–870.

143. Louie AY, Huber MM, Ahrens ET, et al. In vivo visualization of gene expression using magnetic resonance imaging. *Nat Biotechnol* 2000;18:321–325.

144. Huber MM, Staubli AB, Kustedjo K, et al. Fluorescently detectable magnetic resonance imaging agents. *Bioconjug Chem* 1998;9:242–249.

145. Moats RA, Fraser SE, Meade TJ. A "smart" magnetic resonance imaging agent that reports on specific enzymatic activity. *Angew Chem Int Ed Engl* 1997;36:728.

146. Benaron DA, Contag PR. Imaging brain structure and function, infection and gene expression in the body using light. *Philos Trans R Soc Lond B Biol Sci* 1997;352:755–761.

147. Boppart SA, Bouma BE, Pitris C. Intraoprative assessment of microsurgery with three-dimensional optical coherence tomography. *Radiology* 1998;208:81–86.

148. Weissleder R, Moore A, Mahmood U, et al. In vivo magnetic resonance imaging of transgene expression. *Nat Med* 2000;6:351–354.

149. Desrosiers MF, Burlinska. Research and development activities in electron paramagnetic resonance dosimetry. *Radiation Physics and Chemistry* 1995;46:1181–1184.

150. Velan SS, Spencer RGS, Zweier JL, et al. Electron paramagnetic resonance oxygen mapping (EPROM): direct visualization of oxygen concentration in tissue. *Magn Reson Med* 2000;43:804–809.

151. Zweier JL. Free radical generation in human endothelial cells exposed to anoxia and reoxygenation. *Transplant Proc* 1998;30:4228–4232.

152. Vanin AF, Liu X, Samouilov A, et al. Redox properties of iron-dithiocarbamates and their nitrosyl derivatives: implications for their use as traps of nitric oxide in biological systems. *Biochim Biophys Acta* 2000;1474:365–377.

153. Zweier JL, Samouilov A, Kuppusamy P. Non-enzymatic nitric oxide synthesis in biological systems. *Biochim Biophys Acta* 1999;1411:259–262.

154. Wei G, Dawson VL, Zweier JL. Role of neuronal and endothelial nitric oxide synthase in nitric oxide generation in the brain following cerebral ischemia. *Biochim Biophys Acta* 1999;1455:23–24.

35

Ethical Aspects of Neuropsychiatric Research with Human Subjects

Debra A. Pinals

Paul S. Appelbaum

Debra A. Pinals and Paul S. Appelbaum: University of Massachusetts Medical School, Worcester, Massachusetts

Considerable world attention in the last century has focused on the ethics of clinical research with human subjects. Coming to the fore after World War II, with the Nuremberg War Crimes Trials, concerns were raised about the potential for abuse of nonvoluntary, uninformed subjects who might be utilized in questionable research. Out of the trials came the Nuremberg Code (1), which formalized ethical principles surrounding research with human subjects. When, in 1953, the United States opened the doors to the Clinical Center of the National Institutes of Health, guiding principles regarding human subject research at that institution were in place to greet the first subjects who enrolled in studies on site (2). By the following decade, in 1964, the World Medical Association developed the Declaration of Helsinki, which was an attempt to modify and expand upon the Nuremberg Code (1). This document classified research into clinical and nontherapeutic categories, and outlined the practice of consent that these types of research would require. The Declaration has since gone through multiple revisions and continues to be a significant guideline for research with human subjects, especially in Europe.

Despite these initial attempts to clarify the ethical principles and practices of human subject research, repeated abuses were widely publicized in the ensuing years: the Tuskegee syphilis study (3), studies involving injection of live cancerous cells in patients without their consent (4), and studies in which subjects were unknowingly exposed to radiation (5) among them. As each story was exposed, it inspired international review of the ethics of human subject research. Minority populations and those who may be vulnerable to exploitive research, such as the mentally ill, evoked particular concern. In the United States, federal commissions and agencies were created to address the concerns. Among the outcomes of these initiatives were regulations affording the government greater control over federally funded research. Institutions were required to develop Institutional Review Boards (IRBs) to review research protocols to protect human subjects and ensure an adequate consent process.

Beginning in the 1990s public concern again grew, as research with patients with mental illness became a focus of media attention (6 ,7 ,8 and 9). As an example, Hilts (6) in a widely publicized media report, described a study in which the use of methylphenidate in research subjects “threw 60 per cent of them into severe psychotic episodes.” Another exposé (8) described research at the University of California at Los Angeles (UCLA) involving outpatients with psychotic disorders who were withdrawn from active antipsychotic medications and observed for signs of relapse over time. One patient ultimately committed suicide more than a year after leaving the study, whereas a second had a significant exacerbation in psychotic symptoms, resulting in threats to kill his parents. In 1994 the federal Office of Protection from Research Risks (OPRR) investigated allegations that the UCLA study's research design and implementation had been unethical (10). Although the OPRR did not find unethical research practices, they questioned the adequacy of the informed consent process for this potentially high-risk study (11).

As a result of the controversy in the United States, federal and state agencies have begun to take a closer look at ethical issues raised in psychiatric research. One of the foci has been the process of informed consent in studies involving subjects who may have impairments in their abilities to make decisions, such as patients with severe mental illnesses. A 1995 report of the Advisory Committee on Human Radiation Experiments (ACHRE) found that approximately half of the studies it examined had “inadequate explanations of risks and discomforts in their consent material and paid no attention to the question of how to deal with subjects who might have impaired capacities to consent to research participation” (5).

In 1995, President Clinton appointed the National Bioethics Advisory Commission (NBAC), in part to address these concerns. In December of 1998, NBAC issued its report entitled *Research Involving Persons with Mental Disorders That May Affect Decision-making Capacity* (12). Among other things, the report recommended that an independent professional should assess a potential subject's capacity to provide informed consent for studies involving more than minimal risk. The report generated a swift and critical response from many psychiatric professionals who expressed concern that the recommendations reflected the misconception that all persons with mental illness have decision-making impairments. Thus, some considered the recommendations too restrictive and stigmatizing of persons with illnesses very much in need of study (13). Charney (14), however, wrote on behalf of the psychiatric research community that the NBAC report provided some valuable contributions to the ongoing debate, and acknowledged that “there is a crisis in confidence in the ethics of psychiatric research” that needed to be addressed. NBAC responded to the criticisms by stating that they envisioned their report as “part of a continuing societal conversation… about what regulations and guidelines should govern research involving persons with mental disorders that may affect their decision-making capacity” (15).

These developments highlighted numerous areas of ethical concern regarding research with human subjects with neuropsychiatric disorders, including subject recruitment, confidentiality, data access, and conflicting roles of investigators acting also as treaters. More recently, certain methodologic practices, such as placebo-controlled studies, drug withdrawal studies, and the so-called “challenge” studies, have attracted particular attention. Related to concerns about methodology, the ability of patients with mental illnesses to provide informed consent to research procedures has probably been one of the most controversial issues surrounding psychiatric clinical research, as highlighted in the NBAC report.

METHODOLOGIC CONTROVERSIES

Part of "35 - Ethical Aspects of Neuropsychiatric Research with Human Subjects "

Placebo Studies

In the mid-1990s, controversy over the use of placebos in research was rekindled; some commentators (16 ,17) contended that placebo use is unethical when standard effective treatments exist. Support for this limitation on the use of placebos stems, in part, from the Declaration of Helsinki (1), which declares that human research subjects have a right to therapeutically proven methods and treatments when available. Nevertheless, use of placebo agents is widespread throughout medical research (18 ,19).

Arguments in favor of the use of placebos in research, including psychiatric research, include the superior ability to assess accurately the efficacy of an experimental medication through the use of double-blind, randomized, placebo-controlled studies. In fact, this type of study design has been touted has “one of the major achievements of modern medicine” (20). The FDA, in considering what constitutes “adequate and well-controlled studies” (required for approval of new medications) states that the placebo-controlled experimental design has important scientific merit in establishing therapeutic efficacy as long as the objectives and the rationale for placebo use are clear (1). However, the FDA cautions that placebos should not be used “where existing treatment is life-prolonging” or if the placebo “exposes patients to a documented serious risk” (1).

Some authors have argued that findings of placebo research are misleading and deceptive (21), and that more or equally reliable findings could be had using active control agents (16 ,22). A central argument against the use of placebos in research on serious mental illnesses is that they are likely to contribute to a relapse of or failure to resolve psychiatric symptoms. One early report suggested that psychosis in and of itself may be biologically toxic to the brain (23), and may lead to short- and long-term adverse consequences. In addition to the potential risks associated with an exacerbation of primary symptoms, there are concerns about the psychosocial effects of relapse on patients. Specifically, prolonged periods of significant psychological distress may be associated with loss of interpersonal relationships, financial losses, and increased risk of suicide.

Despite the risks, by now it is relatively clear that the use of nonactive agents as a means of control has scientific merit when effective treatments for a particular illness are not yet known. When effective treatments do exist, a placebo comparison may still allow investigators to establish efficacy, learn more about the natural course of the illness, and compare side effect profiles of active agents against nonactive compounds. Studies examining the efficacy of a known medication and an experimental medication may—because of flaws in the studies’ design or implementation—not clearly differentiate between the two, and drugs with only minimal potential may be seen as more worthwhile than they are (24). Moreover, historic controls are not an adequate substitute for placebos because the apparent increase in the prevalence of experimental subjects who may already be resistant to treatment with standard medications (25 ,26). Thus, the analysis of the complexities of psychiatric illness and decisions regarding the risks and benefits of existing compounds compared with novel agents would be limited by only having the existing active comparison agent as a reference point (20 ,27). Furthermore, it has been argued that brief placebo periods may be conducted safely, particularly for inpatients (28), and do not appear to lead to long-term negative prognostic effects (29).

Drug Withdrawal Studies

Medication discontinuation studies in psychiatric research have become another point of ethical contention. The scientific rationale for drug discontinuation has included the desire to examine the pathophysiology and course of underlying illnesses when patients are in an unmedicated state. Furthermore, assessment of the clinical and neurochemical effects of medications in some cases can be more legitimately interpreted in a given individual after a period of drug washout, as the potential therapeutic or adverse effects of the initial treatment may present a confounding variable, making interpretation difficult (20 ,30).

However, significant concern about the use of this approach has been raised both in the scientific community and in the lay press. For example, recent literature suggests that chronic patients may have a poorer response to treatment or deleterious effects should they be taken off medication and experience relapses (31 ,32 and 33). In patients with bipolar disorder, concern has been raised that the clinical state following withdrawal of maintenance medication may actually be distinct from what it would have been had the natural course of the illness progressed without treatment at all (34). Furthermore, the risk of relapse itself has been of significant concern. In fact, a meta-analysis of the effects of drug discontinuation in schizophrenia demonstrated a relapse rate of 53% during an average 9.7-month follow-up (35) compared to a 16% relapse rate for those patients remaining on their medication. Despite the greater relapse risk, patients who experienced a worsening of their symptoms when off medications were able to return to baseline following reinitiation of treatment. Relapse risk may be particularly high when medications are discontinued abruptly (36 ,37). Questions have been raised about whether inconsistent use of neuroleptics may result in a higher risk of tardive dyskinesia (38 ,39).

Much as with placebos, the debate about the potential for neurotoxic damage as a result of experiencing psychosis itself (23 ,40) has raised further ethical questions about drug discontinuation studies. Although, the theoretic long- and short-term risks of psychosis have been widely cited, others have argued that the data on the risks of brief psychosis occurring during research studies are not clear (41 ,42). Furthermore, in the case of psychotic disorders, continuous treatment with neuroleptics is not without its own risks, including some risk of relapse and the risk of serious side effects (38). About 30% of patients will have no significant response to neuroleptics, and some patients can remain without relapse even after years of being off medications (43). The risks of ongoing treatment and potential adverse sequelae of withdrawing medications must be weighed in all psychiatric research. In this way, risks to human subjects may be minimized and drug withdrawal conducted when essential.

Challenge Studies

Another of the controversial research techniques that has undergone public scrutiny are provocation or “challenge” studies. These terms refer to experiments in which patients and sometimes healthy control subjects are exposed to drugs that exacerbate or create psychiatric symptoms. Provocation studies are not unique to psychiatry. In general clinical research, provocation studies have been conducted to induce pain, nausea and/or vomiting, bronchoconstriction, tachycardia, cognitive impairment, and even sepsis (44). These studies share the same basic goal of allowing investigators to learn more about symptom expression and potential therapeutic interventions. Although widely used in medical research, their use in studies examining psychiatric illnesses seems to have captured the interest of lay persons, advocacy groups, the media, and even policy makers. One theory is that these types of studies may be more common in neurobiological research, where less is known about the diseases being studied and animal models are sparse (45).

One of the hottest debates most recently has involved the use of ketamine, an NMDA receptor antagonist, and an approved anesthetic agent, to provoke psychotic symptom exacerbation in patients with schizophrenia and produce transient psychotic states in well control subjects. Tishler and Gordon (46) expressed concern that giving a healthy control or nonpsychotic person ketamine might present a risk of producing illness, given the “biological stressor of the experimental procedure and psychological stressor of psychosis [induced by the pharmacologic challenge].” In a review of all North American schizophrenia subjects who underwent ketamine challenge studies, Carpenter (47) concluded that the ketamine-induced increase in psychosis was mild to moderate and brief, any anxiety induced was mild and brief, and there was no evidence of ongoing negative consequences for subjects. It is noteworthy, as Carpenter points out, that the controversy surrounding the ketamine challenge study has been raised when results with fewer than 50 patients have been published. The media outcry against this type of study has led to trepidation to continue this novel avenue of scientific research. Yet, other authors have suggested that symptom provocation studies, beginning with early research involving amphetamine loading and including the more recent symptom induction studies, have contributed significantly to our understanding of psychiatric disease, at a cost of inducing only transient psychotic states with no long-term adverse effects (48 ,49) or evidence of altered disease course (50).

Although the data suggesting the safety of current challenge studies are encouraging, ethical implementation of such studies is complex because of the potential for negative consequences, even if transient or remote. It has been argued that these types of studies might be ethically justifiable if the underlying scientific principle is sound, if the effects are

not thought to be long-term or severe, and if subjects have the capacity to participate as “knowing, voluntary partners in the research enterprise” (49).

INFORMED CONSENT AND THE CAPACITY FOR CONSENT IN PERSONS WITH NEUROPSYCHIATRIC DISORDERS

Part of "35 - Ethical Aspects of Neuropsychiatric Research with Human Subjects "

Challenge, placebo, and drug withdrawal studies always raise questions of research ethics, but the controversy is heightened when the subjects involved suffer from mental illnesses. At the heart of the debate is the concern that these subjects, more than other human research subjects, have significant deficits in their abilities to provide informed consent, so that they may enter studies without full understanding of the inherent risks. Unfortunately, although this has become the focus of political and media attention, there is often a lack of understanding of what informed consent is, and what the literature shows regarding the capacity of mentally ill subjects to give informed consent.

The doctrine of informed consent is built from a complex interrelationship of medicolegal and ethical principles. Generally, informed consent, whether to research or treatment, is broken down into three parts: *voluntariness, disclosure*, and *competence* (51). Voluntariness implies that research subjects must be acting of their own free will when they agree to participate in research. Disclosure provides information on the basis of which potential subjects may make an informed choice. In research settings, disclosures must generally include such details as the nature and purpose of the study, as well as the potential risks and benefits involved in the study. Other information provided includes disclosure of the right to discontinue participation in the study, who will have access to the data, the differences between participation in research and routine treatment, and the availability of compensation should harm ensue as a result of the study.

Informed consent also requires task-specific competence. Competence consists of four separate elements (52). First, subjects must be able to *evidence a choice* regarding the decision at hand. The choice need not be expressed verbally, but subjects must be able to communicate their preference in some way. Also, the choice must be sustained over time. The inability to maintain a consistent choice over time might reflect significant mental status deficits such as those seen in psychotic ambivalence or delirious states.

Additionally, subjects must have a *factual understanding of the information* that has been presented to them. The degree of factual understanding required for competence is unclear, and there is no threshold value of how much information must be understood in order to be considered to have “enough” factual understanding. Furthermore, acceptable levels of understanding may vary depending on the risks involved in a proposed research study.

Subjects must also be able to *rationally manipulate the information* in a way that is not impaired by symptoms of their illness. They must demonstrate ability to reason through the information presented to come up with a logical decision, which need not be the decision that the person assessing competence would make. Patients with or without psychosis who have impairments in their reasoning, in addition to their primary symptoms (e.g., cognitive deficits, concrete thinking, or inability to abstract), might have difficulties in this regard.

Finally, subjects must have a *realistic appreciation of their situation*. Patients with schizophrenia, for example, who do not believe they are ill will have a limited appreciation of why they are being enrolled in a study examining that particular illness. The appreciation must include some awareness of the fact that the study involves research and not treatment, and so may be of no direct benefit to the individual.

Understanding of the capacities of persons with mental illnesses to consent to research has historically relied on data gathered from studies looking at competence to consent to treatment. Recent years have seen an expansion of the previously limited literature on competence to consent in research settings.

With regard to the ability to communicate a choice, although sometimes taken for granted, studies have shown that a proportion of patients will have difficulties in this area. In a study by Appelbaum, Mirkin, and Bateman, 9% of community mental health center patients who were contacted to participate in a study were found to be mute or catatonic (53). Eighteen percent of primarily depressed inpatients were unable to make a decision in vignettes that required some problem solving (54). The risk of simply excluding these persons from studies is that their inability to communicate a choice may reflect a degree of illness that is worthy of study, and their exclusion might skew the results of research based on altered group composition. Therefore, there may be value in considering whether proxy decision makers might in certain circumstances and with appropriate safeguards, enter into research those subjects who are unable to express a consistent choice.

Studies have also examined the capacity of patients to understand information. For example, Grossman and Summers (55) found that patients with schizophrenia understood only about half the information presented to them regarding the risks and benefits of a fictitious medication, and thus concluded that these patients may have difficulty providing true informed consent. The degree of psychopathology may affect learning of new information in schizophrenics (56). Kleinman and colleagues (57) suggested that a formalized informing process increased schizophrenic patient understanding of tardive dyskinesia. In a frequently cited study of 41 patients with affective disorders who were potential subjects of a sleep EEG study, Roth and colleagues (58) found that only about 50% of the subjects understood more than two-thirds of the information presented to them through a formal consent process, whereas a significant minority

of patients (about 25%) understood half or less of the information. Benson and associates (59) showed that patients with schizophrenia demonstrated greater impairment in understanding specific psychiatric research purposes and methodology in comparison to psychiatric patients with less severe psychopathology. Comparing the capacity of stable patients with schizophrenia and healthy volunteers to understand a low-risk study involving a magnetic resonance imaging test for research purposes, Pinals and co-workers (60) found no difference in understanding of consent forms between groups. Of note, neither group on average was able to correctly answer 100% of the questions on a brief questionnaire related to information on the consent form. Another study using a questionnaire relating to research protocols found that out of 49 patients with schizophrenia, 53% required a second trial at the questionnaire after re-education about the protocols to achieve a score of 100%, and 37% of subjects required three or more trials (61). The authors concluded that with an adequate informed consent process, research subjects with schizophrenia were able to comprehend consent form information.

Impairment of the ability to appreciate the nature of one's situation and potential consequences may have particular relevance in psychiatric disorders where insight into one's illness is often compromised. In a classic report, Soskis (62) found that 68% of schizophrenic subjects did not recognize the reason they were receiving treatment compared to 13% of medically ill patients. In an earlier study looking at patient appreciation of their participation in research, Appelbaum and associates (63) showed that more than half of the psychiatric patients interviewed failed to comprehend the research nature of some component of the methodology of the research in which they were participating. The authors called the subjects' tendency to view research as a therapeutic process, when in fact there may be no benefit to the subject at all, the "therapeutic misconception."

With regard to the ability of psychiatric patients to rationally manipulate information pertaining to research, Stanley and colleagues (64) reported that the degree of psychopathology in patients with mental illness did not appear to influence their willingness to participate in hypothetical research compared to nonpsychiatrically ill subjects. In that study, patients tended to agree to low-risk/high-benefit hypothetical studies more than high-risk/low-benefit studies. In a subsequent study, Stanley and associates (65) found that approximately one-third of patients with mixed psychiatric diagnoses refused low-risk/high-benefit hypothetical study enrollment, whereas about 40% of patients agreed to participate in a hypothetical study of high risk/low benefit. Garety and associates (66) found that subjects with schizophrenia or delusional disorder requested less information before reaching a decision and were quicker to change their estimates of the likelihood of an adverse event compared to nondelusional psychiatric patients and normal controls. In a study by Sachs and co-workers (67), persons with dementia were noted not to perform as well as nondemented elderly subjects in providing logical reasons for their decisions to participate in hypothetical research protocols.

Probably the most extensive data examining competency to make treatment decisions was reported by Grisso and Appelbaum (68 ,69) from the MacArthur Treatment Competence Study. This study utilized standardized instruments designed to assess capacities to make treatment decisions, and involved the assessment of multiple components of competence (understanding, appreciation, and reasoning) and the use of several subject groups. Deficits were most pronounced in patients with schizophrenia, and slightly more patients with depression were likely to have deficits than controls. Because the majority of all subjects performed well on measures of competence, the study underscored the notion that subjects cannot be presumed incompetent by virtue of mental illness alone.

Carpenter and associates (70) recently reported their findings examining how psychopathology and cognition affect decisional capacity. They used a modification of the MacArthur study instruments (MacCAT-CR: MacArthur Competency Assessment Tool-Clinical Research) (71) to examine making decision abilities relevant to research. In this study, 30 research subjects with schizophrenia did not perform as well as healthy controls in decision making, and performance was strongly related to cognitive impairments and somewhat related to symptomatology. However, the study found that a weeklong educational intervention that provided information regarding the hypothetical study led to improved decisional capacity such that scores of schizophrenic subjects were not significantly different from the well control group. In another recently published study, Appelbaum and associates (72) assessed the capacities of depressed patients to consent to research utilizing the MacCAT-CR. In this study, female outpatients with major depression did not show impairments in their decision-making capacities related to research. This study further demonstrated the utility of instruments such as the MacCAT-CR as a means of assessing decisional capacity as part of the broader informed consent process in an actual research study.

COMMENTARY

Part of "35 - Ethical Aspects of Neuropsychiatric Research with Human Subjects "

Although ethics in human subject research has long been the focus of attention, awareness of the ethical dilemmas has been heightened in recent years. Despite calls for a moratorium on all nontherapeutic, "high-risk" experiments, including drug washout and challenge studies (73), adverse events appear to be much less common than the public may have been led to believe. What can be gleaned from the current debate is that researchers must attend to the concerns raised, both to maintain public trust and ensure the

ethical integrity of research itself. As Bonnie (74) noted, the challenge is to create generally accepted guidelines on safeguards for subjects without compromising the pursuit of important knowledge or threatening the integral partnership of mental health advocates, persons with mental illness, and researchers.

The research community has made several efforts to tackle these issues. The American College of Neuropsychopharmacology (ACNP) has developed guidelines on ethical practices related to neuropsychopharmacologic research. Highlighted are the needs to: (a) ensure appropriateness of the study and its design; (b) minimize risk to subjects and maximize benefit to subjects or to the population of patients with the illnesses under study; (c) ensure informed consent, while paying particular attention to the needs of those subjects who may have decision-making impairments; and (d) protect confidentiality (75). The NIMH has established new rules for "high-risk" studies, including the creation of a special Human Subject Research Workgroup of the National Advisory Mental Health Council (NAMHC), which will review study protocols involving challenge methodology or drug withdrawal studies (76). After several meetings with representatives of the NIMH, in 1995 the National Alliance for the Mentally Ill (NAMI) adopted "Policies on Strengthened Standards for Protection of Individual with Severe Mental Illnesses who Participate in Human Subjects Research" (77). Among these policies are a recognition of the "critical necessity" of human subject research, and recommendations for protection of persons with cognitive impairments, clearer standards for consent protocols, and specialized training for members of IRBs that review studies involving neuropsychiatric disorders. Measures to ensure that ethical issues are addressed have been developed, including the Research Protocol Ethics Assessment Tool (RePEAT), which may assist in the planning of experimental protocols (78).

In addition to these efforts, there is a growing consensus on mechanisms for the ethical conduct of human subject research. It has been suggested that not pursuing placebo and drug withdrawal studies would be unethical, given all there is to learn from them regarding the pathophysiology, natural course, and treatment of severe mental illness (79). That said, it also is clear that specific approaches can be utilized in order to ensure that this research is conducted safely and ethically. For example, there may be some studies of pathophysiology in which subjects may be maintained on a low but effective neuroleptic dose, without interfering with the acquisition of valid data (38). For neuroleptic withdrawal in patients with schizophrenia, slow rather than abrupt tapering with careful ongoing monitoring may mitigate potential for bad outcomes (40). Drug-free phases may best be conducted while subjects are in an inpatient setting, or while they are very closely monitored as outpatients (38). In this way, if symptoms begin to re-emerge, subjects may be quickly and effectively treated before a bad outcome ensues. "Exit criteria" should be established *a priori* to determine when patients will be restarted on their medications (41). In addition, alternative treatments (such as adjuvant medications, psychotherapy, and rehabilitation treatments) during the placebo phase may be beneficial without compromising research design. Patients at known risk of catastrophic responses to relapse should be excluded from the subject pool. Finally, study subjects must be given the opportunity to provide informed consent, pose questions, and withdraw from study participation at any time. Those patients with initial decision-making deficits and those who may become decisionally impaired during the study will require special measures of protection, addressed in the following.

When challenge studies are proposed, again the scientific merit of the protocol must be weighed against its potential risks. Several suggestions have been made that may offer protections to subjects. For example, Tishler and Gordon (46) have suggested a careful recruitment process that would include detailed disclosure of the inherent risks, review of compensation for participation, and screening prospective normal subjects for the presence of or vulnerability to develop psychiatric illness. Miller and Rosenstein (49) indicated that: (a) the study should have clear scientific merit, (b) subjects with specific clinical vulnerabilities may need to be excluded from participation, (c) selected methods should minimize risks, (d) subjects should have access to careful monitoring and follow-up, and (e) informed consent disclosure should make clear that the challenge study is distinct from other studies in which the subject may be enrolled (80).

With regard to the consent process and the potential for decision-making impairments of mentally ill research subjects, existing literature provides only rudimentary guidance in identifying groups at high risk of impairment. A substantial number of persons with severe neuropsychiatric illnesses may have impairments in their decision-making ability related to research consent. Yet, the data also have shown that many of these persons will retain abilities to make decisions that affect their lives, and thus it is misleading to presume them incompetent by virtue of their diagnoses without adequate assessment.

Although policies will need to offer protection for those who may have decision-making impairments, excessive burdens must be avoided if advancement of knowledge is to continue. By thwarting attempts to conduct bold and novel studies, society runs the risk of limiting knowledge of the very populations who may be most in need of such research. The many subjects who have participated in neurobiological research willingly, even when the risk is high and the potential for benefit is low, testify to the desperation that some of these patients may feel regarding their illnesses. Brody has similarly commented on the justification for use of mentally infirm adults in nontherapeutic research, even if the research presents greater than minimal risk (1), because of the need

to study these complex illnesses. With these caveats, areas worthy of further consideration include disclosure practices, identification and assessment of subject competence, and questions of threshold levels of competence (81 ,82).

In all of these arenas, existing IRBs seem to be in a strong position to provide the scrutiny required. Unfortunately, many people believe that IRBs have become little more than clearinghouses for consent forms, rather than committees designed for careful review of all aspects of research ethics (83). In an attempt to deal with this concern, the NBAC report proposed the establishment of a special standing panel to review certain protocols that way present a greater risk to subjects (12). There are, of course, negative aspects of a shift from currently accepted local IRB authority to a federal agency far removed from where the study would take place (84). Regardless of the reviewing body, if the methodology appears questionable, persons with specialized knowledge in these areas should be consulted to address the questions raised. Attention to the minimization of potential risks of studies is also an important part of the mission of an IRB. With regard to the consent process, the IRB, in addition to reviewing consent forms, should be able to monitor investigator disclosure and determine the level of required subject competence based on a standardized evaluation of the risks and potential benefits involved in a proposed study.

Investigators may have other ways of advancing our current approaches to consent to research. For example, current literature has demonstrated that a modification of disclosure procedures may facilitate subject understanding and enhanced learning (57 ,59 ,60 ,70 ,85 ,86 ,87 ,88 and 89). Even with such efforts, however, there will always be potential subjects who will lack capacity, in one or more of its realms, to provide valid informed consent to participate in research. When patients are participating in studies of greater risk, a higher standard of competence should be required. The investigator and the IRB could work together to decide when formal capacity assessments are indicated (90).

After the inherent risks and competence needs are determined, a sliding scale of options regarding capacity assessment might be implemented. For example, in a low-risk study, one might consider a straightforward consent form and clinical assessment of competence, perhaps aided by a questionnaire specifically geared to the study at hand. As the stakes increase, formalized assessment instruments, such as the MacCAT-CR, might be adapted to the study in question. Oldham and colleagues (13), in their response to the NBAC report, suggested that “formal capacity assessments should be required for subjects when there is reason to believe that a mental or emotional state or a primary or secondary brain dysfunction may interfere with decision making.” They also suggest that, given the inherent potential for investigator bias, for research that presents more than a minor increase over minimal risk, independent evaluators, who function separately from the research team, could ascertain subject capacities. It is unclear at this time how feasible such an approach would be, but it may merit exploration.

Once subjects are identified who clearly lack capacity to consent or who may come to lack capacity as research progresses (such as patients with Alzheimer's disease or patients with schizophrenia enrolled in placebo-controlled studies), additional protections might be implemented to allow such persons to participate in research. Pursuit of a legal determination of incompetence and the appointment of a guardian to make decisions for the subject appears to be utilized rarely, in part because of the impracticalities and cost involved (51). The use of a durable power of attorney or advance directive might, however, allow a substitute decision maker to make decisions that the patient would have made during periods of greater competence (91 ,92 ,93 and 94).

Human subject research will always require careful scrutiny. Our history has shown that even well intentioned investigators may not be able to assess ethical aspects of the research they are undertaking objectively. Additionally, potential research subjects may enroll in studies for a variety of reasons, conscious and unconscious, without a full awareness or appreciation of the risks they are undertaking. Nevertheless, the current focus on ethical issues related to research should serve to heighten the awareness of the research team, including both investigators and subjects, regarding measures that can be taken to allow scientific advancement while protecting potentially vulnerable populations.

ACKNOWLEDGMENT

Part of "35 - Ethical Aspects of Neuropsychiatric Research with Human Subjects "

Dr. Pinals has served on a speaker's bureau for Janssen Pharmaceutica.

REFERENCES

1. Brody BA. *The Ethics of biomedical research: an international perspective*. Oxford University Press, New York, 1998.

2. Frankel MS. The development of policy guidelines governing human experimentation in the United States: a case study of public policy-making for science and technology. *Ethics Sci Med* 1975;2:43–59.

3. Howard-Jones N. Human experimentation in historical and ethical perspectives. *Soc Sci Med* 1982;16:1429–1448.

4. Beecher HK. Ethics and clinical research. *N Engl J Med* 1966;274:1354–1360.

5. Advisory Committee on Human Radiation Experiments. *Advisory Committee on Human Radiation Experiments, Final Report*. Washington, DC: US Government Printing Office, 1995.

6. Hilts PJ. Psychiatric researchers under fire. *The New York Times* 1998, May 19.

7. Davis R. Holes are growing in medical testing's safety nets. *USA Today* 1998, June 8.

8. Willwerth J. Madness in fine print. *Time* 1994, November 7.

9. Ethics in neurobiological research with human subjects. *The Journal of the California Alliance for the Mentally Ill*, 1994;5:1–69.

10. Appelbaum PS. Drug-free research in schizophrenia: an overview of the controversy. *IRB* 1996;18:1–5.

11. Office of Protection from Research Risks. *Evaluation of human subject protections in schizophrenia research conducted by the University of California, Los Angeles*, 1994, May 11.

12. National Bioethics Advisory Commission. *Research involving persons with mental disorders that may affect decision-making capacity: Vol 1. Report and recommendations of the National Bioethics Advisory Commission, Rockville, MD*. Rockville, MD: National Bioethics Advisory Commission, 1998.

13. Oldham JM, Haimowitz S, Delano SJ. Protection of persons with mental disorders from research risk: a response to the report of the National Bioethics Advisory Commission. *Arch Gen Psychiatry* 1999;56:688–693.

14. Charney DS. The National Bioethics Advisory Commission report: the response of the psychiatric research community is critical to restoring public trust. *Arch Gen Psychiatry* 1999;56:699–700.

15. Childress JF, Shapiro HT. Almost persuaded: reactions to Oldham et al. *Arch Gen Psychiatry* 1999;56:697–698.

16. Rothman KJ, Michels KB. The continuing unethical use of placebo controls. *N Engl J Med* 1994;331:394–398.

17. Cagliano S, Traversa G. The use of placebo controls. (Letter) *N Engl J Med* 1995;332:60.

18. Tugwell, P, Bombardier C, Gent M, et al. Low-dose cyclosporin versus placebo in patients with rheumatoid arthritis. *Lancet* 1990;335:1051–1055.

19. Packer M, Narahara KA, Elkayan U, et al. Double-blind, placebo-controlled study of the efficacy of flosequinan in patients with chronic heart failure. *J Am Coll Cardiol* 1993;22:65–72.

20. Kane JM, Borenstein M. The use of placebos in psychiatric research. In: Shamoo AE, ed. *Ethics in neurobiological research with human subjects: the Baltimore Conference on Ethics*. Gordon and Breach Publishers, 1997:207.

21. Bok S. The ethics of giving placebos. *Sci Am* 1974;231:17–23.

22. Weijer C. Placebo-controlled trials in schizophrenia. Are they ethical? Are they necessary? *Schiz Res* 1999;35:211–218.

23. Wyatt RJ. Neuroleptics and the natural course of schizophrenia. *Schiz Bull* 1991;17:325–351.

24. Young SN, Annable L, Stat D. The Use of placebos in psychiatry: a response to the draft document prepared by the tri-council working group. *J Psychiatry Neurosci* 1996;21:235–238.

25. Addington D. The use of placebos in clinical trial for acute schizophrenia. *Can J Psychiatry* 1995;40:171–176.

26. Quitkin FM. Placebos, drug effects, and study design: a clinician's guide. *Am J Psychiatry* 1999;156:829–836.

27. Streiner DL. Placebo-controlled trials: when are they needed? *Schiz Res* 1999;35:201–210.

28. Kumra S, Briguglio C, Lenane M, et al. Including children and adolescents with schizophrenia in medication-free research. *Am J Psychiatry* 1999;56:1065–1068.

29. Curson DA, Hirsch SR, Platt SD, et al. Does short-term placebo treatment of chronic schizophrenia produce long-term harm? *Br Med J* 1986;293:726–728.

30. Carpenter WT, Tamminga CA. Why neuroleptic withdrawal in schizophrenia? *Arch Gen Psychiatry* 1995;52:192–193.

31. Post RM, Leverich GS, Altshuler L, et al. Lithium-discontinuation-induced refractoriness: preliminary observations. *Am J Psychiatry* 1992;49:1727–1729.

32. Johnson DAW, Pasterski G, Ludlow JM, et al. The discontinuance of maintenance neuroleptic therapy in chronic schizophrenia patients: drug and social consequences. *Acta Psychiatr Scand* 1983;67:339–352.

33. Szymanski S, Lieberman JA, Alvir JM, et al. Gender differences in onset of illness, treatment response, course, and biologic indexes in first-episode schizophrenic patients. *Am J Psychiatry* 1995;152:698–703.

34. Suppes T, Baldessarini RJ, Faedda GL, et al. Discontinuing maintenance treatment in bipolar manic-depression: risks and implications. *Harvard Rev Psychiatry* 1993;1:131–144.

35. Gilbert PL, Harris MJ, McAdams A, et al. Neuroleptic withdrawal in schizophrenia: a review of the literature. *Arch Gen Psychiatry* 1995;52:173–188.

36. Baldessarini RJ, Viguera AC. Neuroleptic withdrawal in schizophrenic patients. *Arch Gen Psychiatry* 1995;52:189–192.

37. Faedda GL, Tondo L, Baldessarini RJ, et al. Outcome after rapid vs gradual discontinuation of lithium treatment in bipolar mood disorders. *Arch Gen Psychiatry* 1993;50:448–455.

38. Jeste DV, Palmer BW, Harris MJ. Neuroleptic discontinuation in clinical and research settings: scientific issues and ethical dilemmas. *Biol Psychiatry* 1999;46:1050–1059.

39. Goldman MB, Luchins DJ. Intermittent neuroleptic therapy and tardive dyskinesia: a literature review. *Hosp Commun Psychiatry* 1984;35:1215–1219.

40. Wyatt RJ. Risks of withdrawing antipsychotic medications. *Arch Gen Psychiatry* 1995; 52:205–208.

41. Carpenter WT. The risk of medication-free research. *Schiz Bull* 1997;23:11–18.

42. Carpenter WT. Schizophrenia research: a challenge for constructive criticism. In: Shamoo AE, ed. *Ethics in neurobiological research with human subjects: The Baltimore Conference on Ethics*. Gordon and Breach Publishers, 1997.

43. Lieberman JA, Kane JM, Sarantakos S, et al. Ethics of drug discontinuation studies in schizophrenia: in reply. *Arch Gen Psychiatry* 1989;46:387.

44. Heath G, Pierce C. Inducing “disease” symptoms in healthy volunteers. *Appl Clin Trials* 1999; April:42–48.

45. Hall LL. Medication discontinuation and symptom provocation in research: a consumer and family perspective. *Biol Psychiatry* 1999;46:1017–1020.

46. Tishler CL, Gordon LB. Ethical parameters of challenge studies inducing psychosis with ketamine. *Ethics Behav* 1999;9:211–217.

47. Carpenter WT. The schizophrenia ketamine challenge study debate. *Biol Psychiatry* 1999;46:1081–1091.

48. D'souza DC, Berman RM, Krystal JH, et al. Symptom provocation studies in psychiatric disorders: scientific value, risks, and future. *Biol Psychiatry* 1999;46:1060–1080.

49. Miller FG, Rosenstein DL. Psychiatric symptom-provoking studies: an ethical appraisal. *Biol Psychiatry* 1997;42:403–409.

50. Warfel D, Lahti AC, Michaelidis T, et al. Follow-up patients who receive challenge with ketamine as part of a research study. [Abstract] *Schiz Res* 1999;36:317–318.

51. Appelbaum PS. Patients’ competence to consent to neurobiological research. *Account Res* 1996;4:241–225.

52. Appelbaum PS, Roth LH. Competency to consent to research. A psychiatric overview. *Arch Gen Psychiatry* 1982;39:951–958.

53. Appelbaum PS, Mirkin SA, Bateman AL. Empirical assessment of competency to consent to psychiatric hospitalization. *Am J Psychiatry* 1981;138:1170–1176.

54. Radford MH, Mann L, Kalucy RS. Psychiatric disturbance and decision-making. *Aust NZ J Psychiatry* 1986;20:210–217.

55. Grossman L, Summers F. A study of the capacity of schizophrenic patients to give informed consent. *Hosp Commun Psychiatry* 1980;31:205–206.

56. Schacter D, Kleinman I, et al. The Effect of psychopathology on the ability of schizophrenic patients to give informed consent. *J Nerv Ment Dis* 1994;182:360–362.

57. Kleinman I, Schacter D, Koritar E. Informed consent and tardive dyskinesia. *Am J Psychiatry* 1989;146:902–904.

58. Roth LH, Lidz CW, Meisel A, et al. Competency to decide about treatment or research: an overview of some empirical data. *Int J Law Psychiatry* 1982;5:29–50.

59. Benson PR, Roth LH, Appelbaum PS, et al. Information disclosure, subject understanding, and informed consent in psychiatric research. *Law Hum Behav* 1988;12:455–476.

60. Pinals DA, Malhotra AK, Breier A, et al. Informed consent in schizophrenia research. *Psychiatric Serv* 1998;49:244.

61. Wirshing DA, Wirshing WC, Marder SR, et al. Informed consent: assessment of comprehension. *Am J Psychiatry* 1998;155:1508–1511.

62. Soskis DA. Schizophrenic and medical inpatients as informed drug consumers. *Arch Gen Psychiatry* 1978;35:645–647.

63. Appelbaum PS, Roth LH, Lidz CW, et al. False hopes and best data: consent to research and the therapeutic misconception. *Hastings Center Report* 1987;17:20–24.

64. Stanley B, Stanley M, Latin A, et al. Preliminary findings on psychiatric patients as research participants: a population at risk? *Am J Psychiatry* 1981;138:669–671.

65. Stanley B, Stanley M, Peselow E, et al. The effects of psychotropic drugs on informed consent. *Psychopharmacol Bull* 1982;18:102–104.

66. Garety PA, Hemsley DR, Wessely S. Reasoning in deluded schizophrenic and paranoid patients: biases in performance on a probabilistic inference task. *J Nerv Mental Dis* 1991;179:194–201.

67. Sachs GA, Stocking CB, Stern R, et al. Ethical aspects of dementia research: informed consent and proxy consent. *Clin Res* 1994;42:403–412.

68. Grisso T, Appelbaum PS. A comparison of standards for assessing patients' capacities to make treatment decisions. *Am J Psychiatry* 1995;152:1033–1037.

69. Grisso T, Appelbaum PS. The MacArthur treatment competence study. III. Abilities of patients to consent to psychiatric and medical treatments. *Law Hum Behav* 1995;19:149–174.

70. Carpenter WT, Gold JM, Lahti AC, et al. Decisional capacity for informed consent in schizophrenia research. *Arch Gen Psychiatry* 2000;57:533–538.

71. Appelbaum PS, Grisso T. The MacArthur Competence Assessment Tool. Clinical Research, 1996.

72. Appelbaum PS, Grisso T, Frank E, et al. Competence of depressed patients for consent to research. *Am J Psychiatry* 1999;156:1380–1384.

73. Shamoo AE. The Unethical use of persons with mental illness in high risk research experiments: testimony to Human Subjects Subcommittee of the National Bioethics Advisory Commission United States Government, July 15, 1997.

74. Bonnie R. Research with cognitively impaired subjects: unfinished business in the regulation of human research. *Arch Gen Psychiatry* 1997;54:105–111.

75. Kupfer DJ. A statement of principles of ethical conduct for neuropsychopharmacologic research in human subjects. Draft Document of ethical guidelines for neuropsychopharmacologic research with human subjects performed by members of the American College of Neuropsychopharmacology (ACNP). April 26, 1999.

76. New NIMH rules for high-risk studies. *Psychiatric News* April 16, 1999, p 12.

77. Shamoo AE, Johnson JR, Honbeg R, et al. NAMI's standards for protection of individuals with severe mental illnesses who participate as human subjects in research. In: Shamoo AE, ed. *Ethics in neurobiological research with human subjects: the Baltimore Conference on Ethics*. The Netherlands: Gordon and Breach, 1997.

78. Roberts LW. Ethical dimensions of psychiatric research: A constructive, criterion-based approach to protocol preparation. The Research Protocol Ethics Assessment Tool (RePEAT). *Biol Psychiatry* 1999;46:1106–1119.

79. Neylan TC, Wright BA, Shelton MD, et al. More on ethics of drug discontinuation studies in schizophrenia. *Arch Gen Psychiatry* 1990;47:192.

80. Rosenstein DL. IRB Review of psychiatric medication discontinuation and symptom-provoking studies. *Biol Psychiatry* 1999;46:1039–1043.

81. Berg JW, Appelbaum PS. Subjects' capacity to consent to neurobiological research. In: Pincus HA, Lieberman JA, Ferris S, eds. *Ethics in psychiatric research: a resource manual for human subjects protection*. Washington, DC: American Psychiatric Association, 1999.

82. Appelbaum PS. Missing the boat: competence and consent in psychiatric research. *Am J Psychiatry* 1998;155:1486–1488.

83. Appelbaum PS. Rethinking the conduct of psychiatric research. *Arch Gen Psychiatry* 1997;54:117–120.

84. Michels R. Are research ethics bad for our mental health? *N Engl J Med* 1999;340:1427–1430.

85. Munetz MR, Roth LH, Cornes CL. Tardive dyskinesia and informed consent: myths and realities. *Bull Am Acad Psychiatry Law* 1982;10:77–88.

86. Silva MC, Sorrell JM. Enhancing comprehension of information for informed consent: a review of empirical research. *IRB: Rev Human Subjects Res* 1988;10:1–5.

87. Tymchuk AJ, Ouslander JG, Rahbar B, et al. Medical decision-making among elderly people in long-term care. *Gerontologist* 1988;28:59–63.

88. Miller R, Willner HS. The two-part consent form: a suggestion for promoting free and informed consent. *N Engl J Med* 1974;290:964–965.

89. Wager E, Tooley PJH, Emanuel MB, et al. How to do it: get patients' consent to enter clinical trials. *Br Med J* 1995;311:734–737.

90. Dresser R. Mentally disabled research subjects: the enduring policy issues. *JAMA* 1996;276:67–72.

91. Sunderland T, Dukoff R. Surrogate decision-making and advanced directives with cognitively impaired research subjects. In: Pincus HA, Lieberman JA, Ferris S (Des). *Ethics in psychiatric research: a resource manual for human subjects protection*. Washington, DC: American Psychiatric Association, 1999.

92. Sachs GA. Advance consent for dementia research. *Alzheimer Dis Assoc Dis* 1994;4:19–27.

93. DeRenzo E. Surrogate decision-making for severely cognitively impaired research subjects: the continuing debate. *Cambridge Quart Health Ethics* 1994;3:539–548.

94. Kapp MB. Proxy decision-making in Alzheimer disease research: durable powers of attorney, guardianship, and other alternatives. *Alzheimer Dis Assoc Dis* 1994;4:28–37.

36

Regulatory Issues

Paul Leber

Paul Leber: Neuro-Pharm Group, LLC, Potomac, Maryland.

For over 60 years, the United States has relied primarily on a federal system of premarket drug product clearance to ensure the quality of the nation's drug supply. When the premarket clearance system was first introduced in 1938 in the aftermath of the Elixir of Sulfanilamide tragedy in which over a hundred patients needlessly died because of a drug manufacturer's carelessness (1), federal law required only that new drugs be tested and shown, prior to marketing, to be "safe for use." Since 1962, however, the law requires that new drug products also be shown to be "effective in use" under the conditions of use recommended in their proposed labeling.

This chapter considers how these two fundamental requirements of the Federal Food, Drug and Cosmetic Act (FFDCA), our national drug regulatory law, are currently interpreted and applied by the Food and Drug Administration (FDA) in its evaluation of new drug products. Issues that are singularly important to the evaluation of products intended for use in the management of psychiatric conditions are identified and explicated.

THE US FEDERAL DRUG REGULATORY SYSTEM: LEGAL BASIS, STRUCTURE, AND MODE OF OPERATION

Part of "36 - Regulatory Issues "

The basic structure and operation of the US federal drug regulatory system are established under the provisions of the FFDCA.

The FFDCA makes it unlawful "…to introduce or deliver for introduction into interstate commerce any new drug, unless an approval of an application… is effective with respect to such drug."

The application to which the Act refers is a New Drug Application (NDA). By law, authority to approve NDAs resides with the Secretary of the Department of Health and Human Services, but the Secretary delegates the actual authority to review and approve NDAs to the Food and Drug Administration (FDA, the agency).

The Act instructs anyone (i.e., a sponsor) seeking to market a new drug product to submit and gain FDA's approval of an NDA for the product prior to marketing it. Importantly, although it is not widely appreciated, NDAs are not approved for drug substances (i.e., chemical entities), *per se*, but for one or more specific "claimed" uses of a specific drug product (i.e., a specific formulation of the drug substance) under a specific set of conditions of use recommended (i.e., described) in the product's proposed labeling.

The Act describes, albeit in rather general terms, the information and reports that each NDA must contain. The details need not concern us, however. Suffice it to say that a sponsor's NDA is required to provide all the information necessary to allow the FDA to determine whether or not the drug product that is the subject of the application meets the standards set out in the Act for a lawfully marketed drug product. These standards address not only matters bearing on the product's safety and effectiveness in clinical use, but on its method of manufacture; chemical identity, purity, and strength; pharmaceutical performance; bioavailability; and proposed labeling.

The Act requires the FDA, in turn, to "file" (i.e., accept for review) any NDA that appears on initial inspection to provide, in a reasonably organized and coherent format, full reports of all the tests necessary to evaluate whether or not the drug product meets the requirements just cited.

Finally, the Act instructs the FDA to review and to approve a sponsor's NDA within 180 days of its submission, *unless*, on review, the agency determines that the reports it contains *fail* to establish that the drug product identified in the application fully complies with the Act's requirements.

THE INVESTIGATIONAL NEW DRUG APPLICATION

Part of "36 - Regulatory Issues "

Because the Act forbids the introduction into interstate commerce of new drugs unless they are the subjects of an approved NDA, the lawful clinical testing of unapproved new drug products would be a practical impossibility if the Act did not provide for an exemption to this ban.

The original FFDCA (1938), accordingly, provided for

precisely such an exemption, known then as a “Notice of Claimed Investigational Exemption for a New Drug.” The exemption is still available, but it is now officially known as an Investigational New Drug (IND) application.

Initially, an investigational exemption could be obtained largely for the asking. Between 1938 and 1963, the sponsor of an IND had only to agree to keep records and clearly label its new drug as to its status as an unapproved investigational new drug, but little else.

With the passage of the Kefauver and Harris amendments of 1962, however, the IND requirements were extensively revised and expanded. Congress was led to alter the requirements for investigational use because of yet another public health disaster involving a drug product. In this case, fortunately, the drug thalidomide, although widely marketed in Europe, was not marketed in the United States. However, the potent teratogen thalidomide was widely distributed under INDs in the United States; worse, when its teratogenicity was recognized and efforts were undertaken to recall the supplies of it that had been distributed, the extent of domestic distribution was not easily determined.

Although very few American women who had received thalidomide under an IND bore children with limb reduction defects, the episode raised substantial concerns about the safety of human research subjects (2 ,3). Thus, Congress amended the Act so as to give the FDA the authority to monitor and control the conduct of clinical drug research within the United States.

Under the 1962 amendments, the agency gained explicit authority not only to establish mandatory prerequisites for the granting of INDs, but also the power to prevent the initiation and/or suspend the conduct of a clinical investigation (i.e., impose a “clinical hold”) being carried out under an IND if and when the agency concludes that an investigation poses an unreasonable or unnecessary risk to human subjects. In 1997, with the passage of the Food and Drug Modernization Act (FDAMA), FDA’s authority under the Act’s IND provisions was clarified and explicated in more detail, but not substantively modified.

In sum, since 1962, the IND serves not only as a license sponsors must obtain to allow them lawfully to ship unapproved new drugs in interstate commerce, but also the device through which the agency monitors and maintains control over the way in which clinical research with new drugs is conducted within the United States.

WHAT AN APPROVABLE NDA MUST DEMONSTRATE

Part of "36 - Regulatory Issues "

The text of the Act, at least insofar as safety and effectiveness standards are concerned, speaks almost entirely to broad goals and generic principles. Responsibility for interpreting the Act and developing, revising, and promulgating the regulations and policies necessary to secure the aims Congress had in mind in drafting the Act are delegated to the FDA.

Safety

Insofar as safety is concerned, the Act demands that a sponsor provide full reports of all tests necessary to establish that its product will be safe for use. The Act instructs the agency to reject a sponsor’s application, if, on review, it determines that the drug has been inadequately tested, or, if tested adequately, the findings of the tests conducted are inadequate to show that the drug, as recommended for use, is safe for use, or show that the drug, again as recommended for use, is unsafe for use.

Efficacy

The Act instructs the FDA to approve an NDA *unless*, on review of the reports submitted, it concludes there is a lack of “substantial evidence” that the drug is effective as claimed in its proposed labeling.

HOW THE AGENCY INTERPRETS THE ACT’S GENERIC REQUIREMENTS

Part of "36 - Regulatory Issues "

Safety for Use

No pharmacologically active drug substance is ever likely to be entirely free of risk. Accordingly, the agency maintains that a regulatory determination that a drug is “safe for use” is, in actuality, a favorable “risk-benefit” determination (i.e., a conclusion, based on the information evaluated, that the risks imposed by the use of the drug are acceptable in light of the benefits it provides).

Risk-benefit assessments, however, are hardly straightforward undertakings. To begin, their reliability is in large part a function of the extent and quality of the information on which they are based. Unfortunately, the information ordinarily available to inform a regulatory risk-benefit assessment is limited in scope; a typical NDA, for example, is approved based on experience gained with a drug product in perhaps 1,000 to 2,000 human subjects *in toto*.

The information that is available, moreover, is in many respects marginal in regard to its aptness; the individuals who participate in drug development programs, although they are reasonably representative of the patient population to which the drug will be administered when marketed, are not fully representative of it. Accordingly, risks uniquely, or more likely to be, associated with the use of a drug in various subgroups of the population, particularly when those subgroups are rare and/or under-represented in the samples studied during the drug’s development, are almost never appreciated, let alone factored into the regulatory risk-benefit determination.

Data bearing on the risks of a drug are collected during premarket testing under conditions of use (e.g., dose, regimen of administration, duration of use, restricted use of concomitant medications, etc.) that vary substantively from those under which a marketed drug is likely to be used. This is of especial concern where duration of use is concerned.

Ordinarily, the bulk of premarketing data are obtained in relatively short-term clinical trials (weeks or months) although the product under development typically will be used, once marketed, over much longer intervals (months to years). As a consequence, a typical drug development program has little, if any, chance of detecting untoward effects of a drug that emerge only after an extended period of exposure.

Regulatory officials are well aware of these limitations, but for practical (i.e., a politic word for economic and political) reasons, must tolerate them. The International Conference on Harmonization (ICH) in which the US participates, for example, has issued a guideline (4) that states that it is ordinarily sufficient to evaluate a new drug, prior to marketing, in no more than 300 to 600 patients for 6 months and 100 patients for a year.

Anyone familiar with the arithmetic of risk estimation will recognize that an experience of “safe passage” on a drug gained in such limited numbers of patients is not very reassuring. The failure to see even one catastrophic or fatal event in a sample of 300 patients only reduces to 5% the chance that the drug investigated causes such unobserved events at a rate no greater than in one of every 100 patients exposed to it (5).

(The notion of “safe passage” in drug safety assessment is the author’s invention; it likens regulatory premarket clinical trials to journeys undertaken on uncharted waters. Much as early seafarers determined which of several routes between two points was safer by comparing the risks of one with another, society determines whether or not a new drug is safe for use from the proportion of patients exposed to it who enjoy safe passage.)

Whatever one’s personal sense of the size of such a risk, it is truly enormous from a public health perspective. Just imagine the horror if a new antidepressant drug product caused a fatality in one in every thousand patients exposed to it. Yet, our society, like those of other developed nations, is seemingly content to market drugs without being able to confidently exclude a 10-fold greater risk.

In recent years, highly publicized withdrawals of new drugs shortly after their approval for marketing because of previously unrecognized risks of use have focused public attention on the limitations of premarket drug safety assessment. Both the agency and regulated industry have, consequently, been urged to develop new and better methods and paradigms to predict the likelihood of new drugs to cause injury (6).

No doubt, both the agency and regulated industry would be delighted to do so if they could. It is somewhat difficult to fathom, however, what kinds of methods will make it possible to identify drug-induced injuries in advance of their occurrence caused by pathogenetic mechanisms that have yet to be recognized, let alone characterized.

On the other hand, it is difficult to deny that the systematic study of a drug product’s capacity to cause effects that have in the past been associated with an increased risk of untoward effects should sometimes be useful. The risk of pharmacokinetic interactions, for example, should be predictable if the major metabolic pathways involved in the elimination of a new drug, its metabolites, and the pathways of elimination of other drug products likely to be coadministered with the new drug are identified and adequately characterized. Knowledge of a drug product’s metabolism also makes it possible to identify individuals in the population that might be at unique risk of suffering injury because of their diminished capacity to metabolize the drug (e.g., 6% to 8% of the white population are “poor” metabolizers of drugs that are CYP 450 2D6 substrates). Presumably, as our knowledge of the human genome expands, our ability to predict drug-induced risks on such grounds will grow.

Efforts to screen drugs prior to marketing for specific properties that predict drug-associated harms are still largely in their infancy, however. Moreover, such approaches have inherent limitations. Their utility is typically predicated on the assumption that the indicator of risk employed (e.g., a capacity to prolong the QTc interval on the surface ECG) reliably and consistently predicts a drug’s capacity to cause harm. As is the case with almost all surrogate indicators, however, there is always a possibility that the association between the surrogate and harm found in one set of circumstances will not hold in another.

In contemplating the development of new approaches to premarket safety assessment, it is important to be mindful that many of our expectations may be unrealistic, even magical. Congress, for example, did not intend that premarket testing would successfully identify every unsafe or unfit drug product. If it had, it would not have authorized the agency (Section 505(e) of the FFDCA), to withdraw approval of any NDA for a drug if new information, not available at the time of approval, becomes available which shows the drug is unsafe for use as labeled. Moreover, it is evident that Congress not only anticipated that new risks would be recognized after a drug’s approval for marketing, but expected that postmarketing surveillance would detect them. Specifically, to ensure that adverse information and reports bearing on the safety of marketed drug products would be collected and made available for evaluation, the Act (Section 505(k)) requires the sponsors of marketed drugs “…to establish and maintain… records, and make… reports to the Secretary, of data relating to clinical experience and other data or information, received or otherwise obtained by [the sponsor]… with respect to… [its drug]… to enable the Secretary to determine, or facilitate a determination,

whether there is or may be ground for… [withdrawing approval of the NDA]."

In actuality, however, only a relatively small proportion of marketed drug products are withdrawn from the market on grounds of being unsafe. This fact is frequently overlooked in the midst of the sensationalist publicity and second-guessing that so often accompanies product withdrawals. In the vast majority of instances, in fact, after their evaluation, adverse reports received on a drug from postmarketing surveillance sources lead, at most, to revisions being made in the product's approved labeling.

The nature of the labeling changes made, and the publicity given to them, is a function of the severity, estimated frequency, potential reversibility, and likelihood of mitigation or avoidance of each newly appreciated risk. From a regulatory perspective, a labeling change is a sufficient legal remedy, even for relatively serious newly appreciated risks, provided it remains possible for the agency to sustain its earlier conclusion, albeit under the newly revised labeling, that the drug product is "safe for use" *as labeled*.

The foregoing discussion reveals just how subjective and tenuous drug safety assessments actually are. Indeed, even if the risks associated with the use of a drug product were known exhaustively and in detail, its risk-benefit assessment would likely remain arguable. Paradoxically, the crux of disputes about risk-benefit determinations less often concerns the seriousness of the harms a drug causes, than the value of the benefits its use provides.

There is good reason for this. Except in those few instances in medical therapeutics were the use of a drug can be shown to reduce the absolute incidence of death or a serious, otherwise irreversible, injury, a substantively meaningful numerical estimate of the value of the benefit provided by a drug lies beyond reach. This occurs because there is a substantive distinction between a drug with an effect and an effective drug treatment.

Given the proper choice of experimental design, an appropriate patient sample, and the use of validated outcome measures, a competently executed clinical experiment can, of course, generate a numerical estimate of the magnitude of a drug's effect on some assessment instrument or scale of measurement. What such a trial cannot do, at least in a way that can be understood in a public sense, is to provide a meaningful measure of the value of the drug's clinical benefit. Unfortunately, it is the clinical benefit provided, and *not* the numerical estimate of the magnitude of a drug's effect as measured on some rating instrument that must be considered in a substantively meaningful risk-benefit assessment.

Thus, in areas of medicine, like psychiatry, where the beneficial effects of a treatment are *not* ordinarily measured by counts of cures or numbers of actual deaths prevented, but in the degree of symptomatic relief afforded by treatment, it is ordinarily impossible to obtain a publicly meaningful, let alone quantitative, estimate of a drug's value.

It is to be acknowledged that there are statistical estimates of "effect size" available (7), but these are intended only to gauge the relative magnitudes of a measured effect and the natural variation of that measured effect among individuals in the population being investigated. Although such "effect size" estimates are necessary to make informed "guesses" about the numbers of patients that must be admitted to a controlled clinical trial to obtain a statistically significant result, they say nothing whatsoever about the value of the effect being measured.

In sum, although a regulatory risk-benefit determination is widely represented to be a reasonable and responsibly considered judgment that derives from a disinterested weighing of the gains and harms known to be associated with the use of a drug, it is much more accurately depicted as a gestalt informed by an inchoate process that mixes, in undetermined proportions, evidence, sentiment, and personal values.

Nonetheless, because the Act requires that marketed drugs be shown to be safe for use prior to marketing, the FDA must take the information and reports presented in a sponsor's NDA, and with the assistance of its consultants and advisors, determine whether a reasonably qualified and informed expert, in possession of the data that are available, could conclude, fairly and responsibly, that the drug is "safe for use" as labeled for use within the meaning of the Act.

Effectiveness in Use

Effectiveness determinations under the Act are, at least in comparison to those involving safety, relatively straightforward, provided, however, that a modicum of agreement exists within the community of qualified medical experts as to: (a) what constitutes a beneficial treatment effect in the area of therapeutics involved, and (b) how that therapeutic effect is to be measured.

Given agreement on these two matters, it is a relatively simple, although not always an easy or quick, task for a sponsor to procure the evidence necessary to meet the Act's substantial evidence standard, provided, of course, that the sponsor's product is truly an effective one. If controversies are extant within the community concerning the condition that is the target of treatment (e.g., its features, diagnostic boundaries, cardinal manifestations, etc.), as they often are in the field of psychiatry, the task of demonstrating a drug's effectiveness in use becomes considerably more complicated but still possible.

The key question that remains to be addressed in this section then, is what constitutes substantial evidence within the meaning of the Act.

Between 1962 and 1997, Section 505[d] of the Act offered the following definition of substantial evidence:

> … the term "substantial evidence" means evidence consisting of adequate and well-controlled investigations, including clinical

> investigations, by experts qualified by scientific training and experience to evaluate the effectiveness the drug involved, on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof.

A number of points about the agency's traditional interpretation of the substantial evidence standard are noteworthy.

The agency has long held that "substantial" evidence must derive, at least in part, from the findings of valid *clinical* experiments. This presumably reflects the view that the value of drugs intended for the treatment of a human disease or condition can only be evaluated meaningfully in tests conducted in human subjects actually afflicted with or at risk of developing that disease or condition.

Agency regulations make clear that for evidence to be deemed substantial, it must, in part, be adduced in scientifically valid experiments (i.e., adequate and well-controlled investigations). Inferences based on scientific theory alone, or on the findings of uncontrolled clinical studies or observations (e.g., case reports, case series, etc.) will not suffice (e.g., see 21 CFR 314.126 (e)).

Neither can clinical judgment or professional opinion, per se, contribute to the body of evidence required to meet the substantial evidence burden. In light of the importance seemingly given to expert opinion in the statutory definition of substantial evidence, this assertion may seem at odds with the Act's requirements. A careful reading of the statutory definition reveals, however, that substantial evidence does not include what experts believe. To the contrary, the reference to experts in the definition of substantial evidence serves only to describe the character of the evidence that can be deemed substantial. (The evidence must be of a kind, quality, and quantity that would allow a disinterested and informed *expert* to conclude, "fairly and responsibly," from the evidence that the drug will have the effect its sponsors claims it has.)

It is also important to be mindful that "substantial" is an arcane legal term. It is frequently, and incorrectly, taken to mean definitive, even compelling, but substantial does not have that connotation within the meaning of the Act. Authorities on the legislative history of the Act (8) note that Congress considered, but decided not to employ, a "preponderant" standard of evidence, choosing, instead, the legally less demanding "substantial" evidence standard. Congress concluded, evidently, that the public health would be better served if the Act allowed the marketing of a drug even if only a minority of qualified experts agreed that it had been shown to be effective. It may help, therefore, to think of substantial evidence as the quanta of evidence sufficient to persuade at least a sizable number, but by no means a majority, of disinterested experts that a drug has been shown to be effective as claimed.

Among all the agency's interpretations of the substantial evidence standard, none, perhaps, has caused more concern and enduring complaint than its determination that positive findings from *more than one* adequate and well-controlled clinical investigation are ordinarily required to establish a drug product's effectiveness in use.

Although the requirement that experimental findings be independently substantiated prior to their formal acceptance is fully consonant with common scientific practice and epistemological principle, the agency's interpretation of the statute in this fashion proved so controversial and so politically vexatious that it eventually led Congress (Food and Drug Administration Modernization Act of 1997 [FDAMA]) to add the following sentence to the Act's definition of substantial evidence.

> If the Secretary determines, based on relevant science, that data from one adequate and well-controlled clinical investigation and confirmatory evidence (obtained prior to or after such investigation) are sufficient to establish effectiveness, the Secretary may consider such data and evidence to constitute substantial evidence for purposes of the preceding sentence.

FDA authorities contend (circa 1999 to 2000) that the revised definition does no more than confirm FDA's long standing authority (its regulatory discretion) to interpret the Act's efficacy provision flexibly and responsibly so as to secure the "aims of Congress" and advance the interests of the public health. In fact, before the Act was amended in 1997, the agency had, on occasion, taken exception to its traditional interpretation of the effectiveness standard, approving an NDA on the basis of findings from but a single adequate and well-controlled clinical study (9).

Agency spokespersons assert, further, that in all but highly unusual circumstances corroborating positive findings from more than one adequate and well-controlled clinical investigation will continue to represent the minimal quanta of evidence sufficient to satisfy the Act's substantial evidence requirement.

Evidently, because Congress failed to explain clearly what it intended by the terms "relevant science" and "confirmatory evidence" in drafting FDAMA, agency officials seemingly enjoy a degree of latitude as to whether and when to rely on the findings of a single controlled clinical investigation. Whether FDA officials will elect to or be able to retain this flexibility in the long run is far from certain, however.

THE EVALUATION OF DRUG PRODUCTS INTENDED FOR USE IN THE MANAGEMENT OF PSYCHIATRIC CONDITIONS

Part of "36 - Regulatory Issues "

The Division of Neuropharmacological Drug Products (DNDP) is the organizational unit within the FDA's Center for Drug Evaluation and Research (CDER) that is responsible for the monitoring of INDs and the evaluation of NDAs for psychiatric drug products.

It is the Division's obligation to review an application, determine whether or not it meets the requirements of the FFDCA as interpreted under FDA's prevailing regulations and policies, and, based on that review, take appropriate regulatory action.

For IND issues, signatory authority on most matters is at the level of the Division Director. Signatory authority is divided where NDAs are concerned. NDAs for new chemical entities (NCEs) are approved or disapproved at the Office level (Office of Drug Evaluation 1); supplemental NDAs (those involving new claimed uses of already marketed drug products) are approved at the Division level.

It is important to be mindful that Divisional policies evolve over time. A reader who wishes to understand whether, and if so, how a particular policy applies to a specific problem or drug product would be prudent, therefore, to seek fresh guidance on the subject directly from an appropriate Division representative.

From a regulatory perspective, the fact that a drug product is intended for the management of a psychiatric, sign, symptom, condition, or disease presents no unique problem.

A drug claim can be advanced for virtually any effect on the structure or function of the body of humans, for the cure or management of a disease or condition, or relief of a sign or symptom. The Act makes no distinction between the value of symptomatic treatments and those that are advanced as cures for a disease. Almost any claim can be made, provided that it can be presented in product labeling in a way that does not make the product's labeling, "false or misleading" in any particular.

In psychiatry, as in most other therapeutic areas, the effectiveness of a new chemical entity will almost always have to be demonstrated in *more than one* adequate and well controlled clinical investigation. It is possible, however, when an application is submitted for a claim closely related to one for which the drug product is already marketed, that a single controlled clinical study with robust and internally consistent findings might suffice. Because reliance on a single study is an exception to ordinary practice, however, the decision whether or not to take this approach will invariably be made on an ad hoc basis.

The agency's regulations (21 CFR 314.126) enumerate five control conditions that may be suitable for the evaluation of the effectiveness of new drug products. In light of the variability in course and outcome among samples of patients assigned the same psychiatric diagnosis, it is highly unlikely that either the historic control or the no control designs would ever be deemed acceptable for the evaluation of a drug for a psychiatric indication.

A randomized controlled and blinded trial employing any one or combination of the three other enumerated control conditions (placebo, graded dose, and standard active drug), provided it produces a statistically significant ($P \leq 0.05$, two-tail) difference favoring the investigational drug over a control condition, will invariably serve as one source of evidence contributing toward the quanta required to establish substantial evidence of a psychiatric drug product's effectiveness in use.

Among all the designs conforming to the requirements enumerated in 21 CFR 314.126, however, one that includes both a placebo and a standard drug control is especially attractive. Not only does such a design allow an estimate of what might have been had no treatment been administered (i.e., the response among placebo-assigned subjects) (10), but it provides a test of the capacity of the sample of patients participating in the experiment to respond to a treatment of established effectiveness (i.e., the response among patients randomized to the active control)

The three- (or more) arm design just described provides an internal means to assess what Modell and Houde (11) describe as an experiment's "assay sensitivity," its capacity to discriminate an inert substance from placebo and one level of an active drug substance from another.

Knowledge of an experiment's "assay sensitivity," or lack thereof, is especially important in the evaluation of psychiatric drug products because a sizable proportion of psychiatric drug trials (e.g., close to 50% or so of antidepressant trials) (12–14) fail to discriminate between active and inert treatments. Obviously, the trial's failure can be discounted if the failure of a study to find a drug placebo difference is owing to the inability of the sample of patients randomized in the study to respond to drug. On the other hand, if the sample of patients enrolled can respond to standard treatment, but not the investigational drug, the trial must be viewed as a source of evidence that speaks against the effectiveness of the investigational drug.

Why the agency encourages the use of an active control arm in a placebo-controlled study, incidentally, is often misunderstood. It bears emphasis that the active control is not included to obtain an estimate of the standard drug's performance relative to that of the investigational drug, but solely to gauge whether or not the experiment has "assay sensitivity."

When an investigational drug appears to have a relatively low therapeutic ratio, the use of fixed-dose graded designs is advantageous because it provides the surest means to identify the dose or dose range most likely to be acceptable for a typical patient. The agency is mindful, of course, that a single dose is unlikely to be the best choice for all patients; nonetheless, for dose evaluation purposes, a fixed-dose design is more likely to be interpretable than a dose titration design. Clinicians often find this assertion counterintuitive, but designs allowing up-titration for therapeutic nonresponse commonly produce an inverted dose-response relationship (i.e., treatment resistant subjects who show poor response are given the highest doses of drug).

Another vexing issue regularly confronting regulators is how best to extrapolate the results of clinical investigations of new psychiatric drug products to labeling claims.

Although the agency's regulations require that the sample of patients evaluated in controlled effectiveness trials be "reasonably representative" of the population of patients for which a drug will be recommended for use, the patients recruited in typical commercially sponsored drug trials are never truly representative of the population, at least not in any formal statistical sampling sense. To the contrary, the choice of patient subject is almost always based on the sponsor's convenience and its desire to maximize the statistical efficiency of its study (15). This sampling strategy is not objectionable from a regulatory perspective when the primary goal of a clinical study is to establish that the investigational drug product has the effect claimed for it in at least some patients with the condition for which the treatment will be marketed.

What claimed uses should be granted to a sponsor of a drug based on the results of such studies is yet another arguable matter. During the author's tenure as DNDP's director, for example, clinical trials conducted in acutely exacerbated schizophrenic patients were deemed sufficient to support a claim for the use of a neuroleptic drug product in the "management of the manifestations of psychotic disorders." In a good faith attempt to adhere to the Act's requirement that product labeling not be false or misleading in any particular, the text of the Indications and Usage section also briefly described the clinical investigations that supported the "antipsychotic" claim, including the nature of the patient samples that had been employed in them and whether or not long-term maintenance trials had been conducted. The strategy employed was intended to reserve for practitioners the right to determine the extent to which the results of a sponsor's effectiveness trials applied to psychotic conditions other than schizophrenia. DNDP's current leadership takes a different view. At DNDP's annual morning session at the NIMH's annual NCDEU meeting (June 2, 2000) held in Boca Raton, Florida, Tom Laughren (group leader for DNDP's Psychopharmacology Unit) announced that henceforth claimed uses for psychiatric drug products would be more narrowly defined (i.e., ordinarily limited to the population of patients actually studied). Thus, drugs shown to be effective in studies enrolling schizophrenic patients will get claims for use in schizophrenia, not psychosis. Incidentally, the newly announced approach to product labeling is perfectly reasonable and certainly consistent with the requirements of law, although it is obviously not the one that the author prefers.

Difficulties arise in the extrapolation of study results in many other areas as well. Drugs are administered to individuals, not diagnoses. There is advantage in knowing, therefore, whether, and if so how and to what extent, various individual patient characteristics (sex, age, race, severity of illness, etc.) and the interactions among them affect response to a drug. Offsetting the interest in obtaining the data necessary to address these issues are the difficulties and costs (both time and money) encountered in obtaining them; recruiting representative patients from even the more important of the patient subpopulations (e.g., children, the elderly, the very ill, etc.) is often exceedingly difficult, and sometimes, just not feasible.

For decades, the FDA took a relatively passive stance in regard to the demands it made on the regulated industry for data that might better inform the use of prescription drug products in children. Groups interested in the welfare of children have lobbied long and hard for making the study of investigational drugs in children a premarket obligation. FDAMA attempted to encourage sponsors to conduct remedial pediatric studies by offering the incentive of a 6-month extension on the patent life of certain drug products. In 1998, however, the agency issued new regulations (21 CFR 314.55 and 21 CFR 201.23), asserting its authority to require sponsors to evaluate new drugs in clinical trials enrolling subjects of pediatric age.

It seems unlikely to the author that the agency would actually refuse to approve an otherwise safe and effective new psychiatric drug product on the grounds that the effects of the drug had not been adequately studied in children, unless, of course, the primary use of the product was likely to be in children (i.e., the drug is intended as a treatment for ADHD).

The fact that the official psychiatric nosology is continuously evolving further complicates the extrapolation of study results to product labeling claims. For example, prior to the promulgation of DSM-III in 1980, sedatives, as they were then known, were granted broad and nonspecific claims for anxiety, anxiety neurosis, etc. As a result of DSM-III, a distinction had to be made between generalized anxiety disorder and panic disorder. In recent years, claims have been further expanded to include not only long-established entities such as obsessive–compulsive disorder (OCD), but also new entities such as social anxiety.

On each occasion that the psychiatric nosology expands or changes, the agency is confronted with a set of new problems including not only how to evaluate claims for newly created entities, but how to conform existing drug product claims to fit with the revised nosology. In dealing with these issues, the agency has to consider whether or not already approved older claims subsume the new entities. (Is a new claim simply a re-expression of a previous one, a claim for a subset of the patients covered by the previously approved claimed use, or an entirely new claim for a previously nonexistent entity?)

The agency also must decide how broad a claim or set of related claims may be. For example, sponsors often seek to define a claimed use in a way that will allow the unique promotion of their drug product (i.e., distinguish their drug from others within the same therapeutic class). When a claim links some diagnostic entity or subtype of entity to a drug's effect and that linkage is irrelevant to the expression of the effect, the claim is considered, "pseudospecific." (The author early in his tenure as DNDP director coined the

term “pseudospecificity”; it was first applied in connection with claims advanced for the use of benzodiazepines in anxious patients suffering from specific medical conditions [e.g., the anxiety of heart disease, cancer, etc.].) Such claims, moreover, are misleading because they seek to promote a distinction without meaning; consequently, they can be rejected by the agency because they can be held to be a violation of the Act’s requirement that a product’s labeling not be false or misleading in any particular.

A claim that a marketed antibiotic is effective for the pneumonia of dementia, for example, even if based on empirical evidence that the drug is effective in curing pneumonias in patients with dementia, is pseudospecific because the linkage between the pneumonia and the diagnosis of the patients treated is of no pharmacologic or biological importance, existing solely because of the sponsor’s decision to select demented patients with pneumonia as subjects for study. A legitimate (i.e., nonpseudospecific) disease-related claim requires a demonstration that the effect of the drug is in some way conditioned on the presence of the diagnosis (i.e., the diagnosis of the disease controls to what extent, if any, the effect of the drug is expressed).

The distinction made by the Division between pseudospecific and legitimate disease-related claims has often proved to be both unpopular and a source of continuing controversy. Claims advanced by sponsors for the use of marketed antipsychotic drug products in the management of psychotic demented patients are a case in point. During the author’s tenure as Director of DNDP, these claims were regularly deemed to be pseudospecific, in the absence of evidence to prove the contrary.

Based on testimony and discussion at a Psychiatric Drug Products Advisory Committee (PDAC) held in March, 2000 on treatments for the management of behavior in dementia, however, it appears that the Division is now inclined to accept “the psychosis of Alzheimer’s disease” as a bona fide entity for which drug product claims may be made.

The seeming reversal in the Division’s prior position, as far as the author can determine, came about because the current membership of the PDAC endorsed the psychosis of Alzheimer’s disease as an entity *sui generis*. Support for the existence of this entity derived primarily from the testimony of psychiatrists who treat demented patients as to what they believe is the true state of nature. The fact that a diagnostic algorithm for the capture of patients with the “psychosis of Alzheimer’s disease” had been recently proposed, seemingly gave further support to the reality of the putative entity. Although there is little doubt that the algorithm endorsed by the PDAC does capture demented individuals who exhibit behaviors that can be deemed psychotic, the fact that it does so in no way establishes that there is, in nature, a unique psychosis of Alzheimer’s disease. Indeed, what the discourse at March meeting of the PDAC revealed is that when a diagnostic nosology is based on a taxonomic system controlled by authoritative figures (i.e., “opinion leaders” in the terminology of marketing departments within the regulated industry), the existence of a diagnostic entity may owe more to politics than biology.

Incidentally, the taxonomic nature of our official psychiatric nosology not only complicates the task of drafting of accurate product labeling, but also contributes to the high failure rate of studies intended to document the effectiveness of psychiatric drug products.

It bears note that DSM-III was developed, at least in part, in the hope that phenotypic similarities among patients would reduce within diagnostic category genotypic variability and, thereby, make psychiatric diagnoses more predictive of course, outcome, and treatment response than were the diagnostic categories based on the pseudoexplanatory, dynamic systems employed in DSM-II (16).

Unfortunately, this goal of DSM-III has yet to be achieved, in part, perhaps, because the effort to improve psychiatric nosology has been confounded by issues and interests that have little to do directly with biologic classification (e.g., a focus on new entities that expand the size of a drug’s potential market, or that advance a professional career, or a political interest, etc.).

Whatever the explanation may ultimately be, however, sample-to-sample variation in both drug and placebo response rates among samples of patients assigned identical psychiatric diagnoses documents that our current psychiatric nosology has poor predictive power, at least insofar as drug responsiveness is concerned. As a consequence, to demonstrate that their products are effective in use, sponsors have had little option but to conduct trials using relatively large sample sizes.

Some critics have complained that the FDA, by allowing sponsors to conduct these “overpowered” trials, has turned the drug regulatory system into an institution that certifies as valuable drug products that provide little, if any, more benefit than placebo (17). Perhaps there is some truth in these allegations, but, then again, who is to say what drug effect is so small as to be dismissible? As discussed, from a strictly regulatory perspective such criticisms are irrelevant. It is important to recall that Congress could have set out other and/or more demanding standards for establishing drug effectiveness, but it did not.

The agency does have an obligation under the law, however, to ensure that its effectiveness determinations are based on fair and reliable estimates of a drug’s measured effects. Accordingly, a large part of the effort expended in the review of an NDA is devoted to excluding the possibility that bias, rather than a true drug effect, accounts for the positive study reports submitted in the application.

Great pains are thus taken to ensure that a sponsor’s claims are supported by valid statistical analyses of outcome measures prospectively identified in each study’s protocol. Each study report is scrutinized for lapses in the execution of a study that might have introduced bias into the estimates

of drug effect it provides. Reviewers search for evidence that the randomization process failed, the treatment mask was penetrated, subjects failed to comply with protocol requirements, etc. The extent, pattern, and timing of premature discontinuations are carefully analyzed in an effort to determine whether the censoring process has created a biased estimate of the drug's effect.

The immediate goal of these efforts is to ensure that the comparison between the investigational and control treatments can be made "ceteris paribus" (all other things being equal).

The overarching aim is to determine whether or not the evidence submitted meets the Act's substantial standard fairly and responsibly.

THE FUTURE OF DRUG REGULATION

Part of "36 - Regulatory Issues "

These are, at once, the best of times and the worst of times for medical therapeutics. On one hand, we have access to more information about the function and body of humans than we have ever had; therefore, we have good reason to hold sanguine expectations for the future in regard to the discovery and development of effective new treatments. On the other hand, the armamentarium stands at increasing risk of being polluted by worthless, even unsafe drugs. Signs of this potential risk abound.

The public is gullible and uncritical where therapeutic claims are concerned, exhibiting not only unrealistic expectations for the countless potential treatments touted as promising by the medical establishment and the pharmaceutical industry, but a willingness to spend billions of dollars on unproved, perhaps unsafe, remedies and nostrums.

Although the FDA's substantive powers under the FFDCA to regulate the drug supply, even after the passage of FDAMA, have been constrained to only a relatively minor extent, FDA's mission has noticeably changed.

Today's agency is expected to work as a "partner" of the regulated industry in a joint effort to expedite the development of new and promising treatments, not merely to keep unfit drugs off the market. It is arguable whether such dual expectations truly advance the interests of the public health, however. As the former head of the FAA noted when she resigned in the wake the ValueJet crash in Florida's Everglades, it is difficult for a federal regulatory agency to serve simultaneously and well the interests of both the consumer and the industry it regulates.

Congress has limited FDA's authority in less obvious ways as well. An egregious and illuminating example of its behavior is the Dietary Supplement Health and Education Act (DSHEA) of 1994. DSHEA has, by fiat, simply removed whole classes of drugs from FDA's jurisdiction and oversight.

By what is as close to an Orwellian *1984*-like maneuver as one is likely to find, the Congress simply declared whole classes of drugs to be something else: nutriceuticals, botanicals, food supplements, etc. These substances escape the premarket drug clearance requirements of the FDDCA because they are deemed by DSHEA *not* to be drugs.

Some may consider the legerdemain of DSHEA to constitute a relatively minor threat to the public health. Vitamins, for example, have long escaped the premarket clearance requirements of the Act, and have not proved dangerous, except perhaps when some are taken at excessive doses. Subsumed within the ill-described mix of botanicals, nutriceuticals, food supplements, and other materials deemed to be "nondrugs" under DSHEA, however, are large numbers of incompletely characterized, pharmacologically active drug substances that have yet to be fully and reliably evaluated for their safety and efficacy.

The lack of data bearing on the safety and efficacy of these drug substances did not deter the Congressional Leadership, however; to the contrary, when Senator Orin Hatch, a sponsor of DSHEA, introduced the bill he asserted that…dietary supplements can help promote health and prevent certain diseases, a fact substantiated by an ever-growing body of scientific studies and other evidence. In my own state of Utah, healthy life-styles, coupled with common use of dietary supplements have made a real difference. Our state is one of the healthiest in the nation and we enjoy one of the lowest incidence rates for cancer and heart disease. In Utah, the use of herbs is a well-accepted practice that has passed from generation to generation (18).

What is chilling here is not Senator Hatch's espousal of his state's virtues and the benefits of herbs, but his seeming willingness to approach an important public health issue in his nonscientific manner. Even more alarming is the inference congressional passage of DSHEA allows about the general state of our society's respect for science and the scientific method. Evidently, a majority of our elected representatives believes the benefits of scientific progress are attainable without the use of sound scientific methods.

If the armamentarium is to be kept reasonably free of unsafe and ineffective drugs, the antidrug regulatory drift of our national politics must be reversed. The FDA has demonstrated that it can, with the limited authority it has under the FFDCA, serve as an effective guardian of the drug supply. It will continue to serve effectively in this role, however, only if it obtains the necessary political support.

Given the current antiregulatory political environment, and the unrealistic expectations of the body politic for near magical breakthroughs in therapeutics, that support will not be forthcoming without the full support of the medical, scientific, and health care communities.

REFERENCES

1. Jackson CO. Doctor Massengill's elixir. In: *Food and drug legislation in the new deal*. Princeton, NJ: Princeton University Press, 1970:151–174.

2. Silverman M, Lee PR. Drug efficacy: the case of the emperor's old drugs. In: *Pills, profits and politics*. Berkeley: University of California Press, 1974:107–137.

3. Silverman M, Lee PR. The impossible dream: search for the absolutely safe drug. In: *Pills, profits and politics*. Berkeley: University of California Press, 1974:81–106.

4. Anonymous. The extent of population exposure to assess clinical safety: for drugs intended for long-term treatment of non-life-threatening conditions. *Fed Reg* 1995;60:11270.

5. Leber P. Postmarketing surveillance of adverse drug effects. In: Lieberman J, Kane J, eds. *Adverse effects of psychotropic drugs*. New York: Guilford Press, 1992:3–12.

6. Wood AAJ. Thrombotic thrombocytopenic purpura and Clopidogrel—a need for new approaches to drug safety. *NEJM* 2000;342:1824–1826.

7. Cohen J. *Statistical power analysis for the behavioral sciences*, rev ed. Orlando, FL: Academic Press, 1977.

8. Hoffman JE. Administrative procedures of the FDA. In: Adams DG, Cooper FM, Kahan JS, eds. *Fundamentals of law and regulations*, vol. 2. Washington, DC: Food and Drug Law Institute, 1977:13–53.

9. Anonymous. *Guidance for industry: providing clinical evidence of effectiveness for human drugs and biological products*. Rockville, MD: Food and Drug Administration, 1998.

10. Leber P. The use of placebo control groups in the assessment of psychiatric drugs: an historical context. *Biol Psychol* 2000;47:699–706.

11. Modell W, Houde RW. Factors influencing clinical evaluation of drugs with special reference to the double-blind technique. *JAMA* 1958;167:2190–2199.

12. Leber PD. Hazards of inference: the active control investigation. *Epilepsia* 1989;30:S57–S63.

13. Leber PD. Is there an alternative to the randomized controlled trial? *Psychopharmacol Bull* 1991;27:3–8.

14. Khan A, Warner H, Brown W. Symptom reduction and suicide risk in patients treated with placebo in antidepressant clinical trials: an analysis of the FDA database. *Arch Gen Psychiatry* 2000;57:311–317.

15. Leber PD, Davis CS. Threats to the validity of clinical trials employing enrichment strategies for sample selection. *Controlled Clin Trials* 1998;19:178–187.

16. Wilson M. DSM-III and the transformation of American psychiatry: a history. *Am J Psychiatry* 1993;150:399–410.

17. Fisher S, Greenberg RP, eds. *From placebo to panacea: putting psychiatric drugs to the test*. New York: John Wiley, 1997.

18. Statement of Sen. Hatch on Introduction of S. 784/April 7, 1993, Appendix 1: Section H. In: Bass IS, Young AL. *Dietary Supplement Health and Education Act: a legislative history and analysis*. Washington, DC: Food and Drug Law Institute, 1996.

37

Pharmacogenomics and Personalized Therapeutics in Psychiatry

Vural Özdemir

Vincenzo S. Basile

Mario Masellis

Pierandrea Muglia

James L. Kennedy

Vural Özdemir: Department of Psychiatry and Pharmacology, Center for Addiction and Health, University of Toronto, Toronto, Ontario, Canada.

Vincenzo S. Basile, Mario Masellis, Pierandrea Muglia, and James L. Kennedy: Department of Psychiatry, Center for Addiction and Mental Health, University of Toronto, Toronto, Ontario, Canada.

The advances in molecular medicine are taking place at a hitherto unprecedented pace. Genetics has come of age and will greatly influence the future of health care and therapeutics in the twenty-first century, in much the same way the breakthroughs in quantum physics and chemistry shaped science and society during the early phases of the twentieth century.

The field of pharmacogenetics was introduced more than 40 years ago to emphasize the role of heredity in person-to-person differences in drug response (1 ,2). The focus of pharmacogenetic investigations has traditionally been unusual and extreme drug responses resulting from a single gene effect. Pharmacogenomics is a recently introduced concept that attempts to explain the hereditary basis of both monogenic as well as subtler and continuous variations in drug responses that are under multigenic control (3). Although the two terms are often used interchangeably, the scope of pharmacogenomic investigations follows a genome-wide approach and also aims to identify novel biological targets for drug discovery, with use of the new affordable high-throughput molecular genetic technologies (4). In theory, pharmacogenomics can assist in clinical decision making to choose the most appropriate medication and dose titration regimen for individual patients. Moreover, the principles of pharmacogenomics are not limited to therapeutics. They can be applied to understand the hereditary basis of differences in sensitivity or resistance to any foreign chemical (xenobiotics) or environmental factor including foodstuffs, pesticides, infectious diseases, and ionizing radiation (5 ,6 and 7).

The Human Genome Project has already provided a draft nucleotide sequence of the human genome by mid-2000 and the nearly complete sequence is projected to be available by 2003. In addition, nucleotide sequence variations among individuals, populations, and species will be available in the near future. It is clear that these advances will soon lead to identification of many genes causing common complex diseases, thereby creating numerous new potential drug targets. This will also present a bioinformatics and data analysis challenge for lead optimization among numerous new chemical entities (NCE) directed to such disease targets for therapeutic purposes. Pharmacogenomics is a hybrid research field that bridges the knowledge gained from the Human Genome Project with existing principles of population genetics, pharmacokinetics, pharmacodynamics, cell physiology, proteomics, and bioinformatics. It is expected that pharmacogenomics will importantly contribute to development of guidelines for rational and personalized drug treatment; it should also expedite the drug discovery, development, and approval process in the pharmaceutical industry (8).

The purpose of this chapter is to introduce pharmacogenomics to those from a clinical psychiatry perspective, and discuss the future research challenges for those who may have prior experience in the field.

HISTORICAL OVERVIEW AND CONCEPTUAL FRAMEWORK FOR PERSONALIZED THERAPEUTICS

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

The marked interindividual variability in psychotropic drug effects was recognized long ago (9 ,10). For example, the

same dose of an antidepressant medication may cause toxicity, efficacious treatment, lack of efficacy, or qualitatively different drug effects among patients and populations (11). Understanding the sources of such variability in dose–response relationships is central to individualized dosage and choice of drugs for therapeutic purposes (12). Interestingly, it has been a commonly held viewpoint that genetics is important mainly for the permanent characteristics of an individual (e.g., stature) or predisposition to certain diseases, rather than variations in drug effects (13). In 1932, Snyder documented one of the first known interactions between heredity and response to xenobiotics: the ability to sense the bitter taste of phenylthiourea, which is under strong genetic control (14). In 1962 Werner Kalow published the first monograph on pharmacogenetics (1). Some argue that the field of pharmacogenetic inquiry dates as early as 510 BC, when Pythagorus in Croton, southern Italy, warned about the "…dangers of some, but not other, individuals who eat the fava bean" (5). The molecular basis of this historic observation was later documented to be hemolytic anemia owing to glucose-6-phosphate dehydrogenase deficiency.

The interest in personalized therapeutics was further fueled by the thalidomide disaster in 1960s. Subsequent observational studies found that drug–drug interactions, and hepatic and renal insufficiency importantly contribute to the risk for adverse drug reactions (15). A series of studies in monozygotic and dizygotic twins firmly established that genetic factors play an important role in metabolism of many drugs, and not only in a few cases of unusual adverse

drug reactions (16). This led to the publication of systematic guidelines on individualization of drug therapy by the American College of Physicians, based on genetic, environmental, and disease-related determinants of person-to-person differences in dose–effect relationships (17).

The pharmacologic drug response is a complex trait and is likely under polygenic control, rather than simple monogenic regulation (18). However, in comparison to disease-related complex traits, there is a well-established theoretic working model describing the relationship between the prescribed drug dosage and clinical drug effects (19). The conceptual framework developed by the seminal works of Sheiner and others allows one to target candidate genes with potential mechanistic relevance to partition the variability in any drug effect into pharmacokinetic and pharmacodynamic components (Fig. 37.1) (20 ,21 ,22 ,23 ,24 and 25).

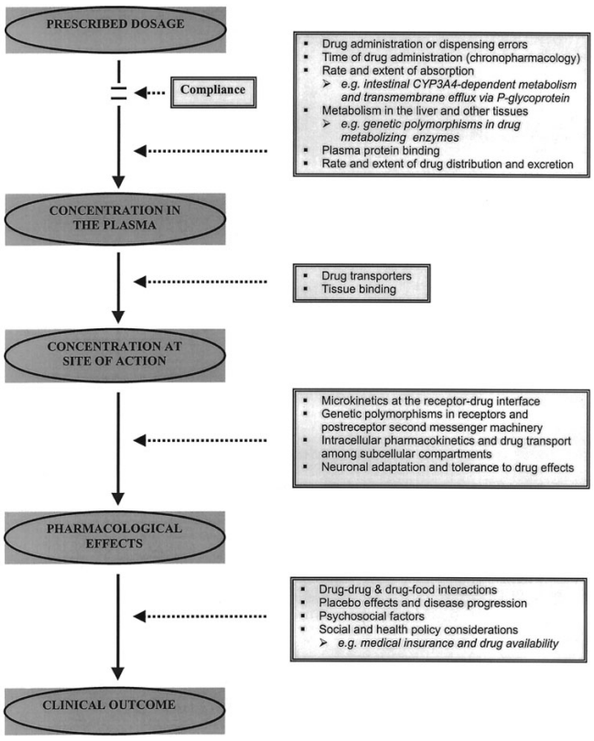


FIGURE 37.1. Pharmacological cascade describing the dose: clinical outcome relationship. Modified from Koch-Weser J. Drug therapy. Serum drug concentrations as therapeutic guides. *N Engl J Med* 1972;287:227–231.

The term pharmacokinetics describes the “drug concentration versus time” relationships in an organism by mathematical formulations of drug absorption from the site of administration, distribution, and elimination by metabolism and/or excretion (25). The pharmacodynamics explains the “drug concentration versus response” relationships and the related biological covariates (e.g., receptors, second messenger systems) (25). At present, pharmacogenomics is extending the early pharmacogenetic studies of drug metabolism to a broader context, to dissect the genetic control at multiple levels of the pharmacokinetic and pharmacodynamic pharmacologic cascade, from drug absorption and transport to drug–receptor interface and beyond. The progress made by pharmacogenomics, in many ways, is akin to developments in the computer industry. The potential benefits of high speed and efficient computing was self-evident early in the days of cumbersome mainframe computers, but it was not until the development of low-cost and proficient microcomputers that computerized information processing could be applied in daily life (analogous to contemporary ultrahigh throughput microarray genotyping technologies).

HUMAN GENETIC VARIATION AND SINGLE NUCLEOTIDE POLYMORPHISMS

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

The biallelic single nucleotide polymorphisms (SNPs) represent the most common DNA sequence variation in the human genome. It is thought that the complete human sequence including the coding regions, introns, and promoters will contain approximately one million SNPs (26). SNPs often result in predictable changes in amino acid sequence and contribute to diversity in protein function. SNPs are valuable biomarkers to elucidate the genetic basis of common complex diseases and pharmacologic traits. In the near term, the new genomic technologies will allow large-scale genetic association studies between numerous SNPs and drug response phenotype(s) in large samples of patients during routine drug treatment and clinical trials.

A pharmacogenetic polymorphism refers to a “…Mendelian or monogenic trait that exists in the population in at least two phenotypes (and presumably at least two genotypes), neither of which is rare; that is, neither of which occurs with a frequency of less than 1 to 2%…” (27). The definition of the minimum frequency threshold (i.e., 1% to 2%) is arbitrary and aims to emphasize that pharmacogenetic polymorphisms are *not* rare and different from those owing to recurrent spontaneous mutations occurring at much lower frequencies. A characteristic feature of pharmacogenetic polymorphisms is that they are usually biologically silent and do not present an evolutionary disadvantage or result in disease. This allows the maintenance of the less frequent phenotype at or above the 1% to 2% frequency level. Their clinical manifestations occur only when exposed to drugs or other xenobiotics, which target the polymorphic gene products.

Evident in the definition of pharmacogenetic polymorphism described in the preceding is the emphasis on phenotype (14). Alternative definitions of pharmacogenetic polymorphisms based on allele frequency also have been suggested (28 ,29). For example, Harris (1980) proposed that a genetic polymorphism occurs if the “…commonest identifiable allele (p) has a frequency no greater than 0.99…” (29). It was pointed out earlier that if the goal of pharmacogenetic studies is to investigate the clinical relevance of pharmacogenetic polymorphisms, phenotype-based definition of polymorphisms might be more applicable (28). On the other hand, if the goal is to use genetic polymorphisms as anthropological tools to study the evolution of the species and differences in response to xenobiotics between populations, a definition incorporating both allelic variations and phenotype may be more appropriate to allow understanding at a molecular and mechanistic level (28).

GENETIC VARIABILITY IN DRUG METABOLISM: CONTRIBUTION TO PHARMACOKINETIC VARIABILITY

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

Drug metabolism is one of the pivotal factors, which contribute to variability in pharmacokinetics. Drug metabolism is generally divided into two phases. Phase 1 reactions involve oxidative, reductive, and hydrolytic reactions, which unmask or introduce a functional group (e.g., a hydroxyl-moiety) to the parent compound. This often results in an increase in polarity of the drug. Phase 2 reactions involve conjugation (e.g., with glucuronic acid) of the metabolite produced in phase 1 reactions, or the parent compound, to more hydrophilic metabolites (30). Although drug metabolism is necessary for the elimination of lipophilic drugs (e.g., psychotropics), it may also be crucial for activation of prodrugs (e.g., codeine).

Phase 1 reactions are mediated, to a large extent, by the CYP enzymes that are mostly found attached to the smooth endoplasmic reticulum of the hepatocytes and other drug metabolizing cells (e.g., enterocyte in the gut) (31). A recent analysis of over 300 drugs from diverse therapeutic classes such as psychotropics, analgesics, and anti-infectious agents found that 56% of them primarily depend on CYPs for their metabolic clearance (32). Among CYPs, the largest contributions are made by CYP3A4 (50%), CYP2D6 (20%), CYP2C9, and CYP2C19 (15%) (32). Some of the drug metabolizing enzymes (e.g., CYP1A2 and CYP2D6) are also expressed in the brain and may potentially play a role in local disposition of psychotropics at the site of action (33 ,34).

CYTOCHROME P450 2D6 GENETIC POLYMORPHISM: A COMMON AND CLASSICAL EXAMPLE OF A MONOGENIC VARIATION IN DRUG METABOLISM

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

CYP2D6 genetic polymorphism is one of the most intensively studied autosomal recessive monogenic defects in drug metabolism. Many psychotropics, including most typical (e.g., perphenazine) and some atypical antipsychotics (e.g., risperidone), tricyclic antidepressants (e.g., nortriptyline), drugs of abuse, some of the serotonin reuptake inhibitors (SSRIs) (e.g., paroxetine), and codeine are metabolized by CYP2D6 (35 ,36 and 37).

Among Whites, approximately 7% of the population are poor metabolizers (PMs), whereas the rest are extensive metabolizers (EMs) for CYP2D6 substrates (35). The prevalence of PMs and the distribution of enzyme activity appear to be fairly consistent across the Western European and North American Whites. On the other hand, the frequency and type of *CYP2D6* alleles vary considerably among different ethnic groups. In Asians, the prevalence of PMs is only 1%, owing to almost complete absence of the nonfunctional alleles (e.g., *CYP2D6*4*) found in Whites (35). However, an often overlooked point in comparisons of pharmacogenetic polymorphisms between populations (e.g., Asians *versus* Whites) is that the key dependent variable is not only the prevalence of PMs but also the distribution of enzyme activity *within* EMs (36). Asian EMs (i.e., 99% of the population) display a significant shift in the distribution of CYP2D6 activity toward lower levels. The molecular basis of a lower CYP2D6 activity in Asian EMs is owing to a $C^{188}\rightarrow T$ base change in exon 1 which leads to $Pro^{34}\rightarrow Ser$ amino acid substitution in a highly conserved region (Pro-Pro-Gly-Pro) characteristic of CYP1 and CYP2 families (37). This allele was named *CYP2D6*10* (51% allele frequency in Chinese) and leads to a 10-fold decrease in activity *in vivo* (35). Thus, there may be discrete interindividual differences in disposition and therapeutic/adverse effects of psychotropics *within* Asians, depending on the gene-dose for the *CYP2D6*10* allele.

Another novel allele with reduced catalytic function, *CYP2D6*17*, occurs at high frequency in many black African populations and African-Americans. However, there appears to be considerable heterogeneity in the *CYP2D6* locus in Black populations. For example, in Ethiopia, only 1.8% were PMs of debrisoquine, 16% carried the *CYP2D6*10B* allele characteristic of Asian populations, and the *CYP2D6*17* was present in 18% of the subjects (38). Importantly, 29% of the Ethiopians carried alleles with duplicated and multiduplicated CYP2D6 genes associated with ultrarapid metabolism of substrates (38). Also, a high percentage of gene duplication and ultrarapid metabolism was found in Saudi Arabia and Spain, presumably owing to the genetic admixture during the earlier Islamic migration originating from some of the North African populations with ultrarapid CYP2D6 activity (36 ,37). The clinical significance of reduced catalytic function associated with the *CYP2D6*17* allele requires further research in patients with African ancestry.

At present, more than 50 *CYP2D6* alleles were described that encode an enzyme with inactive, decreased, increased, or normal catalytic function. The number of functional *CYP2D6* genes correlates with drug and metabolite concentrations in the plasma, as aptly documented using nortriptyline as a model substrate (Fig. 37.2) (39). In general, PMs are at risk for drug toxicity on treatment with medications predominantly inactivated by metabolism via CYP2D6. On the other hand, in prodrugs, which need to be converted to their active form by CYP2D6, opposite clinical consequences may occur in PMs. For example, codeine does not produce analgesic effects in PMs or after treatment of EMs with CYP2D6 inhibitors such as quinidine. Among EMs, those with duplicated or multiduplicated CYP2D6 genes and ultrarapid metabolism may develop subtherapeutic plasma concentrations and inadequate clinical response (40). Although high doses would be necessary in such patients, an alternative strategy would be to use low subtherapeutic doses of the CYP2D6 inhibitor quinidine, especially for drugs with high acquisition costs, to attain therapeutic plasma concentrations (41). Overall, routine genotyping for *CYP2D6* may be useful to avoid drug toxicity in PMs, to ascertain the pharmacokinetic mechanism of resistance to some psychotropics, as well as for differential diagnosis of noncompliance versus ultrarapid CYP2D6 activity. In addition, the markedly increased metabolite formation in patients with multiduplicated *CYP2D6* genes may potentially lead to qualitatively different and unexpected drug effects and toxicity (40).

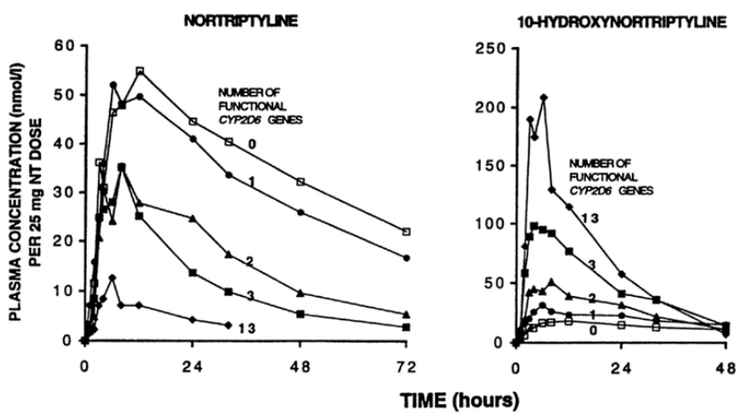


FIGURE 37.2. Average plasma concentrations of nortriptyline and 10-hydroxynortriptyline after a 25-mg single oral dose in White healthy volunteers with 0, 1, 2, 3, and 13 functional copies of the CYP2D6 gene. Note that the concentration of nortriptyline and its metabolite 10-hydroxynortriptyline are inversely affected by the number of functional CYP2D6 gene. Reprinted with permission from Dalén P, Dahl ML, Ruiz ML et al. 10-Hydroxylation of nortriptyline in white persons with 0, 1, 2, 3, and 13 functional CYP2D6 genes. *Clin Pharmacol Ther* 1998;63:444–452.

It is noteworthy that CYP2D6 with identical pharmacologic and molecular properties was identified in microsomal fractions in the brain. Hence, CYP2D6 may potentially contribute to local clearance of psychotropics at the site of action (42). Moreover, CYP2D6 in the brain is functionally associated with the dopamine transporter and shares similarities

in substrates and inhibitors (e.g., d-amphetamine), suggesting a role in dopaminergic neurotransmission (42). Differences in personality traits between EMs and PMs were noted in both Swedish and Spanish healthy White subjects, also suggesting that there may be an endogenous substrate for CYP2D6 in the brain (42).

The common polymorphic drug metabolizing enzymes in humans and their major variant alleles are presented in Table 37.1 (37). A worldwide web page with detailed descriptions of new alleles, nomenclature and useful references can be found at (http://www.imm.ki.se/CYPalleles/).

Enzyme	Major Variant Alleles	Mutation	Consequences for Enzyme Function	Allele Frequencies (%)			
				Caucasians	Asians	Black Africans	Ethiopians and Saudi Arabians
CYP2A6	*CYP2A6*2*	Leu160His	Inactive enzyme	1–3	0	ND	ND
	*CYP2A6*del*	Gene deletion	No enzyme	1	15	ND	ND
CYP2C9	*CYP2C9*2*	Arg144Cys	Reduced affinity for P450 oxidoreductase	8–13	0	ND	ND
	*CYP2C9*3*	Ile359Leu	Altered substrate specificity	6–9	2–3	ND	ND
CYP2C19	*CYP2C19*2*	Aberrant splice site	Inactive enzyme	13	23–32	13	14–15
	*CYP2C19*3*	Premature stop codon	Inactive enzyme	0	6–10	ND	0–2
CYP2D6	*CYP2D6*2xN*	Gene duplication or multiduplication	Increased enzyme activity	1–5	0–2	2	10–16
	*CYP2D6*4*	Defective splicing	Inactive enzyme	12–21	1	2	1–4
	*CYP2D6*5*	Gene deletion	No enzyme	2–7	6	4	1–3
	*CYP2D6*10*	Pro34Ser, Ser486Thr	Unstable enzyme	1–2	51	6	3–9
	*CYP2D6*17*	Thr107Ile, Arg296Cys, Ser486Thr	Reduced affinity for substrates	0	ND	34	3–9

Reprinted with permission from Ingelman-Sundberg et al. *Trends Pharmacol Sci* 1999;20:342–349.
ND, not determined.

TABLE 37.1. HUMAN POLYMORPHIC CYTOCHROME P450 ENZYMES AND THE GLOBAL DISTRIBUTION OF THEIR MAJOR VARIANT ALLELES

CYP3A4: A NONPOLYMORPHIC VARIATION IN DRUG METABOLISM

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

The term “polymorphic metabolism” is often perceived as an alarming indication of marked variability in drug disposition. Although this assertion is correct to a certain extent, it does not imply that a nonpolymorphic drug-metabolizing enzyme is associated with reduced variability. For example, CYP3A4 is the most abundant CYP isoform in the adult human liver with large interindividual variability in its expression. *In vivo*, CYP3A4 activity displays at least 20-fold

difference in the population (43). Yet, the distribution of CYP3A4 catalytic activity is unimodal and nonpolymorphic in many populations.

CYP3A4 contributes to disposition of more than 60 frequently prescribed therapeutic agents with diverse chemical structures including antilipidemics, benzodiazepines, HIV protease inhibitors, immunosuppressants, and macrolide antibiotics (43 ,44). CYP3A4 also plays an important role for the metabolism of endogenous steroids (e.g., testosterone) as well as activation of dietary mycotoxins (e.g., aflatoxin B1) (43). The prediction of CYP3A4-mediated drug metabolism is complicated by the presence of at least two distinct pools of the CYP3A4 protein, in the liver and intestine, whose expressions appear to be regulated independently. The appreciation of marked variability in CYP3A4 activity is critical for individualized treatment with CYP3A4 substrates, to forecast drug–drug interactions mediated by CYP3A4, and to identify the factors predisposing to long-term toxicity (e.g., prostate and liver cancer) associated with variable metabolism of steroid hormones and procarcinogens (44).

CYP3A4 expression can be markedly induced *in vivo* during chronic treatment with drugs such as the antibiotic rifampicin, anticonvulsant carbamazepine, and glucocorticoid dexamethasone (43). Conversely, CYP3A4 catalytic activity can be inhibited potently by commonly used drugs including the azole antifungal agents (e.g., ketoconazole) and the macrolide antibiotics (e.g., erythromycin), or by foodstuffs such as grapefruit juice (44). For example, excessive sedation or psychomotor impairment can occur after oral administration of benzodiazepines with a low bioavailability (e.g., triazolam) or some nonbenzodiazepine (e.g., buspirone) hypnosedatives together with grapefruit juice (44). Many patients with a mental health problem also use nonpsychotropic medications. In such cases, clinically significant hypotension may be observed during treatment with dihydropyridine calcium channel antagonists (e.g., nifedipine) and CYP3A4 inhibitors.

Studies in monozygotic and dizygotic twins indicate a high heritability (H_2 = 0.88) of CYP3A4 activity (45); however, there has been relatively little progress in identification of the molecular genetic underpinnings of heterogeneity in *CYP3A4* expression. Recently, a novel allele, *CYP3A4*2*, causing a Ser222Pro change was found in Whites at a frequency of 2.7%, but this allele was absent in Black and Chinese subjects (46). The *CYP3A4*2* displays a substrate-dependent diminished metabolic clearance; for instance, nifedipine (but not testosterone) intrinsic clearance is impaired (46). Because functional polymorphisms in the promoter or the coding region of CYP3A4 do not appear to be very common, it is likely that CYP3A4 activity represents a complex trait regulated by multiple interacting genetic loci in the genome (47).

GENETIC VARIABILITY IN RECEPTORS AND DRUG TRANSPORTERS: CONTRIBUTION TO PHARMACODYNAMIC VARIABILITY

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

Historically, the pharmacogenetic factors related to drug efficacy and safety were mainly studied in the context of drug metabolism (48). In the past decade, the increasing application of mathematical models for “concentration *versus* effect” relationships during routine drug development clearly documented the marked interindividual variability in pharmacodynamics (23). It is estimated that approximately 30,000 proteins with diverse structures are expressed in the human brain, many of which may serve as potential drug targets (49). Recent studies of SNPs in genes relevant for psychotropic pharmacodynamics indicate that human genetic variation in drug receptors and transporters may significantly contribute to overall variance in response to drugs.

Dopamine D3 receptor *(DRD3)* gene is expressed in the basal ganglia and is thought to play a role in locomotion. Three independent studies found that the Ser9Gly polymorphism in the N-terminal extracellular domain of the DRD3 is associated with an increased propensity to develop tardive dyskinesia in patients treated with typical antipsychotics (Fig. 37.3) (50 ,51 and 52). Pharmacogenetic polymorphisms in the dopamine D4 receptor *(DRD4)* gene, especially the hypervariable exon III 48 bp variable number of tandem repeat (VNTR), have been studied intensively in relation to antipsychotic response to clozapine (53). Although the exon III 48 bp VNTR in *DRD4* does not appear to be a major contributor to clinical outcome during clozapine treatment, it is possible that haplotype analyses incorporating several variants in different locations would be necessary before definitive conclusions can be drawn for the importance of *DRD4* in antipsychotic response. Moreover, studies involving the serotonin receptor gene polymorphisms suggest an association between the 5-HT2A receptor and response to clozapine (53 ,54 and 55). Other neurotransmitters such as norepinephrine, acetylcholine, and glutamate may also contribute to antipsychotic drug effects, but genetic variation in these receptors has not been investigated following a pharmacogenomic perspective (18). A detailed review of polymorphisms in dopamine and serotonin receptor genes, their relevance for response to clozapine and other atypical antipsychotics, and methodologic considerations for application of molecular approaches to psychiatric genetics are available elsewhere (53).

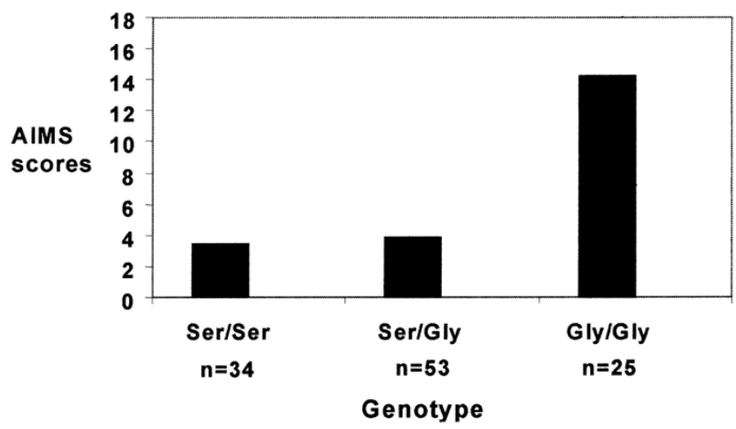


FIGURE 37.3. Average Abnormal Involuntary Movement Scale (AIMS) scores in 112 schizophrenic patients previously treated with typical antipsychotics and genotyped for the serine to glycine polymorphism in the N-terminal extracellular domain of the dopamine D3 (DRD3) receptor. A post hoc Student-Newman-Keuls test revealed a higher average AIMS score in patients homozygous for the glycine allele of the DRD3 gene, compared to those with a heterozygous or homozygous genotype for the serine allele. The analysis of variance results were corrected for age, gender, ethnicity and pairwise comparisons (F = 8.25, df = 2, P < 0.0005 [P < 0.0015, Bonferroni corrected]). Reprinted with permission from Basile VS, Masellis M, Badri F, et al. Association of the MscI polymorphism of the dopamine D3 receptor gene with tardive dyskinesia in schizophrenia. *Neuropsychopharmacology* 1999;21:17–27.

Because the first report more than 60 years ago on the use of amphetamine in children with attention deficit hyperactivity disorder (ADHD), it became clear that approximately 75% of ADHD cases show clinically significant improvement after d-amphetamine or methylphenidate treatment. Although the precise mechanism of action of these stimulant agents still remains elusive, their interaction with the dopamine transporter may contribute to their therapeutic

effects in ADHD. The VNTR polymorphism of the dopamine transporter gene appears to influence the response to methylphenidate, based on a preliminary study in 30 African-American children with ADHD (56). These limited data, however, do not allow generalizations on genetic determinants of response to pharmacologic interventions in ADHD at the present time.

The marked temporal delay in therapeutic effects is a well-known phenomenon with antidepressant agents. Therefore, it is advantageous to identify beforehand the subpopulation of patients who are unlikely to respond to a given medication so that various augmentation efforts can be initiated promptly. The high-affinity serotonin transporter (5-HTT) is a prime target for the serotonin reuptake inhibitor antidepressants (SSRIs). A functional polymorphic variant of the *5-HTT* gene characterized by a 44-bp insertion in its promoter region leads to differences in the amount of *5-HTT* transcript and the extent of 5-HT reuptake (57). Clinical studies suggest that the 44-bp insertion polymorphism of the *5-HTT* gene influences the antidepressant response to SSRIs including fluvoxamine and paroxetine (57). Further studies with other SSRIs and classical tricyclic antidepressants are called for to assess the overall clinical significance of *5-HTT* promotor polymorphism(s).

P-glycoprotein encoded by the *MDR1* gene is another drug transporter that affects transmembrane efflux and intracellular or tissue availability of numerous drugs. For example, amitriptyline (but not fluoxetine) can penetrate the brain more readily in knockout mice that do not express p-glycoprotein (58). Hence, differences in *MDR1* expression owing to genetic polymorphisms or secondary to chronic antidepressant treatment may explain treatment-resistance to amitriptyline in patients who otherwise attain therapeutic plasma drug concentrations.

The study of pharmacogenetic polymorphisms in drug targets is a relatively new but rapidly expanding research area. It is likely that molecular genetic profiling of patients for SNPs or other types of human genetic variation in both pharmacokinetic and pharmacodynamic targets will bring psychiatric genetics and clinical pharmacology one step closer to achieve the ultimate goal of individualized therapeutics.

PHARMACOGENOMICS AND DRUG DISCOVERY

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

The drug discovery in psychiatry was initially based on serendipity. The identification of lithium in 1949 and chlorpromazine in 1950s are two well-known examples where putative mechanisms of action were elucidated after the drugs were shown to be efficacious. The newer drug discovery paradigms have depended on the synthesis and identification of novel compounds through combinatorial chemistry

and screening for biological activity against known receptors or other biological targets with established endogenous ligands or substrates (59 ,60). With the Human Genome Project approaching to its completion, essentially all human genes will be available as potential drug targets. The challenge in drug discovery will then be to discern the function and therapeutic utility of these genes and their expressed products.

The experimental paradigms used by pharmacogenomics borrow substantially from the field of population genetics and the methodology used in earlier genetic studies of common complex diseases (60 ,61). For example, linkage and association studies are two well-known strategies to identify the genes causing a specific disease or variability in drug effects. The linkage design was traditionally used to test the relationship between inheritance of a complex disease phenotype within family members and microsatellite markers comprised of five or less short tandem repeats of DNA. The increasing availability of SNP markers and the ability to genotype the entire genome of large segments of patient populations with ultrahigh throughput methods such as the DNA microarrays, often referred to as “DNA chips,” now allow the application of genetic linkage or association designs to elucidate the genes responsible for variations in therapeutic response and toxicity. On the other hand, the obvious difficulties in administering drugs to different family members and obtaining relevant data on drug response phenotype may pose a constraint on application of linkage design to pharmacogenomics.

DNA microarray is an emerging powerful technological breakthrough that enables the study of global gene expression patterns and sequence variations at a genome level (62). In essence, DNA microarray is an extension of the Southern blot procedure and is comprised of different cDNAs or oligonucleotides etched systematically on a solid surface such as silica or glass plate. Each DNA species on the array represents a specific gene or expressed sequence tag, which is used to identify different SNPs or transcripts by hybridization and fluorescence detection. Microarrays with 10,000 or more genes are now available for use in clinical research or trials. An important application of microarrays is monitoring of temporal changes in gene expression during drug treatment or patients versus healthy individuals. The premise in these studies is that patterns of gene expression may serve as indirect clues about disease-causing genes or drug targets. Moreover, the effects of drugs with established efficacy on global gene expression patterns may provide a guidepost, or a “genetic signature,” against which the new drug candidates can be validated (63). Other new genomic technologies such as genotyping by mass spectrometry also are being developed. Collectively, pharmacogenomics adds another dimension to contemporary drug discovery efforts because it aims to identify novel drug targets in the *entire* human genome without *a priori* assumptions on disease pathogenesis or drug targets, thereby presenting an opportunity to unlock unprecedented novel mechanisms of drug action (4).

PHARMACOGENOMICS AND DRUG DEVELOPMENT

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

After the discovery of an NCE with therapeutic potential, the next step involves clinical testing in healthy volunteers and relevant target patient populations. For every clinician, an appreciation of the drug development process is important to make evidence-based choices among therapeutic alternatives and to be aware of the shortcomings of the data presented to support the efficacy and safety of new medications.

The drug discovery and development is a high-risk venture. Typically, it takes 8 to 12 years to introduce a new drug from discovery to clinical practice, with costs often approaching $100 to $300 million. On the other hand, it is estimated that approximately 90% or more of NCEs under development fail to meet the regulatory approval for clinical use (64). It is well known to most pharmaceutical scientists that the art of timely and cost-effective drug development rests on early identification and removal of drug candidates with poor efficacy and safety. It is conceivable that some of these NCEs may in fact have a favorable efficacy and safety profile in certain genetically determined subpopulations. Through proper design of clinical trials and using low-cost high-throughput genetic analyses, pharmacogenomics eventually can allow patenting of such “failed” NCEs in discrete patient populations and reinstate their market potential. In addition, a genetic test predicting drug effects would be considered another pharmaceutical product and an additional financial incentive for drug developers. For example, clozapine was recognized as a potential antipsychotic drug in early 1970s. The occurrence of agranulocytosis in several cases caused the termination of further development of clozapine in treatment-resistant patient populations until the late 1980s. The presence of genetic or other predictors of agranulocytosis would have expedited the development of clozapine and prevented the inconvenience of periodical hematologic monitoring.

Pharmacogenomics is also relevant for better use of medications that are already in routine clinical use (phase 4 drug development). A recent meta-analysis of prospective studies from 1966 to 1996 found that the incidence of serious and fatal adverse drug reactions in United States was 6.7% and 0.32%, respectively; ranking between the fourth to sixth leading cause of death, ahead of pneumonia and diabetes (65). Importantly, the adverse drug reactions in the latter study occurred during treatment with usual doses of drugs that already met the regulatory requirements for clinical use, and excluded cases owing to intentional or accidental overdose, errors in drug administration, or noncompliance.

It is likely that the proportion of such patients who are inadequately treated may further increase after accounting for therapeutic failures secondary to ultrarapid drug metabolism, for instance, and mismatches between the pharmacodynamic attributes of medications and drug targets in individual patients (23 ,37). Evidently, the existing pharmacotherapy system based on the traditional trial-and-error approach is unable to deliver personalized drug treatment and health care.

From the perspective of patients, healthcare providers, and managed care organizations, an increased probability of therapeutic response through genetic testing would reduce duration of inpatient hospital care, frequency and inconvenience of repeated physician visits owing to treatment resistance, and thus, easily offset the costs of pharmacogenomic-based drug development. For NCEs that readily meet the regulatory requirements with large efficacy margins (i.e., “blockbuster drugs”) over placebo or the existing standard treatment modalities, there may be less financial incentive—on the manufacturers’ part—to identify different subpopulations with differing drug effects, because this may potentially decrease the market share of their newly introduced medication. Therefore, although pharmacogenomics provides a clear rationale for improved drug discovery and personalized therapeutics, it will likely need enforcement by regulatory agencies before it can be utilized in routine clinical practice and pharmaceutical industry. This may in turn require amendments to existing regulatory policies for drug development. Also, as a result of the global harmonization attempts to standardize drug development and approval process, the settings of future clinical trials will not only be limited to Western society; therefore, it would be critical and advantageous to plan such policy amendments at a multinational level.

ETHICAL AND HEALTH POLICY CONSIDERATIONS

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

Pharmacogenomics arrived at the heels of the molecular biology revolution in 1980s and early 1990s. Although there is much optimism for a more efficient drug discovery and development process, there is also increasing concern about the implementation of genetic testing at various levels of the medical practice and its repercussions for health policy and managed care organizations. Confidentiality of the genetic test results is critical because it has important bearings on finding employment and obtaining life, health, or disability insurance. There is an urgent need to amend the existing medical curriculum to educate future clinical personnel for genetic counseling and fundamentals of molecular medicine. About 5% of the budget for the Human Genome Project is reserved to address these social and ethical issues.

CONCLUSION

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

Personalized therapeutics has been a preoccupation in clinical psychiatry for many decades. The traditional gene-by-gene approach to explain variability in drug response has been a mainstay in most pharmacogenetic investigations to date. However, the biological underpinnings of drug response are complex and often involve contributions by multiple genes, environmental factors, drug–drug and drug–food interactions, to mention a few (Fig. 37.4) (18 ,66 ,67 ,68 ,69 ,70 and 71). Clearly, a genome-wide approach will be an important advance in understanding the variability in drug efficacy and safety. To this end, sequence variations in the genome are only the first level of complexity. More intriguing and challenging is establishing the significance of differences in global gene expression patterns in relation to drug effects and targets. High throughput and genome-wide transcript profiling for differentially regulated mRNA species in disease, normal physiology, and after drug treatment offer an additional dynamic perspective for drug discovery. The development of protein chips may permit further explorations of functional genomics in the context of psychopharmacology. We may soon be surprised that the mechanism of action of some psychotropics may in fact rest on targets entirely different than what the conventional pharmacologic wisdom suggests (e.g., the monoamine hypothesis for antidepressant drug effects).

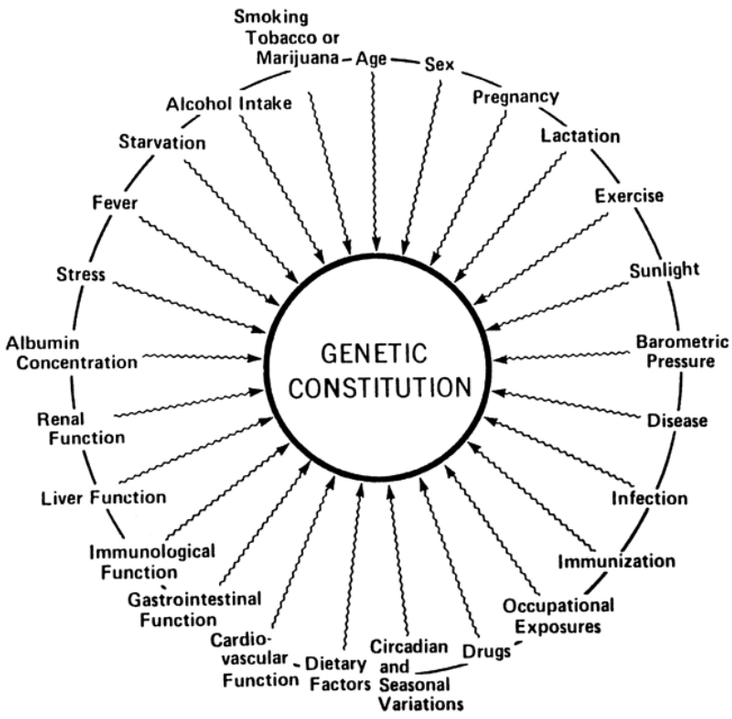


FIGURE 37.4. The interaction of genetic and environmental factors that may influence drug response in humans. Reprinted with permission from Vesell ES. On the significance of host factors that affect drug disposition. *Clin Pharmacol Ther* 1982;31:1–7.

At the present time, however, it is not clear whether and to what extent the genomic hypotheses can be tested within the framework of the available clinical trial methodology. For example, the sample size in most phase 3 clinical trials does not usually exceed 3,000 to 4,000 patients. Genome-wide association studies and statistical correction for multiple testing will require sample sizes well beyond the current resources of any single pharmaceutical company or an academic laboratory. Ideally, pharmacogenomics should be used for the prospective design of phase 3 clinical trials and not to salvage an NCE that proved to be ineffective or unsafe at the end of phase 3 investigations. Care should be taken for adequate representation of each subpopulation identified by genetic markers. The information obtained by genomic methods should ultimately be translated into discrete product labeling information. Otherwise, it is uncertain whether the off-label data available in the form of scientific publications will transform routine clinical practice and lead to personalized therapeutics. Also, genomic data are fundamentally different than the traditional covariates (e.g., weight, age) that have been used to explain variability in drug efficacy and safety, and thus require special considerations. Clear regulatory guidelines and new collaborations between academic institutions and the pharmaceutical industry, both at the level of basic and clinical research, are called for to implement pharmacogenomics in the drug development process and evaluate its significance for achieving drug safety, efficacy and effectiveness (72 ,73 and 74).

Pharmacogenomics emerged in late 1990s by coalescence of traditional methodologies used in human genetics, common complex diseases and pharmacogenetics, together with the impetus provided by novel genomic technologies developed as part of the Human Genome Project. The collection of genomic data is being more feasible by increasing accessibility and decreasing costs of molecular genetic analyses. Pharmacogenomics has far-reaching implications in medicine and biology and can be applied to various facets of therapeutics from drug discovery and neuroimaging to drug–drug and drug–food interactions (5 ,6 ,75). The road from pharmacogenomics to personalized therapeutics is arduous and challenging but the technology is now in place to validate the utility of pharmacogenomics in routine clinical practice and pharmacotherapy (76).

ACKNOWLEDGMENTS

Part of "37 - Pharmacogenomics and Personalized Therapeutics in Psychiatry "

V. Özdemir is the recipient of a postdoctoral fellowship from the Ontario Mental Health Foundation and a NARSAD young investigator award. M. Masellis is the recipient of a research studentship from the Faculty of Medicine, University of Toronto. P. Muglia is supported by a postdoctoral fellowship from the Canadian Institutes of Health Research and the Schizophrenia Society of Canada. J.L. Kennedy is supported by grants-in-aid from the Canadian Institutes of Health Research and a NARSAD independent investigator award. The authors thank Professors Werner Kalow and Laszlo Endrenyi for many insightful discussions and continuing support and encouragement.

REFERENCES

1. Kalow W. *Pharmacogenetics: heredity and the response to drugs*. Philadelphia: WB Saunders, 1962.

2. Motulsky AG. Drug reactions, enzymes and biochemical genetics. *J Am Med Assn* 1957;165:835–837.

3. Grant DM. Pharmacogenomics and the changing face of clinical pharmacology. *Can J Clin Pharmacol* 1999;6:131–132.

4. Bailey DS, Bondar A, Furness LM. Pharmacogenomics: it’s not just pharmacogenetics. *Curr Opin Biotechnol* 1998;9:595–601.

5. Nebert DW. Pharmacogenetics and pharmacogenomics: why is this relevant to the clinical geneticist? *Clin Genet* 1999;56:247–258.

6. Patterson RE, Eaton DL, Potter JD. The genetic revolution: change and challenge for the dietetics profession. *J Am Diet Assoc* 1999;99:1412–1420.

7. Weber WW. Influence of heredity on human sensitivity to environmental chemicals. *Environ Mol Mutagen* 1995;25:102–114.

8. Lichter JB, Kurth JH. The impact of pharmacogenetics on the future of healthcare. *Curr Opin Biotechnol* 1997;8:692–695.

9. Sjöqvist F. The past, present and future of clinical pharmacology. *Eur J Clin Pharmacol* 1999;55:553–557.

10. Baldessarini RJ, Lipinski JF. Risks versus benefits of antipsychotic drugs. *N Engl J Med* 1973;289:427–428.

11. Brøsen K. Drug-metabolizing enzymes and therapeutic drug monitoring in psychiatry. *Ther Drug Monit* 1996;18:393–396.

12. Sheiner LB. Learning versus confirming in clinical drug development. *Clin Pharmacol Ther* 1997;61:275–291.

13. Weber WW. *Pharmacogenetics*. New York: Oxford University Press, 1997.

14. Snyder LH. Studies in human inheritance. IX. The inheritance of taste deficiency in man. *Ohio J Sci* 1932;32:436–468.

15. Reidenberg MM. Clinical pharmacology: the scientific basis of therapeutics. *Clin Pharmacol Ther* 1999;66:2–8.

16. Vesell ES. Reflections from distant cuvettes. *Drug Metab Rev* 1996;28:493–511.

17. Reidenberg MM. Individualization of drug therapy. *Med Clin N Am* 1974;58:905–1162.

18. Meltzer HY. Genetics and etiology of schizophrenia and bipolar disorder. *Biol Psychiatry* 2000;47:171–173.

19. Kennedy JL. Schizophrenia genetics: the quest for an anchor. *Am J Psychiatry* 1996;153:1513–1514.

20. Koch-Weser J. Serum drug concentrations as therapeutic guides. *N Engl J Med* 1972;287:227–231.

21. Holford NH, Sheiner LB. Understanding the dose-effect relationship: clinical application of pharmacokinetic-pharmacodynamic models. *Clin Pharmacokinet* 1981;6:429–453.

22. Greenblatt DJ, Harmatz JS. Kinetic-dynamic modeling in clinical psychopharmacology. *J Clin Psychopharmacol* 1993;13:231–234.

23. Levy G. Predicting effective drug concentrations for individual patients. Determinants of pharmacodynamic variability. *Clin Pharmacokinet* 1998;34:323–333.

24. Burke MJ, Preskorn SH. Therapeutic drug monitoring of antidepressants: cost implications and relevance to clinical practice. *Clin Pharmacokinet* 1999;37:147–165.

25. Rowland M, Tozer TN. *Clinical pharmacokinetics. Concepts and applications*, third ed. Baltimore: Williams & Wilkins, 1995.

26. Halushka MK, Fan JB, Bentley K, et al. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat Genet* 1999;22:239–247.

27. Friedrich V, Motulsky AG. *Human genetics: problems and approaches*, second ed. New York: Springer-Verlag, 1986.

28. Jackson PR, Boobis AR, Tucker GT. Phenotype or genotype? *Br J Clin Pharmacol* 1991;31:119–120.

29. Harris H. *Principles of human biochemical genetics*, third ed. New York: Elsevier/North Holland Biomedical, 1980.

30. Eaton DL, Bammler TK. Concise review of the glutathione S-transferases and their significance to toxicology. *Toxicol Sci* 1999;49:156–164.

31. Morgan ET, Sewer MB, Iber H, et al. Physiological and pathophysiological regulation of cytochrome P450. *Drug Metab Dispos* 1998;26:1232–1240.

32. Bertz RJ, Granneman GR. Use of in vitro and in vivo data to estimate the likelihood of metabolic pharmacokinetic interactions. *Clin Pharmacokinet* 1997;32:210–258.

33. Tyndale RF, Li Y, Li NY, et al. Characterization of cytochrome P-450 2D1 activity in rat brain: high-affinity kinetics for dextromethorphan. *Drug Metab Dispos* 1999;27:924–930.

34. Farin FM, Omiecinski CJ. Regiospecific expression of cytochrome P-450s and microsomal epoxide hydrolase in human brain tissue. *J Toxicol Environ Health* 1993;40:317–335.

35. Bertilsson L. Geographical/interracial differences in polymorphic drug oxidation. Current state of knowledge of cytochromes P450 (CYP) 2D6 and 2C19. *Clin Pharmacokinet* 1995;29:192–209.

36. Kalow W. Interethnic variation of drug metabolism. *Trends Pharmacol Sci* 1991;12:102–107.

37. Ingelman-Sundberg M, Oscarson M, McLellan RA. Polymorphic human cytochrome P450 enzymes: an opportunity for individualized drug treatment. *Trends Pharmacol Sci* 1999;20:342–349.

38. Aklillu E, Persson I, Bertilsson L, et al. Frequent distribution of ultrarapid metabolizers of debrisoquine in an Ethiopian population carrying duplicated and multiduplicated functional CYP2D6 alleles. *J Pharmacol Exp Ther* 1996;278:441–446.

39. Dalen P, Dahl ML, Ruiz ML, et al. 10-Hydroxylation of nortriptyline in white persons with 0, 1, 2, 3, and 13 functional CYP2D6 genes. *Clin Pharmacol Ther* 1998;63:444–452.

40. Bertilsson L, Åberg-Wistedt A, Gustafsson LL, et al. Extremely rapid hydroxylation of debrisoquine: A case report with implication for treatment with nortriptyline and other tricyclic antidepressants. *Ther Drug Monit* 1985;7:478–480.

41. Dalen P, Dahl M, Andersson K, et al. Inhibition of debrisoquine hydroxylation with quinidine in subjects with three or more functional CYP2D6 genes. *Br J Clin Pharmacol* 2000;49:180–184.

42. Llerena A, Edman G, Cobaleda J, et al. Relationship between personality and debrisoquine hydroxylation capacity. Suggestion of an endogenous neuroactive substrate or product of the cytochrome P4502D6. *Acta Psychiatr Scand* 1993;87:23–28.

43. Wilkinson GR. Cytochrome P4503A (CYP3A) metabolism: prediction of in vivo activity in humans. *J Pharmacokinet Biopharmacol* 1996;24:475–490.

44. Dresser GK, Spence JD, Bailey DG. Pharmacokinetic-pharmacodynamic consequences and clinical relevance of cytochrome P450 3A4 inhibition. *Clin Pharmacokinet* 2000;38:41–57.

45. Penno MB, Dvorchik BH, Vesell ES. Genetic variation in rates of antipyrine metabolite formation: A study in uninduced twins. *Proc Natl Acad Sci USA* 1981;78:5193–5196.

46. Sata F, Sapone A, Elizondo G, et al. CYP3A4 allelic variants with amino acid substitutions in exons 7 and 12: evidence for an allelic variant with altered catalytic activity. *Clin Pharmacol Ther* 2000;67:48–56.

47. Özdemir V, Kalow W, Tang BK, et al. Evaluation of the genetic contribution to CYP3A4 activity in vivo: a repeated drug administration method. *Pharmacogenetics* 2000:10:373–388.

48. Weber WW. Populations and genetic polymorphisms. *Mol Diagn* 1999;4:299–307.

49. Propping P, Nothen MM. Genetic variation of CNS receptors—a new perspective for pharmacogenetics. *Pharmacogenetics* 1995;5:318–325.

50. Basile VS, Masellis M, Badri F, et al. Association of the MscI polymorphism of the dopamine D3 receptor gene with tardive dyskinesia in schizophrenia. *Neuropsychopharmacology* 1999;21:17–27.

51. Segman R, Neeman T, Heresco-Levy U, et al. Genotypic association between the dopamine D3 receptor and tardive dyskinesia in chronic schizophrenia. *Mol Psychiatry* 1999;4:247–253.

52. Steen VM, Lovlie R, MacEwan T, et al. Dopamine D3-receptor gene variant and susceptibility to tardive dyskinesia in schizophrenic patients. *Mol Psychiatry* 1997;2:139–145.

53. Masellis M, Basile VS, Ozdemir V, et al. Pharmacogenetics of antipsychotic treatment: lessons learned from clozapine. *Biol Psychiatry* 2000;47:252–266.

54. Masellis M, Basile V, Meltzer HY, et al. Serotonin subtype 2 receptor genes and clinical response to clozapine in schizophrenia patients. *Neuropsychopharmacology* 1998;19:123–132.

55. Masellis M, Paterson AD, Badri F, et al. Genetic variation of 5-HT2A receptor and response to clozapine. *Lancet* 1995;346:1108.

56. Winsberg BG, Comings DE. Association of the dopamine transporter gene *(DAT1)* with poor methylphenidate response. *J Am Acad Child Adolesc Psychiatry* 1999;38:1474–1477.

57. Catalano M. The challenges of psychopharmacogenetics. *Am J Hum Genet* 1999;65:606–610.

58. Uhr M, Steckler T, Yassouridis A, et al. Penetration of amitriptyline, but not of fluoxetine, into brain is enhanced in mice with blood-brain barrier deficiency due to Mdr1a p-glycoprotein gene disruption. *Neuropsychopharmacology* 2000;22:380–387.

59. Feltus MS, Gardner DM. Second generation antipsychotics for schizophrenia. *Can J Clin Pharmacol* 1999;6:187–195.

60. Kleyn PW, Vesell ES. Genetic variation as a guide to drug development. *Science* 1998;281:1820–1821.

61. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994;265:2037–2048.

62. Hacia JG, Brody LC, Collins FS. Applications of DNA chips for genomic analysis. *Mol Psychiatry* 1998;3:483–492.

63. Debouck C, Goodfellow PN. DNA microarrays in drug discovery and development. *Nat Genet* 1999;21:48–50.

64. Drews J. Genomic sciences and the medicine of tomorrow. *Nat Biotechnol* 1996;14:1516–1518.

65. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: meta-analysis of prospective studies. *JAMA* 1998;279:1200–1205.

66. Alfaro CL, Lam YW, Simpson J, et al. CYP2D6 inhibition by fluoxetine, paroxetine, sertraline, and venlafaxine in a crossover study: intraindividual variability and plasma concentration correlations. *J Clin Pharmacol* 2000;40:58–66.

67. Kashuba AD, Nafziger AN. Physiological changes during the menstrual cycle and their effects on the pharmacokinetics and pharmacodynamics of drugs. *Clin Pharmacokinet* 1998;34:203–218.

68. Albers LJ, Reist C, Helmeste D, et al. Paroxetine shifts imipramine metabolism. *Psychiatry Res* 1996;59:189–196.

69. Lin KM, Poland RE, Wan YJ, et al. The evolving science of pharmacogenetics: clinical and ethnic perspectives. *Psychopharmacol Bull* 1996;32:205–217.

70. Lin KM, Anderson D, Poland RE. Ethnicity and psychopharmacology. Bridging the gap. *Psychiatr Clin North Am* 1995;18:635–647.

71. Vesell ES. On the significance of host factors that affect drug disposition. *Clin Pharmacol Ther* 1982;31:1–7.

72. Hodgson J, Marshall A. Pharmacogenomics: will the regulators approve? *Nat Biotechnol* 1998;16:243–246.

73. Lebowitz BD, Rudorfer MV. Treatment research at the millennium: from efficacy to effectiveness. *J Clin Psychopharmacol* 1998;18:1.

74. Jobe PC, Adams-Curtis LE, Burks TF, et al. The essential role of integrative biomedical sciences in protecting and contributing to the health and well-being of our nation. *Physiologist* 1994;37:79–86.

75. Venkatakrishnan K, von Moltke LL, Greenblatt DJ. Effects of the antifungal agents on oxidative drug metabolism. Clinical relevance. *Clin Pharmacokinet* 2000;38:111–180.

76. Brockmoller J. Pharmacogenomics—science fiction come true. *Int J Clin Pharmacol Ther* 1999;37:317–318.

38

Pharmacokinetics, Pharmacodynamics, and Drug Disposition

David J. Greenblatt

Lisa L. Von Moltke

Jerold S. Harmatz

Richard I. Shader

D. J. Greenblatt, L. L. von Moltke, J. S. Harmatz, and R. I. Shader: Department of Pharmacology and Experimental Therapeutics, Tufts University School of Medicine, and Division of Clinical Pharmacology, New England Medical Center, Boston, Massachusetts.

During the last decade, the application of pharmacokinetic and pharmacodynamic modeling techniques has become an increasingly important aspect of contemporary clinical psychopharmacology (1 ,2 ,3 ,4 and 5). These techniques have been applied during the process of development of new drug entities as well as for the improved understanding of the clinical actions of drugs that are already marketed. Techniques for the study of drug metabolism *in vitro* have advanced substantially during the last decade, and now are an integral component of preclinical drug development and the link to subsequent clinical studies of drug metabolism and disposition. Kinetic-dynamic modeling techniques have been combined with *in vitro* metabolism procedures and *in vitro–in vivo* mathematical scaling models to provide insight into the general problem of pharmacokinetic drug interactions in clinical psychopharmacology (6 ,7 ,8 and 9).

This chapter reviews some advances in pharmacokinetics, pharmacodynamics, and drug metabolism, along with methodologic applications to selected problems in clinical psychopharmacology.

POPULATION PHARMACOKINETICS

Part of "38 - Pharmacokinetics, Pharmacodynamics, and Drug Disposition "

Principles

Pharmacokinetic studies based on a traditional intensive-design model are usually conducted using carefully selected volunteer subjects, a controlled experimental design, and collection of multiple blood samples. After measurement of drug and metabolite concentrations in all samples, pharmacokinetic models are applied to determine parameters such as elimination half-life, volume of distribution, and clearance. During the new drug development process, a series of pharmacokinetic studies are conducted to determine the influence of major disease states or experimental conditions hypothesized to affect drug disposition. Such factors might include age, gender, body weight, ethnicity, hepatic and renal disease, coadministration of food, and various drug interactions. Classical pharmacokinetic studies can quantitate the effects of anticipated influences on drug disposition under controlled circumstances, but cannot identify the unexpected factors affecting pharmacokinetics. A number of examples of altered drug pharmacokinetics became apparent in the patient care setting only in the postmarketing phase of extensive clinical use. Examples include the digoxin-quinidine interaction, altered drug metabolism due to cimetidine, and the ketoconazole-terfenadine interaction.

Population pharmacokinetic methodology has developed as an approach to detect and quantify unexpected influences on drug pharmacokinetics (10 ,11 ,12 ,13 ,14 ,15 ,16 ,17 and 18). Population pharmacokinetic studies, in contrast to classical or traditional pharmacokinetic studies, focus on the central tendency of a pharmacokinetic parameter across an entire population, and identify deviations from that central tendency in a subgroup of individual patients. One software program widely applied to population pharmacokinetic problems is the nonlinear mixed-effects model (NONMEM). Analysis of clinical data using a population approach allows pharmacokinetic parameters to be determined directly in patient populations of interest and allows evaluation of the influence of various patient characteristics on pharmacokinetics. Because the number of blood samples that need to be collected per subject is small, this approach is often suitable for patient groups unable to participate in traditional pharmacokinetic studies requiring multiple blood samples (e.g., neonates,

children, critically ill patients, or individuals who are not able to provide informed consent) (19). In many cases the population approach has yielded pharmacokinetic parameter estimates similar to those delineated in classical pharmacokinetic studies of the same drug.

Application: Methylphenidate Pharmacokinetics

The population approach is illustrated in a study of methylphenidate (MP) pharmacokinetics in children (20). This is a patient group for whom the multiple-sample pharmacokinetic study design may not be appropriate for ethical and practical reasons. Participating subjects were 273 children aged 5 to 18 years having a primary diagnosis of attention-deficit/hyperactivity disorder (ADHD). They had been receiving MP at a fixed dosage level for at least 4 weeks, and were under treatment for at least 3 months. The treating physician for each patient judged MP to be clinically effective.

Children meeting the eligibility criteria had an initial screening visit, at which one parent or a legal guardian provided written informed consent, and the child provided assent. Demographic characteristics were recorded, including the dosage of MP, the usual times for individual doses, and the duration of treatment.

The second visit, which followed shortly, was a blood-sampling day. Each child, accompanied by parent or guardian, arrived at the investigator's office 30 to 60 minutes prior to blood sampling. The time and size of the last MP dose, and of any other medication received that day or during the prior 2 weeks, were recorded. A 5-mL whole blood sample was obtained by venipuncture. This sample was immediately centrifuged, and a 2-mL aliquot of plasma was removed for subsequent determination of MP concentrations by a liquid chromatography/mass spectroscopy/mass spectroscopy (LC/MS/MS) assay.

Analysis of Data

The identified independent variables were age, sex, body weight, size of each dose, and time of sample relative to the most recent dose. Since only single samples were available for all but 16 of these children, the contribution of within-subject variance to overall variability in outcome could not be assessed. The pharmacokinetic model was a one-compartment model with first-order absorption and first-order elimination, under the assumption that all subjects were at steady state (Fig. 38.1).

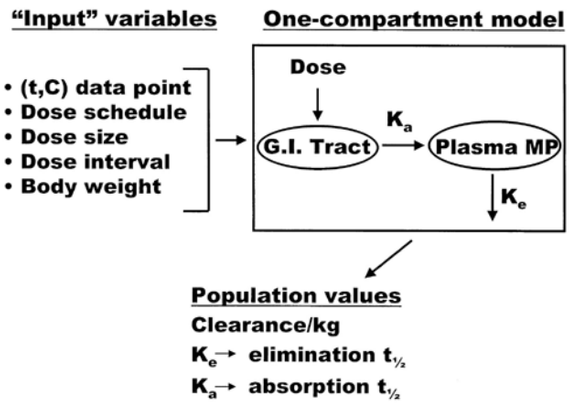


FIGURE 38.1. Population pharmacokinetic model for methylphenidate (MP). A series of data points, each consisting of the time (*t*) after the first dose of the day and the plasma MP concentration (*C*) at that time, was available from 273 subjects (one data point per subject). Each of these was linked to that subject's individual dose schedule, size of each dose, interval between doses, and body weight. These variables were entered into a one-compartment pharmacokinetic model with first-order absorption and first-order elimination, as shown. Using nonlinear regression, the process yielded "typical" population values of clearance per kilogram body weight, the elimination rate constant (K_e), and the absorption rate constant (K_a).

The overall model was specifically modified for each of the 273 subjects to incorporate the individually applicable independent variables, as well as the dosage schedule (b.i.d. or t.i.d.). Individual values of continuous variables (*t* = time sample taken relative to the first dose; *C* = plasma MP concentration) were fitted to a single set of iterated variables using unweighted nonlinear regression (Fig. 38.1). When the time between first and second doses, or between second and third doses, was not available, the mean value was assigned based on cases in which the data were available. For the b.i.d. dosage, the mean interval was 4.3 hours. For the t.i.d. dosage, the mean intervals were 4.1 and 3.7 hours, respectively. As is customary, clearance was assumed to be proportional to body weight.

Results

The total daily dose of MP was significantly lower in subjects receiving MP b.i.d. (n = 109) compared to subjects on a t.i.d. schedule (n = 164); the mean total daily dosages in the two groups were 25 and 39.3 mg, respectively (p <.001). Within each group, clinicians' choices of total daily dosages were influenced by body weight, as mean total daily dose increased significantly with higher body weights. However, the association of body weight with mean plasma concentration was not significant for the b.i.d. dosage group, and of only borderline significance (.05< p <.1) for the t.i.d group. This finding is consistent with the underlying assumption that clearance is proportional to body weight.

Age was significantly correlated with body weight (r^2 = 0.54, p <.001) and with height (r^2 = 0.77). Height and body weight also were significantly correlated (r^2 = 0.77).

An acceptable estimate of absorption rate constant could be derived only for the b.i.d. dosing data. The iterated parameter

estimate was 1.192/h, corresponding to an absorption half-life of 34.9 minutes. This estimate was then fixed, and the entire data set analyzed to determine clearance per kilogram of body weight, and the first-order elimination rate constant. The iterated estimates were 0.154/h for elimination rate constant, corresponding to an elimination half-life of 4.5 hours (relative standard error: 23%). For clearance, the estimate was 90.7 mL/min/kg (relative standard error: 9%). The overall *r*-square was 0.43 (Fig. 38.2). There were no evident differences in pharmacokinetics attributable to gender. Figure 38.3 shows predicted plasma MP concentration curves for b.i.d. and t.i.d. dosage schedules, based on the population estimates.

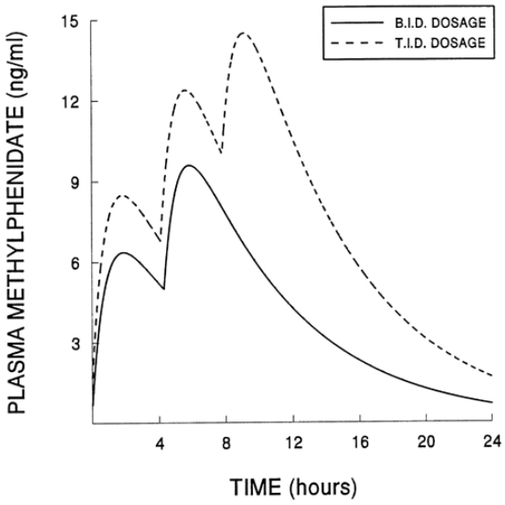


FIGURE 38.3. Predicted plasma methylphenidate concentration curves for b.i.d. and t.i.d. dosage schedules, based on parameter estimates from the population analysis, together with mean values of input variables (body weight, size of doses, intervals between doses). (From Shader RI, Harmatz JS, Oesterheld JR, et al. Population pharmacokinetics of methylphenidate in children with attention-deficit hyperactivity disorder. *J Clin Pharmacol* 1999;39:775–785, with permission).

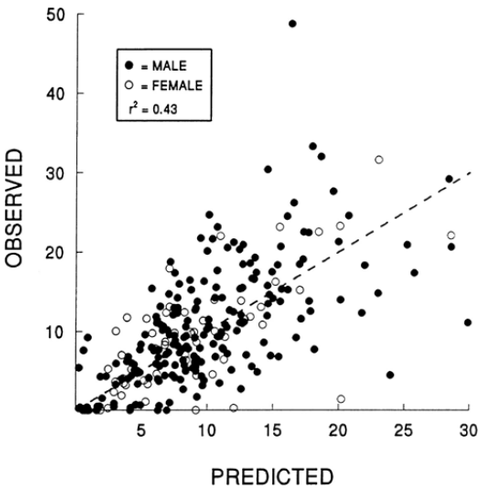


FIGURE 38.2. Overall relation of observed and predicted plasma methylphenidate concentrations (ng/ml). The *r*-square value of 0.43 indicates that the model accounts for 43% of the overall variance in plasma concentrations. (From Shader RI, Harmatz JS, Oesterheld JR, et al. Population pharmacokinetics of methylphenidate in children with attention-deficit hyperactivity disorder. *J Clin Pharmacol* 1999;39:775–785, with permission.)

Implications

Pharmacokinetically based approaches to the treatment of ADHD with MP are not clearly established (21 ,22 ,23 ,24 and 25). In the present study of prescribing patterns in particular clinical practices, the mean prescribed per dose amount for the whole study population was 0.335 mg/kg per dose (range = 0.044–0.568), and 36% of the children received between 0.25 and 0.35 mg/kg per dose. The mean total daily dose was 0.98 mg/kg/day for the entire sample, and increased significantly in association with larger body weight. This may reflect the clinicians' considering body weight in their choice of total daily dosage, or it may be that the dose was titrated according to response, which in turn was influenced by associations among concentration, clearance, and weight.

The pharmacokinetic model explained 43% of the variability in plasma MP concentrations during typical naturalistic therapy. The model fit equally well for both genders. Assuming that clearance is proportional to body weight in the context of intercorrelated age and weight allows age, weight, and daily dosage to be used to predict plasma concentrations of MP during clinical use in children. These findings support the value of prescribing MP on a weight-adjusted basis.

Our typical population value of elimination half-life was 4.5 hours, with a confidence interval of 3.1 to 8.1 hours. This estimate somewhat exceeds the usual range of half-life values reported in single-dose kinetic studies of MP (25 ,26). This could reflect the relatively small number of plasma samples from the terminal phase of the plasma concentration curve, upon which reliable estimates of beta are dependent. MP kinetics may also have a previously unrecognized dose-dependent component, in which estimated values of half-life are larger at steady state than following a single dose.

The single-sample approach described in this study allows relatively noninvasive assessment of pharmacokinetic parameters in a group of children and adolescents under naturalistic circumstances of usual clinical use, when blood sampling is not otherwise clinically indicated. This approach in general can be applied to other special populations such as neonates, the elderly, or individuals with serious medical disease.

KINETIC-DYNAMIC MODELING

Part of "38 - Pharmacokinetics, Pharmacodynamics, and Drug Disposition "

Principles

Pharmacokinetics is the discipline that applies mathematical models to describe and predict the time course of drug concentrations in body fluids, whereas pharmacodynamics refers to the time course and intensity of drug effects on the organism, whether human or experimental animal (Fig. 38.4). Both have evolved as the techniques for measuring drug concentrations, and drug effects have become more accurate and sensitive. Evolving in parallel is kinetic-dynamic modeling, in which the variable of time is incorporated into the relationship of effect to concentration (Fig. 38.4) (27 ,28 ,29 ,30 ,31 and 32). A concentration-effect relationship is, in principle, the most clinically relevant, because it potentially validates the clinical rationale for measuring drug concentrations in serum or plasma.

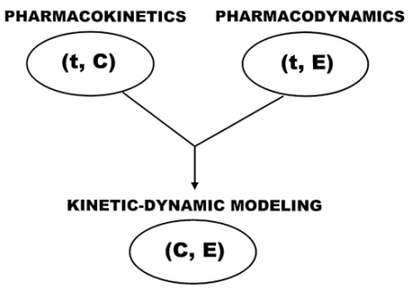


FIGURE 38.4. Schematic relation between pharmacokinetics, pharmacodynamics, and kinetic-dynamic modeling, based on the status of the variables of time (*t*), concentration (*C*), and effect (*E*). Note that kinetic-dynamic modeling incorporates both pharmacokinetics and pharmacodynamics, with time subsumed into the relation of concentration and effect.

A kinetic-dynamic study in clinical psychopharmacology typically involves medication administration (usually under placebo-controlled, double-blind laboratory conditions) followed by quantitation of both drug concentration and clinical effect at multiple times after dosing. Measures of effect necessarily depend on the type of drug under study. For sedative-anxiolytic drugs such as benzodiazepines, effects of interest may include subjective or observer ratings of sedation and mood; semiobjective measures of psychomotor performance, reaction time, or memory; or objective effect measures such as the EEG or saccadic eye movement velocity. The various measures differ substantially in their relevance to the principal therapeutic actions of the drug, the stability of the measure in terms of response to placebo or changes caused by practice or adaptation, the objective or subjective nature of the quantitative assessment, and the comparability of results across different investigators and different laboratories (Table 38.1). The extent to which the various pharmacodynamic measures provide unique information, as opposed to being overlapping or redundant, is not clearly established.

Classification (with Examples)	Relation to Primary Therapeutic Action	Effect of Placebo	Effect of Adaptation/Practice	Need for "Blind" Conditions	Approach to Quantitation
Subjective					
Global clinical ratings; targeted rating scales	Close	Yes	Yes	Yes	Transformation of ratings into numbers
Semi-objective					
Psychomotor function tests; memory tests	May be linked to adverse effect profile	Yes	Yes	Yes	Test outcomes are quantitative
Objective					
Electroencephalography	Not established	No	No	No	Fully objective computer-determined quantitation

GABA, γ-aminobutyric acid.

TABLE 38.1. PHARMACODYNAMIC ENDPOINTS APPLICABLE TO STUDIES OF GABA-BENZODIAZEPINE AGONISTS

Pharmacokinetic and pharmacodynamic relationships initially are evaluated separately, and the relationship of effect versus concentration at corresponding times is examined graphically and mathematically. Effect measures are usually expressed as change scores: the net effect (*E*) at postdosage time *t* is calculated as the absolute effect at this time (E_t) minus the predose baseline value (E_o), that is, $E = E_t - E_o$. Several mathematical relationships between effect and concentration (*E* versus *C*), often termed "link" models, are of theoretical and practical importance (5 ,32). The "sigmoid E_{max}" model, incorporates a value of E_{max}, the maximum pharmacodynamic effect, and EC_{50} is the "50% effective concentration," the concentration that is associated with half of the maximum effect (Fig. 38.5). The exponent *A* reflects the "steepness" of the concentration-response relationship in its ascending portion. The biological importance of *A* is not established.

Sigmoid E_{max}

$$E = \frac{E_{max} \cdot C^A}{C^A + EC_{50}{}^A}$$

Exponential

$$E = m \cdot C^A$$

Linear

$$E = m \cdot C$$

FIGURE 38.5. Three mathematical relationships between concentration (*C*) and change in pharmacodynamic effect (*E*) that are commonly applied in kinetic-dynamic modeling procedures. For the sigmoid E_{max} model, E_{max} is maximum pharmacodynamic effect, EC_{50} is the concentration producing a value of *E* equal to 50% of E_{max}, and *A* is an exponent. For the exponential and linear models, *m* is a slope factor.

A concentration-effect relationship that is consistent with the sigmoid E_{max} model may be of mechanistic importance, because drug-receptor interactions often fit the same model. The E_{max} and EC_{50} values allow inferences about questions such as the relative potency or efficacy of drugs producing the same clinical effect, individual differences in drug sensitivity, the mechanism of action of pharmacologic potentiators or antagonists, and the possible clinical role of new medications.

The sigmoid E_{max} model does not necessarily apply to all concentration-effect data (32). When experimental data are not consistent with the model, the corresponding misapplication of the sigmoid E_{max} relationship can lead to misleading conclusions about E_{max} and EC_{50}. Some data sets are consistent with less complex models, such as exponential or linear equations (Fig. 38.5); in these cases, the concepts of E_{max} and EC_{50} are not applicable. Kinetic-dynamic modeling is further complicated when drug concentrations measured in serum or plasma do not reflect the concentration at the site of action, which is sometimes termed the "effect site." This is illustrated by the data described below.

Application: Kinetics And Dynamics Of Intravenous Lorazepam

In this study the benzodiazepine derivative lorazepam was administered intravenously according to a complex bolus-infusion scheme (33). On the morning of the study day, a rapid intravenous dose of lorazepam, 2 mg, was administered into an antecubital vein, coincident with the start of a zero-order infusion at a rate of 2 µg/kg/h. The infusion continued for 4 hours and then was terminated. Venous blood samples were drawn from the arm contralateral to the site of the infusion prior to drug administration and at multiple time points during 24 hours after the start of lorazepam infusion. Samples were centrifuged, and the plasma separated and frozen until the time of assay. The EEG was used as the principal pharmacodynamic outcome measure (Table 38.1). The EEG was recorded prior to lorazepam administration, and at times corresponding to blood samples. EEG data were digitized over the power spectrum from 4 to 30 cycles per second (Hz), and analyzed by fast Fourier transform to determine amplitude in the total spectrum (4 to 30 Hz) and in the beta (12 to 30 Hz) frequency range (33 ,34 and 35). Concentrations of lorazepam in plasma samples were determined by gas-chromatography with electron-capture detection.

Analysis of Data

The relative EEG beta amplitudes (beta divided by total, expressed as percent) in the predose recordings were used as the baseline. All values after lorazepam administration were expressed as the increment or decrement over the mean predose baseline value, with values averaged across eight recording sites. The EEG change values were subsequently used as pharmacodynamic effect (*E*) measures in kinetic-dynamic modeling procedures described below. For pharmacokinetic modeling, the relation of plasma lorazepam concentration (*C*) to time (*t*) was assumed to be consistent with a two-compartment model (Fig. 38.6 and Fig. 38.7).

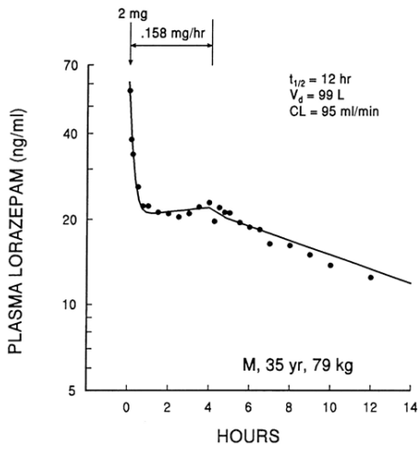


FIGURE 38.7. Plasma lorazepam concentrations (*solid circles*) together with the pharmacokinetic function determined by nonlinear regression (*solid line*), in a representative volunteer subject. Shown are the derived pharmacokinetic variables of elimination half-life ($t_{1/2}$), volume of distribution (V_d), and clearance (CL).

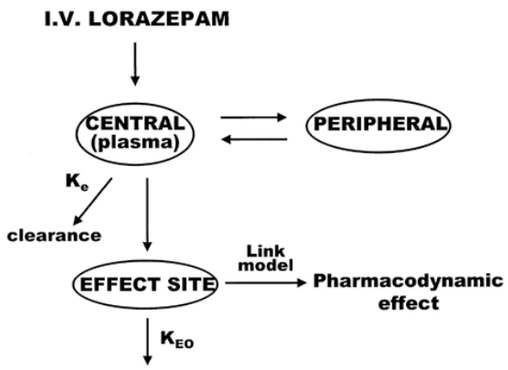


FIGURE 38.6. Schematic representation of the kinetic-dynamic model for the lorazepam study. Intravenous lorazepam was assumed to have kinetic behavior consistent with a two-compartment model: reversible distribution to a peripheral compartment, and first-order elimination (clearance) from the central compartment (rate constant: K_e). Lorazepam in plasma was postulated to equilibrate with a hypothetical effect site, from which the exit rate constant is K_{EO}. Finally, effect-site concentrations were presumed to be the principal determinant of pharmacodynamic effect, via a kinetic-dynamic link model as shown in Fig. 38.5.

Examination of plots of pharmacodynamic EEG effect versus plasma lorazepam concentration (*E* vs. *C*) indicated counterclockwise hysteresis (see below), suggesting a delay in equilibration of lorazepam between plasma and the site of pharmacodynamic action in brain. This equilibration effect has been described in previous clinical and experimental studies of lorazepam (34 ,36 ,37 ,38 and 39). Accordingly the relationship was modified to incorporate a distinct “effect site,” at which the hypothetical lorazepam concentration is C_E (Fig. 38.6). The apparent rate constant for drug disappearance from the effect compartment is k_{EO}; this rate constant determines the apparent half-life of drug equilibration between

plasma and effect site. Under these assumptions, the relation of E to C_E was postulated to be consistent with a sigmoid E_{max} model (Fig. 38.5).

Results

Kinetic variables for lorazepam were similar to those reported in previous single-dose studies of lorazepam pharmacokinetics (34 ,40 ,41 ,42 ,43 ,44 and 45). Overall mean values were volume of distribution, 1.7 L/kg; elimination half-life, 14 hours; clearance, 1.44 mL/min/kg. The bolus-infusion scheme rapidly produced mean plasma lorazepam concentrations in the range of 18 to 19 ng/mL, values close to the mean predicted value of 24 ng/mL.

Lorazepam infusion produced significant increases in EEG beta amplitude throughout the 24-hour duration of the study. The maximum change over baseline was measured at 0.25 to 0.75 hours after the initiation of lorazepam dosage, whereas the maximum plasma concentration was measured immediately after the loading dose (Fig. 38.8). The effect-site model eliminated the hysteresis, with a mean equilibration half-life of 8.8 minutes (Fig. 38.9).

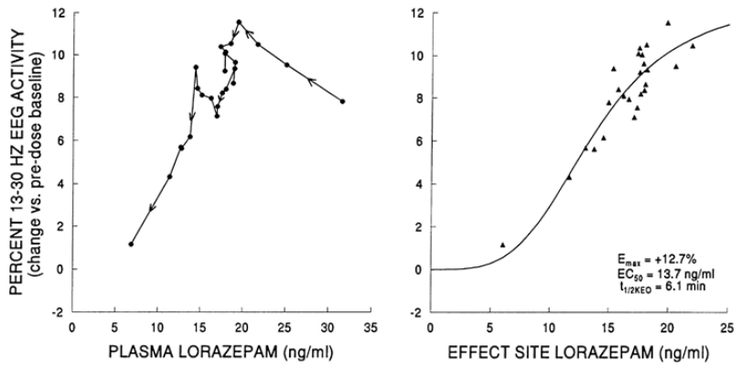


FIGURE 38.9. Left: Mean values of plasma lorazepam concentration versus pharmacodynamic EEG effect at corresponding times, with *arrows* indicating the direction of increasing time. As indicated in Fig. 38.8, the maximum EEG effect is delayed, and does not correspond in time to the maximum plasma concentration. *Right*: The scheme shown in Fig. 38.6 was applied to the data points, with the link model being the sigmoid E_{max} relationship shown in Fig. 38.5. The data points (*closed triangles*) are the hypothetical effect site concentrations and pharmacodynamic effect values at corresponding times. The *solid line* is the link model function determined by nonlinear regression, yielding the indicated values of E_{max} and EC_{50}. The overall mean equilibration half-life was 8.8 minutes.

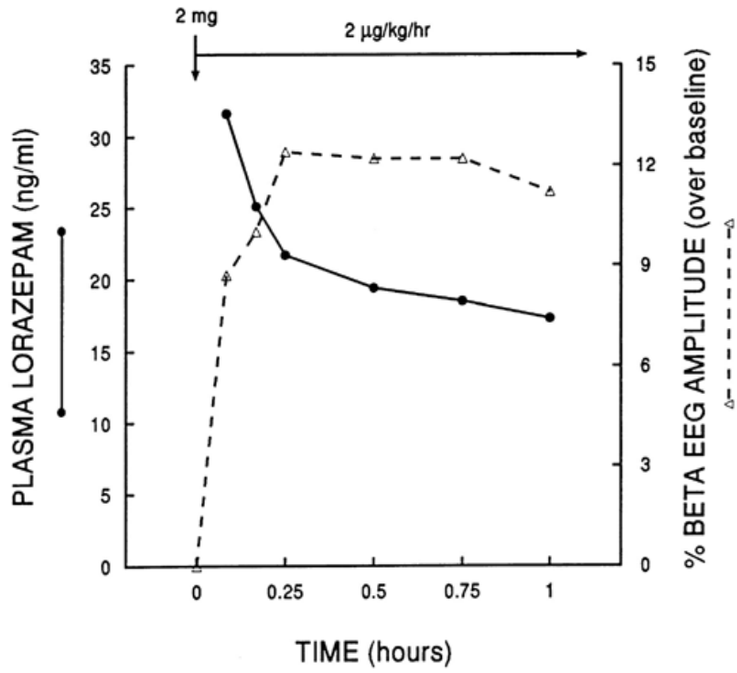


FIGURE 38.8. Mean values of plasma lorazepam concentration, and of EEG beta amplitude, during the first hour of the study. Note that pharmacodynamic EEG effects are delayed following the peak value of lorazepam in plasma.

Implications

Maximum EEG effects of lorazepam were significantly delayed following the initial intravenous bolus dose. Previous single-dose pharmacodynamic studies of lorazepam, using the EEG or other methods for quantitation of benzodiazepine

effect, consistently demonstrate a delay in attainment of maximum drug effect compared to attainment of peak concentrations in plasma (34 ,36 ,37). After rapid intravenous dosage, for example, maximum effects may be delayed for an average of 30 minutes after dosage. Experimental studies of the time-course of whole-brain concentrations of lorazepam, or of the degree of benzodiazepine receptor occupancy, indicate that the delay is attributable to the slow physical entry of lorazepam into brain tissue, probably because of the relatively low lipid solubility of lorazepam (34 ,38 ,39). The delay was mathematically consistent with a kinetic-dynamic model incorporating a hypothetical “effect site” distinct from the central compartment. The half-life for equilibration between plasma and the effect compartment was approximately 9 minutes. This matches clinical experience indicating that intravenous lorazepam cannot easily be used in situations requiring minute-to-minute titration of sedative or amnestic effects (40). Nonetheless, intravenous lorazepam can be used for the treatment of status epilepticus, although its onset of action may be somewhat slower than that of intravenous diazepam (46 ,47).

CYTOCHROMES P-450 IN PSYCHOPHARMACOLOGY: THE IMPORTANCE OF P-450-3A ISOFORMS

Part of "38 - Pharmacokinetics, Pharmacodynamics, and Drug Disposition "

The cytochrome P-450 (CYP) superfamily of drug metabolizing enzymes is now established as being of primary importance for the metabolism and clearance of most drugs used in psychopharmacology and in other areas of clinical therapeutics (6 ,7 ,8 and 9 ,48 ,49 ,50 ,51 ,52 ,53 ,54 and 55) (Fig. 38.10). For the CYP isoforms most relevant to human drug metabolism, each has its own distinct pattern of relative abundance, anatomic location, mechanism of regulation, substrate specificity, and susceptibility to inhibition and induction by other drugs or foreign chemicals (Table 38.2). The expression and *in vivo* function of at least two CYP isoforms (CYP2D6 and CYP2C19) are regulated by a genetic polymorphism, such that some members of a population fail to express “normal” levels of enzyme or expresses poorly functional protein (56 ,57 ,58 ,59 ,60 ,61 and 62). Individuals identified as “CYP2D6 poor metabolizers,” as an example, have very low clearance of drugs that are major substrates for biotransformation by CYP2D6 (such as desipramine, nortriptyline, venlafaxine, tramadol, and dextromethorphan). Such individuals are at risk for developing high and potentially toxic plasma concentrations of these

substrate drugs despite dosages in the usual therapeutic range.

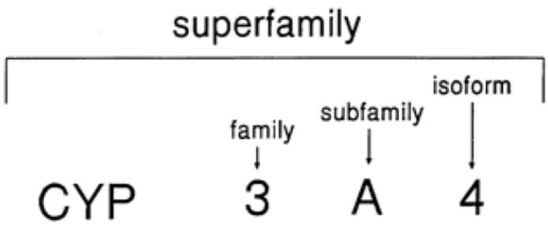


FIGURE 38.10. Nomenclature system for the cytochrome P-450 (CYP) superfamily of enzymes. Following the CYP designation, the number-letter-number sequence indicates the family, subfamily, and specific isoform.

CYP Isoform	Relative Hepatic Abundance	Genetic Polymorphism	Representative Substrates
1A2	13%	?	Caffeine, theophylline, tacrine
2B6	<1%	–	Bupropion, propofol
2C9	15%	+	S-warfarin, phenytoin, tolbutamide, NSAIDs
2C19	4%	+	S-mephenytoin, omeprazole (partial contributor to many others)
2D6	2%	+	Some psychotropic and cardiac drugs
2E1	7%	–	Chlorzoxazone, some inhaled anesthetics
3A4/5	29%[a]	–	Many (see also Table 38.3)

CYP, cytochrome P-450; NSAID, nonsteroidal antiinflammatory drug.
[a]Also present in gastrointestinal tract mucosa.

TABLE 38.2. OVERVIEW OF HUMAN CYTOCHROMES P-450

The CYP3A Isoforms

The overall importance of the CYP3A subfamily of drug-metabolizing enzymes, particularly in the field of psychopharmacology, has become increasingly evident over the last decade (6 ,7 ,8 and 9 ,63 ,64 ,65 ,66 ,67 ,68 and 69) (Table 38.3). The CYP3A isoforms are the most abundant of the CYPs, accounting on average for approximately 29% of identified cytochrome P-450 in liver (70) (Table 38.2). Within the CYP3A subfamily, CYP3A4 is the most important in the adult human, in terms of drug-metabolizing activity as well as quantitative dominance. CYP3A5, another CYP3A isoform, is also detected in varying amounts in some human livers and in esophagus, but quantities of CYP3A5 are less than quantities of CYP3A4. It is not established to what extent hepatic CYP3A5 is of clinical significance for drug-metabolizing activity. CYP3A7 is principally a fetal-specific isoform. The location and sequence of the genetic element responsible for CYP3A4 expression have been identified, as well as a regulatory segment located on the 5′ flanking region corresponding to the CYP3A gene.

Contribution of CYP3A to Net Clearance		
Complete or Nearly Complete	**Partial**	**Small**
Midazolam	Diazepam	Fluoxetine
Triazolam	Desmethyldiazepam	Sertraline
Alprazolam	Flunitrazepam	Nortriptyline
Bromazepam	Clonazepam	
Nefazodone	Zolpidem	
Trazodone	Citalopram	
Reboxetine	Haloperidol	
Buspirone	Clozapine	
Gepirone	Olanzapine	
Adinazolam	Mirtazapine	
Quetiapine	Amitriptyline	
Sildenafil	Imipramine	

TABLE 38.3. PSYCHOTROPIC DRUG SUBSTRATES FOR HUMAN CYP3A

CYP3A4 typically functions as a high-capacity, low-affinity enzyme. Its high substrate capacity is a consequence of both the relatively high value of maximum reaction velocity (V_{max}, expressed in nanomoles of product produced per unit time per milligram of protein) in a Michaelis-Menten relationship, as well as the high quantitative abundance of the protein in hepatic tissue. The low-affinity characteristic is reflected in the high K_m value (substrate concentration corresponding to 50% of V_{max}) in a Michaelis-Menten relationship. One consequence is that CYP3A-mediated metabolism usually is not “saturable” at substrate concentrations within the therapeutic range, because this range is likely to be far below the reaction K_m. Furthermore, in situations in which CYP3A is one of several cytochromes contributing to metabolism [e.g., amitriptyline *N*-demethylation (71), citalopram *N*-demethylation (72), or zolpidem hydroxylation (73)], the relative importance of CYP3A will increase at higher substrate concentrations. However, this is not invariably true. Nefazodone is a CYP3A substrate, but K_m values for production of the various metabolites are relatively low (74), and kinetics are nonlinear (75). Midazolam has a low K_m for the principal pathway (76 ,77), and there is evidence of nonlinear kinetics at higher concentrations in humans (78).

Significant quantities of CYP3A exist in gastrointestinal (GI) tract mucosa (65 ,69 ,79). The quantitative expression/

activity of GI tract CYP3A is not correlated with its expression/activity in liver, even though the expressed protein is identical at the two sites. For a number of moderate or high-clearance CYP3A substrates (e.g., midazolam and triazolam), GI tract metabolism contributes importantly to presystemic extraction (first-pass metabolism) after oral dosage (79 ,80 and 81); incomplete oral bioavailability therefore results from a combination of GI tract and hepatic presystemic extraction (Fig. 38.11). For low-clearance CYP3A substrates having oral bioavailability in the range of 80% to 90% or greater (e.g., alprazolam), the contribution of the GI tract is apparently small.

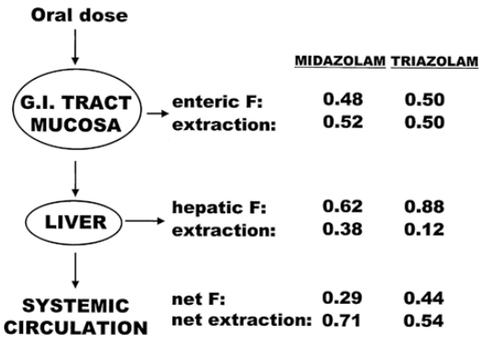


FIGURE 38.11. Relative contributions of CYP3A enzymes present in gastrointestinal (GI) tract mucosa, and in the liver, to net bioavailability (*F*) of orally administered midazolam and triazolam. Both of these compounds have net *F* values of less than 50% (29% for midazolam, 44% for triazolam). Both compounds undergo approximately 50% extraction during passage through the G.I. tract mucosa. However midazolam undergoes another 38% extraction across the liver, compared to only 12% for triazolam.

Inhibition and induction by other drugs or chemicals may modify CYP3A activity both *in vitro* and *in vivo* (Table 38.4). Identification of these compounds is of clear clinical importance, because it may allow anticipating of drug interactions that may be either potentially hazardous or of therapeutic benefit (6 ,7 ,8 and 9 ,65 ,69 ,79 ,80 ,81 ,82 ,83 ,84 and 85). Inhibiting drugs may also be used for investigating the relative contribution of CYP3A to net clearance, or for distinguishing the contribution of hepatic and GI tract CYP3A to overall presystemic extraction (81). Among the most potent CYP3A inhibitors are the azole antifungal agents (ketoconazole, itraconazole, fluconazole), the antidepressants nefazodone and fluvoxamine, and the calcium channel antagonists verapamil and diltiazem. These compounds produce “reversible” inhibition, by a competitive, noncompetitive, or mixed mechanism. Other potent inhibitors, such as the macrolide antimicrobials erythromycin and clarithromycin produce “mechanism-based” inhibition via a metabolic intermediate that complexes with and inactivates the CYP3A enzyme (86 ,87). The HIV protease inhibitor ritonavir and the nonnucleoside reverse transcriptase inhibitor delavirdine also are potent CYP3A inhibitors (88 ,89 ,90 and 91). A component of grapefruit juice inhibits CYP3A in the GI tract (92). Inducers of CYP3A include carbamazepine, rifampin, phenobarbital, nevirapine, dexamethasone, St. John’s wort, and possibly venlafaxine. Ritonavir is an inducer as well as an inhibitor, yielding a net effect on CYP3A metabolism that is difficult to predict (88 ,89 ,90 and 91 ,93 ,94 and 95).

Drug	Inhibition of:	Induction of:
Azole antifungals		
Ketoconazole	CYP3A	
Itraconazole	CYP3A	
Fluconazole	CYP3A, 2C9	
Terbinafine	CYP2D6	
Antidepressants		
Fluoxetine	CYP2D6	
Paroxetine	CYP2D6	
Fluvoxamine	CYP1A2, 2C19, 3A	
Nefazodone	CYP3A	
St. John's wort		CYP3A
Antipsychotics		
Perphenazine	CYP2D6	
Anticonvulsants		
Carbamazepine		CYP3A
Antithrombotics		
Ticlopidine	CYP2D6, 2C19	
Antiinfectives		
Erythromycin	CYP3A	
Clarithromycin	CYP3A	
Ciprofloxacin	CYP1A2	
Rifampin		CYP3A
Viral protease inhibitors		
Ritonavir	CYP3A	CYP3A
Nonnucleoside reverse transcriptase inhibitors		
Delavirdine	CYP3A	
Nevirapine		CYP3A
Cardiovascular agents		
Quinidine	CYP2D6	
Diltiazem	CYP3A	
Verapamil	CYP3A	
Antiulcer agents		
Cimetidine	CYP3A	
Omeprazole	CYP2C19	

TABLE 38.4. REPRESENTATIVE DRUGS HAVING CLINICALLY IMPORTANT EFFECTS ON THE HUMAN CYP ENZYMES

Variability among individuals in CYP3A activity is substantial, even when relatively homogeneous groups of healthy subjects are studied. A consistent finding in population studies of CYP3A activity is that distributions are unimodal, without evidence of genetic polymorphic regulation (96 ,97). However, several studies of CYP3A substrates have demonstrated a small number of individuals with unusually low clearance (96 ,98). The explanation for these observations is unclear, but the genomic determinants of such individual variations in clinical CYP3A activity have become

an active research topic. A number of genetic variants or single nucleotide polymorphisms in either the promoter or coding regions of the human CYP3A4 gene have recently been described (99 ,100 ,101 ,102 ,103 ,104 and 105). One of the promoter region polymorphisms, designated as CYP3A4*1B, is more prevalent in the African-American as opposed to the Caucasian populations. However, there is no evidence to indicate that any of the identified CYP3A4 variants accounts for individual variation in clearance of CYP3A substrates.

In Vitro Models of Drug Metabolism

In vitro systems now are extensively utilized to provide presumptive answers to fundamental clinical questions regarding drug metabolism and drug interactions, and to guide the planning of clinical pharmacokinetic studies (6 ,7 ,8 and 9 ,48 ,49 ,50 ,51 ,52 ,53 ,54 and 55 ,71 ,72 ,73 and 74 ,82 ,83 and 84). If drug X is biotransformed in humans to metabolite Y, two core questions occur: (a) What CYP isoform or isoforms mediate the biotransformation of X to Y? (b) What CYP isoforms do X or Y themselves either induce or inhibit?

Human liver microsomes generally are an important component of currently utilized *in vitro* systems. These preparations contain the various human CYPs in proportion to their abundance in human liver *in vivo*. If biotransformation of a specific substrate to its initial metabolite or metabolites can be replicated in microsomal preparations (Fig. 38.12), inhibition of that reaction by a relatively specific chemical inhibitor can be used as evidence supporting the contribution of the corresponding cytochrome. Chemical inhibitors can also be used in clinical studies, but the *in vitro* model has the advantages of lower cost, more rapid implementation, no risk of human drug exposure, the availability of a greater number of potential chemical inhibitors for research purposes, and the possibility of determining both the quantitative and qualitative contributions of specific cytochromes. Antibodies with relatively specific inhibitory activity against individual human cytochromes can also be used to support or confirm data from *in vitro* chemical inhibition studies (106 ,107). *In vitro* approaches have been strengthened with the availability of microsomes containing pure human cytochromes as expressed by cDNA-transfected human lymphoblastoid cells (108 ,109 and 110). These heterologously expressed pure cytochromes further support definitive identification of cytochromes mediating a specific reaction *in vivo*.

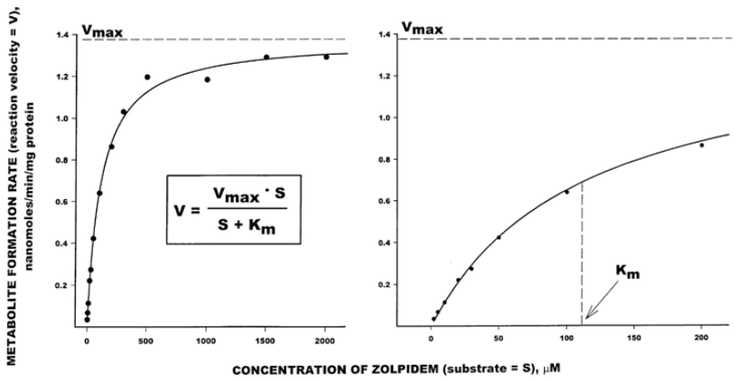


FIGURE 38.12. Example of an *in vitro* metabolism study using human liver microsomes (73). The substrate in this study was zolpidem, of which varying concentrations were incubated with liver microsomes and appropriate reaction cofactors. At each concentration of zolpidem, the rate of formation (*V*) of the principal metabolite of zolpidem (termed the M-3 metabolite) was determined. The relation between substrate concentration (*S*) and reaction velocity (*V*) was analyzed by nonlinear regression to determine the maximum reaction velocity (V_{max}) and the substrate concentration (K_m) producing a reaction velocity of 50% of V_{max}. **Left**: A substrate concentration range up to 2,000 μM. **Right**: The lower range of concentrations shown on an expanded scale.

The quantitative inhibitory potency of a series of drugs and their metabolites against specific index reactions can also be determined using human liver microsomes *in vitro* (6 ,7). The first of two general approaches uses fixed concentrations of the index substrate co-incubated with varying concentrations of the inhibitor in question. The relation of decrement in metabolite formation rate to inhibitor concentration is used to estimate a 50% inhibitory concentration (IC_{50}). This procedure is expeditious and relatively inexpensive, and the numbers can be used to compare the potency of a series of inhibitors (Fig. 38.13). IC_{50} values themselves are not dependent on knowledge of the specific biochemical mechanism of inhibition. However, IC_{50} values depend on substrate concentration when inhibition is competitive, and

can be applied to *in vitro–in vivo* scaling only when the mechanism of inhibition is noncompetitive (111 ,112). Inhibitory potency can also be estimated by determining the inhibition constant (K_i), a number that reflects inhibitory activity in reciprocal fashion. Estimation of K_i requires the study of multiple substrate concentrations and multiple inhibitor concentrations and therefore involves more work, time, and expense. The numerical value of K_i depends on the specific biochemical mechanism of inhibition, which may be unknown. Nonetheless, the inhibitor K_i, is independent of substrate concentration and can be used under some defined circumstances for quantitative *in vitro–in vivo* scaling of drug interactions. In general, K_i is always less than or equal to IC_{50}; K_i will be essentially equal to IC_{50} if inhibition is noncompetitive, or if inhibition is competitive and the substrate concentration is far below the reaction K_m. Both K_i and IC_{50} should provide similar or identical rank-order estimates of relative inhibitory potency for a series of inhibitors of a specific reaction. When the inhibition is purely competitive, K_i values for a specific inhibitor should theoretically be identical across different substrates metabolized by that particular CYP. However, this principle is not supported by experimental data, probably because of *in vitro* experimental artifacts, and because actual biochemical mechanisms of inhibition are not purely competitive. Therefore, absolute values of K_i or IC_{50} cannot be assumed to cross different substrates for the same cytochrome, although the relative rank order of inhibitory potency should be maintained.

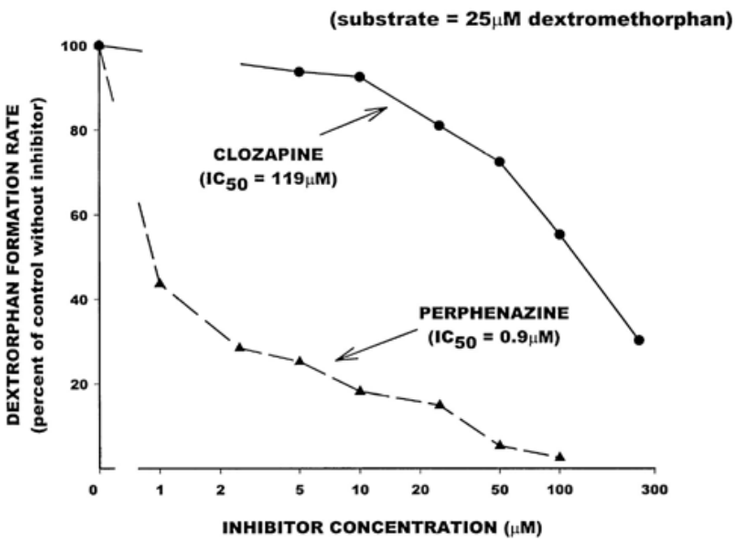


FIGURE 38.13. Example of an *in vitro* study of inhibition of CYP2D6 by two antipsychotic agents, perphenazine and clozapine. A fixed concentration of the substrate dextromethorphan was incubated with liver microsomes, appropriate cofactors, and varying concentrations of perphenazine or clozapine. Rates of formation of dextrorphan (mediated by CYP2D6) with inhibitor present were expressed as a ratio versus the control velocity with no inhibitor. The relation of the velocity ratio to inhibitor concentration can be used to calculate a 50% inhibitory concentration (IC_{50}). The results indicate that perphenazine (IC_{50} = 0.9 μM) is likely to be a clinically important inhibitor of CYP2D6, whereas clozapine (IC_{50} = 119 μM) is a weak inhibitor.

Limitations and drawbacks of *in vitro* systems should be recognized. *In vitro* studies generally utilize substrate concentrations that are one or more orders of magnitude higher than those encountered clinically, even if extensive partitioning of lipophilic drugs from plasma into liver tissue is accounted for. If mathematical models and parameter estimates are valid, the outcome of studies of higher concentrations can be extrapolated down to clinically relevant substrate concentrations. However, a clinically important “high-affinity” metabolic reaction (i.e., one with a low K_m value) could be missed if low substrate concentrations cannot be accurately measured due to limitations of assay sensitivity. The specificity of chemical inhibitor probes is of concern for *in vitro* and *in vivo* studies, because all inhibitors ultimately become nonspecific at higher concentrations. Finally, data from cDNA-expressed human cytochromes can be misinterpreted unless they are considered in the correct context. Pure cytochrome studies can yield quantitative data on the activity of one or more particular cytochromes as a mediator of a specific reaction. However, this information cannot be extrapolated to an estimate of the relative activity of different cytochromes, either *in vivo* or in liver microsomes *in vitro*, without a parallel estimate of the relative quantitative abundance of the cytochromes in question. That is, the importance of a specific cytochrome depends on both activity and abundance.

DRUG INTERACTIONS IN PSYCHOPHARMACOLOGY

Part of "38 - Pharmacokinetics, Pharmacodynamics, and Drug Disposition "

During the last two decades the general problem of pharmacokinetic drug interactions has received increased attention. New classes of medications introduced into clinical practice over this period include the selective serotonin reuptake inhibitor (SSRI) and related mixed mechanism antidepressants, the azole antifungal agents, newer macrolide antimicrobial agents, and the highly active antiretroviral therapies (HAARTs) used against HIV infection and AIDS (Table 38.4). These and other classes of medications have had a major beneficial impact on the therapy on some serious and life-threatening illnesses, but many of the drugs have the secondary pharmacologic property of inducing or inhibiting the human CYP enzymes, thereby raising concerns about pharmacokinetic drug interactions during multiple drug therapy.

One major objective of the drug development process is to generate data on drug interactions, so that treating physicians have the information necessary for safe clinical therapy involving multiple medications. However, the number of possible drug interactions is very large, and time and resources available for implementation of controlled clinical pharmacokinetic studies are inevitably limited.

Some needed studies will therefore be postponed until after a new drug is marketed, and some studies may be bypassed altogether. As discussed above, *in vitro* data are becoming increasingly important as a resource for identifying probable, possible, or unlikely drug interactions, and thereby encouraging rational planning and allocation of resources to more definitive clinical studies.

Pharmacokinetic Versus Pharmacodynamic Drug Interactions

A pharmacokinetic interaction implies that the drug producing the interaction (the “perpetrator”) causes a change in the metabolic clearance of the drug being affected by the interaction (the “victim”), in turn either decreasing or increasing concentrations of the victim drug in plasma and presumably also at the site of action. This change may or may not alter the clinical activity of the victim drug. One pharmacokinetic interaction variant involves modification by the perpetrator of the victim drug’s access to its pharmacologic receptor site, without changing the systemic clearance or plasma levels of the victim. A familiar example is the antagonism of benzodiazepine agonist activity by flumazenil; a less familiar example is benzodiazepine receptor antagonism by ketoconazole (113).

A pharmacodynamic interaction involves either inhibition or enhancement of the clinical effects of the victim drug as a result of similar or identical end-organ actions. Examples are the increase or decrease of the sedative-hypnotic actions of benzodiazepines due to coadministration of ethanol or caffeine, respectively.

Mechanisms of Inhibition Versus Induction of Metabolism

Drug interactions due to inhibition as opposed to induction of CYP-mediated metabolism involve mechanistically different processes. Chemical inhibition is an immediate phenomenon that becomes evident as soon as the inhibitor comes in contact with the enzyme, and is in principle reversible when the inhibitor is no longer present [an exception is “mechanism-based” inhibition (86)]. The magnitude of inhibition depends on the inhibitor concentration at the site of the enzyme in relation to the intrinsic potency of the inhibitor. *In vitro* systems can be used to develop quantitative estimates of inhibitory potency, such as the inhibition constant (K_i) or the 50% inhibitory concentration (IC_{50}) (Fig. 38.13). However, application of K_i or IC_{50} values from *in vitro* systems to quantitative predictions of drug interactions *in vivo* is not straightforward, and requires knowledge of the effective concentration of inhibitor that is available to the enzyme. A generally applicable scheme for relating total or unbound plasma concentrations of inhibitor to effective enzyme-available concentration has not been established. In any case, the theoretical assumption that unbound plasma concentrations are equal to enzyme-available intrahepatic concentrations is incorrect in reality, and may yield underestimates of observed *in vivo* drug interactions by as much as an order of magnitude or more (8 ,83 ,85).

Induction of CYP-mediated metabolism requires prior exposure to a chemical inducer, which signals the synthetic mechanisms to upregulate the production of one or more CYP isoforms (114 ,115 ,116 ,117 and 118). This process takes time, and the increase in CYP activity is of slow onset following initiation of exposure to the inducer, and slowly reverts to baseline after the inducer is removed. Increased CYP expression/activity due to chemical induction therefore reflects prior but not necessarily current exposure to the inducer. The extent of CYP induction probably depends on the dosage (concentration) of the inducer and on the duration of exposure. Induction, unlike inhibition, is not easily studied *in vitro*, because induction requires intact cellular protein synthesis mechanisms as are available in cell culture models.

Inducers and inhibitors of CYP3A can be expected to influence both hepatic and gastrointestinal CYP3A, although not necessarily to the same extent. Very strong inhibitors (such as ketoconazole) or very strong inducers (such as rifampin) will produce substantial changes in both hepatic and gastrointestinal CYP3A. A uniquely complex situation arises for ritonavir, which is both an inhibitor and inducer of CYP3A. Interactions of ritonavir with CYP3A substrate drugs will be time dependent. Initial exposure will produce CYP3A inhibition, but as the duration of exposure proceeds, CYP3A induction may offset the inhibitory effects of acute exposure. The net outcome typically is unpredictable and variable across individuals (93 ,94 and 95).

Perspectives on the Clinical Importance of Drug Interactions

Based on the prevalence of polypharmacy in clinical practice, noninteractions of drugs are far more common than interactions. Coadministration of two drugs usually produces no detectable pharmacokinetic or pharmacodynamic interaction, and the pharmacokinetic disposition and clinical activity of each drug proceed independently of each other. A less common outcome of drug coadministration is a kinetic interaction that could be detected in controlled laboratory circumstances, but that is not clinically important in usual therapeutic circumstances because (a) the interaction, whether or not statistically significant, is not large enough to produce a clinically important change in dynamics of the victim drug; (b) the therapeutic index of the victim drug is large enough so that even a substantial change in plasma levels will not alter therapeutic effects or toxicity; or (c) kinetics and response to the victim drug are so variable that changes in plasma levels due to the drug interaction are far less important than inherent variability. Even less common are clinically important interactions that require

modification in dosage of the perpetrator, the victim, or both. Finally, the most unusual consequence of a drug interaction is a hazardous and contraindicated combination, as in the case of ketoconazole and terfenadine. These situations are rare, but unfortunately receive excessive attention in the public media.

Many secondary sources and compendia are available as summary guides to the extensive literature on drug interactions, but these sources do not necessarily assist clinicians in deciding which interactions should generate serious concern in the course of drug therapy. A useful general guideline is that drug interactions are more likely to be important when (a) the perpetrator drug is a powerful inducer or inhibitor, and produces a very large change in the kinetics and plasma levels of the victim drug; or (b) the therapeutic index of the victim is narrow. Case (a) is exemplified by powerful inducers or inhibitors of CYP3A (ketoconazole, ritonavir, rifampin) coadministered with CYP3A substrates, or powerful inhibitors of CYP2D6 (quinidine, fluoxetine, paroxetine) coadministered with CYP2D6 substrates. Case (b) is exemplified by victim drugs such as phenytoin, warfarin, and digoxin, for which small changes in plasma levels could have important clinical consequences.

Application of Kinetic-Dynamic Methods to Study Drug Interactions

Drug interaction study protocols often incorporate pharmacodynamic endpoints to allow estimating the clinical consequences of drug interactions along with the usual pharmacokinetic outcome measures. The level of complexity of an integrated kinetic-dynamic study depends on the nature of the pharmacodynamic actions of the drug under study as well as the type of pharmacodynamic outcome measures that are required. A number of methodologic principles and dilemmas are illustrated by kinetic-dynamic design options for drug interaction studies involving sedative-hypnotic and anxiolytic drugs acting on the γ-aminobutyric acid (GABA)-benzodiazepine receptor system.

Biotransformation of the benzodiazepine triazolam is dependent on the activity of human CYP3A isoforms (119). Metabolism is strongly inhibited *in vitro* and *in vivo* by CYP3A inhibitors such as ketoconazole, itraconazole, ritonavir, and nefazodone (95 ,119 ,120 ,121 and 122). Some, but not all, of the macrolide antimicrobial agents also are CYP3A inhibitors via "mechanism-based" inhibition, in which the parent compound binds to the metabolically active site on the CYP3A enzyme, yielding a metabolic intermediate that inactivates the enzyme (86 ,87). We tested the inhibitory potency of four macrolide antimicrobial agents [troleandomycin (TAO), erythromycin, clarithromycin, azithromycin] versus triazolam hydroxylation using human liver microsomes *in vitro* (123). Appropriate mean IC_{50} values were TAO, 3.6 μM; erythromycin, 30 μM; and clarithromycin, 28 μM. These values indicate that all three compounds produce substantial *in vitro* inhibition of triazolam hydroxylation and have the potential to produce a significant interaction with triazolam *in vivo*. However, azithromycin was a very weak inhibitor of triazolam *in vitro* (IC_{50} >250 μM), and is anticipated to produce no significant interaction *in vivo*.

In a clinical pharmacokinetic-pharmacodynamic study (123), a series of healthy volunteers were exposed to the following treatment conditions:

A. Triazolam placebo plus macrolide placebo
B. Triazolam (0.125 mg) plus macrolide placebo
C. Triazolam (0.125 mg) plus azithromycin
D. Triazolam (0.125 mg) plus erythromycin
E. Triazolam (0.125 mg) plus clarithromycin

Dosage schedules of the coadministered macrolides were chosen to be consistent with usual dosage recommendations. The five trials were randomized in sequence, and the treatment conditions were double-blind.

Following each dose of triazolam (or placebo to match triazolam), multiple venous blood samples were drawn over a period of 24 hours, and multiple pharmacodynamic testing procedures were performed. Triazolam plasma concentrations were determined by gas chromatography with electron capture detection (Fig. 38.14).

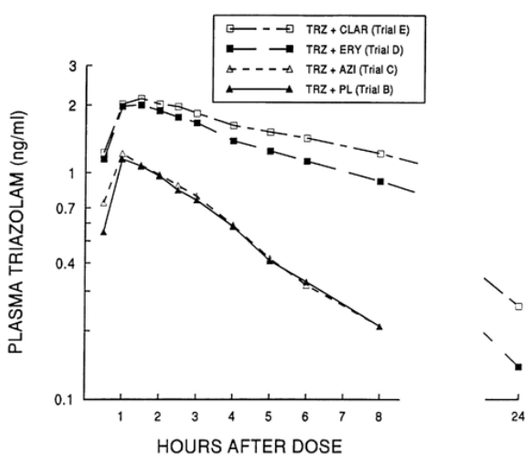


FIGURE 38.14. Mean plasma triazolam concentrations following single 0.125-mg doses of triazolam during trials B, C, D, and E. Note that coadministration of triazolam with azithromycin (AZI, trial C) produced plasma levels nearly identical to triazolam administered with placebo (PL, trial B). However, coadministration with erythromycin (ERY, trial D) or clarithromycin (CLAR, trial E) produced a large increase in plasma triazolam concentrations. (Adapted in part from Greenblatt DJ, von Moltke LL, Harmatz JS, et al. Inhibition of triazolam clearance by macrolide antimicrobial agents: in vitro correlates and dynamic consequences. *Clin Pharmacol Ther* 1998;64:278–285, with permission.)

Mean clearance of triazolam during trials B and C was nearly identical (413 and 416 mL/min, respectively); that is, coadministration of azithromycin had no effect on the pharmacokinetics of triazolam (Fig. 38.14). However, triazolam clearance was significantly reduced to 146 mL/min by erythromycin (trial D), and to 95 mL/min by clarithromycin (trial E) (Fig. 38.14). Thus the *in vivo* kinetic results are highly consistent with the *in vitro* data.

The pharmacodynamic data indicated that the benzodiazepine agonist effects of triazolam plus placebo (trial B), and of triazolam plus azithromycin (trial C) were similar to each other, and greater than the effects of placebo plus placebo (trial A). However, coadministration of triazolam with erythromycin (trial D) or with clarithromycin (trial E) augmented the pharmacodynamic effects of triazolam when compared to trials B or C. The outcome was similar whether based on subjective measures, a semi-objective measure (the Digit-Symbol Substitution Test, DSST), or the fully objective measure (the EEG) (Fig. 38.15). Kinetic-dynamic modeling indicated that the increase in benzodiazepine agonist effects of triazolam caused by coadministration of erythromycin or clarithromycin was fully consistent with the increase in triazolam plasma concentrations (Fig. 38.16).

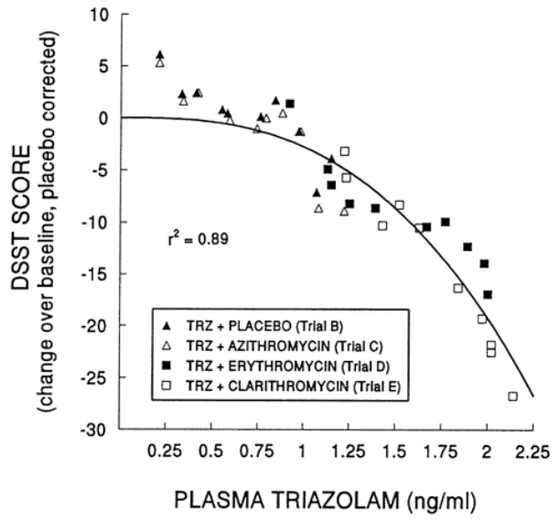


FIGURE 38.16. Relation of mean plasma triazolam concentrations to mean changes over baseline in DSST score at the corresponding times. The *solid line* represents the kinetic-dynamic model relationship based on an exponential function as shown in Fig. 38.5. (From Greenblatt DJ, von Moltke LL, Harmatz JS, et al. Inhibition of triazolam clearance by macrolide antimicrobial agents: in vitro correlates and dynamic consequences. *Clin Pharmacol Ther* 1998;64:278–285, with permission.)

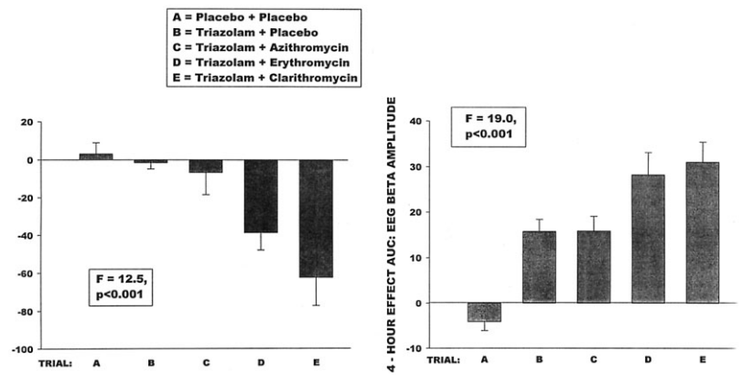


FIGURE 38.15. Mean (± standard error, SE) 4-hour pharmacodynamic effect areas for the digit-symbol substitution test (DSST) score (*left*), and for the EEG beta amplitude (*right*), during the five trials. Note that decrements in DSST score, and increases in EEG beta amplitude, were very similar between trials B and C, whereas effects were significantly enhanced during trials D and E.

COMMENT

Part of "38 - Pharmacokinetics, Pharmacodynamics, and Drug Disposition "

Pharmacokinetic drug interactions in clinical psychopharmacology are assuming increasing importance as polypharmacy becomes more common, and more drugs with enzyme-inducing or -inhibiting properties are introduced into clinical practice. Contemporary approaches to the basic and clinical investigation of drug interactions and their pharmacodynamic consequences are illustrated in this chapter. It is evident that technologic and conceptual advances in pharmacokinetics, pharmacodynamics, and drug metabolism may be usefully applied to the evaluation of drug interactions. An ideal approach would incorporate the collaborative participation of individuals representing expertise in molecular pharmacology, cytochrome biochemistry, *in vitro* metabolism, clinical pharmacokinetics-pharmacodynamics, and clinical therapeutics.

ACKNOWLEDGMENTS

Part of "38 - Pharmacokinetics, Pharmacodynamics, and Drug Disposition "

The work was supported by grants MH-34223, MH-01237, DA-05258, DA-13209, MH-58435, DK-58496, and RR-00054 from the U.S. Department of Health and Human Services.

REFERENCES

1. Stanski DR. Pharmacodynamic modeling of anesthetic EEG drug effects. *Annu Rev Pharmacol Toxicol* 1992;32:423–447.

2. Derendorf H, Meibohm B. Modeling of pharmacokinetic/pharmacodynamic (PK/PD) relationships: concepts and perspectives. *Pharm Res* 1999;16:176–185.

3. Bellissant E, Sébille V, Paintaud G. Methodological issues in pharmacokinetic-pharmacodynamic modelling. *Clin Pharmacokinet* 1998;35:151–166.

4. Mandema JW, Danhof M. Electroencephalogram effect measures and relationships between pharmacokinetics and pharmacodynamics of centrally acting drugs. *Clin Pharmacokinet* 1992;23:191–215.

5. Laurijssens BE, Greenblatt DJ. Pharmacokinetic-pharmacodynamic relationships for benzodiazepines. *Clin Pharmacokinet* 1996;30:52–76.

6. Greenblatt DJ, von Moltke LL, Harmatz JS, et al. Human cytochromes and some newer antidepressants: kinetics, metabolism, and drug interactions. *J Clin Psychopharmacol* 1999;19(suppl 1):23S–35S.

7. Greenblatt DJ, von Moltke LL, Harmatz JS, et al. Drug interactions with newer antidepressants: role of human cytochromes P450. *J Clin Psychiatry* 1998;59(suppl 15):19–27.

8. von Moltke LL, Greenblatt DJ, Schmider J, et al. In vitro approaches to predicting drug interactions in vivo. *Biochem Pharmacol* 1998;55:113–122.

9. Bertz RJ, Granneman GR. Use of in vitro and in vivo data to estimate the likelihood of metabolic pharmacokinetic interactions. *Clin Pharmacokinet* 1997;32:210–258.

10. Whiting B, Kelman AW, Grevel J. Population pharmacokinetics. Theory and clinical application. *Clin Pharmacokinet* 1986;11:387–401.

11. Vozeh S, Steimer JL, Rowland M, et al. The use of population pharmacokinetics in drug development. *Clin Pharmacokinet* 1996;30:81–93.

12. Maitre PO, Bührer M, Thomson D, et al. A three-step approach combining Bayesian regression and NONMEM population analysis: application to midazolam. *J Pharmacokinet Biopharm* 1991;19:377–384.

13. Thomson AH, Whiting B. Bayesian parameter estimation and population pharmacokinetics. *Clin Pharmacokinet* 1992;22:447–467.

14. Jonsson EN, Wade JR, Karlsson MO. Comparison of some practical sampling strategies for population pharmacokinetic studies. *J Pharmacokinet Biopharm* 1996;24:245–263.

15. Carter AA, Rosenbaum SE, Dudley MN. Review of methods in population pharmacokinetics. *Clin Res Regul Affairs* 1995;12:1–21.

16. Samara E, Granneman R. Role of population pharmacokinetics in drug development: a pharmaceutical industry perspective. *Clin Pharmacokinet* 1997;32:294–312.

17. Sheiner LB, Ludden TM. Population pharmacokinetics/dynamics. *Annu Rev Pharmacol Toxicol* 1992;32:185–209.

18. Aarons L. Population pharmacokinetics: theory and practice. *Br J Clin Pharmacol* 1991;32:669–670.

19. de Gatta MMF, García MJ, Lanao JM, et al. Bayesian forecasting in paediatric populations. *Clin Pharmacokinet* 1996;31:325–330.

20. Shader RI, Harmatz JS, Oesterheld JR, et al. Population pharmacokinetics of methylphenidate in children with attention-deficit hyperactivity disorder. *J Clin Pharmacol* 1999;39:775–785.

21. Safer DJ. Central stimulant treatment of childhood attention deficit hyperactivity disorder. *CNS Drugs* 1997;7:264–272.

22. Elia J, Ambrosini PJ, Rapoport JL. Treatment of attention-deficit-hyperactivity disorder. *N Engl J Med* 1999;340:780–788.

23. Cyr M, Brown CS. Current drug therapy recommendations for the treatment of attention deficit hyperactivity disorder. *Drugs* 1998;56:215–233.

24. Wilens TE, Biederman J, Spencer TJ. Pharmacotherapy of attention deficit hyperactivity disorder in adults. *CNS Drugs* 1998;9:347–356.

25. Kimko HC, Cross JT, Abernethy DR. Pharmacokinetics and clinical effectiveness of methylphenidate. *Clin Pharmacokinet* 1999;37:457–470.

26. Srinivas NR, Hubbard JW, Quinn D, et al. Enantioselective pharmacokinetics and pharmacodynamics of dl-threo-methylphenidate in children with attention deficit hyperactivity disorder. *Clin Pharmacol Ther* 1992;52:561–568.

27. Schwinghammer TL, Kroboth PD. Basic concepts in pharmacodynamic modeling. *J Clin Pharmacol* 1988;28:388–394.

28. Holford NH, Sheiner LB. Understanding the dose-effect relationship: clinical application of pharmacokinetic-pharmacodynamic models. *Clin Pharmacokinet* 1981;6:429–453.

29. Swerdlow BN, Holley FO. Intravenous anaesthetic agents. Pharmacokinetic-pharmacodynamic relationships. *Clin Pharmacokinet* 1987;12:79–110.

30. Dingemanse J, Danhof M, Breimer DD. Pharmacokineticpharmacodynamic modeling of CNS drug effects: an overview. *Pharmacol Ther* 1988;38:1–52.

31. Campbell DB. The use of kinetic-dynamic interactions in the evaluation of drugs. *Psychopharmacology* 1990;100:433–450.

32. Greenblatt DJ, Harmatz JS. Kinetic-dynamic modeling in clinical psychopharmacology. *J Clin Psychopharmacol* 1993;13:231–234.

33. Greenblatt DJ, von Moltke LL, Ehrenberg BL, et al. Kinetics and dynamics of lorazepam during and after continuous intravenous infusion. *Crit Care Med* 2000;28:2750–2757.

34. Greenblatt DJ, Ehrenberg BL, Gunderman J, et al. Kinetic and dynamic study of intravenous lorazepam: comparison with intravenous diazepam. *J Pharmacol Exp Ther* 1989;250:134–140.

35. Greenblatt DJ, Ehrenberg BL, Gunderman J, et al. Pharmacokinetic and electroencephalographic study of intravenous diazepam, midazolam, and placebo. *Clin Pharmacol Ther* 1989;45:356–365.

36. Gupta SK, Ellinwood EH, Nikaido AM, et al. Simultaneous modeling of the pharmacokinetic and pharmacodynamic properties of benzodiazepines. I: Lorazepam. *J Pharmacokinet Biopharm* 1990;18:89–102.

37. Tedeschi G, Smith AT, Dhillon S, et al. Rate of entrance of benzodiazepines into the brain determined by eye movement recording. *Br J Clin Pharmacol* 1983;15:103–107.

38. Greenblatt DJ, Sethy VH. Benzodiazepine concentrations in brain directly reflect receptor occupancy: studies of diazepam, lorazepam, and oxazepam. *Psychopharmacology* 1990;102:373–378.

39. Walton NY, Treiman DM. Lorazepam treatment of experimental status epilepticus in the rat: relevance to clinical practice. *Neurology* 1990;40:990–994.

40. Ameer B, Greenblatt DJ. Lorazepam: a review of its clinical pharmacological properties and therapeutic uses. *Drugs* 1981;21:161–200.

41. Greenblatt DJ. Clinical pharmacokinetics of oxazepam and lorazepam. *Clin Pharmacokinet* 1981;6:88–105.

42. Ochs HR, Greenblatt DJ, Knüchel M. Kinetics of diazepam, midazolam, and lorazepam in cigarette smokers. *Chest* 1985;87:223–226.

43. Abernethy DR, Greenblatt DJ, Ameer B, et al. Probenecid impairment of acetaminophen and lorazepam clearance: direct inhibition of ether glucuronide formation. *J Pharmacol Exp Ther* 1985;234:345–349.

44. Abernethy DR, Greenblatt DJ, Divoll M, et al. Differential effect of cimetidine on drug oxidation (antipyrine and diazepam) versus conjugation (acetaminophen and lorazepam): prevention of acetaminophen toxicity by cimetidine. *J Pharmacol Exp Ther* 1983;224:508–513.

45. Greenblatt DJ, Allen MD, Locniskar A, et al. Lorazepam kinetics in the elderly. *Clin Pharmacol Ther* 1979;26:103–113.

46. Treiman DM. The role of benzodiazepines in the management of status epilepticus. *Neurology* 1990;40(suppl 2):32–42.

47. Lowenstein DH, Alldredge BK. Status epilepticus. *N Engl J Med* 1998;338:970–976.

48. Clarke SE. In vitro assessment of human cytochrome P450. *Xenobiotica* 1998;28:1167–1202.

49. Smith G, Stubbins MJ, Harries LW, et al. Molecular genetics of the human cytochrome P450 monooxygenase superfamily. *Xenobiotica* 1998;28:1129–1165.

50. Smith DA, Abel SM, Hyland R, et al. Human cytochrome P450s: selectivity and measurement in vivo. *Xenobiotica* 1998;28:1095–1128.

51. Nelson DR, Koymans L, Kamataki T, et al. P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* 1996;6:1–42.

52. Park BK, Pirmohamed M, Kitteringham NR. The role of cytochrome P450 enzymes in hepatic and extrahepatic human drug toxicity. *Pharmacol Ther* 1995;68:385–424.

53. Wrighton SA, Stevens JC. The human hepatic cytochromes P450 involved in drug metabolism. *Crit Rev Toxicol* 1992;22:1–21.

54. Glue P, Clement RP. Cytochrome P450 enzymes and drug metabolism—basic concepts and methods of assessment. *Cell Mol Neurobiol* 1999;19:309–323.

55. Parkinson A. An overview of current cytochrome P450 technology for assessing the safety and efficacy of new materials. *Toxicol Pathol* 1996;24:45–57.

56. Kroemer HK, Eichelbaum M. Molecular bases and clinical consequences of genetic cytochrome P450 2D6 polymorphism. *Life Sci* 1995;56:2285–2298.

57. Nebert DW. Polymorphisms in drug-metabolizing enzymes: what is their clinical relevance and why do they exist? *Am J Hum Genet* 1997;60:265–271.

58. Ingelman-Sundberg M, Oscarson M, McLellan RA. Polymorphic human cytochrome P450 enzymes: an opportunity for individualized drug treatment. *Trends Pharmacol Sci* 1999;20:342–349.

59. Bertilsson L. Geographical/interracial differences in polymorphic drug oxidation: current state of knowledge of cytochromes P450 (CYP) 2D6 and 2C19. *Clin Pharmacokinet* 1995;29:192–209.

60. Bertilsson L, Dahl M-L. Polymorphic drug oxidation. *CNS Drugs* 1996;3:200–223.

61. Fromm MF, Kroemer HK, Eichelbaum M. Impact of P450 genetic polymorphism on the first-pass extraction of cardiovascular and neuroactive drugs. *Advanced Drug Deliv Rev* 1997;27:171–199.

62. Gonzalez FJ, Meyer UA. Molecular genetics of the debrisoquin-sparteine polymorphism. *Clin Pharmacol Ther* 1991;50:233–238.

63. Guengerich FP. Cytochrome P-450 3A4: regulation and role in drug metabolism. *Annu Rev Pharmacol Toxicol* 1999;39:1–7.

64. Dresser GK, Spence JD, Bailey DG. Pharmacokinetic-pharmacodynamic consequences and clinical relevance of cytochrome P450 3A4 inhibition. *Clin Pharmacokinet* 2000;38:41–57.

65. Thummel KE, Wilkinson GR. In vitro and in vivo drug interactions involving human CYP3A. *Annu Rev Pharmacol Toxicol* 1998;38:389–430.

66. de Wildt SN, Kearns GL, Leeder JS, et al. Cytochrome P450 3A: ontogeny and drug disposition. *Clin Pharmacokinet* 1999;37:485–505.

67. Ketter TA, Flockhart DA, Post RM, et al. The emerging role of cytochrome P450 3A in psychopharmacology. *J Clin Psychopharmacol* 1995;15:387–398.

68. Maurel P. The CYP3A family. In: Ionnides C, ed. *Cytochromes P450*. Boca Raton, FL: CRC Press, 1996:241–270.

69. von Moltke LL, Greenblatt DJ, Schmider J, et al. Metabolism of drugs by cytochrome P450 3A isoforms: implications for drug interactions in psychopharmacology. *Clin Pharmacokinet* 1995;29(suppl 1):33–43.

70. Shimada T, Yamazaki H, Mimura M, et al. Interindividual variations in human liver cytochrome P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. *J Pharmacol Exp Ther* 1994;270:414–423.

71. Venkatakrishnan K, Greenblatt DJ, von Moltke LL, et al. Five distinct human cytochromes mediate amitriptyline N-demethylation in vitro: dominance of CYP 2C19 and 3A4. *J Clin Pharmacol* 1998;38:112–121.

72. von Moltke LL, Greenblatt DJ, Grassi JM, et al. Citalopram and desmethylcitalopram in vitro: human cytochromes mediating transformation, and cytochrome inhibitory effects. *Biol Psychiatry* 1999;46:839–849.

73. von Moltke LL, Greenblatt DJ, Granda BW, et al. Zolpidem metabolism in vitro: responsible cytochromes, chemical inhibitors, and in vivo correlations. *Br J Clin Pharmacol* 1999;48:89–97.

74. von Moltke LL, Greenblatt DJ, Granda BW, et al. Nefazodone, meta-chlorophenylpiperazine, and their metabolites in vitro: cytochromes mediating transformation, and P450–3A4 inhibitory actions. *Psychopharmacology* 1999;145:113–122.

75. Greene DS, Barbhaiya RH. Clinical pharmacokinetics of nefazodone. *Clin Pharmacokinet* 1997;33:260–275.

76. von Moltke LL, Greenblatt DJ, Schmider J, et al. Midazolam hydroxylation by human liver microsomes in vitro: inhibition by fluoxetine, norfluoxetine, and by azole antifungal agents. *J Clin Pharmacol* 1996;36:783–791.

77. Perloff MD, von Moltke LL, Court MH, et al. Midazolam and triazolam biotransformation in mouse and human liver microsomes: relative contribution of CYP3A and CYP2C9 isoforms. *J Pharmacol Exp Ther* 2000;292:618–628.

78. Bornemann LD, Min BH, Crews T, et al. Dose dependent pharmacokinetics of midazolam. *Eur J Clin Pharmacol* 1985;29:91–95.

79. Hall SD, Thummel KE, Watkins PB, et al. Molecular and physical mechanisms of first-pass extraction. *Drug Metab Dispos* 1999;27:161–166.

80. Yuan R, Flockhart DA, Balian JD. Pharmacokinetic and pharmacodynamic consequences of metabolism-based drug interactions with alprazolam, midazolam, and triazolam. *J Clin Pharmacol* 1999;39:1109–1125.

81. Tsunoda SM, Velez RL, von Moltke LL, et al. Differentiation of intestinal and hepatic cytochrome P450 3A activity with use of midazolam as an in vivo probe: effect of ketoconazole. *Clin Pharmacol Ther* 1999;66:461–471.

82. Lin JH, Lu AY. Inhibition and induction of cytochrome P450 and the clinical implications. *Clin Pharmacokinet* 1998;35:361–390.

83. Venkatakrishnan K, von Moltke LL, Greenblatt DJ. Effects of the antifungal agents on oxidative drug metabolism in humans: clinical relevance. *Clin Pharmacokinet* 2000;38:111–180.

84. Ito K, Iwatsubo T, Kanamitsu S, et al. Prediction of pharmacokinetic alterations caused by drug-drug interactions: metabolic interaction in the liver. *Pharmacol Rev* 1998;50:387–412.

85. Greenblatt DJ, von Moltke LL. Sedative-hypnotic and anxiolytic agents. In: Levy RH, Thummel KE, Trager WF, et al., eds. *Metabolic drug interactions*. Philadelphia: Lippincott Williams & Wilkins, 2000:259–270.

86. Silverman R. Mechanism-based enzyme inactivators. *Methods Enzymol* 1995;249:240–283.

87. Gillum JG, Israel DS, Polk RE. Pharmacokinetic drug interactions with antimicrobial agents. *Clin Pharmacokinet* 1993;25:450–482.

88. Barry M, Mulcahy F, Merry C, et al. Pharmacokinetics and potential interactions amongst antiretroviral agents used to treat patients with HIV infection. *Clin Pharmacokinet* 1999;36:289–304.

89. Malaty LI, Kuper JJ. Drug interactions of HIV protease inhibitors. *Drug Safety* 1999;20:147–169.

90. Tseng AL, Foisy MM. Significant interactions with new antiretrovirals and psychotropic drugs. *Ann Pharmacother* 1999;33:461–473.

91. Hsu A, Granneman GR, Bertz RJ. Ritonavir. Clinical pharmacokinetics and interactions with other anti-HIV agents. *Clin Pharmacokinet* 1998;35:275–291.

92. Bailey DG, Malcom J, Arnold O, et al. Grapefruit juice–drug interactions. *Br J Clin Pharmacol* 1998;46:101–110.

93. Greenblatt DJ, von Moltke LL, Harmatz JS, et al. Alprazolam-ritonavir interaction: implications for product labeling. *Clin Pharmacol Ther* 2000;67:335–341.

94. Greenblatt DJ, von Moltke LL, Daily JP, et al. Extensive impairment of triazolam and alprazolam clearance by short-term low-dose ritonavir: the clinical dilemma of concurrent inhibition and induction. *J Clin Psychopharmacol* 1999;19:293–296.

95. Greenblatt DJ, von Moltke LL, Harmatz JS, et al. Differential impairment of triazolam and zolpidem clearance by ritonavir. *J AIDS* 2000;24:129–136.

96. Kassai A, Toth G, Eichelbaum M, et al. No evidence of a genetic polymorphism in the oxidative metabolism of midazolam. *Clin Pharmacokinet* 1988;15:319–325.

97. Friedman H, Greenblatt DJ, Burstein ES, et al. Population study of triazolam pharmacokinetics. *Br J Clin Pharmacol* 1986;22:639–642.

98. Greenblatt DJ, Divoll M, Abernethy DR, et al. Reduced clearance of triazolam in old age: relation to antipyrine oxidizing capacity. *Br J Clin Pharmacol* 1983;15:303–309.

99. Felix CA, Walker AH, Lange BJ, et al. Association of CYP3A4 genotype with treatment-related leukemia. *Proc Natl Acad Sci USA* 1998;95:13176–13181.

100. Rebbeck TR, Jaffe JM, Walker AH, et al. Modification of clinical presentation of prostate tumors by a novel genetic variant in CYP3A4. *J Natl Cancer Inst* 1998;90:1225–1229.

101. Ball SE, Scatina A, Kao J, et al. Population distribution and effects on drug metabolism of a genetic variant in the 5′ promoter region of CYP3A4. *Clin Pharmacol Ther* 1999;66:288–294.

102. Sata F, Sapone A, Elizondo G, Stocker P, et al. CYP3A4 allelic variants with amino acid substitutions in exons 7 and 12: Evidence for an allelic variant with altered catalytic activity. *Clin Pharmacol Ther* 2000;67:48–56.

103. von Moltke LL, Tran TH, Cotreau MM, et al. Unusually low clearance of two CYP3A4 substrates, alprazolam and trazodone, in a volunteer subject with wild-type CYP3A promoter region. *J Clin Pharmacol* 2000;40:200–204.

104. Westlind A, Löfberg L, Tindberg N, et al. Interindividual differences in hepatic expression of CYP3A4: relationship to genetic polymorphism in the 5′-upstream regulatory region. *Biochem Biophys Res Commun* 1999;259:201–205.

105. Wandel C, Witte JS, Hall JM, et al. CYP3A activity in African American and European American men: population differences and functional effect of CYP3A4*1B 5′-promoter region polymorphism. *Clin Pharmacol Ther* 2000;68:82–91.

106. Gelboin HV, Krausz KW, Gonzalez FJ, et al. Inhibitory monoclonal antibodies to human cytochrome P450 enzymes: a new avenue for drug discovery. *Trends Pharmacol Sci* 1999;20:432–438.

107. Shou M, Lu T, Krausz KW, et al. Use of inhibitory monoclonal antibodies to assess the contribution of cytochromes P450 to human drug metabolism. *Eur J Pharmacol* 2000;394:199–209.

108. Gonzalez FJ, Korzekwa KR. Cytochromes P450 expression systems. *Annu Rev Pharmacol Toxicol* 1995;35:369–390.

109. Crespi CL, Miller VP. The use of heterologously expressed drug metabolizing enzymes—state of the art and prospects for the future. *Pharmacol Ther* 1999;84:121–131.

110. Crespi CL, Penman BW. Use of cDNA-expressed human cytochrome P450 enzymes to study potential drug-drug interactions. *Adv Pharmacol* 1997;43:171–188.

111. Halpert JR. Structural basis of selective cytochrome P450 inhibition. *Annu Rev Pharmacol Toxicol* 1995;35:29–53.

112. Segel IH. *Enzyme kinetics*. New York: Wiley, 1975.

113. Fahey JM, Pritchard GA, von Moltke LL, et al. The effects of ketoconazole on triazolam pharmacokinetics, pharmacodynamics and benzodiazepine receptor binding in mice. *J Pharmacol Exp Ther* 1998;285:271–276.

114. Barry M, Feely J. Enzyme induction and inhibition. *Pharmacol Ther* 1990;48:71–94.

115. Denison MS, Whitlock JP. Xenobiotic-inducible transcription of cytochromes P450 genes. *J Biol Chem* 1995;270:18175–18178.

116. Bock KW, Lipp H-P, Bock-Hennig BS. Induction of drug-metabolizing enzymes by xenobiotics. *Xenobiotica* 1990;20:1101–1111.

117. Waxman DJ, Azaroff L. Phenobarbital induction of cytochrome P-450 gene expression. *Biochem J* 1992;281:577–592.

118. Park BK, Kitteringham NR, Piromohamed M, et al. Relevance of induction of human drug-metabolizing enzymes: pharmacological and toxicological implications. *Br J Clin Pharmacol* 1996;41:477–491.

119. von Moltke LL, Greenblatt DJ, Harmatz JS, et al. Triazolam biotransformation by human liver microsomes in vitro: effects of metabolic inhibitors, and clinical confirmation of a predicted interaction with ketoconazole. *J Pharmacol Exp Ther* 1996;276:370–379.

120. von Moltke LL, Greenblatt DJ, Duan SX, et al. Inhibition of triazolam hydroxylation by ketoconazole, itraconazole, hydroxyitraconazole and fluconazole in vitro. *Pharm Pharmacol Commun* 1998;4:443–445.

121. von Moltke LL, Greenblatt DJ, Grassi JM, et al. Protease inhibitors as inhibitors of human cytochromes P450: high risk associated with ritonavir. *J Clin Pharmacol* 1998;38:106–111.

122. Greenblatt DJ, Wright CE, von Moltke LL, et al. Ketoconazole inhibition of triazolam and alprazolam clearance: differential kinetic and dynamic consequences. *Clin Pharmacol Ther* 1998;64:237–247.

123. Greenblatt DJ, von Moltke LL, Harmatz JS, et al. Inhibition of triazolam clearance by macrolide antimicrobial agents: in vitro correlates and dynamic consequences. *Clin Pharmacol Ther* 1998;64:278–285.

39

The Role of Pharmaceuticals in Mental Health Care Outcomes

Ramin Mahmoud

Chris M. Kozma

C. E. Reeder

Amy Grogg

Brian Meissner

Ramin Mahmoud and Amy Grogg: Janssen Research Foundation, Titusville, New Jersey.

Chris M. Kozma: Strategic Outcomes Service of Care Sciences, Inc., Philadelphia, Pennsylvania.

C. E. Reeder and Brian Meissner: College of Pharmacy, University of South Carolina, Columbia, South Carolina.

This chapter discusses the role of pharmaceutical outcome evaluations in mental health care. The first section discusses the importance of pharmaceutical outcome evaluations. The second section describes techniques used in economic evaluations of pharmaceuticals (i.e., pharmacoeconomic methods). The final section discusses how mental health care outcomes data may be used in practice.

WHY OUTCOMES?

Part of "39 - The Role of Pharmaceuticals in Mental Health Care Outcomes "

The idea that outcomes associated with the provision of health care are important is not new. In the 1960s Avedis Donabedian (1) presented health outcomes as changes in health status that were attributable to antecedent health care. For many years, however, evaluations of health care focused on the structure or process of care. As health care moves into the new millennium, financing of health care is evolving from individual providers being solely responsible for patient outcomes to an environment where payers, institutions, and providers are being held accountable for quality and cost of care. As financing of health care has moved to a more centralized locus of control, evaluation of outcomes has become more feasible and desirable. In a book titled *Who Shall Live*, Victor Fuchs (2) discussed three factors that can be balanced in our health care system: costs, quality, and access. Over time the pendulum swings from one to another of these dimensions. If costs containment goes too far, then quality or access may suffer. Likewise, it is possible in today's technologically driven environment to provide a level of quality that is affordable to only a small segment of our society. In an environment of limited resources and high demand for health care, a quality, cost, access trade-off is essential. The issue then becomes how do we define and measure quality so that these trade-offs can be made in pursuit of efficiency and/or equity. Many believe that quality should be defined and measured in terms of patient outcomes.
In simple terms, outcomes are the "end results." Improvement in outcomes is a primary reason for medical intervention, including use of pharmacologic therapies. There is a belief that the use of pharmaceuticals will have a positive impact on the "end results" of patient care. Historically, this belief is self-evident in the treatment of mental health disorders with the use of drugs to treat psychoses and depression—conditions for which treatment was revolutionized by pharmaceuticals (3 ,4). However, with newer, more costly pharmaceuticals, such as selective serotonin reuptake inhibitors (SSRIs) and atypical antipsychotic agents, many payers and health professionals have questioned the value that is received for the resources expended on these agents. Consequently, numerous studies have been directed at these issues. For example, SSRIs have been compared to tricyclic antidepressants (TCAs) for the treatment of depression (5 ,6 ,7 and 8). Payers are interested not only in the most efficacious antidepressant but also in which agent should be chosen as first- or second-line therapy, the appropriate course (length) of therapy, and whether initial therapy should be augmented with an additional agent (9 ,10 ,11 and 12). Given the wide variations in care and the possibility that alternative treatments could lead to similar outcomes, particularly under the less than ideal conditions of "usual care" practice, health care payers are concerned about which treatments regimens lead to the most efficient outcome. In a resource-constrained environment, it is both reasonable and responsible to ask this question:

Would the resources expended for one alternative be more efficient if devoted to another alternative? One useful way to address this question is with data on patient and cost outcomes.

There is an inextricable but sometimes complex relationship between quality of care and outcomes. Outcomes data are one way of evaluating quality. It is noteworthy, however, that establishing quality “thresholds” requires a value judgment. What is acceptable quality to one person may be unacceptable to another at any given level of cost. Outcomes data cannot provide answers to questions that require fundamental value judgments. On a macro level, an analysis using measures of quality and cost will not define the percentage of gross domestic product (GDP) that a nation should spend to achieve a certain level of quality in health care, but rather will provide tools and information to assist decision makers in efficiently allocating scarce resources. On a micro level, there is no specific quality of life score on an instrument that indicates if or when a drug product should be reimbursed. There is no single clinical measure that indicates that a patient is in perfect health. All outcomes data require the interpretation and evaluation of a medical decision maker. Outcomes data provide one more, albeit in many cases relevant, piece of information on which to base decisions.

In the previous discussion, outcomes were defined in abstract terms (i.e., changes in health status, end results). While there is an inherent belief that many pharmaceuticals improve outcomes, for outcomes to be documented and improved, this terminology must be defined in more concrete terms. One conceptualization of health care outcomes is the economic, clinical, humanistic outcomes model (ECHO) (13). This conceptualization portrays health outcomes along three dimensions. Clinical outcomes are outcomes related to the effects of medical treatments or disease on medical events such as hospitalization or death (i.e., end results). Economic outcomes are usually expressed as costs (e.g., dollars) associated with an intervention, and are often considered as ratios of costs to some measure of the consequences of a disease and its treatment. Humanistic outcomes are measures of the impact of disease or treatment on patient’s lives. In addition to these outcomes, there are many intermediate variables that are important when measuring the effects of a disease or treatment. These variables are referred to in the published literature by many names including process variables, surrogate outcomes, or intermediate variables. In many cases making a clear distinction between these consequences of pharmaceutical use is probably not necessary; however, when reviewing literature, it is important to consider whether a consequence is a “true” outcome or an intermediate variable. For example, a score on a depression inventory is probably closer to an intermediate variable, whereas events such as rehospitalization or suicide reflect the “end results” or outcomes one would like to prevent.

Mental health care is expensive. For example, it is estimated that $44 billion is spent annually on the treatment of depression and $100 billion is spent annually on the treatment of Alzheimer’s disease (14 ,15). The cost to treat schizophrenia has been estimated at $33 billion per year, accounting for 22% of dollars spent to treat all categories of mental illness and 2.5% of total health care expenditures (16 ,17). Increasing competition for scarce resources encourages decision makers to use outcomes data to evaluate the effectiveness and efficiency of treatment options for depression, Alzheimer’s disease, and other mental health disorders. These issues are not new to health care providers, but the development of drug formularies as mechanisms to control costs has generated a need for outcomes studies to evaluate the benefits obtained from new pharmacologic agents. Shortly after SSRIs were released on the market, questions arose regarding whether health care outcomes were better for patients treated with SSRIs than for patients treated with traditional TCA therapy (5 ,6 and 7). This new class of pharmaceuticals was more costly than prior standard therapy (the TCAs), prescribed for a wide variety of patients, and had (in clinical trials) fewer side effects. Although the products were shown to be superior in some domains in clinical trials, there was a practical question regarding whether these benefits were realized in real-world practice and if so, at what net incremental cost (18). Do patients treated with SSRIs consume less acute care services, require fewer specialist visits, or have lower suicide rates? Is the total cost per acute depressive episode (successfully treated case) therefore lower with the newer products despite higher drug acquisition costs? Does treatment with SSRIs cost more, but provide better humanistic outcomes such as quality of life or quality-adjusted life years? These are the types of questions that outcomes research and pharmacoeconomic evaluations attempt to answer. This chapter does not specifically address these questions, but rather uses these questions to illustrate issues in outcomes research. It is noteworthy to recognize that not all the outcome questions of interest are likely to be addressed in a single study; rather, answers will come from an evaluation of a body of literature.

Most health care professionals are familiar with the clinical aspects of the treatment of mental health diseases. Given the substantial clinical information in the remainder of this text, this chapter focuses on evaluation of the economic and humanistic outcomes related to pharmaceutical use. Specifically, the techniques of pharmacoeconomics will be reviewed as well as the instruments for evaluating humanistic outcomes in mental health care populations.

PHARMACOECONOMICS

Part of "39 - The Role of Pharmaceuticals in Mental Health Care Outcomes "

In the current health care environment, many decisions are driven by costs. At a minimum, health care systems are looking for systematic methods for reducing costs. Although

the fraction of the health care dollar spent on pharmaceuticals is low, it is clear that as both the pressure to reduce costs and the percentage of health care dollars spent on pharmaceuticals grow, so does interest in the costs of medications. Economists, however, are quick to point out that the acquisition cost of the pharmaceutical is not the most appropriate unit of analysis. It is possible that the acquisition cost of many pharmaceuticals may be offset by reductions in other more expensive forms of care. If the use of an expensive atypical antipsychotic leads to reductions in hospitalizations, then the "value" of the pharmaceutical from a total cost perspective is greater than the acquisition cost of the pharmaceutical. This is a key idea behind pharmacoeconomics. Pharmacoeconomics provides a set of techniques that allow consideration of the costs and consequences of alternative pharmaceutical therapies (19).

Studies are typically categorized by whether they consider costs, outcomes, or both cost and outcomes. In addition, studies can also be categorized by whether or not they consider alternatives. For example, traditional clinical trials focus on comparing the consequences of alternatives when one of the alternatives is typically a placebo. Although placebo comparison is highly relevant from the perspective of a regulatory agency striving to meet its special mandate, from the perspective of many health care decision makers a comparison with placebo is meaningful only if it is a relevant treatment alternative. Pharmacoeconomic studies best provide a comparison of relevant alternatives.

Studies that evaluate only cost for one alternative are referred to as cost descriptions. Other studies may also consider consequences. In these cases the study would describe both the costs and consequences of a single alternative leading to a cost-outcome description. If two alternatives are compared but only costs are considered, then the study is a cost evaluation. However, the primary concern of pharmacoeconomics is the comparison of both costs and consequences simultaneously for two for more relevant alternatives. There are four specific techniques that are typically used when conducting pharmacoeconomic studies (20):

- Cost-minimization analysis
- Cost-effectiveness analysis
- Cost-utility analysis
- Cost-benefit analysis

In each of these cases the numerators are the costs of inputs for a given decision. For example, if the total cost of care for the treatment of depression is considered, input costs might include cost of drug, cost of physician visits (family practitioner, internist, and specialist), behavioral therapy, hospitalization, and emergency department use. (Cost is discussed in greater detail below.) Next, the appropriate outcomes or consequence must be specified in the denominator.

Cost-minimization analysis assumes, not always explicitly, that the outcomes are equal. If this is a valid assumption, then the decision is based entirely on the costs of the inputs. The classic example of a cost minimization analysis is the use of generic versus branded products. If the chemical entities and formulations are identical, then there is no reason to suspect that the outcomes associated with the use of either product would be different. In this case, the decision is based solely on the costs of the inputs. The difficulty with cost-minimization analysis is establishing that outcomes are equal. Even in the case of generic pharmaceuticals there are examples where alternative formulations have been questioned. Additionally, products may be equivalent on some outcomes such as clinically significant improvement in depression, but not with regard to others such as side-effect profiles.

Cost-effectiveness assesses the consequences in natural units. These natural units may include outcomes such as years of life saved, hospitalizations avoided, or scores on a symptom scale. Jonsson et al. (21) used the Mini-Mental State Examination as a mechanism to assess time in a nonsevere disease state. This information was incorporated into a Markov state transition model to compare the cost-effectiveness of newer medications for the treatment of Alzheimer's disease to standard care. In many cases it is possible to develop several cost-effectiveness ratios for a comparison of relevant treatment alternatives. For example, in a comparison of atypical and conventional antipsychotics, cost-effectiveness ratios such as cost per hospitalization avoided, cost per symptom free day, or cost per schizophrenic exacerbation might all have meaning.

A recent cost-effectiveness study for the treatment of depression provides an excellent example of how decision makers can utilize these tools to best allocate scarce resources. Nuijten and colleagues (8) developed a Markov process to model the cost-effectiveness of long-term treatment with a new antidepressant compared to standard treatment with TCAs. The outcomes were time without depression, direct costs, and indirect costs (lost workdays). Clinical data were obtained from the published literature and costs were measured from the perspective of the German health care system. The new antidepressant was found to be associated with a 1.5 months longer time without depression than the TCA and with less cost to the health care system. In this case the new drug was less costly and more effective; thus by definition it is a more cost-effective choice.

Cost-utility analysis is a special case of cost-effectiveness analysis, in which the denominator is quality-adjusted life years (or something conceptually similar). The quality-adjusted life year may, for example, be calculated using patient utilities (from zero to one) for being in a given health state (or series of health states) and multiply them by the number of years of life expected in each health state. This analysis benefits from combining length of life with quality of life. For example, people using different antipsychotic medications may have similar life expectancy. However, if patients taking some antipsychotics have fewer side effects or greater

efficacy, they may experience an improved quality of life. As a result, the quality of the remaining years of life may not be equal for the two treatments. Cost-utility ratios factor this quality difference into the analysis. A typical ratio presented for cost-utility analysis is a cost per quality-adjusted life year (QALY).

Cost-utility analyses have the ability to compare QALYs over multiple treatment regimens. Revicki et al. (22) completed a cost-utility analysis in a managed care setting and compared outcomes of two SSRIs, a TCA, and a stepped approach that began with a TCA that was replaced with an SSRI if the TCA treatment failed. The outcomes measures were lifetime medical costs, QALYs, and cost per QALY gained. The analysis for a base case found that lifetime medical costs ranged from $15,348 to $16,669 per patient, that QALYs gained ranged from 14.32 to 14.64, and that cost per QALY gained ranged from $2,555/QALY to $6,346/QALY. The model allowed certain factors, such as compliance, to be varied.

The final pharmacoeconomic method is cost-benefit analysis. Cost-benefit analysis values the denominator in dollars and calculates a return on investment. Cost-benefit analysis allows comparison of alternatives that lead to dissimilar outcomes. Should a hospital open a gift shop or provide a vaccination program for influenza? The answer would be provided in the following terms: for every dollar invested in a gift shop, there is a return of $1.13; for every dollar invested in a vaccination program, there is a return of $1.25. Thus the vaccination program would be the more attractive investment. A disadvantage of cost-benefit analysis is that it requires all consequences to be valued in dollars. For example, suppose use of a medical alternative increases survival by 1 year. How do you put a cost on 1 year of life? In health care, valuing in dollars such consequences as life years gained and disability days avoided may be considered difficult or unacceptable by many people.

These four methods form the cornerstone of pharmacoeconomics. There are, however, many issues that affect the conduct and interpretation of pharmacoeconomic analyses. Some of the principal issues are discussed in the following sections.

Costs

Valuation of costs in pharmacoeconomic analyses can be difficult. There are two primary issues: first, which costs should be included in an analysis, and second, how should those costs be valued. The costs of inputs in a pharmacoeconomic analysis typically include direct medical costs, direct nonmedical costs, and indirect costs. Direct medical costs include costs such as physician visits, hospitalization, emergency department use, and pharmaceuticals. Examples of direct medical costs associated with the treatment of Alzheimer’s disease include diagnostic tests, medications, and efforts to monitor or treat side effects, acute hospital care, physicians’ services, home health care, and nursing home care (23 ,24). In other words, direct medical costs include any costs that are directly related to medical treatment. Direct nonmedical costs include items such as cost of transportation to the doctor’s office and cost of child care while the parent is hospitalized. An example of a direct nonmedical costs for the treatment of Alzheimer’s disease is in-home day care (25). Examples of indirect costs are costs that arise from lost work or lost patient or caregiver productivity. In conducting a cost analysis, the first challenge is to decide which costs are relevant for the comparison. The issue of perspective (who pays?) becomes critical. Once this hurdle has been cleared, then the issue arises of how costs will be assigned. For example, if prescription costs during hospitalization are included in a pharmacoeconomic analysis, how should the basis for costs be established? Should it be based on actual acquisition costs, charges, or cost-to-charge ratios as a percentage of the entire hospital bill? Clearly, these types of valuation decisions need to be disclosed and discussed in pharmacoeconomic studies.

Analyzing the results of pharmacoeconomic studies requires the evaluator to assess the types of costs included in each study. Large variations in results can be attributable to different cost components (25). In the medical treatment of Alzheimer’s disease, the acquisition cost of the medication is only a small percentage of the total cost. One major component of treating Alzheimer’s disease is the indirect costs absorbed by family members. Researchers encounter difficulties in estimating the cost of such informal care. As a result, investigators may account for those costs absorbed by family members in different ways, which may contribute to varying conclusions (26).

There may also be intangible costs. These costs include things such as pain and suffering. Intangible costs are even more difficult to value in monetary terms. In diseases where intangible costs are significant, it is important to recognize whether any effort has been made to account for these costs. In many cases they are not included. Additional analysis may need to be considered to make fair decisions.

One mechanism to quantify intangible costs is a willingness-to-pay approach. O'Brien et al. (27) performed a willingness-to-pay evaluation in a group of patients with mild to moderate depression. The study was designed to compare a new antidepressant with TCAs. The drugs had similar efficacy but different adverse event profiles. Participants were asked to rank a series of adverse effects and then to quantify the maximum amount they would pay for a new drug that reduced each adverse event. On average, participants were willing to pay an additional $14 per month to reduce the risk of blurred vision from 10% to 5%. When asked their willingness to pay to avoid multiple simultaneous side effects, the range was $23 to $77 per month.

If costs, or benefits, are analyzed over time periods that exceed 1 year it is necessary to apply discounting. Discounting costs is a concept that reflects the “time value of money.”

A dollar today is worth more than a dollar received in the future. Discounting reduces the value of dollars that will be realized more than a year in the future to reflect a present value. A similar concept applies to health benefits.

Perspective

One of the major factors that influences pharmacoeconomic analyses is the perspective taken when conducting the analysis. Using the earlier cost example, if a study is conducted from the perspective of a hospital, the use of actual costs may be appropriate. However, if the same study were conducted from the perspective of a managed care organization, charge data may be more relevant. The perspective of a pharmacoeconomic analysis should always be disclosed in a publication. Given that there are many possible perspectives, it is insightful to evaluate studies from a broad perspective. The societal perspective is the broadest and takes into consideration all costs and consequences relevant to society. For various reasons, this perspective is often used as a "reference case," to permit comparisons across studies that may otherwise use differing perspectives. When measuring the impact of pharmaceuticals on mental health disorders, the perspectives of both society and providers are important (14 ,28 ,29). Because of the desire to serve the needs of health system decision makers, the perspective taken in many studies is that of the payer. A substantial number of patients prescribed antipsychotic medications have their health care paid by Medicaid (30). Therefore, the perspective of Medicaid is important when evaluating the cost of schizophrenia treatment. The payer perspective, however, may not include all costs relevant to society. For example, lost productivity may not be relevant from the Medicaid perspective. It is critical that the study perspective is disclosed when a pharmacoeconomic analysis is published or evaluated. Generalizing results from a specific setting to a different setting is unwise because the relevant costs and outcomes vary between settings.

CONDUCTING PHARMACEUTICAL OUTCOMES RESEARCH

Part of "39 - The Role of Pharmaceuticals in Mental Health Care Outcomes "

There have been several texts and journal articles describing the steps for conducting pharmaceutical outcomes research and pharmacoeconomics (19 ,20). This section does not reproduce these lists of steps, but rather presents some of the issues from these materials that are pertinent to the evaluation of mental health applications.

Decision Analysis

Decision analysis is a systematic approach to structuring decisions over time. Decision trees are developed with branches representing alternative decisions or probabilistic relationships. For example, a decision may involve a choice between two drugs. These drugs may either have side effects or not, and treatment may be either successful or not successful (Fig. 39.1).

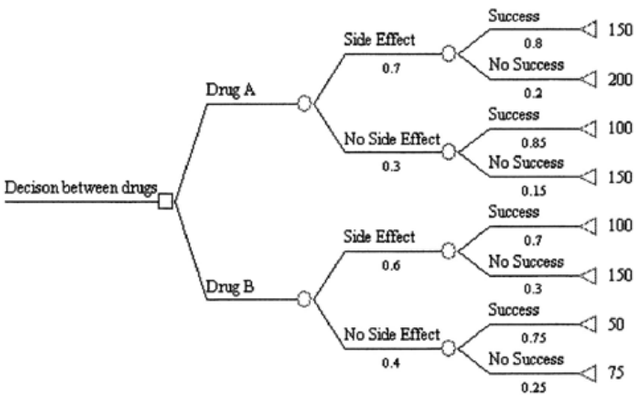


FIGURE 39.1. Example of a decision tree.

For any of the nodes where a chance relationship exists (represented by circles), the probabilities associated with that event are shown (e.g., the probability of a side effect associated with drug A is 0.7). When the outcomes are valued in dollars (i.e., the cost of following a particular path), the expected values that are calculated from a decision tree analysis can provide cost estimates that are used in the numerator of a pharmacoeconomic ratio. Expected values are calculated by summing the product of the probabilities and the costs. The expected values are shown in Fig. 39.2 .

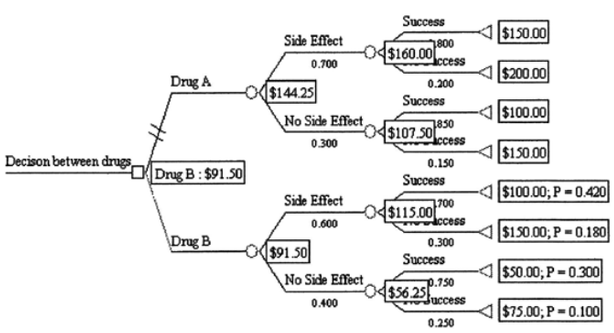


FIGURE 39.2. A solved decision tree.

In this case, assuming that the cost at the end of each branch represents the total cost of care for the selected drug, the expected costs associated with the use of drug A is $144.25 and $91.50 for drug B. If outcomes were assumed to be equal (i.e., cost-minimization) then the least expensive alternative would be drug B. However, outcomes are not equal. As can be seen from the decision tree, the probability of success while using drug A is greater than for drug B. If the path probabilities for the successful branches are summed for each alternative, it can be seen that the probability of successful treatment with drug A is 0.815, and the probability of successful treatment with drug B is 0.720. Therefore, the total cost of treatment while on drug A is higher, but so is the effectiveness (i.e., chance for successful treatment). The real question is one of cost-effectiveness. Is the additional cost of drug A worth the additional benefits? This would be assessed with an incremental cost effectiveness ratio:

$$\text{Incremental cost effectiveness ratio} = \frac{\text{Cost of Drug A} - \text{Cost of Drug B}}{\text{Effectiveness Drug A} - \text{Effectiveness Drug B}}$$

$$\text{Incremental cost effectiveness ratio} = \frac{144.25 - 91.50}{0.815 - 0.720}$$

$$\text{Incremental cost effectiveness ratio} = \$555.26/\text{each additional successful case}$$

Is an additional successful case worth paying an additional $555.26? This is a value judgment that depends on the situation surrounding the decision. Again, outcomes research does not answer the question regarding which product should be used; it simply provides information regarding the efficiency with which these two products produce a desired outcome.

This method, while powerful, is often limited by the availability of data to drive the model and the forced simplification of models, which often results from limited driving data. It is important when reading published modeling exercises such as decision trees or Markov models to evaluate them carefully, as results are highly dependent on the specifics of the structure selected (and how closely it reflects clinical reality) and on the data selected for input. The latter problem can and should properly be addressed by sensitivity analysis, for which there are several techniques. The former problem can be tested by the careful evaluation of experts or the demonstrated ability of a model to predict measurable outcomes, a relatively uncommon exercise in pharmacoeconomics.

In general, the reader should look for some effort to discuss or examine “parameter” uncertainty, “model structure” uncertainty, and “model process” uncertainty. With regard to parameter uncertainty (the term *parameter* refers, for example, to estimates of probabilities or cost or health outcome), univariate analysis alone is often inadequate, and some attempt at multivariate evaluation is desirable. There are different formal approaches to evaluation of cost-effectiveness uncertainty using either frequentist or Bayesian approaches to generation of confidence (or “credible”) regions, including simulation and the delta method. Model structure uncertainty refers to the separate uncertainty about the manner in which parameters should properly be combined (e.g., Are effects linear or nonlinear fashion? Are effects additive or multiplicative?). An approach to evaluation of this kind of uncertainty is to simply examine results for different plausible alternatives. Model process uncertainty refers to the fact that different analysts may come to different conclusions due to a spectrum of differences in approach. There are other key methodologic pitfalls in proper conduct of studies of this type (e.g., adhering to use of incremental cost-effectiveness ratios), a fact that highlights the need for readers to evaluate each study carefully.

The purpose of this discussion is not to fully explain the sometimes complex process of conducting a decision analysis, but rather to suggest the usefulness of this tool in outcomes research. It is also to encourage the reader to develop the critical skills necessary to evaluate studies of this type, much as similar skills have been developed for evaluation of ordinary controlled clinical trials. The reader is referred to *Clinical Decision Analysis* by Weinstein and Fineberg (31) for a more complete description of the process of conducting decision analyses, and to *Cost-Effectiveness in Health and Medicine*, edited by Gold (32) for in-depth discussion of many important issues in cost-effectiveness analysis.

In a final example, we refer to a study using decision analysis to permit evaluation of a systematically developed time-ordered series of events. It incorporates previously published clinical trial data into a model to estimate long-term effects. In this study, Dardennes et al. (33) developed a decision analysis model to compare outcomes and costs of treating major depression with an SSRI, a TCA, or serotonin norepinephrine reuptake inhibitor (SNRI). The perspective of the study was that of a national health care system, and the clinical outcomes data used in the model were derived from published meta-analyses. The analysis found that the SSRI and TCA had comparable efficacy but dissimilar tolerance profiles and that the SNRI had both efficacy and tolerance advantages compared to the SSRI. Direct cost data (hospitalization, medication, physician visits, and laboratory tests) and the efficacy data from the model were entered into a decision tree. The decision tree analysis provided estimates of the expected cost of treatment per

depressive episode that could be used by the health service in its treatment approval process.

Data Sources

Data for pharmaceutical outcomes studies can come from several sources. Many pharmaceutical companies routinely include pharmaceutical outcome (other than just clinical) measurements in their development trials. In addition, postmarketing studies with direct comparisons to relevant alternatives (i.e., intended to serve the needs of pharmaceutical users rather than regulators) are becoming more common. However, conducting single studies that contain all pharmaceutical outcomes of interest for a given product are expensive, and data collection of all relevant information is difficult. Therefore, many pharmaceutical outcome studies contribute to the body of knowledge by evaluating components of the overall picture. Additionally, many pharmacoeconomic analyses are based on models. These models typically use published literature, expert opinion, or data from administrative or encounter databases to get information on probabilities and costs.

The impact of this component approach to building an understanding of pharmaceutical outcomes is that data come from many sources ranging from experimental and nonexperimental research designs to expert opinion and models based on data from multiple and frequently diverse sources. Therefore, when reviewing pharmaceutical outcomes research, it is critical to understand the potential impact of the source of information on the results.

A frequent source of outcomes data in mental health research is randomized clinical trials conducted by the pharmaceutical industry. These trials, however, are often placebo controlled and typically contain (as expected) mostly clinical information. In mental health, however, patient self-reported items (i.e., humanistic measures) are frequently included. There are also many studies that rely on chart review and quasi-experimentation to document differences in resource use for patients using various pharmaceutical agents. Examples include recent comparisons of tricyclic antidepressants and SSRIs, and atypical versus conventional antipsychotic agents. Many of these studies were retrospective and were conducted through chart reviews or administrative data using quasi-experimental techniques. Finally, economic models have been built using published data or expert opinion, where data were not available.

Historically, randomized controlled trials have been the “gold standard” (5). Studies of this type allow the efficacy and safety of a drug to be established. Unfortunately, some of the strengths of such studies can also be a source of less commonly recognized weaknesses. This is a result of the artificial treatment environment purposefully created in efficacy trials, and may be particularly an issue in mental health because of the wide gap recognized to exist between the “optimal” care provided in such trials and the realistic patterns of care experienced by most patients. The primary care provider deals with other issues that influence the effects of a medication such as side effects, dose titration, and out-of-pocket expenses. As a result, real-world effects can be difficult to extrapolate from ordinary clinical trials. This issue is discussed further below. Randomized trials failed to differentiate the SSRIs and TCAs, except for their side-effect profiles. However, SSRIs may have some advantage over TCAs in the primary care practice setting (6 ,28 ,29). In summary, data for building evidence about the value of pharmaceutical outcomes in mental health has been drawn from a number of sources using a variety of experimental and nonexperimental designs. Each of these sources of data and type of experimentation affect the degree of evidence obtained. Review of pharmaceutical outcomes research in mental health care requires careful consideration of the source and strength of the evidence presented.

HUMANISTIC MEASURES

Part of "39 - The Role of Pharmaceuticals in Mental Health Care Outcomes "

Humanistic measures assess how disease or treatment affects patients. Humanistic measures are most important from the perspective of the patient. A primary goal for treatment of any disease should be for patients to function normally, have an acceptable quality of life, and be satisfied with their treatment. This is especially true for mental health disorders where impacts on both physical and social functioning may be significant. In many cases, patients and their friends and families might best judge the success of treatment. Until recently, humanistic measures have taken a back seat to traditional clinical measures and to some extent economic measures. This is in part due to greater variability from patient self-reported measures compared to standard clinical measures (34). The development of valid and reliable instruments is a relatively recent phenomenon. The most common conceptualization of humanistic outcomes used in the evaluation of pharmaceuticals is health-related quality of life.

Health-related quality of life encompasses factors such as functional status, physiologic status, social and emotional well-being, and life satisfaction (35). Health-related quality of life information allows health care providers and payers to make decisions based not only on clinical effectiveness, or costs but also on effects that are important to patients. Measurement of health-related quality of life may be especially important in chronic diseases for which we have no cure. There are many humanistic measures available for assessing mental health disorders. Generic and disease-specific instruments are available for a variety of disorders. Discussion of every instrument is not feasible; however, a few examples are provided.

Generic instruments are global in content and cover a number of dimensions relevant to overall health-related quality of life. One of the most widely used generic health-related

quality of life instruments is the Medical Outcomes Study Short Form 36 (MOS SF-36). The MOS SF-36 captures eight dimensions of health-related quality of life: physical functioning, role limitations due to physical functioning, bodily pain, general health, vitality, social functioning, role limitations due to emotional problems, and mental health (36).

A modified version of the Sickness Impact Profile has also been developed for use in patients with mental illnesses.

Disease-specific health-related quality of life instruments focus on dimensions that are most relevant to the particular disease and are therefore more sensitive to subtle changes in the disease or its treatment. Examples of disease specific instruments used in schizophrenia include the Quality of life Scale (37), the Social Performance Schedule (38), and the Quality of life Interview (39). Several review papers have been published on the use of quality of life instruments in mental health conditions (40 ,41 and 42). These articles highlight that quality of life measurement in mental health conditions, and in particular specific drug comparison, is a developing science. Many of the available studies are observational or cross-sectional. However, quality of life measurement is increasingly being built into clinical trials. There continues to be debate regarding mental health patients' ability to complete quality of life questionnaires, highlighting the importance of population-specific assessment of instrument validity and reliability. However, several articles have shown that it is possible for patients with severe mental health problems to successfully complete these forms (43 ,44). Lenert's group (45) has shown that even when posed the conceptually challenging task of the standard gamble, patients with mental illness have been able to perform adequately. It is important to note that although agreement is not universal, there are many researchers who believe that in principle the best source for measuring patients' quality of life or preferences is the patients themselves whenever such measurement is possible. Much has been published on this subject and in particular on the issue of whose values to use in creation of reference case analyses (to be used for comparisons across studies). But rarely is the view of the health care provider or other proxy considered superior to that of either the patient or of society in general.

Health-related quality of life has many applications in the treatment of mental health disorders. For example, Simon et al. (46) completed an analysis that evaluated clinical effectiveness, health-related quality of life, and economics of treating depression. The study took place within a staff model HMO and utilized net costs. Patients starting new antidepressant therapy were randomized to an SSRI or TCA for 24 months. The primary care providers were allowed to adjust doses and medications or discontinue medications as they deemed appropriate. The quality of life outcome was measured using the Medical Outcomes Study SF-36 Health Survey at 6, 9, 12, 18, and 24 months. The results indicated no significant difference in quality of life or severity of depression when comparing treatment groups (46). Other studies have evaluated health-related quality of life in the treatment of depression and utilized similar generic rating scales (47).

Numerous Alzheimer's disease–specific quality of life tools are available. However, there is a lack of understanding of how to quantify changes in scores. It is important to note that the cognitive impairment of Alzheimer's disease at times requires the administration of the tool to a care provider. The tools assess functions such as daily activities, memory, emotional well-being, and other aspects such as finances (48).

One other area of humanistic measurement concerns the relationship between humanistic and economic outcomes. Cost utility analysis uses patient preferences in the form of utilities to combine cost information with patient preferences. Utilities are usually measured by three techniques: rating scales, the standard gamble, or time trade-off technique (49). Utility scores differ from quality of life measurements. While some quality of life instruments can be used to capture utilities, most cannot. Utilities are a measure of overall patient well-being that lie on a scale between 0 and 1. Utilities typically measure the difference in patient's preferences between perfect health and impaired health states. At present, utilities have been measured for only a few health states. Unfortunately, utility values are difficult and expensive to measure. They require detailed patient interviews with large numbers of subjects with and without the disease. Additionally, there are many questions about patient's ability to give reliable and valid responses. While these techniques have been used in mental health care, more widespread use is dependent on the development of reliable and valid measures of utility or preference for alternative health states in mental health diseases.

USING OUTCOMES DATA IN PRACTICE

Part of "39 - The Role of Pharmaceuticals in Mental Health Care Outcomes "

The use of outcomes data in practice is not about applying the results of a single study. Instead, using outcomes data typically requires synthesis across a body of literature. Outcomes data, and in particular economic and humanistic data, offer additional pieces of information that should be incorporated into decisions. Economic, clinical, and humanistic data are all needed to make fair evaluations of pharmaceutical products and services. In reality, however, decisions will be made even if all these data are not available.

Uses of outcomes data in practice include reimbursement decisions, internal practice decisions, external or regulatory decisions, and marketing of pharmaceutical products. Pharmacy and therapeutics committees are using outcomes data as a component of the formulary decision. Where these decisions were once made almost entirely on clinical parameters, the use of economic and humanistic data is becoming more common.

In practice, the use of terminology such as evidence-based medicine or treatment guidelines has its roots in outcomes evaluations. The evaluation of a body of literature to make decisions about best practice is the goal of evidence-based medicine. Evidence-based medicine involves explicit use of what can be identified as the best evidence in making decisions about the care of both individual patients and populations of patients (Fig. 39.1) (50). This philosophy extends into treatment guidelines that are often established by expert panels that have reviewed the available evidence in the literature regarding effectiveness of alternative treatments. While these efforts rely most heavily on clinical information, economic and humanistic data are being included in these considerations.

Outcomes data is beginning to be considered in the accreditation of health care organizations. Although the measures currently used are more process than outcomes oriented, the evolution toward outcomes can be seen. The National Committee for Quality Assurance (NCQA) conducts accreditation of managed care organizations and has a specific program for behavioral health accreditation. NCQA also sponsors the Health Plan Employer Data and Information Set (HEDIS) report, which is a set of standardized performance measures designed to assist consumers with decisions about purchasing health care coverage. HEDIS 2000 includes several measures relative to mental health care. These measures are organized into several categories. Under the effectiveness of care category, two measures are included: follow-up after hospitalization for mental illness, and antidepressant medication management. In the use of services category, mental health care related measures include mental health utilization, inpatient discharges and average length of stay, and mental health utilization–percentage of subjects receiving services. These measures are evolving to require managed care organizations to consider the outcomes of care they provide. As HEDIS measures continue to evolve, they are expected to raise the quality of health care.

Some of the more sophisticated users of outcomes data may be pharmaceutical companies. Most major pharmaceutical manufacturers are investing resources in departments that focus on the collection and analysis of outcomes data for their products. Although these data are frequently used in the marketing of pharmaceutical products, they are also providing information about the developing science of outcomes measurement.

Efficacy and Effectiveness

The evolution of the use of data for decision making is interesting. Health care organizations have evolved from requiring evidence of efficacy to effectiveness to efficiency. *Efficacy* is defined as how well an intervention can ideally work. This is characterized by the types of ordinary clinical trials that are commonly performed for regulatory purposes. This type of trial generally creates what is believed to be ideal conditions of treatment and limits treatment to optimal patients, conditions that can be difficult or impossible to duplicate in regular practice. Blind, prospective randomization of an adequate number of patients to a study in which outcomes are assessed by raters blind to treatment is intended to minimize observer bias and confounding, and maximize internal validity. Accordingly, clinical trials are excellent for providing confidence that there is a causal relationship between drug use and the measured endpoint. Once confidence in this relationship is established, however, questions of use in the real world arise, which beg the question of effectiveness. Efficacy results, generally using highly select, often healthier, patient populations (not least because informed consent is required) under different practice conditions (tertiary vs. primary care, frequency of follow-up, compliance issues, insurance issues, dosing/titration regimens, etc.), may simply not predict patient outcomes in usual practice. This leads to the question of *effectiveness*, defined as the extent to which health improvements are achieved in real practice settings. Does a pharmaceutical product work under real-world conditions? Is it really influencing outcomes that are important to patients, payers, and clinicians? An experimental research approach to the question of effectiveness is a type of trial referred to variously as “pragmatic trials” or “effectiveness trials.” Such trials employ random treatment assignment to address confounding, but sacrifice a degree of internal validity in an effort to increase generalizability to the real world. This necessitates, for example, inclusion of as many patients as possible (minimizing exclusions), using ordinary practice settings, avoiding protocol-mandated interference in patient care, and permitting the effects of cost and payment mechanisms. The application of this useful type of trial in mental health has been described, but rarely implemented (51). Alternatively, nonexperimental research designs (e.g., cohort studies, retrospective database analyses, etc.) can often shed light on real-world outcomes when a clinical trial is impractical for reasons such as sample size, informed consent, duration of follow-up, etc. Usually, such studies must take special care to address issues of bias and confounding. When real-world data from either effectiveness trials or nonexperimental research are not available, as is often the case, the evidence basis for mental health decision makers is limited to either attempting to generalize from efficacy data or relying on some form of expert opinion.

The increasing expenditures associated with mental health disease states require decision makers to evaluate the full impact of treatment alternatives. The evaluation should include the appropriate variables to fully evaluate patient outcomes (including quality of life); an adequate evaluation of all relevant costs, which permits capture of potential offsets of simple drug acquisition costs; and consideration of issues of efficacy vs. effectiveness. The tools of pharmacoeconomics

and outcomes research provide decision makers with a mechanism for attempting to quantify and balance these factors to assist in the allocation of scarce mental health resources.

CONCLUSION

Part of "39 - The Role of Pharmaceuticals in Mental Health Care Outcomes "

Measurement of economic, clinical, and humanistic outcomes is an important tool for establishing the value of competing mental health care programs and treatments. Ultimately, measures of quality of care in relation to commensurate costs should aid decisions about which programs to implement and which treatments to reimburse. Although no single study is likely to provide an answer, careful evaluation of the economic, clinical, and humanistic outcomes literature may assist decision makers in making more informed decisions. Many of the economic and clinical studies conducted to date use descriptive designs or apply modeling techniques based on the best source of available data. As the science behind outcomes measurement evolves, the level of sophistication of the information provided will improve. Information on treatment outcomes can contribute significantly to decisions that affect the quality of care received by mental health patients.

REFERENCES

1. Donabedian A. Evaluating the quality of medical care. *Milbank Memorial Fund Q* 1966;44(3, part 2):166–206.

2. Fuchs VR. *Who shall live?: health, economics and social choice*. New York: Basic Books, 1975.

3. Lave JR, Frank RG, Schulberg HC, et al. Cost-effectiveness of treatments for major depression in primary care practice. *Arch Gen Psychiatry* 1998;55(7):645–651.

4. Kamlet MS, Paul N, Greenhouse J, et al. Cost utility analysis of maintenance treatment for recurrent depression. *Control Clin Trials* 1995;16(1):17–40.

5. Stewart A. Choosing an antidepressant: effectiveness based pharmacoeconomics. *J Affect Disord* 1998;48(2–3):125–133.

6. Forder J, Kavannagh S, Fenyo A. A comparison of the cost-effectiveness of sertraline versus tricyclic antidepressants in primary care. *J Affect Disord* 1996;38(2–3):97–111.

7. Revicki DA, Brown RE, Palmer W, et al. Modelling the cost effectiveness of antidepressant treatment in primary care. *Pharmacoeconomics* 1995;8(6):524–540.

8. Nuijten MJ, Hardens M, Souetre E. A Markov process analysis comparing the cost effectiveness of maintenance therapy with citalopram versus standard therapy in major depression. *Pharmacoeconomics* 1995;8(2):159–168.

9. Simon GE, VonKorff M, Heiligenstein JH, et al. Initial antidepressant choice in primary care. Effectiveness and cost of fluoxetine vs tricyclic antidepressants. *JAMA* 1996;275(24):1897–1902.

10. Kind P, Sorensen J. Modeling the cost-effectiveness of the prophylactic use of SSRIs in the treatment of depression. *Int Clin Psychopharmacol* 1995;10(suppl 1):41–48.

11. Sclar DA, Skaer TL, Robinson LM, et al. Economic outcomes with antidepressant pharmacotherapy: a retrospective intent-to-treat analysis. *J Clin Psychiatry* 1998;59(suppl 2):13–17.

12. Tome MB, Isaac MT. Cost-benefit and cost-effectiveness analysis of the rapid onset of selective serotonin reuptake inhibitors by augmentation. *Int J Psychiatry Med* 1997;27(4):377–390.

13. Kozma CM, Reeder CE, Schulz RM. Economic, clinical and humanistic outcomes: a planning model for pharmacoeconomic research. *Clin Ther* 1993;15(6):1121–1132.

14. Sclar DA, Skaer TL, Robinson LM, et al. Economic appraisal of antidepressant pharmacotherapy: critical review of the literature and future directions. *Depress Anxiety* 1998;8(suppl 1):121–127.

15. Ernst RL, Hay JW. The US economic and social costs of Alzheimer's disease revisited. *Am J Public Health* 1994;84(8):1261.

16. Buckley PF. Treatment of schizophrenia: let's talk dollars and sense. *Am J Man Care* 1998;4:369–383.

17. Glazer WM, Johnstone BM. Pharmacoeconomic evaluation of antipsychotic therapy for schizophrenia. *J Clin Psychiatry* 1997;58(suppl 10):50–54.

18. Mitchell J, Greenberg J, Finch K, et al. Effectiveness and economic impact of antidepressant medications: a review. *Am J Manag Care* 1997;3(2):323–330.

19. Bootman JL, Townsend RJ, McGhan WF. *Principles of pharmacoeconomics*. Cincinnati: Harvey Whitney, 1991.

20. Drummond MF, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford: Oxfor University Press, 1989.

21. Jonsson L, Lindgen P, Wimo A, et al. The cost-effectiveness of donepezil therapy in Swedish patients with Alzheimer's disease: a Markov model. *Clin Ther* 1999;21(7):1230–1240.

22. Revicki DA, Brown RE, Keller MB, et al. Cost-effectiveness of newer antidepressants compared with tricyclic antidepressants in managed care settings. *J Clin Psychiatry* 1997;58(2):47–58.

23. Neumann PJ, Hermann RC, Berenbaum PA, et al. Methods of cost-effectiveness analysis in the assessment of new drugs for Alzheimer's disease. *Psychiatr Serv* 1997;48(11):1440–1444.

24. Meek PD, McKeithan K, Schumock GT. Economic considerations in Alzheimer's disease. *Pharmacotherapy* 1998;18(2 pt 2):68–73.

25. Small GW, Donohue JA, Brooks RL. An economic evaluation of donepezil in the treatment of Alzheimer's disease. *Clin Ther* 1998;20(4):838–850.

26. Molnar FJ, Dalziel WB. The pharmacoeconomics of dementia therapies. Bringing the clinical, research and economic perspectives together. *Drugs Aging* 1997;10(3):219–233.

27. O'Brien BJ, Novosel S, Torrance G, et al. Assessing the economic value of a new antidepressant. A willingness-to-pay approach. *Pharmacoeconomics* 1995;8(1):34–45.

28. Boyer P, Danion JM, Bisserbe JC, et al. Clinical and economic comparison of sertaline and fluoxetine in the treatment of depression. A 6-month double-blind study in a primary-care setting in France. *Pharmacoeconomics* 1998;13(1 pt 2):157–169.

29. Sclar DA, Robinson LM, Skaer TL, et al. Antidepressant pharmacotherapy: economic evaluation of fluoxetine, paroxetine and sertraline in a health maintenance organization. *J Int Med Res* 1995;23(6):395–412.

30. Nightengale BS, Crumley JM, Liao J, et al. Economic outcomes of antipsychotic agents in a Medicaid population: traditional agents vs. risperidone. *Psychopharmacol Bull* 1998;34(3):373–382.

31. Weinstein MC, Fineberg HV. *Clinical decision analysis*. Philadelphia: WB Saunders, 1980.

32. Gold MR, ed. *Cost-effectiveness in health and medicine*. Oxford: Oxford University Press, 1996.

33. Dardennes R, Berdeaux G, Lafuma A, et al. Comparison of the cost-effectiveness of milnacipran (a SNRI) with TCAs and SSRIs: a modeling approach. *Eur Psychiatry* 1999;14(3):152–162.

34. Busschbach JJ, Brouwer WB, Van Der Donk A, et al. An outline for a cost-effectiveness analysis of a drug for patients with Alzheimer's disease. *Pharmacoeconomics* 1998;13(1 pt 1):21–34.

35. MacKeigan LD, Pathak DV. Overview of health-related quality of life measures. *Am J Hosp Pharm* 1992;49:2236–2245.

36. Ware JE, Snow KK, Kosinski MA, et al. *SF-36 Health Survey Manual and Interpretation Guide*. Boston: Nimrod, 1993.

37. Heinrichs DW, Hanlon TE, Carpenter WT. The Quality of Life Scale: an instrument for rating schizophrenic deficit symptoms. *Schizophr Bull* 1984;10:388–398.

38. Baker F, Intagliata J. Quality of life in the evaluation of community support systems. *Eval Program Plann* 1982;5:69–79.

39. Lehman A. The well-being of chronic mental patients: assessing their quality of life. *Arch Gen Psychiatry* 1983;40:369–373.

40. Awad AG, Voruganti LNP, Heslegrave RJ. Measuring quality of life in patients with schizophrenia. *Pharmacoeconomics* 1997;11(1):32–47.

41. Revicki DA, Murray M. Assessing health-related quality of life outcomes of drug treatments for psychiatric disorders. *CNS Drugs* 1994;1:465–476.

42. Whalley D, McKenna SP. Measuring quality of life in patients with depression or anxiety. *Pharmacoeconomics* 1995;8:305–315.

43. Honigfeld G, Patin J. A 2-year clinical and economic follow-up of patients on clozapine. *Hosp Community Psychiatry* 1990;41:882–885.

44. Van Putten T, May PRA, Marder SR, et al. Subjective response to antipsychotic drugs. *Arch Gen Psychiatry* 1981;77:1417–1426.

45. Lee TT, Ziegler JK, Sommi R, et al. Comparison of preferences for health outcomes in schizophrenia among stakeholder groups. *J Psychiatr Res* 2000;34(3):201–210.

46. Simon GE, Heiligenstein J, Revicki D, et al. Long-term outcomes of initial antidepressant drug choice in a “real world” randomized trial. *Arch Fam Med* 1999;8(4):319–325.

47. Revicki DA, Simon GE, Chan K, et al. Depression, health-related quality of life, and medical cost outcomes of receiving recommended levels of antidepressant treatment. *J Fam Pract* 1998;47(6):446–452.

48. Kerner DN, Patterson TL, Grant I, et al. Validity of the Quality of Well-Being Scale for patients with Alzheimer’s disease. *J Aging Health* 1998;10(1):44–61.

49. Torrance GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 1986;5:1–30.

50. Sackett DL, Rosenburg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn’t. *BMJ* 1996;312(7023):71–72.

51. Mahmoud R, Engelhart L, Ollendorf D, et al. The Risperidone Outcomes Study of Effectiveness (ROSE): a model for evaluating treatment strategies in typical psychiatric practice. *J Clin Psychiatry* 1999;60(suppl 3):42–48.

40

Issues in Clinical Trial Designs

John M. Kane

J. M. Kane: Department of Psychiatry, Hillside Hospital, Glen Oaks, New York; Department of Psychiatry and Neuroscience, Albert Einstein College of Medicine, Bronx, New York.

The introduction of the randomized, double-blind, clinical trial was one of the major advances in the development of medical science. In the arena of psychotropic drug development this approach has proven to be of enormous value in advancing a field in which laboratory tests and strictly objective methods for diagnosis and outcome assessment are not currently available.

Designing trials in the treatment of schizophrenia highlights some of these challenges. Schizophrenia is a complex illness affecting to varying degrees a range of functions, including cognition, affect, behavior, mood, and motivation. The fact that this disorder affects so many different domains, varying from individual to individual, and to some extent within individuals over time, makes development of pharmacologic treatments even more challenging. Although there are core features of schizophrenia that involve perception (hallucinations), cognition (attention, working memory, etc.), motivation (avolition), inferential reasoning (delusions), and affect (blunted or inappropriate), there is no pathognomonic sign or symptom of the disease. This has important implications for the diagnostic process, which is also complicated by the fact that the evaluation of some core features (e.g., hallucinations and delusions) relies solely on subjective reporting, the accuracy of which is potentially influenced by the very symptoms themselves as well as by other social situational and personality variables.

In addition, the fact that such an array of domains and functions is disturbed in this illness creates a challenge for drug development. The tendency has been to conduct an array of assessments to evaluate drug effects in a number of domains concurrently, when in fact different domains may require different study designs, patient selection criteria, durations of treatment, etc. In the future, more attention will be given to those issues, and it is possible that multiple treatments will be studied rather with the goal of finding combinations able to improve outcome across a variety of domains. It is hoped that new treatments will be developed with a focus on specific domains such as negative symptoms and cognitive dysfunction. Although a better understanding of basic mechanisms should facilitate further treatment advances, our current knowledge of pathophysiology remains limited. Advances in imaging techniques and pharmacogenomics are also important potential developments on the horizon that could have enormous impact on drug development and clinical evaluation.

Each area of psychotropic drug development has its own challenges in terms of rates of spontaneous remission, placebo response, patterns of relapse, domains of assessment, etc., but, in general, challenges of design and methodology involve issues that cut across the diagnostic domains.

DESIGN ISSUES

Part of "40 - Issues in Clinical Trial Designs "

There are a number of critical issues in general design that need to be addressed in both the individual study as well as the particular program of drug development. A drug development program needs to be comprehensive as well as adaptive so that early results can inform subsequent evaluation. Although even when a drug is marketed there are still limitations in the amount of knowledge available to clinicians, several fundamental questions should have been at least partially addressed: (a) What benefits are likely to result from the drug? (b) What are its risks? (c) What dosage is indicated? (d) How does the new drug compare to alternative treatments? (e) Are there specific patients most likely to benefit from the drug?

There are a number of specific concerns that should be addressed when designing clinical trials of psychotropic drugs. Some of the most salient issues include dose finding; efficacy vs. placebo; efficacy vs. a standard reference compound; acute and long-term adverse effects; continuation and maintenance treatment efficacy; and relative efficacy or adverse effects in specific subgroups (e.g., early-phase illness, late-phase illness, refractory patients).

Dose-finding tolerability studies involving antipsychotic medications generally call for involvement of target patient populations earlier in the process than with other classes of drugs because it is difficult to ethically justify administering

these drugs to healthy volunteers for more than a week or two, and patterns of tolerance may be quite different in patients versus healthy volunteers.

It is not always possible to accurately predict clinical dosage requirements from preclinical studies; therefore, it is important to establish a full range of tolerable dosages in order to provide an appropriate range for efficacy studies. Drug development programs have been delayed and at times abandoned because of inadequate dose-finding efforts in the early stages of development (1). In addition, it is not unusual for dosage recommendations to change after a drug is marketed.

It is also important to have sufficient data on absorption, elimination, metabolism, and drug–drug interactions, to inform trial design.

Treatment trials generally fall into three broad categories: acute, continuation, and maintenance (or relapse prevention). Sometimes attempts are made to study two or even three phases in the same trial, but controversy surrounds the need to rerandomize patients before drawing conclusions about relative efficacy in maintenance-phase treatment. Patient characteristics may vary somewhat in terms of desirability within specific trials, but overall the following issues should be considered.

Patient Characteristics

It is important to be clear on whether or not patients are in a state of acute relapse or exacerbation as opposed to partial remission or a “stable plateau” of chronic symptomatology. At times investigators will withdraw patients from ongoing treatment in order to transition them to a clinical trial, resulting in some symptom exacerbation. The importance of these different approaches is that they may result in patients with very different degrees of drug responsiveness, different patterns of baseline symptomatology, and varying degrees of “stability” in baseline symptomatology.

The ideal sample of patients is probably those who have not already been partially treated so that the full degree and time course of response can be determined. However, given the way that subjects must be ascertained and recruited for trials, it is likely that some treatment will have already been administered. The fact that participants have been partially treated or are in a chronic symptomatic state does not necessarily preclude the detection of a subsequent, clinically significant drug effect, but it is likely that the nature and magnitude of the effect will be altered.

The subjectivity of many components of symptomatology in psychiatric disorders creates special challenges. Given the fact that many symptoms are subjective and cannot be confirmed or quantified using objective measures, the assessment of baseline status can be difficult. Clinicians are particularly familiar with patients suffering from psychoses who are more open and explicit about pretreatment psychopathology once they begin to improve. Some patients may not appear eligible for a trial or be willing or able to give informed consent until they are partially treated.

A variety of subject characteristics should be considered in terms of inclusion and exclusion criteria. Specific decisions will be influenced by the nature and goals of the particular trial.

Age is often a basis for exclusion (either too young or too old). Age can certainly affect pharmacokinetics of particular drugs. The elderly are more likely to have comorbid medical conditions and be more sensitive to some adverse effects, and there are a variety of issues when young patients are included in trials. These and other factors have led to a paucity of subjects at the extreme age ranges in clinical trials. However, there has recently been increased recognition of the need for more early data on diverse age groups, and mechanisms are being implemented to encourage their inclusion in clinical trials.

Gender can be an important variable, and women are often underrepresented in clinical trials.

Ethnicity may have implications for drug metabolism and tolerance. In addition, as pharmacogenomic strategies are developed to extend clinical trial data, more accurate documentation of race will be critical.

Marital status can be a proxy for psychosocial adjustment and illness course, and may therefore be of prognostic significance.

Weight and body mass index have become an increasing concern from a public health standpoint and because of the considerable weight gain observed with some psychotropic drugs and in particular several new-generation antipsychotic medications (2).

Diagnostic subtype can be important in helping to characterize those patients most likely to benefit from specific treatments. Duration of illness and the duration of the current episode can be important in helping to define populations in terms of drug responsivity as well as long-term course and outcome. A particular problem in many trials is categorizing patients’ histories in terms of drug responsiveness. A current episode duration of more than 2 or 3 weeks could suggest that the patient is poorly or only partially responsive to the treatments that have already been administered, or, alternatively that some other factor is complicating treatment response (e.g., noncompliance, comorbid conditions, etc.). It can often be difficult to time the onset of illness or of a specific episode. As putative novel compounds are developed, it may become increasingly important to test these agents in patients who have not already been chronically exposed to other medications.

The specific type and severity of signs and symptoms required for entry into a trial will vary depending on the overall goals. Usually a minimal threshold of severity is established for core symptoms of interest. It is hoped that studies will also focus on patients selected on the basis of significant residual or secondary symptoms if they are associated with subjective distress and/or functional impairment.

If trials are designed to focus specifically on patients who were nonresponders or intolerant to other treatments, explicit criteria should be developed to identify such groups. There is debate as to whether or not a prospective trial is necessary to confirm treatment refractoriness, but this is certainly the most conservative approach because it also addresses to some extent the potential change in treatment milieu and attention resulting from participation in a research trial. In addition, there is enormous variability in the quality of retrospective assessment of treatment response.

Drug washout is a challenge in acutely ill patients. If some exacerbation in symptoms occurs, this complicates establishment of a baseline as well as adding to ethical concerns and management issues. On the other hand, absence of a washout means a true “baseline” is not achieved, assuming that there has been some degree of response and or adverse effects from the prior treatment. The use of a concurrent placebo group in the treatment trial mitigates these concerns to some extent, but does not eliminate them entirely. The type, dosage, and half-life of prior treatments will influence how long a washout is necessary to prevent potential withdrawal effects from influencing baseline ratings. Whether or not a washout takes place (and how long it is) can have implications for assessing the effects of subsequent treatment. The effects of withdrawal are neither consistent nor predictable, which complicates establishment of an appropriate baseline.

Premorbid social adjustment is a variable that does have prognostic significance, particularly in schizophrenia. Poor adjustment is associated with poorer outcome, and may be an indicator of those patients in whom early neurodevelopmental abnormalities or prodromal symptoms were more severe.

Comorbid psychiatric disorders should be evaluated and documented. Though there are insufficient data to determine what influence comorbid conditions are likely to have on overall response to psychotropic medications, common comorbid conditions should be studied at some point to help assure generalizability and to inform clinical practice. In addition, some studies tentatively suggest that different medications may have more or less impact on measures of, for example, substance abuse, suggesting that this could also be an important outcome measure in appropriate populations.

In studying antipsychotic medications it is important to document the presence and severity of any preexisting movement disorders in order to have an adequate baseline assessment and to ensure that a preexisting condition (or withdrawal effect) is not attributed to subsequent treatment.

It is essential that patients be assessed for their capacity to give informed consent. It is beyond the scope of this chapter to discuss this in great detail, but patients should be able to describe and explain in their own words the research in which they are agreeing to participate, its goals, its experimental aspects, and its potential risks and benefits. They must understand that they have the right to withdraw at any time and that they will not be penalized in any way if they choose to do so.

Trial Design

One of the most critical and difficult aspects of trial design is weighing and balancing what is ideal and what is feasible. An ideal trial for which patients cannot be recruited or in which they cannot be retained will not achieve its goals. In addition, though many questions ultimately need to be addressed, it is usually impossible to adequately address multiple questions in a single trial.

The duration of a trial will be influenced by whether or not a placebo group is included. The longer the duration, the more difficult to justify the retention of patients on placebo, and the higher the dropout rate, the less useful are the data.

The time course of response to psychotropic medication is generally variable. The modal time frame of response has to be factored into trial design in order to allow estimates of statistical power. In the acute treatment of schizophrenia, for example, most patients will experience at least half of the ultimate degree of improvement within the first 4 to 6 weeks (assuming that there was not an inordinately long titration phase). In many studies a significant drug effect is seen after only 1 to 2 weeks; however, different signs and symptoms are likely to have a different time course of response. For example, agitation is likely to respond more rapidly than delusions or thought disorder. In addition, there may be a subgroup of patients who are slower to respond, and for such patients longer trials may be needed. If a between-drug comparison of the full extent of response is ultimately important, then much longer trials are needed (e.g., 6 months or longer), and this begins to encompass the continuation phase of treatment. As more and more domains of outcome are of interest in clinical trials (such as primary negative symptoms or cognitive dysfunction in schizophrenia), it will be important to better characterize the time course of response for these variables in order to establish minimum and optimum durations of trials for these purposes. Estimates of expected degrees of improvement in various domains will be critical for statistical power calculations.

The Role Of Placebos

The decision as to whether or not to use a placebo in short-term, acute trials remains a topic of considerable controversy, and some dynamic tension continues to exist between “regulatory” requirements, investigators, institutional review boards, patients and families, and other interested parties. There are a number of important arguments that can be made against the routine use of a placebo in clinical trials. Rothman and Michels (3) argue that when an effective

treatment exists for a particular disease, the use of a placebo is inappropriate on both logical and ethical bases. However, the argument suggests that the use of a placebo is appropriate in cases when an effective treatment is not available. A problem remains in how to define effectiveness. The use of the term *effective* in this context is not necessarily identical to the current use of *effectiveness* as differentiated from *efficacy*.

In a complex disease such as schizophrenia, we continue to struggle with establishing the most meaningful definitions of efficacy and effectiveness. If we define response narrowly in terms of positive symptoms, then certainly some response to conventional agents is expected. In the case of severe deficit symptoms or in patients who have proven refractory to other drugs, the issue is less clear.

A particular problem arises when response to a proven effective treatment (or so-called gold standard) can vary enormously from trial to trial and in some cases be rather low, or when response to a placebo is generally high (4).

The argument is often made that in developing new drugs to treat a condition for which effective treatments are already available, the question should not be is the new drug superior to placebo but rather is the new drug superior to an already available agent. Unfortunately, given the nature of the diseases and the adverse effects associated with some psychotropic drugs, a new drug could be superior in one domain and inferior in another, while being a very valuable addition to the therapeutic armamentarium. The use of placebo controls can still be important to determine whether or not in some domains a drug is inferior, but still better than a placebo, or whether its inferiority in one domain is such that it would change the overall effectiveness equation.

To provide an example, suppose drug A were somewhat less effective than drug B in controlling acute symptoms, but some patients did quite well on drug A. At the same time, drug B was associated with serious side effects that might result in a substantial number of patients discontinuing the medication within a short period of time. Would we prefer to have drug A available to treat those patients who benefited from it, while then giving drug B to those who don't. Before approving drug A, we would want to be certain that it was superior to a placebo, though inferior to drug B in the particular domain of acute response.

There are a host of issues relating to the use of placebos that have been discussed in more detail elsewhere. As Lavori (5) has emphasized, the data sets available from current placebo-controlled trials are usually "heavily truncated, differentially by treatment groups, and certainly nonrandomly." He argues that most investigators "use ad hoc statistically unjustifiable maneuvers such as last observation carried forward (LOCF)" and that "the interpretation of positive results in the context of badly truncated data requires unverifiable assumptions, external to the observed data of the study."

Another important consideration in the use of a placebo is the potential harm resulting from a delay in instituting active treatment. This is a difficult question to adequately address; however, there have been some attempts to examine the consequences, both short- and long-term, of receiving a placebo in the context of short-term trials. Overall, there do not appear to be demonstratable deleterious effects of participating in short-term trials (6 ,7). The issue of lengthy delays (i.e., 6 months or longer) in implementing treatment has been a topic of discussion in first-episode schizophrenia patients, with some authors suggesting that the longer duration of untreated psychosis is associated with poor outcome. In one patient cohort, this effect was reported in short-term outcome (8), but the effect was no longer evident in long-term follow-up (9). Short-term clinical trials usually involve durations of 4 to 8 weeks. Therefore, it is important to recognize potential differences in consequences between brief delays and relatively long delays in treatment. Lavori (5) argues that because assessments in placebo-treated patients are usually truncated because of high dropout rates, we do not know the full consequences of exposure to a placebo. The field would certainly benefit from more intent-to-treat analyses as well as long-term follow-up of patients who were involved in placebo-controlled trials.

Designs involving the treatment of patients who have failed on other treatments are another challenge. One could argue that placebo controls are more acceptable in this context because there is no effective treatment. However, it is usually the case that these patients have demonstrated some benefit from standard, albeit inadequate, treatment. Therefore, the appropriate comparison would be the new treatment versus standard treatment, with the only outcome of interest being the superiority of the former.

The decision as to whether or not to use placebo or active controls or both in a particular trial is not an easy one. There are complex issues that need to be considered, and it is hoped that further knowledge involving the determinants of heterogeneity in response will facilitate more rational and acceptable trial designs (10).

A related problem is the use of rescue medication. Balancing the desire to retain subjects and the desire to prevent harm and not withhold effective treatment is a critical issue. To what extent should other medications be available for those participants who would otherwise be dropped from a trial due to lack of efficacy and need for alternative treatment? Extensive use of rescue medication can make it difficult to accurately assess the drug effect (even though use of rescue medication can be a telling outcome in and of itself). The use of adjunctive medication to treat adverse effects that occur in the course of a trial can also be a concern (e.g., the use of antiparkinsonian medication). Here, too, rates of utilization can be an important outcome measure, yet at the same time the additional medication might have other undesirable effects (e.g., cognitive impairment).

A number of novel designs have not been widely used, and to some extent there is a disincentive to utilize them, particularly in a regulatory context.

Crossover designs have been suggested as one alternative, although some exposure to a placebo is still involved. A patient receives a potentially active compound and if response occurs, crossover to a placebo takes place. If response does not occur, the placebo phase is not required. The placebo phase in this context helps to determine whether or not the response to medication was a true drug effect or not. It is argued that this design has the advantage of each patient serving as his or her own control, allowing all patients to eventually receive active medication and increasing statistical power.

The applicability of this design varies depending on the nature of the disorder being studied, the time course of response, and the vulnerability to relapse or symptom exacerbation once active treatment is replaced by a placebo. For example, this design may be more informative in rapid cycling bipolar patients (11) than in the context of an acute treatment trial in other disorders. Also, this trial does not eliminate exposure to a placebo. From an ethical standpoint, how do we weigh the delay in providing active treatment against the withdrawal of effective treatment once a response occurs, with the outcome of interest being an exacerbation of symptoms?

Other alternative designs include adaptive allocation strategies. The intent of this approach is to reduce the number of subjects exposed to placebo, ineffective, or toxic treatments. This is achieved by altering the probability of a participant's receiving one treatment or another based on the probability established to that point in the trial of which treatment is associated with the best outcome. These designs are difficult to conduct, and they require knowledge of the results of completed subjects in order to allocate treatment for the next subject. In addition, the response criteria have to be clearly established a priori. The design becomes more complicated when three or more arms are included in a trial. Some studies have utilized such designs with success (12). The ultimate goal of reducing the number of subjects exposed to inferior treatments can be achieved; ultimately, however, the number of subjects required will depend on the effect size of interest. (For further discussion see ref. 13 .)

Active Controls

Comparisons between experimental treatments and active controls require careful consideration in terms of specific drugs, dosage, adverse effects profiles, titration requirements, etc. If a dose of the comparator is too low, efficacy could be less than possible, and if the dosage is too high, then adverse effects may occur more often. This issue is often a particular concern in industry-sponsored studies, where marketing issues often influence the choice of comparator and even its dose. This highlights the potential value of studying a range of doses of both the comparator and the experimental drug. Though this is costly, the information can be particularly valuable in informing clinical practice. Unfortunately, this is rarely done (14). To some extent, this results from unfounded assumptions that we have good data on dose-response relationships with drugs that have been in widespread use. Often that is not the case. In addition, dosage requirements will vary depending on the population. For example, in schizophrenia, first-episode patients in general respond to lower doses than multiepisode patients, and acute treatment usually requires higher doses than maintenance treatment.

Another design that is being increasingly utilized is the adjunctive or add-on strategy. This is particularly useful when subjects with partial or inadequate response are the focus of interest. Rather than switching participants from the unsatisfactory treatment to a new treatment, participants are randomized to an added placebo or added experimental treatment. In this approach, no drug withdrawal is necessary and the question of interest is whether or not the new treatment provides additional benefit.

The potential disadvantages of such a design include drug–drug interactions, particularly if a novel effect is anticipated from the adjunctive treatment. Will this be influenced by the original treatment (e.g., different receptor binding profiles)? This approach is particularly relevant when monotherapy is the exception rather than the rule. This type of design has been employed in the development of anticonvulsant medications (15).

Continuation Treatment

After improvement in acute symptomatology, there is a period of consolidation and stabilization often referred to as the continuation phase. It is assumed that discontinuation of medication during this period would be associated with a higher risk of relapse than subsequent discontinuation. It is difficult to specify when the transition from continuation treatment to maintenance (or prophylactic) treatment occurs, but at least 6 months is a reasonably conservative threshold. The question arises as to how to characterize those patients who have experienced clinically significant improvement, but continue to have more than mild symptoms. In such patients, the continuation phase could become indefinite rather than transitioning to maintenance treatment. This is a semantic distinction because the goal of maintenance treatment is to prevent a relapse or reexacerbation of psychotic signs and symptoms.

A continuation versus discontinuation design can be a sensitive test for drug effect. However, ethically, consent and protection issues are a major concern when any degree of worsening becomes an outcome measure. If such designs are considered, strategies such as sequential analyses or planned interim analysis would be important in terminating the study at the earliest appropriate time.

Maintenance Treatment

In any potentially recurring or chronic illness, the issue of long-term treatment is critical (15). Clearly, the more information on natural history and untreated course that is available from whatever source, the better in helping to define the goals and objectives of maintenance treatment. However, as is often the case, long-term outcome data in such a context are likely to be unavailable, and when comparisons are made with historical data there have often been changes in diagnostic criteria, ascertainment techniques, or other factors that would limit generalizability.

In considering the role of maintenance treatment, frequency, severity, and potential consequences of relapse are critical. Is maintenance treatment justified if a relapse is unlikely to occur for several years? This will be influenced not only by the consequences of a potential relapse, but also by the potential consequences of the prophylactic treatment itself.

In this context, the appropriateness both from a scientific and ethical standpoint of including a placebo control is an enormous concern. The fact that relapse rates on active medication and placebo can vary enormously from one study, one site, or one population to another is an important consideration. Some would argue that an active comparison involving an experimental medication could result in as many or more relapses than could occur in a placebo-controlled trial given the sample size needed to avoid a type II error. Concerns similar to those raised previously apply here as well in terms of multiple domains of outcome and benefit-to-risk ratio. If drug A had a significantly higher relapse rate than drug B but was much safer and more likely to be taken on an ongoing basis, would this drug be utilized if it were shown to be superior to a placebo? How much worse than standard treatment and how much better than placebo would a drug have to be in order to decide one way or the other? This is an unresolved issue in terms of regulatory, scientific, and ethical concerns.

Many of the issues raised previously in the discussion of acute treatment apply here as well. Patient characteristics, age, sex, ethnicity, age at onset of illness, duration of current or most recent episode, baseline psychopathology, comorbid conditions, etc. are all important issues. Even premorbid psychosocial adjustment has been shown to have some predictive power in relapse prevention studies in schizophrenia (8 ,17).

Issues such as reference comparator, dosage, route of administration, concomitant treatments (both pharmacologic and nonpharmacologic), a priori relapse or exacerbation criteria, duration, and strategies to enhance and measure compliance are all important in designing such studies.

The duration of such trials is critical in achieving overall goals. Results can be quite different during the first year of maintenance treatment as compared to the second, with relapse rates often being higher in the first year following recovery from an acute episode as compared to the second year (18). At the same time, in some studies involving dosage reduction, relapse rates were higher in the second year than in the first (19).

This discussion also relates to the issue of time course of relapse in establishing appropriate durations for maintenance trials. In schizophrenia, for example, based on historical data most relapses do not occur for several months after complete drug discontinuation in stable outpatients. One context where time course of potential noncompliance and time course of relapse was such that trial designs were probably inadequate to find meaningful differences was in the comparison of oral and depot medications. A number of double-blind controlled trials were conducted in which patients were randomly assigned to depot or oral medications and therefore had to receive both injections and tablets, one of which was a placebo. The duration of all but one of these trials was 1 year. In general, they failed to find the significant differences that had been expected given high rates of noncompliance in schizophrenia and high rates of relapse following drug discontinuation. However, meta-analysis of these studies supports the value of long-acting injectable preparations (20).

It is likely that the less than expected effects were due to an inadequate duration. Given the fact that subjects agreeing to receive both injections and tablets in a double-blind design are on the more compliant end of the spectrum, one would not expect noncompliance to occur rapidly. In fact, it could take many weeks or months, particularly given the frequent assessments and the psychosocial support involved in being part of a research project. Because the relapse that ensues after complete discontinuation of medications is not likely to occur for several months, it would be unrealistic to expect to observe a difference between depot and oral medication in such a study if the duration was only 1 year (21). The only such study that lasted 2 years found no difference between treatments in the first year, but evidence of clear separation in the second (22). However, the sample size was inadequate to have sufficient statistical power to establish a significant difference, even in the second year.

The role of nonpharmacologic treatments and environmental factors in long-term studies is also important. There is clear evidence that application of nonsomatic interventions can have significant impact on relapse rates among individuals receiving pharmacotherapy. Although ideally nonpharmacologic treatment should be controlled, if it is not there should be documentation of availability and utilization so that potential confounds can be identified.

Another important issue in the design of maintenance trials is whether or not rerandomization following recovery from an acute episode is necessary to demonstrate efficacy in the maintenance phase. In some drug development programs, those patients who respond in the context of an acute trial will be followed and relapse rates reported in comparison to a reference drug. This design provides data

from only those patients who responded to each drug acutely. The argument is made that to demonstrate efficacy in relapse prevention, patients should be rerandomized or the study should be started after patients have been stabilized on any drug. This then allows conclusions to be drawn regarding prophylactic efficacy among patients in general, not just those who responded to an acute trial of a particular drug. (In addition, it is important to recognize high rates of attrition for other causes in acute treatment trials.) This is not to say that there is no value in collecting long-term continuation data on a particular medication, because these data are important in setting the stage for subsequent evaluation and comparisons.

As more domains of interest are examined in schizophrenia, it is necessary to consider the specific designs required to establish efficacy and particular outcome measures. In recent clinical trials, attempts have been made to collect data on an array of measures when at times important confounds can compromise interpretation. For example, in schizophrenia, primary negative symptoms are difficult to study in the context of an acute treatment trial that has selected patients on the basis of having clinically significant positive symptoms. Trials need to be conducted in patients selected on the basis of having residual negative symptoms not complicated by acute positive symptoms or significant extrapyramidal side effects. Remarkably few such studies have been done.

Similar concerns surround the issue of cognitive dysfunction. Newer antipsychotics show some promise in improving measures of cognitive function (23). However, studying these measures requires designs specific to their optimum assessment. In addition, the ultimate question in measuring cognitive performance will be what impact these changes have on functioning, either psychosocial or vocational, level of care, family burden, etc. To date, such studies have not been conducted, and it is premature to conclude that measurable differences on specific cognitive tests will translate into meaningful differences in functioning.

SELECTION OF PARTICIPANTS IN CLINICAL TRIALS

Part of "40 - Issues in Clinical Trial Designs "

The issues discussed in the previous paragraph serve as examples of how patient selection becomes a critical focus in expanding our knowledge of specific drug effects.

Effectiveness Research

Increasing attention has been focused on the fact that traditional randomized clinical trials often include highly selected patients who may not be representative of the population at large. As new medications are used in routine clinical practice, there is often a considerable gap in the knowledge base needed to inform decision making. For example, many patients with schizophrenia have comorbid conditions (e.g., substance abuse) that could influence dosing patterns, adverse effects, overall response rates, compliance, drug interactions, etc. The pharmaceutical industry does not necessarily have an incentive to conduct effectiveness research, as the narrowly defined clinical trial is the most useful and probably cost-effective approach to the drug approval process. In addition, including patients with comorbid psychiatric and medical conditions can potentially increase rates of apparent adverse effects where attribution can be difficult.

At the same time, mechanisms should be sought for conducting effectiveness trials, which are extremely important in informing clinical practice and public policy decisions.

Approaches to Subject Selection

Diagnosis and Phenomenologic Characterization

At present, diagnostic classification is an important element in patient selection. Although nosology shifts over time, it is important to incorporate into the selection criteria the use of an established diagnostic system with proven validity and established reliability. Ideally, research should involve a more systematic and formal diagnostic process than simply relying on a hospital chart diagnosis. Formal evaluation instruments are available for specific diagnostic systems. Although the use of the complete interviews may be overly time-consuming and not cost-effective for some types of research, at minimum a checklist indicating how patients met specific diagnostic criteria should be completed.

As discussed previously, diagnostic subtype has not been a consistent predictor of drug response; however, as classification systems improve and, it is hoped, subtypes become more meaningful, this element will have increasing importance in clinical trial design.

Because many psychotropic drugs are effective across a range of illnesses, a phenomenologic approach to characterizing pharmacologic effect could be reasonable. Although issues of reliability and generalizability would have to be carefully addressed, it is hoped that further research will lead to advances in this perspective.

Biological Classification

Although diseases such as schizophrenia have been characterized by a broad array of biologic abnormalities, there are as yet no well-validated biological classification systems that have proven to be useful in clinical trials or in drug development. This may be largely due to lack of systematic effects in this direction rather than an absence of potentially informative relationships. As further advances take place in diverse perspectives ranging from neuroimaging to pharmacogenomics, it is just a matter of time before biological classification becomes a critical ingredient in this context.

At present, many of the findings are based on group differences and are not necessarily appropriate as selection criteria for clinical trials. In addition, a variety of concerns including sensitivity and specificity will need to be addressed in further developing this perspective.

PHARMACOKINETIC ISSUES

Part of "40 - Issues in Clinical Trial Designs "

The more knowledge available about pharmacokinetics and metabolism (including activity of metabolites) before large-scale clinical trials are designed, the better. Understanding potential relationships between blood levels and therapeutic response as well as adverse effects can be very helpful in optimizing treatment outcome. However, relevant data are often inadequate before critical decisions about dose and dosing schedules are made. If more attention were given to these issues earlier, clinicians would have to struggle less with establishing appropriate treatment strategies. Advances in brain imaging have set the stage for useful investigation during early stages of drug development; however, here, too, few systematic efforts have been made to take advantage of the potential of such studies to help establish optimum strategies for clinical trials.

Clinicians value the availability of different delivery methods for psychotropic medications, given the challenges of both acute and long-term treatment. Oral, liquid, intramuscular, and long-acting forms should be developed and tested in clinical trials as early as possible. Different clinical trial designs may be necessary with different preparations intended for different levels of acuity or phases of treatment. Here, too, the more information available about pharmacokinetic and pharmacodynamic issues, the better.

Given the heterogeneity of clinical response and the enormous variability in drug absorption and metabolism, randomly assigning patients to different plasma levels of interest can be a powerful tool in establishing dose-response relationships and optimum dosing guidelines. Though more difficult than the standard trial, such studies are feasible, but rarely done (24).

ASSESSMENT OF THERAPEUTIC EFFECTS AND CLINICAL CHANGE

Part of "40 - Issues in Clinical Trial Designs "

There are many established instruments for the assessment of psychopathology in clinical trials. In some cases, these instruments have been utilized for many years. However, there continues to be a dearth of new scale development. This is partially due to the tedious nature of the development process and the reluctance of many sponsors of clinical trials to utilize a new instrument. As new drugs are developed with potentially different spectrums of activity, it would be useful to have new scales designed to be sensitive to specific therapeutic effects. This is particularly appropriate since many of the original assessment scales were validated by proving sensitive to the effects of specific classes of psychotropic medications. (For detailed discussions of specific instruments for clinical assessment see ref. 25 and ref. 26 .)

As outcome measures of interest become more broad, an array of separate supplemental instruments are being employed to measure quality of life, social and vocational adjustment, cognitive functioning, and substance abuse. In designing assessment batteries, it is important to choose instruments with proven reliability and validity as well as instruments that are likely to be sensitive to the kind of treatment effect being measured. Meaningful clinical effects should be identified with specific measures of change in order to ensure that the sample size provides adequate statistical power.

As increasing numbers of assessments are employed, it is also important to recognize the burden created for patients and raters. Careful thought should go into selecting the most informative measures and planning a data analysis program with a priori primary and secondary hypotheses.

PROBLEMS IN ASSESSMENT

Part of "40 - Issues in Clinical Trial Designs "

Because psychiatric disorders are often complex, multifaceted diseases and some key symptoms are purely subjective, the techniques used for assessment can be critical. Information regarding psychopathology is most frequently obtained from direct patient interview and observation, though information from other sources (e.g., family, nurses) is sometimes used. Patient report can be impeded by intentional concealment, lack of insight, paranoid ideation, and the overall acuity and severity of the illness. It is not uncommon for psychiatric patients to reveal more psychopathology as they begin to respond to treatment than they did prior to its initiation. The reliability and validity of different sources of information in assessing specific domains have not been adequately studied. In many trials assessors who are not familiar with the patient on an ongoing basis are asked to rate psychopathology. Although these ratings can be sensitive to treatment effects, it is likely that a person who has ongoing contact with the patient in a treatment context will provide a more accurate assessment. Here, too, research comparing different rater allocation strategies would be helpful in determining which is most valid and cost-effective. It is critical to have the same rater evaluating the patient throughout the trial whenever possible. Despite establishing high degrees of interrater reliability, this kind of continuity is important.

The timing of assessments and the time frame chosen for a given assessment should be determined by the goals in the study. In general, when rating psychopathology, the previous week is a reasonable time frame. Patients are less likely to accurately recall specific symptoms that are more

remote in time. The time frame used for a particular assessment does not need to coincide with the interval between assessments. In a long-term trial it is not necessary to rate patients weekly. But when they are assessed, the previous week can be the focus of the assessment.

ASSESSMENT OF ADVERSE EXPERIENCES

Part of "40 - Issues in Clinical Trial Designs "

The two major goals of drug development—to enhance therapeutic efficacy and to improve tolerability—go hand in hand, particularly in the case of psychotropic medications, where many side effects of psychotropic drugs overlap clinical signs and symptoms of psychiatric illnesses. Given the frequent long-term nature of psychotropic drug treatment, adverse effects become critical in influencing compliance and determining the overall benefit-to-risk ratio.

In general, the methods for detecting adverse events have been given far less attention than the methods for evaluating efficacy. Controversy exists as to the most valid means of accurately estimating the incidence of adverse effects. Many clinical trials rely on patient self-report, with some specific queries or rating scales used to assess known adverse effects (e.g., extrapyramidal side effects or tardive dyskinesia) that are outcomes of interest. Given the subjective nature of many adverse events, there is a concern that detailed, specific queries across a broad range of possible symptoms will result in the elicitation of far more symptoms than an unstructured approach.

A methodologic comparison study (27) suggested that the general elicitation of adverse events is more practical and appropriate for routine clinical trials than a comprehensive and lengthy interview. At the same time the field needs to acknowledge the possibility of inordinate delays in recognizing the frequency of specific adverse events such as the sexual dysfunction associated with selective serotonin reuptake inhibitor (SSRI) antidepressants.

There is a strong argument for the use of data and safety monitoring boards when large and/or long-term studies are involved or high-risk treatments are being studied.

CONCLUSION

Part of "40 - Issues in Clinical Trial Designs "

The clinical trial remains the mainstay of treatment development. It is always hoped that further advances will evolve more rapidly than they do, but there is reason for considerable optimism that over the next decade there will be important advances in predicting and understanding psychotropic drug response whether via functional neuroimaging, pharmacogenomics, or other potential developments. In addition, it is hoped that increasing emphasis on studying a broader array of functionally meaningful outcome measures in the context of better informed benefit-to-risk assessment and documentation of cost-effectiveness will lead to clinical trial designs to better address the full range of public health issues.

REFERENCES

1. Klein DF. Improvement of phase III psychotropic drug trials by intensive phase II work. *Neuropsychopharmacology* 1991;4:251–257.

2. Allison DB, Mentore JL, Heo M, et al. Antipsychotic-induced weight gain: a comprehensive research synthesis. *Am J Psychiatry* 1999;156(11):1686–1696.

3. Rothman KJ, Michels KB. The continuing unethical use of placebo controls. *N Engl J Med* 1994;331:394–398.

4. Leber P. The placebo control in clinical trials (a view from the FDA). *Psychopharmacol Bull* 1986;22:30–32.

5. Lavori PW. Placebo control groups in randomized treatment trials: a statistician's perspective. *Biol Psychiatry* 2000;47:717–723.

6. Curson DA, Hirsch SR, Platt SD, et al. Does short term placebo treatment of chronic schizophrenia produce long term harm? *Br Med J* 1986;293:726–728.

7. Wyatt RJ, Hinter ID, Bartko JJ. The long-term effects of placebo in patients with chronic schizophrenia. *Biol Psychiatry* 1999;46:1092–1105.

8. Loebel AD, Lieberman JA, Alvir JMJ, et al. Duration of psychosis and outcome in first episode schizophrenia. *Am J Psychiatry* 1992;149:1183–1188.

9. Robinson D, Woerner MG, Alvir JMJ, et al. Predictors of relapse following response to first episode schizophrenia or schizoaffective disorder. *Arch Gen Psychiatry* 1999;56:241–247.

10. Mattocks KM, Horwitz RI. Placebos, active control groups, and the unpredictability paradox. *Biol Psychiatry* 2000;47:693–698.

11. Post RM, L'Herrou T, Luckenbaugh DA, et al. Statistical approaches to trial durations in episodic affective illness. *Psychiatry Res* 1998;78:71–87.

12. Tamura RN, Farris DE, Andersen JS, et al. A case study of an adaptive clinical trial in the treatment of outpatients with depressive disorder. *J Am Stat Assoc* 1994;89:768–776.

13. Rosenberger WF, Hu F. Bootstrap methods for adaptive designs. *Stat Med* 1999;18:1757–1767.

14. Zimbroff DL, Kane JM, Tamminga CA, et al. Control, dose-response study of sertindole and haloperidol in the treatment of schizophrenia. *Am J Psychiatry* 1997;154:782–791.

15. Brodie MJ. Antiepileptic drugs, clinical trials, and the marketplace. *Lancet* 1996;347:777–779.

16. Greenhouse JP, Stangl D, Kupfer DJ, et al. Methodological issues in maintenance therapy clinical trials. *Arch Gen Psychiatry* 1991;48:313–318.

17. Kane JM, Rifkin A, Quitkin F, et al. Fluphenazine versus placebo in patients with remitted, acute first episode schizophrenia. *Arch Gen Psychiatry* 1982;39:70–73.

18. Hogarty GE. Prevention of relapse in chronic schizophrenic patients. *J Clin Psychiatry* 1993;54(suppl 3):18–23.

19. Marder SR, Van Putten T, Mintz J, et al. Low and conventional dose maintenance therapy with fluphenazine decanoate. Two-year outcome. *Arch Gen Psychiatry* 1987;44:518–521.

20. Davis JM, Barter JT, Kane JM. Antipsychotic drugs. In: Kaplan HI, Sadock BJ, eds. *Comprehensive textbook of psychiatry*. Baltimore: Williams & Wilkins, 1989:1591–1626.

21. Kane JM, Borenstein M. Compliance in the long-term treatment of schizophrenia. *Psychopharmacol Bull* 1985;21:23–27.

22. Hogarty JE, Schooler NR, Ulrich R, et al. Fluphenazine and social therapy in the after care of schizophrenic patients: relapse analysis of a two-year controlled study of fluphenazine decanoate and fluphenazine hydrochloride. *Arch Gen Psychiatry* 1979;36:1283–1294.

23. Meltzer HY, McGurk S. The effects of clozapine, risperidone, and olanzapine on cognitive function in schizophrenia. *Schizophr Bull* 1999;25(2):233–255.

24. Peck CC. The randomized concentration-controlled clinical trial: an information rich alternative to the randomized placebo-controlled trial. *Clin Pharmacol Ther* 1990;47:126.

25. Barnes RE, Nelson HE, eds. *The assessment of psychosis: a practical handbook*. London: Chapman and Hall Medical, 1994.

26. American Psychiatric Association. *Handbook of psychiatric measures*. Washington, DC: American Psychiatric Association, 2000.

27. Rabkin JG, Markowitz JS, Ocepak-Welikson K, et al. General versus systematic inquiry about emergent clinical events with SAFTEE: implication for clinical research. *J Clin Psychiatry* 1992;12:3–10.

Section V
Disorders of Development

Joseph T. Coyle

Disorders of Development - Introduction

Until recently, the developmental aspects of psychiatry were at best tenuously connected to the rapidly advancing neuroscience research informing adult psychiatry. To a significant extent, this reflected the historical accident that child psychiatry evolved out of the child guidance movement, which had a psychosocial orientation and was not associated with medical schools. Nevertheless, the last decade has witnessed a remarkable advancement in the appreciation of the neurobiologic underpinnings of behavioral disorders affecting children.

With increasing evidence of genetic risk factors for psychiatric disorders, the developmental features that transform genetic risk to phenotype have become of particular interest to psychiatric research, especially with regard to prevention. Thus, the seeds of Alzheimer's disease are sown early in the formation of the nervous system, not in the seventh decade of life. Furthermore, family studies are disclosing the early manifestations of serious psychiatric illness including affective disorders, anxiety disorders, and schizophrenia in children, raising the question of appropriate pharmacologic treatments. To this end, the Food and Drug Administration (FDA) is now requiring the pharmaceutical industry to carry out controlled studies of the efficacy of all drugs that might be used in the treatment of children, and the National Institute of Mental Health (NIMH) has funded the Research Units on Pediatric Psychopharmacology (RUPP) to provide the infrastructure to support clinical trials of psychotropic medications in children.

The chapters in this section demonstrate the scientific vigor and rigor that are transforming pediatric neuropsychopharmacology. This is especially so for those disorders that have traditionally been at the borderlands between psychiatry and developmental pediatrics that now provide fertile grounds for linking behavioral pathology to specific developmental processes.

Wassink and his colleagues review the very promising and rapidly advancing area of the genetics of autism and related pervasive development disorders. From being misperceived as being caused by poor maternal care (“refrigerator mother”), autism is now known to be highly heritable, resulting from the interaction of several genes. McDougle reviews the evidence that psychotropic medications can attenuate specific subsets of symptoms and pathologic behaviors that occur in the pervasive development disorders. For too long, this area of the psychopharmacologic management of developmental disorders has rested on anecdotes and hunches; but, increasingly now, these issues are being addressed in well-controlled clinical trials.

Attention-deficit/hyperactivity disorder (ADHD) is one of the most common psychiatric diagnosis in children. Long the subject of criticism by those who oppose psychopharmacology, recent research has elucidated the pathophysiology of the disorder, thereby establishing face validity for the diagnosis. Faraone and Biederman provide a thorough review of the area with special emphasis on the genetics of ADHD. In a related and often co-morbid clinical condition—learning disorders—much progress has been made in understanding the neurobiologic mechanisms as well as characterizing effective interventions, as reviewed by Conners and Schulte.

Psychosis is the most extreme manifestation of psychiatric illness and in children can lead to diagnostic confusion. Joshi and Towbin provide a lucid analysis of the causes of psychosis and how to treat them. Finally, Harris provides an overview of the emerging area of behavioral phenotypes of neurodevelopmental disorders. Careful analysis has differentiated subtypes of developmental disorders in which

specific behaviors can be linked to specific genes or groups of genes in the case of deletions or reduplications. These advances have important implication for the field of behavioral genetics.

The results from clinical trials with psychotropic medications in children disabuse us of the simplistic notion that children are simply small adults, who should exhibit comparable responses to treatment. Nevertheless, the advances in pediatric neuropsychopharmacology raise important questions about the interaction of family environment and social risk factors that must be considered and addressed along with or in place of psychotropic medications.

41

The Molecular and Cellular Genetics of Autism

Thomas H. Wassink

James S. Sutcliffe

Veronica J. Vieland

Joseph Piven

Thomas H. Wassink: Department of Psychiatry, University of Iowa College of Medicine, Iowa City, Iowa.

James S. Sutcliffe: Program in Human Genetics, Department of Molecular Physiology and Biophysics, Vanderbilt University Medical Center, Nashville, Tennessee.

Veronica J. Vieland: Department of Biostatistics, University of Iowa College of Public Health, Iowa City, Iowa

Joseph Piven: North Carolina Mental Retardation Research Center, University of North Carolina, Chapel Hill, North Carolina.

Autism is a behavioral syndrome that is generally associated with lifelong impairment and often confers a substantial burden on the families of affected individuals. Epidemiologic research over the past two decades has demonstrated a significant role for hereditary factors in the etiology of autism, stimulating an aggressive search for susceptibility genes. This chapter summarizes these efforts to elucidate the genetic basis of this severe neurodevelopmental disorder.

THE PHENOTYPE

Part of "41 - The Molecular and Cellular Genetics of Autism "

The three core symptom domains of autism are excessive ritualistic and repetitive behaviors, deficits in communication, and abnormal social interaction. These domains encompass a broad spectrum of behavioral and cognitive abnormalities such as speech delay, echolalia, decreased spontaneous affection, reduced eye-to-eye gaze, motor stereotypies, rigid adherence to routine and environment, and others. The onset of autism is variable, typically manifesting at age when the normal complements, such as speech and prosocial activity, to the disturbed behaviors are expected to develop (usually by the age of 2). Symptoms change with development and generally continue throughout life. When first described in 1943, infantile autism was narrowly defined, referring to a population of children with severe manifestations of all core features and generally higher IQ levels. Kanner, for example, in his original descriptions of autism, included elaborate stereotyped behaviors (e.g., complex rituals and marked distress in response to change in routine or environment) as an essential, pathognomonic component. The severity of this component, however, has gradually been relaxed so that current criteria require the presence of only milder behaviors in this domain.

Similarly, the concept of autism itself has been broadened and now includes the group of syndromes referred to as pervasive developmental disorders (PDDs). Specific PDDs, in addition to autistic disorder, include Asperger's syndrome, pervasive developmental disorder not otherwise specified (NOS), disintegrative disorder, and Rett syndrome. The validity of these diagnostic distinctions, however, is open to question. While the gene that causes Rett syndrome has recently been identified (1), and disintegrative disorder involves a clear loss of function that is likely to arise from a distinct mechanism, there is little evidence to support the remaining PDDs being etiologically distinct.

The existence of multiple symptom domains and the spectrum of related disorders demonstrate the complexity of the autism phenotype. The core symptom domains exist in varying combinations of severity across affected individuals, and it is unclear whether these clinically defined domains represent distinct genetic entities. Less severe symptomatology is represented as a distinct diagnostic entity, but again whether this is genetically justified is not known.

Mental Retardation

Adding to the phenotypic complexity is the wide range of IQs associated with autism. Although some individuals with autism have normal or even exceptional IQs, 70% to 80% are mentally retarded (2). Approximately one-third of those with mental retardation (MR) are in the mildly affected range, the remainder have moderate to severe deficits, and gender proportionality differs across IQ groups (see below).

Biological Correlates

Investigators have searched for biological correlates of autism hoping to better define and categorize the phenotype. Hyperserotonemia was the first biological abnormality to be reported, found by some in up to one-fourth of autistic individuals (3). There is evidence to suggest that hyperserotonemia in autism may be familial (4 ,5), and that elevated platelet serotonin levels may index genetic liability for autism (6). Dysmorphic facial features have also been investigated, with one positive finding coming from a report linking autism to an early developmental abnormality in the branchial arches (7). Epilepsy occurs in 15% to 20% of individuals with autism (8), much more frequently than expected by chance, though whether the presence of epilepsy defines an etiologically meaningful autistic subgroup is unclear.

Two other biological traits that have been investigated are head size and brain morphology. Enlarged head size was noted by Kanner in seven of the first 11 children he described in 1943. Subsequent studies have revealed that approximately 20% of autistic individuals have macrocephaly (> 98th percentile for head circumference) (9 ,10), and the few published postmortem studies report that the brains of autistic individuals are larger and heavier (megalencephalic) than those of normal controls (11 ,12). Retrospective longitudinal studies of head circumference also suggest that while some enlargement may take place before birth, an increased rate of growth appears to occur during the early postnatal period (10 ,13). Magnetic resonance imaging (MRI) studies confirm that brain volume in autism is increased (14 ,15). This enlargement, rather than being generalized, may be confined to discrete structures (16). One recent report, for example, found evidence linking abnormalities in caudate volume to ritualistic-repetitive behaviors in subjects with autism, a finding that is similar to reported relationships in Tourette's syndrome and obsessive-compulsive disorder (17).

As with the core autistic symptomatology, however, these biological correlates are highly variable across individuals, and none is yet able to independently identify cases or define meaningful subgroups. As our ability to measure these correlates becomes more precise, however, their value is likely to increase. They may then serve the purpose of adding power to genetic studies by increasing phenotypic information.

EPIDEMIOLOGY AND GENETIC MECHANISMS

Part of "41 - The Molecular and Cellular Genetics of Autism "

Prevalence

The estimated prevalence of autism has increased since the mid-1980s from 3 to 5 cases per 10,000 to a current estimate of 6 to 10 per 10,000 (2 ,18). This increase is most likely attributable to changing diagnostic practices and increased ascertainment. Epidemiologic studies in the 1970s and early 1980s were based primarily on Kanner's strict diagnostic criteria. With the incorporation of less stringent criteria into the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) and the International Classification of Diseases (ICD), some individuals are now diagnosed with autism who would not have been previously (18). The prevalence of the other PDDs, such as Asperger's disorder and PDD NOS, has not been studied as thoroughly as that of autism. Estimates, therefore, vary widely, though median figures from extant studies suggest a rate 1.5 to 2 times that of autism (2). Thus, taken together, the PDDs may affect 15 to 25 per 10,000 school-aged children. Given that autism is a lifelong condition, the prevalence in adults is likely to be similar to that found in children.

Family and Twin Studies

Family and twin studies help to determine the pattern and strength of the heritability of a disorder. The recurrence risk of autism for siblings of autistic probands is approximately 4% to 5% (19), translating to a sibling relative risk (sibling recurrence risk/population prevalence) of roughly 50 to 100. The risk to second- and third-degree relatives drops off dramatically to less than 1% (20). Twin studies, which compare concordance rates between monozygotic (MZ) and dizygotic (DZ) twins, estimate the heritability for autism to be greater than 90% (21 ,22).

Gender Differences

All epidemiologic studies of autism demonstrate a male preponderance of the disorder. The overall ratio of males to females is approximately 4:1, though this varies with IQ, approaching 6:1 in normal IQ groups and being less than 2:1 in moderate to severe MR groups (2).

Associated Medical Conditions

A host of medical conditions have been reported to cause occasional cases of autism, including neurofibromatosis, tuberous sclerosis (TS), phenylketonuria, rubella, cerebral palsy, trisomy 21, and epilepsy (23). For most of these disorders, however, whether they occur in autism more frequently than expected by chance is unclear. TS has the strongest association; its population prevalence is 1/10,000, and up to 25% of individuals with TS meet diagnostic criteria for autism or PDD (24) (discussed in more detail below). Overall, it has been estimated that approximately 5% of autistic individuals have an associated medical condition that may play an etiologic role in the development of the disorder (2).

Environmental Determinants

Investigators have repeatedly postulated that *in utero* events might predispose a fetus to the development of autism. Early twin studies, for example, suggested that obstetric complications differentiated autistic twins from nonautistic co-twins (25). Subsequent examination of these and other data, however, has shown that the obstetric complications are typically quite minor, the association between autism and complications is weak (26), and that the causality may be inverted—an impaired fetus may actually predispose to obstetric complications instead of complications having affected the fetus (27). Similarly, some studies have reported associations between viral infections [i.e., rubella (28), cytomegalovirus] during pregnancy or season of birth and the subsequent development of autism. The weight of evidence, however, either fails to support such associations or suggests that they account for only a small minority of autism cases (29 ,30). Thus, although perinatal factors are reasonably inferred in rare instances (e.g., encephalitis), in most cases they appear to have either a negligible effect or a minor effect of undetermined significance.

Chromosomal Abnormalities

Estimates of the frequency of chromosomal abnormalities in autism vary widely. Early studies reported rates as high as 20% (31), though recent surveys have reported lower frequencies ranging from 3% to 8% (32–34; Wassink et al., submitted), with the fragile X mutation accounting for one-third to one-half of these. These rates may increase, however, as more sophisticated molecular cytogenetic techniques are applied. Up to 10% of unexplained cases of MR, for example, have been found to be associated with cytogenetic abnormalities detectable only by recently developed subtelomeric probes, and similar abnormalities may be found in autism as well. The most common chromosomal abnormalities currently associated with autism include the fragile X mutation, other sex chromosome abnormalities, and abnormalities of 15q11-q13 (the Prader-Willi/Angelman syndrome (PW/AS) region).

Genetic Mechanisms

Thus, although a small proportion of cases of autism are due to chromosomal abnormalities or medical conditions, the vast majority are likely to be multifactorial, arising from an as yet unknown environmental component superimposed on a strong genetic predisposition. The heritability for autism of 90% exceeds that of other common psychiatric disorders such as schizophrenia, bipolar disorder, or alcoholism. The mode of heritability, like other psychiatric disorders, appears to be complex. Autism pedigrees have not been reported that demonstrate mendelian segregation (unless the broader autism phenotype is included—see below), and the differential gender distribution across IQs suggests genetic heterogeneity. The rapidly diminishing relative risk from first- to second- to third-degree relatives, combined with the >4:1 MZ:DZ concordance ratio, indicates that autism is likely to be due to multiple genes interacting in variable combinations in additive, multiplicative, epistatic, or as yet unknown fashions (35). Estimates of the number of genes involved have ranged from at least three (36) to more than 15 (37). Furthermore, other disorders composed of isolated components of the autism phenotype (e.g., specific language impairment) are themselves considered to be due to multiple, interacting genes, making it likely that the genetics of autism will be complicated as well.

GENETIC INVESTIGATIONS OF AUTISM

Part of "41 - The Molecular and Cellular Genetics of Autism "

Early genetic investigations of autism were hampered by a number of constraints, including small sample sizes, inconsistent diagnostic criteria, and limited molecular tools. The development of standardized diagnostic criteria and advanced molecular tools, such as high-quality, densely spaced genetic markers, FISH (fluorescent in situ hybridization) chromosomal probes, and high-throughput sequencing, is beginning to overcome these constraints. Multicenter collaborations can now gather large, consistently characterized samples, genome-wide screens are practical, sequence data are available for focused genetic investigations, and chromosomal studies are more exact and informative. These advances are reflected in the recent surge of genetic investigations of autism, which are summarized below.

Genome-Wide and Focused Linkage and Association Studies

Four genome-wide linkage studies of autism have been published to date (37 ,38 ,39 and 40). All these studies have examined families containing at least two affected siblings [affected sibling pair (ASP) families] and are summarized in Table 41.1 . The strongest single finding to emerge from these screens is a multipoint heterogeneity logarithm of odds (LOD) score of 3.0 on chromosome 13q at 55.3 centimorgans (cM) reported by the Collaborative Linkage Study of Autism (CLSA) (39), whereas the most replicated finding consists of support for linkage to chromosome 7q.

Research Group	Sample Characteristics				Findings		
	Sample	Markers	Chrom	Markers	cM	MLS	Comments
IMGSAC (38)	87 ASPs	354	7q	D7S530	134.6	2.53	Multipoint analyses were calculated using ASPEX under an additive model that assumed no dominance variance; in a subsample of 56 ASPs, all from the UK, the MLS at D7S530 was 3.55
			4p	D4S412	4.8	1.55	
			16q	D16S407	17.3	1.51	
			22q	D22S264	5.0	1.39	
			10q	D10S197	51.9	1.36	
			14q	D14S70	32.9	0.99	
			19q	D19S49	48.2	0.99	
PARIS5 (40)	51 ASPs	264	6q	D6S283	132.8	2.23	The MLS was maximized over the "possible triangle" using MAPMAKER/SIBS
			19q	D19S226	24.1	1.37	
			15q	D15S118	41.1	1.10	
			7q	D7S486	135.3	0.83	
Stanford (37)	147 ASPs	519	1p	D1S1675	149.2	2.15	This study used ASPEX to calculate a multiplicative model that allowed for dominance variance; this study also found a diffuse excess of allele sharing across the genome, suggesting a large number (≥15) of genes of small effect
			17p	D17S1876	10.7	1.21	
			7p	D7S2564	41.7	1.01	
			18q	D18S878		1.00	
			7q	D7S684	147.2	0.62	
			7q	D7S1804	137.0	0.93	
			13q	D13S800	55.3	0.68	
CLSA (39)	75 ASPs	349	13q	D13S800	55.3	3.00	This study used GENEHUNTER to calculate an MLS that allowed for heterogeneity
			13q	D13S217	17.2	2.30	
			7q	D7S1813	104.0	2.20	
			4q	D4S2368	167.6	1.52	
			4q	D4S3248	72.5	1.33	
			11q	D11S968	147.8	1.22	
			16q	D16S516	100.4	1.03	
			8q	D8S1477	60.3	1.02	
			15q	D15S975	13.1	0.50	
			7q	D7S1824	149.9	0.79	

ASP, affected sibling pair; cM, centimorgan; MLS, multipoint LOD score.

TABLE 41.1. GENOME-WIDE LINKAGE STUDIES OF AUTISM

Some caution must be taken when comparing these studies, however, because none of them report exactly the same statistic. The CLSA (39) reported a maximum multipoint LOD score (MLS), calculated using the program GENEHUNTER (41) and based on a likelihood that allowed explicitly for heterogeneity (42). The Stanford group also reported an MLS (37), but the underlying likelihood was parameterized in terms of a multiplicative model allowing for dominance variance, and calculated using ASPEX (43 ,44). The International Molecular Genetic Study of Autism

Consortium (IMGSAC) also used ASPEX to calculated the MLS, but under an additive model that assumed no dominance variance (38). The Paris Autism Research International Sibpair Study (PARISS) used a related MLS, maximized over the "possible triangle" (45), using MAPMAKER/SIBS (40). While all these statistical approaches are related to one another (Huang and Vieland, in press), they may involve estimation of somewhat different numbers of parameters, or "degrees of freedom." The most appropriate use of these data in aggregate, therefore, is not to directly compare numerical results, but rather to look for regions that have either very strong support for linkage within individual studies or that have recurrent support across studies.

Focused genetic studies have examined smaller chromosomal regions chosen for one of three reasons: (a) the region showed evidence of linkage in a genome-wide screen; (b) the region contains a high rate of chromosomal abnormalities associated with autism; or (c) a "candidate" gene of interest, chosen because of its potential biological or developmental relevance to autism, is located in the region. The samples for these focused studies include both ASP families and trios (proband and both parents), and are summarized in Table 41.2 .

	Study Characteristics				Findings		
Research Group	Ref	Sample	Markers Tested (Significant) Chrom	Gene	Marker	Result	Test
IMGSAC	51	91 ASPs, 8 trios	2 (0) 17q11	*HTT*			Linkage
			7 (0) 15q11-q13				Linkage
Duke	47	76 ASPs, 32 trios	7 7q		D7S2527	MLS = 1.77	Linkage
	48	63 multiplex	14 15q11-q13		D15S217	MLS = 1.78	Linkage
	148	54 trios, 36 ASPs, 33 other	4 (1) 15q11-q13		GABRB3	$p = 0.03$	Linkage disequilibrium
Stanford	149	147 ASPs	8 (0) 15q11-q13			MLS < 0.0	Linkage
	150	97 ASPs	10 (0) 6p	HLA region		MLS < 0.0	Linkage
Other	151	125 trios, 6 ASPs	9 (1) 15q11-q13		GABRB3 15SCA-2	$p = .0014$	Linkage disequilibrium
	49	86 trios	2 (1) 17q11	*HTT*	promoter	$p = .018$	Multiallelic TDT*
	50	117 trios	2 (1) 17q11	*HTT*	promoter	$p = .049$	Multiallelic TDT*
	152	53 trios	1 (0) 10q21	*HTR7*			TDT
	153	35 multiplex	4 (0) Xq27	*FMR-1*			Linkage
	154	85 probands, 90 controls	3 (0) 17q11	*NF-1*			Association
	155	10 probands	15q11-q13	*UBE3A*	exon screening		No mutations
	156	50 probands, 50 controls	5 (1) 11p15	*HRAS-1*	*Bam* H1	$p = .008$	χ^2
	157	55 probands, 55 controls includes sample from (169)	2 (2) 11p15	*HRAS-1*	*Bam* H1	$p < .05$	χ^2
					Nsi 1	$p < .01$	χ^2
	158	48 probands, 50 controls	1 (1) 11p15	*HRAS-1*	*Msp* 1	$p = .034$	χ^2
	159	66 probands, 89 controls	1 (0) 11p15	*TH*			χ^2
	160	100 probands, 100 controls	2 (1) 7q36	*EN2*	*Pvu* II	$p < .01$	χ^2
	161	72 probands, 72 controls	2 (0) 11p15	*Ins*			χ^2
			1 (0)	*IgF*			χ^2
	137	19 trios, 62 controls	haplotypes 6p21	HLA-III	C4B null allele	$p = .03$	χ^2
	138	21 trios, 62 controls	haplotypes 6p21	HLA-III	DRβ1 haplotypes	$p < .001$	χ^2
	140	45 probands, 79 controls	haplotypes 6p21	HLA-III	C4B null allele	$p < .001$	χ^2
					DRβ1 haplotypes	$p < .01$	χ^2
	130	44 trios, 6 probands, 79 controls	haplotypes 6p21	HLA-III	DRβ1 HVR-3	$p < .001$	χ^2

*Significant results are for opposite alleles.
ASP, affected sibling pair; CM, centimorgan; MLS, multipoint LOD score; TDT, transmission disequilibrium test.

TABLE 41.2. FOCUSED LINKAGE AND ASSOCIATION STUDIES OF AUTISM

Chromosome 7

The most replicated evidence for linkage is to chromosome 7q (Fig. 41.1). The IMGSAC, examining 87 ASPs, reported a LOD of 2.53 at D7S530 (134.6 cM), which increased to 3.55 in a subset of 66 United Kingdom ASPs (38). A recent second-stage analysis of 125 ASPs by this same group reported, in poster format, a multipoint MLS near D7S2533 (140.5 cM) of 3.63 (46). The CLSA, examining 75 ASPs, reported an MLS of 2.2 at D7S813 (104 cM) (39), whereas the PARISS (40) and Stanford (37) studies reported modestly positive results (MLS = 0.83 and 0.93, respectively).

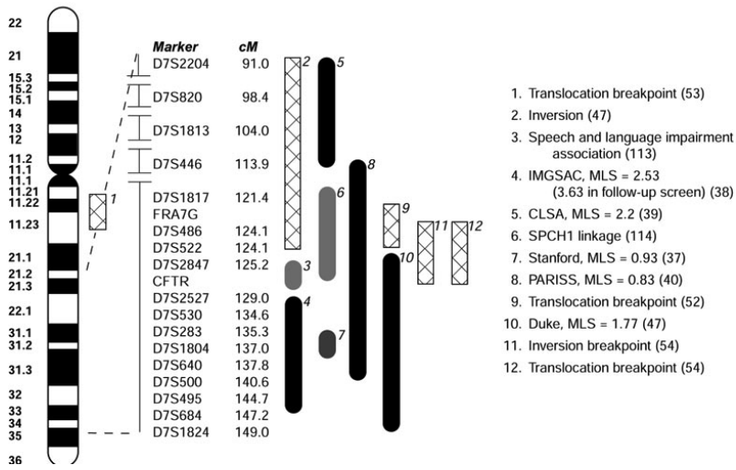


FIGURE 41.1. Chromosome 7 findings in autism. *Rectangles* represent chromosome abnormalities, *black ovals* indicate regions with evidence for linkage or association in autism, and *gray ovals* indicate regions with evidence for linkage or association in related disorders. MLS, maximum multipoint LOD score within the region indicated.

A subsequent focused examination of nine 7q markers was performed in 76 multiplex families and 32 trios by the Duke/University of South Carolina (USC) group (47). They found a peak multipoint MLS of 1.77 at D7S2527 (129.0 cM) using an additive model calculated with ASPEX. This group also found high rates of recombination throughout 7q and evidence of linkage disequilibrium at D7S1824

(149.9 cM), reporting that the linkage, excess recombination, and linkage disequilibrium appeared to be due primarily to paternal and not maternal effects.

Chromosome 15q11-Q13

Because of numerous reports of cytogenetic abnormalities (discussed below), chromosome 15q11-q13 has been intensively examined in linkage and association studies. None of the currently published genome-wide screens, which all included tightly spaced 15q11-q13 markers, identified linkage to the region. The Duke/USC group screened fourteen 15q11-q13 markers in their families and reported a maximum MLS of 1.78 at D15S217 (48). Another group has reported highly significant linkage disequilibrium with the marker GABRB3 155CA-2, a dinucleotide repeat polymorphism that was not typed in any of the genome-wide screens (Buxbaum et al., submitted). This result was obtained when the data were pooled with a number of other focused studies that included this marker.

Candidate Gene Studies

Genes tested as candidates for involvement in autism include genes involved in neurotransmission (i.e., *HTT*, *TH*, *DBH*), genes contributing to related disorders (i.e., *FMR-1*, the fragile X syndrome (FXS) gene, and *NF-1*, a neurofibromatosis gene), and genes thought to be involved in brain development (i.e., *EN2*, *HRAS-1*) (Table 41.2). Though positive results have occasionally been reported, replication has been limited. Cook et al. (49), for example, reported preferential transmission of the small allele of the *HTT* promoter polymorphism to autistic probands. Klauck et al. (50), attempting replication, reported preferential transmission of the larger allele, and the IMGSAC reported no association with either allele (51). No other candidate genes tested thus far have found consistent support.

CHROMOSOMAL ABNORMALITIES

Part of "41 - The Molecular and Cellular Genetics of Autism "

Studies of cytogenetic abnormalities can complement molecular approaches by identifying genes whose effects are either too small to be detected by linkage or are obscured by epigenetic phenomena. The X chromosome and chromosome 15q11-q13, for example, have both received scrutiny because they are frequent sites, relative to other regions, of cytogenetic abnormalities in individuals with autism. Additionally, linked regions in complex disorders are typically broad, and chromosomal abnormalities occurring in linked regions can help to pinpoint disease susceptibility genes. This can be done either by cloning the chromosomal break points and identifying disrupted genes, or by overlaying deleted regions across individuals in order to delineate a minimal deleted region that might harbor a disease susceptibility gene.

The prevalence of chromosomal abnormalities in autism has been discussed above (see Epidemiology and Genetic Mechanisms). This section, therefore, focuses on abnormalities in specific regions: 7q, 15q11-q13, the sex chromosomes, and the fragile X mutation.

Chromosome 7 Cytogenetic Abnormalities

A number of chromosome 7 cytogenetic abnormalities in close proximity to the 7q linkage findings have been identified in individuals with autism (Fig. 41.1). In one family, a paracentric inversion, inv(7)(q22q31.2), is carried by two brothers, a sister, and their mother (47). The brothers appear to have autistic disorder, the sister has expressive language disorder, and the mother has neither of these abnormalities. Another autistic individual has been described with a translocation t(7;13)(q31.2;q21) (52). The break points for these abnormalities have been cloned, and identification and screening of nearby candidate genes is in progress. Two autistic twins have been found to have translocations t(7;q20)(q11.2;p11.2) (53). The chromosome 7 break point and the gene it disrupts, named “autism-related gene 1” (*ARG1*), have been cloned; the gene is novel with an unknown function, spans an 800-kilobase (kb) genomic region, and is highly expressed in fetal and adult brain (53). Lastly, one individual with autism and another with a specific developmental disorder of speech and language (SDDSL) have been reported, both of whom have chromosomal abnormalities involving 7q31 [autism, 46,XY, inv(7)(p12.2q31.3); and SDDSL, 46,XY,t(2;7)(p23;q31.3)] (54).

Chromosome 15q11-Q13

Chromosome 15q11-13 is the most frequent site of autosomal abnormalities in autism. In a recent chromosomal survey, 6 (2.2%) of 278 autistic subjects referred for cytogenetic studies had a gross abnormality of chromosome 15 (Wassink et al., submitted). These abnormalities most commonly involve duplication of maternal DNA, typically as either interstitial duplications (55 ,56 ,57 ,58 ,59 ,60 and 61) or inverted duplicated isodicentric marker chromosomes [inv dup(15)] (62 ,63 ,64 and 65). The duplications that produce illness typically extend into the Prader-Willi syndrome (PWS) and Angelman syndrome (AS) critical region, as smaller duplications are generally asymptomatic (64 ,66). Complementing these data, individuals identified because of 15q11-q13 duplications frequently have autistic features (67). Deletions of 15q11-q13, though less frequent, have also been reported in autistic individuals, with the deleted material usually of paternal origin (68 ,69). Figure 41.2 summarizes these data, displaying the relevant genes and markers as well as a putative autistic disorder region.

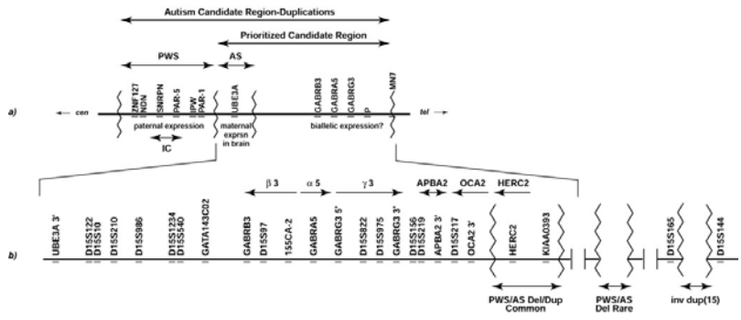


FIGURE 41.2. Schematic map of 15q11-q13 autism candidate region. The *upper part* of the figure is a low-resolution schematic representation of the 15q11-q13 interval deleted in Prader-Willi syndrome (PWS) and Angelman syndrome (AS) and duplicated in cases of autism. PWS and AS critical regions are indicated by *arrows* over the map, with relevant imprinted genes indicated in their respective regions. A prioritized autism region excludes the imprinted PWS domain based on the apparent maternal specificity of 15q11-q13 duplications [interstitial or inv dup(15) markers] in association with autism. The lower expanded region reveals gene and marker order and transcriptional orientation within the candidate region. Large break-point regions are depicted for the primary distal PWS/AS deletion break point, as well as a less common (˜10%) PWS/AS break point and a break-point interval associated with inv dup(15) marker chromosomes.

In addition to these gross abnormalities, cytogenetic abnormalities at the molecular level are also being reported. Cook et al. (70), while screening autism trios across 15q11-q13, identified a nonautistic mother in whom a duplication had arisen *de novo* on her paternally derived homologue (70). The duplication was transmitted to one child who

developed autism and a second child with atypical autism, whereas a third child who did not receive the duplication remained unaffected. In the CLSA genomic screen, which examined 75 ASP families (152 affected individuals) and included two 15q11-q13 markers, six individuals (3.9% of probands) from four families (5.3% of families) were found to have either maternal duplications or paternal deletions at one or both of these markers (Wassink et al., submitted). Forty-five unrelated autism trios were subsequently screened using 12 polymorphic 15q markers, with three (6.1%) autistic probands found to have similar abnormalities (Wassink et al., submitted).

The apparent gender specificity of the 15q11-q13 abnormalities is presumably attributable to imprinting, an epigenetic mechanism by which only one of a gene's two inherited alleles is expressed, with expression determined by the allele's gender of origin (71). The two primary 15q11-q13 syndromes, PWS and AS, are oppositely imprinted mental retardation syndromes that bear some phenotypic similarities to autism (72). In brain tissue (but not elsewhere), *UBE3A*, the AS gene, is expressed predominantly from the maternally derived allele. Disrupted expression of the maternal *UBE3A*, therefore, produces AS, whereas disruption of the paternally derived allele produces no discernible abnormal phenotype (73). PWS genes, conversely, are paternally expressed; the predominant cause of PWS, therefore, is disrupted expression of the paternal copy of the small nuclear ribonucleoprotein polypeptide N (SNRPN) gene and other contiguous genes (74). Another recently identified element of imprinting is the presence of antisense transcripts for imprinted genes. These segments of RNA are oppositely imprinted complements to an imprinted gene's coding sequence (75). *UBE3A*, for example, has an antisense transcript that is expressed solely from the paternally derived allele (71). This antisense transcript may play a role in the suppression of the nonexpressed allele, and mutations in this transcript, therefore, could contribute to some cases of AS (76). Thus, just as imprinting plays a pivotal role in PWS and AS, it is likely to significantly influence the effect of 15q11-q13 abnormalities in autism as well.

The 15q11-q13 abnormalities themselves may be due to the presence of repeated or duplicated homologous genomic segments that exist in multiple copies throughout this region (77). Duplications of large genomic segments are associated with chromosomal abnormalities in a number of specific syndromes (78 ,79). One such repeated segment (duplicon) appears near each of three most common PWS/AS duplication breakpoints (80). Another, located centromeric to the PWS/AS critical region, is repeated an increased and variable number of times in PWS/AS individuals (81). Though this does not imply a traditional repeat expansion mechanism, it may be that these repeats predispose the region to recombination abnormalities or "mistakes" (77), a finding with support from data showing increased rates of recombination across 15q11-q13 in subjects with either PWS/AS (82) or autism (48).

Interestingly, one other chromosomal region that shares many of these genomic features is 22q11 (77). Chromosome 22q11 contains large repeated segments that contribute to a high rate of deletions and duplications (83). These chromosomal abnormalities are associated with a constellation of syndromes grouped under the umbrella term *CATCH-22* (84). One of these syndromes, velocardiofacial syndrome (VCFS), is associated with a high rate of schizophrenia (85), a psychiatric disorder that, like autism, is felt to arise from disturbed brain development. In addition, a small but significant percentage of individuals with schizophrenia have now been shown to have microabnormalities of 22q11 (86). Thus, schizophrenia and autism may share a common genomic mechanism for a subgroup of cases, and insights from one disorder may inform investigations into the other.

Fragile X

The association of autism with the fragile X syndrome (FXS) was first suggested nearly 20 years ago (87 ,88). The fragile X phenotype is frequently characterized by behaviors that can resemble the core symptom domains of autism such as language abnormalities, decreased nonverbal communication, social isolation, and repetitive motor behaviors such as rocking and hand biting (89). In support of this association, early chromosomal investigations reported a rate of the fragile X mutation [fra(X)(q27.3)] in autism that approached 20% (31 ,90).

Recent studies, however, have questioned the strength of the link between these two disorders. Current surveys estimate the frequency of fragile X in autism to be 2% to 4% (91 ,92 and 93), similar to the rate of fragile X in the general MR population (94). This is more common than other types of chromosomal abnormalities in autism, though not necessarily disproportionately so. Likewise, there are subtle but significant differences between the behavioral phenotypes of the two disorders. Autism is characterized by social indifference and deficits in the perception of emotion, whereas individuals with FXS experience social anxiety and gaze avoidance with no attendant impairment of emotional perception (95).

Genetically, FXS is a disorder of unstable DNA caused by a trinucleotide repeat that expands as it is transmitted to successive generations (96). The repeat is located at Xq27.3 in the 5′ UTR (untranslated region) of the *FMR1* gene (89). Once this expanded region crosses a threshold (approximately 200 repeats), it becomes susceptible to methylation, which inhibits transcription of *FMR1*. FMRP, the *FMR1* protein, is an RNA binding protein that appears to act as a chaperone for transport of RNA from the nucleus to the cytoplasm (97). FMRP is expressed in numerous tissues including fetal brain. Intracellularly, it is found in the nucleus near the nucleolus and in cytoplasm in association

with ribosomes. It may function, therefore, as a chaperone molecule in the transportation of messenger RNA (mRNA) from the nucleus to the cytoplasm (98). How dysfunction of this protein gives rise to FXS, however, remains unclear.

Other Sex Chromosome Abnormalities

The possibility of a sex chromosome–related genetic effect is suggested by the preponderance of males affected by autism (2). In accord with this, a sizable number of sex chromosome abnormalities, in addition to the fragile X mutation, have been reported in subjects with autism (99). In a recent survey of a clinical population, six out of 265 (2.3%) autistic individuals referred for cytogenetic testing were found to have abnormalities of the sex chromosomes other than fragile X (Wassink et al., submitted). In addition, two X-linked disorders, Turner syndrome and Rett syndrome, have phenotypic features that are similar to some of the core features of autism. Skuse et al. (100) reported that Turner syndrome (45,X) females with maternally derived X chromosomes had diminished verbal skills and social cognition compared to those with paternally derived Xs. Molecular studies implicated a paternally imprinted disease locus that escapes X-inactivation in distal Xp22.3 (101). This paternal imprinting could explain why karyotypically normal males (who have a maternally derived X) are more vulnerable to developmental disorders of language and social cognition, such as autism, than females. Skuse et al. also reported three Turner syndrome females, all with maternally inherited X chromosomes, who had been diagnosed with autism. Two more XO autistic individuals have recently been reported, one with a maternally derived X (102) and the other with an X of unknown origin (Wassink et al., submitted).

The gene for Rett syndrome was recently identified on Xq28 (1). Rett syndrome, considered to be a subtype of PDD, is a disorder occurring only in girls that is characterized by mental retardation, loss of speech, and stereotypic hand movements after 1 to 2 years of normal development. The gene for Rett syndrome, (*MECP2*), is widely expressed and codes for a DNA binding protein that regulates gene expression (1).

Linkage screens of the X chromosome in autism have generally been negative, excluding genes of even small effect (37 ,103), and have contributed to a reluctance to examine the sex chromosomes for autism disease genes. The evidence from sex chromosome abnormalities and from X-linked disorders with phenotypic similarities, however, suggests that such pessimism is premature, and that the X and Y chromosomes should continue to be a focus of attention in autism.

BROADER AUTISM PHENOTYPE

Part of "41 - The Molecular and Cellular Genetics of Autism "

In addition to describing the hereditary basis of autism, family and twin studies have demonstrated, in nonautistic relatives of autistic probands, the presence of milder traits that are qualitatively similar to the defining features of autism. These collective traits, referred to as the “broader autism phenotype” (BAP), were first observed by Kanner in parents of autistic children. Bailey et al., replicating and extending findings from the original Folstein and Rutter (104) twin study, found a substantially higher concordance rate for the presence of mild social and communication deficits in MZ versus DZ twin pairs (92% versus 10%) (21). These results are supported by several family studies using the family history method of assessment (105 ,106). In the London Autism Family Study, Bolton et al. (105) reported that familial aggregation of the BAP was associated with proband verbal IQ. In the Iowa Autism Family Study (Piven and Palmer, submitted) familial aggregation of the BAP was higher in relatives from families with two autistic siblings (multiple-incidence families) than in families ascertained through a single autistic child.

A more detailed examination of the BAP has been accomplished through direct assessment of relatives. Relatives from multiple-incidence families, for example, were found to have (a) elevated rates of personality characteristics such as aloofness and rigidity, (b) diminished pragmatic language and speech abilities, (c) fewer quality friendships, and (d) decreased scores on a number of specific cognitive measures (107 ,108 and 109).

Investigation of the BAP may, by clarifying the range of phenotypic expression of the underlying genetic liability to autism, provide a complementary approach to traditional linkage that increases power to detect genes by identifying more affected individuals, thereby enabling extension of typically small autism pedigrees. Understanding the boundaries and nature of the BAP may also help our efforts to detect genes in autism by enabling focused investigation of specific BAP components (e.g., language deficits, ritualistic-repetitive behaviors, or cognitive deficits) that may map on to separate genes that together cause the full syndrome of autism. This approach to disaggregating complex phenotypes has proved successful in dyslexia, where separate linkages were found to single-word reading and phonemic awareness (110). Clearly, clarification of the genetically relevant aspects of both the autism and the broader autism phenotype is an important strategy to pursue in our search for genes in this disorder.

RELATED DISORDERS

Part of "41 - The Molecular and Cellular Genetics of Autism "

Autism is characterized by dysfunction in three symptom domains: language; social interaction; and repetitive, stereotyped movements and behaviors. As autism is a heterogeneous, genetically complex disorder, it may be that each of these domains has unique, independent genetic determinants. Studying disorders that resemble these individual domains, therefore, may provide insight into their etiology in

autism. There are also related disorders, such as tuberous sclerosis, and domains of investigation, such as immunogenetics, that may provide insight into autism.

Disorders of Language

Specific language impairment (SLI) is a disorder characterized by isolated impairment of language skills, and may be characterized by grammatical impairment, word finding difficulties, or an underlying perceptual deficit (111). Though traditionally considered distinct disorders, SLI has been found to be common in autistic individuals and to occur at higher rates than expected in their nonautistic relatives (108). Conversely, an increased rate of autistic disorder has recently been found in siblings of children with SLI (112). Tying this in to chromosome 7, an association study found significant associations between two 7q31 genetic markers and a group of SLI trios (113). Also, a family has been identified with a severe speech and language disorder characterized by deficits in grammar, expressive language, articulation, and coordination of orofacial musculature (114). A genome-wide screen of this three-generation pedigree found a maximum LOD score of 6.62 at a marker in 7q31, with fine mapping narrowing the region to a 5.6-cM interval (SPCH1 locus). These findings, therefore, may represent localizations of heritable components of the autism phenotype and are of particular interest given the evidence for linkage of autism to this same region of chromosome 7.

Disorders of Repetitive and Stereotyped Behaviors

The phenomenologic overlap between autism, obsessive-compulsive disorder (OCD), and Tourette's syndrome has led some to wonder whether these disorders have etiologic mechanisms in common. For example, caudate volumes have been found to be abnormal in both Tourette's (115) and OCD (116). Therefore, we examined caudate volumes in an MRI study of autistic children and found enlargement of the caudate that was correlated to ritualistic, stereotyped behaviors but not to social or communication deficits (17). In an earlier family study, we reported higher familial aggregation of autism and the BAP in families ascertained through a Kanner proband (more severe ritualistic behavior) versus more broadly defined (DSM, third edition, revised) probands (117). Findings such as these suggest that traits such as stereotypies or ritualistic behavior may have unique genetic determinants that, when combined with genes that give rise to other traits such as language or communication deficits, could give rise to the syndrome of autism.

Disorders of Social Activity

Examples of disorders that involve significant social deficits include Turner syndrome and the fragile X syndrome, both of which have been discussed above (see Other Sex Chromosome Abnormalities). Additional research related to social deficits that may have relevance to autism comes from studies of various neuropeptide systems. For example, nematode worms that lack receptors for neuropeptide Y become strikingly isolated in situations where they would normally congregate with other worms (118). Genetic variability in receptors for oxytocin/vasopressin in mice and other rodents is also associated with clear variability in social behavior (119). Thus, though there is significant evolutionary distance between worms, rodents, and humans, these transmitter systems may merit closer examination in individuals with autism.

Tuberous Sclerosis

Tuberous sclerosis complex (TSC) is a neurocutaneous disorder characterized by benign tumors affecting numerous organs, most commonly the brain, eyes, skin, kidneys, and heart (120), with a population prevalence estimated at 1/10,000 (121). The occurrence of autism and other behavioral and psychiatric disturbances in the context of TSC has long been recognized (122). Clinic-based and epidemiologic studies of autism in TSC suggest that up to 25% of individuals with TSC meet diagnostic criteria for autism and over 40% meet criteria for any PDD (24 ,123). Conversely, 1% to 3% of autistic individuals will have TSC (124), though this rate approaches 10% for autistic individuals with seizure disorders (24 ,123).

TSC is an autosomal-dominant disorder caused by mutation in one of either two genes, *TSC1* or *TSC2* (125). *TSC2* is located on chromosome 16p13 and codes for the protein tuberin, whereas *TSC1* is located on 9q34 and codes for the protein hamartin. Tuberin has numerous functions, acting as a tumor suppressor or a chaperone, and having an influence on cell cycle passage. Dysfunction of tuberin may result in constitutive activation of RAP1, a protein that regulates DNA synthesis and cellular transition, thereby producing excessive proliferation and impaired differentiation of a variety of cell types. Hamartin is one of the proteins for which tuberin acts as a cytosolic chaperone. Other than an ill-defined role in tumor suppression, the function of hamartin remains unknown (126). Approximately two-thirds of TSC cases are sporadic and one-third familial. Half of the familial cases and 75% to 80% of sporadic cases arise from mutations of *TSC2*, with the remainder attributed to *TSC1*. Two studies have shown that TSC due to *TSC2* mutations is more likely to be associated with either mental retardation or intellectual impairment than TSC due to *TSC1* mutations (127 ,128). Despite the strong association between TSC and autism, however, the mechanistic link between the two disorders remains unclear. Autism in the context of TSC may arise directly from the TSC mutations, from the tubers they produce, or from some other as yet undiscovered mechanism. One group, for example, has reported an association between the presence of temporal lobe tubers and autism (129), though this finding has not been replicated (24).

Immunogenetics

A number of investigators have suggested that some cases of autism may be attributable to interactions between infections, the immune system, and genetic factors (130). Subjects with autism have been shown to have deficits in the number and function of various immune cell subtypes (131 ,132 ,133 ,134 ,135 and 136). A series of investigations, therefore, performed primarily by one research group, has investigated specific components of the major histocompatibility complex (MHC) on chromosome 6p21 for association with autism (130 ,137 ,138 ,139 and 140) (Table 41.2). The samples in these studies were generally small and overlapping, and associations emerged primarily when probands were compared to a population control group as opposed to a parent-based test. Nonetheless, the authors report consistently significant findings that await replication by others.

FUTURE DIRECTIONS

Part of "41 - The Molecular and Cellular Genetics of Autism "

Alternative Sampling Designs

As noted above, all of the genome-wide linkage screens performed to date have focused almost exclusively on affected sibling-pair (ASP) pedigrees. This is in part due to the difficulty of gathering extended autism pedigrees. Also, however, because the mode of inheritance for autism is unknown, many researchers appear to be more comfortable with the simple “model-free” linkage analysis methods available for ASP data than with the more complicated methods required for the analysis of general pedigree structures. Yet, limiting pedigrees solely to ASPs may be unnecessarily and detrimentally restrictive.

The optimal pedigree structure for the detection of linkage depends on the true, underlying genetic model for the trait, which in this case is unknown. Large, multiplex pedigrees, for example, are optimal when a trait is transmitted as a rare dominant but tend to be uninformative if the trait is a common recessive. Furthermore, an ASP data set contains only two primary pieces of information—the number of sibling pairs sharing 1 allele identical by descent and the number of pairs sharing 2 alleles—thereby limiting the complexity of any model that can be fit to ASP data. By limiting the families that are gathered to ASPs, the ability to detect linkage under all possible modes of transmission may be similarly limited. It could be argued, therefore, that precisely because the mode of inheritance for autism is unknown, the optimal strategy would be to ascertain all types of potentially informative pedigree structures. This could include large multigenerational extended pedigrees and moderate-sized pedigrees with more than just two affected individuals, in addition to ASPs.

The difficulty is that there are few larger pedigrees for autism, thus highlighting the potential benefit and utility of the broader autism phenotype. As noted above, classifying individuals with familial autism–related traits as “affected” may increase the prevalence of extended pedigrees and reveal patterns of segregation within those pedigrees beyond what is seen for autism itself. Segregation of such traits (and their underlying genetic diatheses), however, will only be detected if extended pedigrees are sought to begin with. A further benefit to using the BAP is that it enables us to diagnose parents, and possibly siblings, of ASPs as well. Thus even within the nuclear families already collected through existing genomic screens, the amount of linkage information may be greatly increased if indeed the BAP is genetically related to autism.

Another sampling design enabled by the BAP is discordant sibling pair (DSP) analysis. The BAP conceptualizes and measures the traits related to autism along continua. Those subjects who exceed severity thresholds on these measures but do not meet criteria for PDDs are considered to have some form of the BAP. These measures also indicate, however, family members who are very *unaffected*, and therefore discordant, with their autistic siblings. If the traits measured by the BAP truly reflect underlying genetic diatheses toward autism, one would expect reduced allele sharing between DSPs at susceptibility loci as opposed to the excess allele sharing expected in ASPs (141).

The major advantage of the DSP method is that, theoretically, it may require far fewer sibling pairs than the ASP method under some circumstances. Risch and Zhang (142) estimated that DSP analysis may provide the same power to detect disease genes as the ASP method with a sample 10- to 40-fold smaller. The primary disadvantage of the DSP method is that DSPs are difficult to find, and require an extensive screening effort to identify enough pairs.

Microarray Technology

DNA chip microarray technology, which enables the simultaneous analysis of tens of thousands of DNA or RNA sequences, may be of significant benefit to autism research in two primary domains. First, an important focus of the current effort to sequence the human genome is the identification of single nucleotide polymorphisms (SNPs) (143). SNPs are sites of single base pair substitutions, are much more common than di-, tri-, or tetranucleotide polymorphisms, and, in contrast to current methodologies, can be easily and rapidly genotyped using DNA chips (144). The hope, therefore, is that DNA SNP chips will be used to rapidly screen a dense marker map, thereby enabling the performance of genome-wide association studies (145). Implementation of DNA chips for this purpose awaits the development of such SNP maps as well as the statistical and computational tools with which to analyze and interpret the resultant data.

The second potential use of microarray technology is to examine gene expression patterns in relevant tissues. DNA chips can be created that recognize all possible mRNAs that a tissue is currently producing (146). Thus, the expression

of thousands of genes from the brains of autistic individuals and controls can be compared in order to detect etiologically meaningful differences. The primary limitations of this method are the collection of brain tissue, which must be done rapidly and according to exact protocols, and interpretation of the data generated from such experiments (147). Altered gene expression, for example, may be due to medication effects as opposed to an etiologically relevant mechanism. The most meaningful studies of gene expression at this time, therefore, may come from animals genetically designed to exhibit autistic behaviors, where the environment and timing of the tissue analysis can be controlled.

Refining the Phenotype Through Biological Correlates of Illness

The primary purpose for investigating the BAP is to enable identification of meaningful genetic subgroups based on the core symptom domains of autism. Similar strategies are being pursued with various biological correlates of the disorder, including serotonin metabolism, minor physical anomalies, head circumference, brain morphology and function, and others.

CONCLUSION

Part of "41 - The Molecular and Cellular Genetics of Autism "

Numerous complementary strategies are currently being employed to attempt to locate autism disease genes. Linkage studies have identified a number of suggestive loci, most notably distal 7q. Interest in this region is supported by findings from language-disorder families and from a small number of 7q chromosomal anomalies in individuals with autism. Chromosomal anomalies also implicate 15q11-q13 as a region of interest, though linkage and association studies of the region have not been as impressive. These findings clearly demonstrate progress in the effort to find regions harboring genes in autism. As sample sizes continue to grow through collaborative efforts and molecular and statistical methods improve, and as complementary strategies such as the use of the BAP emerge, it seems plausible that in the near future more disease loci will be uncovered and specific autism disease genes may be identified.

REFERENCES

1. Amir RE, Van den Veyver IB, Wan M, et al. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genet* 1999;23:185–188.

2. Fombonne E. The epidemiology of autism: a review. *Psychol Med* 1999;29:769–86.

3. Leventhal BL, Cook EH Jr, Morford M, et al. Relationships of whole blood serotonin and plasma norepinephrine within families. *J Autism Dev Disord* 1990;20:499–511.

4. Kuperman S, Beeghly JH, Burns TL, et al. Serotonin relationships of autistic probands and their first-degree relatives. *J Am Acad Child Psychiatry* 1985;24:186–190.

5. Abramson RK, Wright HH, Carpenter R, et al. Elevated blood serotonin in autistic probands and their first-degree relatives. *J Autism Dev Disord* 1989;19:397–407.

6. Piven J, Tsai GC, Nehme E, et al. Platelet serotonin, a possible marker for familial autism. *J Autism Dev Disord* 1991;21:51–59.

7. Rodier PM, Bryson SE, Welch JP. Minor malformations and physical measurements in autism: data from Nova Scotia. *Teratology* 1997;55:319–325.

8. Volkmar FR, Nelson DS. Seizure disorders in autism. *J Am Acad Child Adolesc Psychiatry* 1990;29:127–129.

9. Fombonne E, Roge B, Claverie J, et al. Microcephaly and macrocephaly in autism. *J Autism Dev Disord* 1999;29:113–119.

10. Lainhart JE, Piven J, Wzorek M, et al. Macrocephaly in children and adults with autism. *J Am Acad Child Adolesc Psychiatry* 1997;36:282–290.

11. Bailey A, Luthert P, Bolton P, et al. Autism and megalencephaly. *Lancet* 1993;341:1225–1226.

12. Bauman ML. Brief report: neuroanatomic observations of the brain in pervasive developmental disorders. *J Autism Dev Disord* 1996;26:199–203.

13. Stevenson RE, Schroer RJ, Skinner C, et al. Autism and macrocephaly. *Lancet* 1997;349:1744–1745.

14. Piven J, Arndt S, Bailey J, et al. An MRI study of brain size in autism. *Am J Psychiatry* 1995;152:1145–1149.

15. Piven J, Saliba K, Bailey J, et al. An MRI study of autism: the cerebellum revisited. *Neurology* 1997;49:546–551.

16. Piven J, Arndt S, Bailey J, et al. Regional brain enlargement in autism: a magnetic resonance imaging study. *J Am Acad Child Adolesc Psychiatry* 1996;35:530–536.

17. Sears LL, Vest C, Mohamed S, et al. An MRI study of the basal ganglia in autism. *Prog Neuropsychopharmacol Biol Psychiatry* 1999;23:613–624.

18. Gillberg C, Wing L. Autism: not an extremely rare disorder. *Acta Psychiatr Scand* 1999;99:399–406.

19. Ritvo ER, Jorde LB, Mason-Brothers A, et al. The UCLA-University of Utah epidemiologic survey of autism: recurrence risk estimates and genetic counseling. *Am J Psychiatry* 1989;146:1032–1036.

20. Szatmari P, Jones MB, Zwaigenbaum L, et al. Genetics of autism: overview and new directions. *J Autism Dev Disord* 1998;28:351–368.

21. Bailey A, Le CA, Gottesman I, et al. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* 1995;25:63–77.

22. Folstein S. Twin and adoption studies in child and adolescent psychiatric disorders. *Curr Opin Pediatr* 1996;8:339–347.

23. Barton M, Volkmar F. How commonly are known medical conditions associated with autism? *J Autism Dev Disord* 1998;28:273–278.

24. Baker P, Piven J, Sato Y. Autism and tuberous sclerosis complex: prevalence and clinical features. *J Autism Dev Disord* 1998;28:279–285.

25. Folstein S, Rutter M. Infantile autism: a genetic study of 21 twin pairs. *J Child Psychol Psychiatry* 1977;18:297–321.

26. Nelson KB. Prenatal and perinatal factors in the etiology of autism. *Pediatrics* 1991;87:761–766.

27. Bolton PF, Murphy M, Macdonald H, et al. Obstetric complications in autism: consequences or causes of the condition? *J Am Acad Child Adolesc Psychiatry* 1997;36:272–281.

28. Chess S. Autism in children with congenital rubella. *J Autism Child Schizophr* 1971;1:33–47.

29. Chess S. Follow-up report on autism in congenital rubella. *J Autism Child Schizophr* 1977;7:69–81.

30. Bolton P, Pickles A, Harrington R, et al. Season of birth: issues, approaches and findings for autism. *J Child Psychol Psychiatry* 1992;33:509–530.

31. Gillberg C, Wahlstrom J. Chromosome abnormalities in infantile autism and other childhood psychoses: a population study of 66 cases. *Dev Med Child Neurol* 1985;27:293–304.

32. Lauritsen M, Mors O, Mortensen PB, et al. Infantile autism and associated autosomal chromosome abnormalities: a register-based study and a literature survey. *J Child Psychol Psychiatry* 1999;40:335–345.

33. Weidmer-Mikhail E, Sheldon S, Ghaziuddin M. Chromosomes in autism and related pervasive developmental disorders: a cytogenetic study. *J Intell Disabil Res* 1998;42:8–12.

34. Konstantareas MM, Homatidis S. Chromosomal abnormalities in a series of children with autistic disorder. *J Autism Dev Disord* 1999;29:275–285.

35. Szatmari P. Heterogeneity and the genetics of autism. *J Psychiatry Neurosci* 1999;24:159–165.

36. Pickles A, Bolton P, Macdonald H, et al. Latent-class analysis of recurrence risks for complex phenotypes with selection and measurement error: a twin and family history study of autism. *Am J Hum Genet* 1995;57:717–276.

37. Risch N, Spiker D, Lotspeich L, et al. A genomic screen of autism: evidence for a multilocus etiology. *Am J Hum Genet* 1999;65:493–507.

38. Anonymous. A full genome screen for autism with evidence for linkage to a region on chromosome 7q. International Molecular Genetic Study of Autism Consortium. *Hum Mol Genet* 1998;7:571–578.

39. Anonymous. An autosomal genomic screen for autism. *Am J Med Genet* 1999;88:609–615.

40. Philippe A, Martinez M, Guilloud-Bataille M, et al. Genome-wide scan for autism susceptibility genes. Paris Autism Research International Sibpair Study. *Hum Mol Genet* 1999;8:805–812.

41. Kruglyak L, Daly MJ, Reeve-Daly MP, et al. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996;58:1347–1363.

42. Smith CAB. Testing the heterogeneity of the recombination fraction values in human genetics. *Ann Hum Genet* 1963;27:175–182.

43. Hauser ER, Boehnke M, Guo SW, et al. Affected-sib-pair interval mapping and exclusion for complex genetic traits: sampling considerations. *Genet Epidemiol* 1996;13:117–137.

44. Hinds D, Risch N. The ASPEX package: affected sib-pair mapping. ftp://lahmed.stanford.edu/pub/aspex 1996.

45. Holmans P. Asymptotic properties of affected-sib-pair linkage analysis. *Am J Hum Genet* 1993;52:362–374.

46. IMGSAC. Search for autism susceptibility loci: Genome screen follow-up and fine mapping of a candidate region on chromosome 7q. *Am J Hum Genet* 1999;65(suppl):A106(abst 560).

47. Ashley-Koch A, Wolpert CM, Menold MM, et al. Genetic studies of autistic disorder and chromosome 7. *Genomics* 1999;61:227–236.

48. Bass MP, Menold MM, Wolpert CM, et al. Genetic studies in autistic disorder and chromosome 15. *Neurogenetics* 2000;2:219–226.

49. Cook EH Jr, Courchesne R, Lord C, et al. Evidence of linkage between the serotonin transporter and autistic disorder. *Mol Psychiatry* 1997;2:247–250.

50. Klauck SM, Poustka F, Benner A, et al. Serotonin transporter (5-HTT) gene variants associated with autism? *Hum Mol Genet* 1997;6:2233–2238.

51. Maestrini E, Lai C, Marlow A, et al. Serotonin transporter (5-HTT) and gamma-aminobutyric acid receptor subunit beta3 (GABRB3) gene polymorphisms are not associated with autism in the IMGSAC families. The International Molecular Genetic Study of Autism Consortium. *Am J Med Genet* 1999;88:492–496.

52. Vincent JB, Herbrick JA, Gurling HMD, et al. Identification of genes at translocation breakpoints on chromosome 7q31 in autistic individuals. *Am J Hum Genet Suppl* 1999;65:A471.

53. Sultana R, Yu J, Raskind W, et al. Cloning of a candidate gene (ARG1) from the breakpoint of t(7;20) in an autistic twin pair. *Am J Hum Genet Suppl* 1999;65:A44.

54. Warburton P, Baird G, Chen W, et al. Support for linkage of autism and specific language impairment to 7q3 from two chromosome rearrangements involving band 7q31. *Am J Med Genet* 2000;96:228–234.

55. Robinson WP, Binkert F, Gine R, et al. Clinical and molecular analysis of five inv dup(15) patients. *Eur J Hum Genet* 1993;1:37–50.

56. Leana-Cox J, Jenkins L, Palmer CG, et al. Molecular cytogenetic analysis of inv dup(15) chromosomes, using probes specific for the Prader-Willi/Angelman syndrome region: clinical implications. *Am J Hum Genet* 1994;54:748–756.

57. Schinzel AA, Brecevic L, Bernasconi F, et al. Intrachromosomal triplication of 15q11-q13. *J Med Genet* 1994;31:798–803.

58. Baker P, Piven J, Schwartz S, et al. Brief report: duplication of chromosome 15q11-13 in two individuals with autistic disorder. *J Autism Dev Disord* 1994;24:529–535.

59. Bundey S, Hardy C, Vickers S, et al. Duplication of the 15q11-13 region in a patient with autism, epilepsy and ataxia. *Dev Med Child Neurol* 1994;36:736–742.

60. Hotopf M, Bolton P. A case of autism associated with partial tetrasomy 15. *J Autism Dev Disord* 1995;25:41–49.

61. Flejter WL, Bennett-Baker PE, Ghaziuddin M, et al. Cytogenetic and molecular analysis of inv dup(15) chromosomes observed in two patients with autistic disorder and mental retardation. *Am J Med Genet* 1996;61:182–187.

62. Schinzel A. Autistic disorder and additional inv dup(15)(pter-q13) chromosome. *Am J Med Genet* 1990;35:447–448.

63. Gillberg C, Steffenburg S, Wahlstrom J, et al. Autism associated with marker chromosome. *J Am Acad Child Adolesc Psychiatry* 1991;30:489–494.

64. Browne CE, Dennis NR, Maher E, et al. Inherited interstitial duplications of proximal 15q: genotype-phenotype correlations. *Am J Hum Genet* 1997;61:1342–1352.

65. Wolpert CM, Menold MM, Bass MP, et al. Three patients with autistic disorder and isodicentric 15 chromosomes: implications for genetic analysis. *Am J Med Genet* 2000;96:365–372.

66. Huang B, Crolla JA, Christian SL, et al. Refined molecular characterization of the breakpoints in small inv dup(15) chromosomes. *Hum Genet* 1997;99:11–17.

67. Rineer S, Finucane B, Simon EW. Autistic symptoms among children and young adults with isodicentric chromosome 15. *Am J Med Genet* 1998;81:428–433.

68. Kerbeshian J, Burd L, Randall T, et al. Autism, profound mental retardation and atypical bipolar disorder in a 33-year-old female with a deletion of 15q12. *J Ment Defic Res* 1990;34:205–210.

69. Sabry MA, Farag TI. Chromosome 15q11-13 region and the autistic disorder. *J Intell Disabil Res* 1998;42:259.

70. Cook EH Jr, Lindgren V, Leventhal BL, et al. Autism or atypical autism in maternally but not paternally derived proximal 15q duplication. *Am J Hum Genet* 1997;60:928–934.

71. Constancia M, Pickard B, Kelsey G, et al. Imprinting mechanisms. *Genome Res* 1998;8:881–900.

72. Nicholls RD, Saitoh S, Horsthemke B. Imprinting in Prader-Willi and Angelman syndromes. *Trends Genet* 1998;14:194–200.

73. Knoll JH, Wagstaff J, Lalande M. Cytogenetic and molecular studies in the Prader-Willi and Angelman syndromes: an overview. *Am J Med Genet* 1993;46:2–6.

74. Ozcelik T, Leff S, Robinson W, et al. Small nuclear ribonucleoprotein polypeptide N (SNRPN), an expressed gene in the Prader-Willi syndrome critical region. *Nat Genet* 1992;2:265–269.

75. Knee R, Murphy PR. Regulation of gene expression by natural antisense RNA transcripts. *Neurochem Int* 1997;31:379–392.

76. Rougeulle C, Cardoso C, Fontes M, et al. An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. *Nat Genet* 1998;19:15–16.

77. Eichler E. Masquerading repeats: paralogous pitfalls of the human genome. *Genome Res* 1998;8:758–762.

78. Dorschner MO, Sybert VP, Weaver M, et al. NF1 microdeletion breakpoints are clustered at flanking repetitive sequences. *Hum Mol Genet* 2000;9:35–46.

79. Potocki L, Chen KS, Park SS, et al. Molecular mechanism for duplication 17p11.2—the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet* 2000;24:84–87.

80. Ji Y, Walkowicz MJ, Buiting K, et al. The ancestral gene for transcribed, low-copy repeats in the Prader-Willi/Angelman region encodes a large protein implicated in protein trafficking, which is deficient in mice with neuromuscular and spermiogenic abnormalities. *Hum Mol Genet* 1999;8:533–542.

81. Ritchie RJ, Mattei MG, Lalande M. A large polymorphic repeat in the pericentromeric region of human chromosome 15q contains three partial gene duplications. *Hum Mol Genet* 1998;7:1253–1260.

82. Robinson WP, Dutly F, Nicholls RD, et al. The mechanisms involved in formation of deletions and duplications of 15q11-q13. *J Med Genet* 1998;35:130–136.

83. Edelmann L, Pandita RK, Morrow BE. Low-copy repeats mediate the common 3-Mb deletion in patients with velo-cardio-facial syndrome. *Am J Hum Genet* 1999;64:1076–1086.

84. Wilson DI, Burn J, Scambler P, et al. DiGeorge syndrome: part of CATCH 22. *J Med Genet* 1993;30:852–856.

85. Bassett AS, Hodgkinson K, Chow EW, et al. 22q11 deletion syndrome in adults with schizophrenia. *Am J Med Genet* 1998;81:328–337.

86. Gothelf D, Frisch A, Munitz H, et al. Velocardiofacial manifestations and microdeletions in schizophrenic inpatients. *Am J Med Genet* 1997;72:455–461.

87. Brown WT, Jenkins EC, Friedman E, et al. Autism is associated with the fragile-X syndrome. *J Autism Dev Disord* 1982;12:303–308.

88. Gillberg C. Identical triplets with infantile autism and the fragile-X syndrome. *Br J Psychiatry* 1983;143:256–260.

89. Kaufmann WE, Reiss AL. Molecular and cellular genetics of fragile X syndrome. *Am J Med Genet* 1999;88:11–24.

90. Blomquist HK, Bohman M, Edvinsson SO, et al. Frequency of the fragile X syndrome in infantile autism. A Swedish multicenter study. *Clin Genet* 1985;27:113–117.

91. Piven J, Gayle J, Landa R, et al. The prevalence of fragile X in a sample of autistic individuals diagnosed using a standardized interview. *J Am Acad Child Adolesc Psychiatry* 1991;30:825–830.

92. Bailey A, Bolton P, Butler L, et al. Prevalence of the fragile X anomaly amongst autistic twins and singletons. *J Child Psychol Psychiatry* 1993;34:673–688.

93. Dykens EM, Volkmar FR. Medical conditions associated with autism. In: Cohen DJ, Volkmar FR, eds. *Handbook of autism and pervasive developmental disorders*. New York: Wiley, 1997:388–410.

94. Hagerman RJ, Wilson P, Staley LW, et al. Evaluation of school children at high risk for fragile X syndrome utilizing buccal cell FMR-1 testing. *Am J Med Genet* 1994;51:474–481.

95. Hagerman RJ, Chudley AE, Knoll JH, et al. Autism in fragile X females. *Am J Med Genet* 1986;23:375–380.

96. de Vries BB, Halley DJ, Oostra BA, et al. The fragile X syndrome. *J Med Genet* 1998;35:579–589.

97. Khandjian EW. Biology of the fragile X mental retardation protein, an RNA-binding protein. *Biochem Cell Biol* 1999;77:331–342.

98. Drouin R, Angers M, Dallaire N, et al. Structural and functional characterization of the human FMR1 promoter reveals similarities with the hnRNP-A2 promoter region. *Hum Mol Genet* 1997;6:2051–2060.

99. Thomas NS, Sharp AJ, Browne CE, et al. Xp deletions associated with autism in three females. *Hum Genet* 1999;104:43–48.

100. Skuse DH, James RS, Bishop DV, et al. Evidence from Turner's syndrome of an imprinted X-linked locus affecting cognitive function. *Nature* 1997;387:705–708.

101. James RS, Coppin B, Dalton P, et al. A study of females with deletions of the short arm of the X chromosome. *Hum Genet* 1998;102:507–516.

102. Donnelly SL, Wolpert CM, Menold MM, et al. Female with autistic disorder and monosomy X (Turner syndrome): parent-of-origin effect of the X chromosome. *Am J Med Genet* 2000;96:312–316.

103. Hallmayer J, Hebert JM, Spiker D, et al. Autism and the X chromosome. Multipoint sib-pair analysis. *Arch Gen Psychiatry* 1996;53:985–989.

104. Folstein S, Rutter M. Infantile autism: a genetic study of 21 twin pairs. *J Child Psychol Psychiatry* 1977;18:297–321.

105. Bolton P, Macdonald H, Pickles A, et al. A case-control family history study of autism. *J Child Psychol Psychiatry* 1994;35:877–900.

106. Piven J, Palmer P, Jacobi D, et al. Broader autism phenotype: evidence from a family history study of multiple-incidence autism families. *Am J Psychiatry* 1997;154:185–190.

107. Piven J, Palmer P. Cognitive deficits in parents from multiple-incidence autism families. *J Child Psychol Psychiatry* 1997;38:1011–1021.

108. Piven J, Palmer P, Landa R, et al. Personality and language characteristics in parents from multiple-incidence autism families. *Am J Med Genet* 1997;74:398–411.

109. Hughes C, Leboyer M, Bouvard M. Executive function in parents of children with autism. *Psychol Med* 1997;27:209–220.

110. Grigorenko EL, Wood FB, Meyer MS, et al. Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *Am J Hum Genet* 1997;60:27–39.

111. Bishop DV. How does the brain learn language? Insights from the study of children with and without language impairment. *Dev Med Child Neurol* 2000;42:133–142.

112. Hafeman L, Tomblin JB. Autism behaviors in the siblings of children with specific language impairment. *Mol Psychiatry* 1999;4(suppl 1):S14.

113. Tomblin JB, Nishimura C, Zhang X, et al. Association of developmental language impairment with loci at 7q3. *Am J Hum Genet Suppl* 1998;63:A312.

114. Fisher SE, Vargha-Khadem F, Watkins KE, et al. Localisation of a gene implicated in a severe speech and language disorder. *Nature Genet* 1999;18:168–170.

115. Hyde TM, Stacey ME, Coppola R, et al. Cerebral morphometric abnormalities in Tourette's syndrome: a quantitative MRI study of monozygotic twins. *Neurology* 1995;45:1176–1182.

116. Rosenberg DR, Keshavan MS, O'Hearn KM, et al. Frontostriatal measurement in treatment-naive children with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1997;54:824–830.

117. Piven J, Gayle J, Chase GA, et al. A family history study of neuropsychiatric disorders in the adult siblings of autistic individuals. *J Am Acad Child Adolesc Psychiatry* 1990;29:177–183.

118. de Bono M, Bargmann CI. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in C. elegans. *Cell* 1998;94:679–689.

119. Insel TR, O'Brien DJ, Leckman JF. Oxytocin, vasopressin, and autism: is there a connection? *Biol Psychiatry* 1999;45:145–157.

120. Gomez MR. *Tuberous sclerosis*. New York: Raven Press, 1988.

121. O'Callaghan FJ, Shiell AW, Osborne JP, et al. Prevalence of tuberous sclerosis estimated by capture-recapture analysis. *Lancet* 1998;351:1490.

122. Hunt A, Dennis J. Psychiatric disorder among children with tuberous sclerosis. *Dev Med Child Neurol* 1987;29:190–198.

123. Smalley SL. Autism and tuberous sclerosis. *J Autism Dev Disord* 1998;28:407–414.

124. Smalley SL, Tanguay PE, Smith M, et al. Autism and tuberous sclerosis. *J Autism Dev Disord* 1992;22:339–355.

125. Crino PB, Henske EP. New developments in the neurobiology of the tuberous sclerosis complex. *Neurology* 1999;53:1384–1390.

126. Nellist M, van Slegtenhorst MA, Goedbloed M, et al. Characterization of the cytosolic tuberin-hamartin complex. Tuberin is a cytosolic chaperone for hamartin. *J Biol Chem* 1999;274:35647–35652.

127. Zhang H, Nanba E, Yamamoto T, et al. Mutational analysis of TSC1 and TSC2 genes in Japanese patients with tuberous sclerosis complex. *J Hum Genet* 1999;44:391–396.

128. Jones AC, Shyamsundar MM, Thomas MW, et al. Comprehensive mutation analysis of TSC1 and TSC2—and phenotypic correlations in 150 families with tuberous sclerosis. *Am J Hum Genet* 1999;64:1305–1315.

129. Bolton PF, Griffiths PD. Association of tuberous sclerosis of temporal lobes with autism and atypical autism. *Lancet* 1997;349:392–395.

130. Warren RP, Odell JD, Warren WL, et al. Strong association of the third hypervariable region of HLA-DR beta 1 with autism. *J Neuroimmunol* 1996;67:97–102.

131. Stubbs EG, Crawford ML. Depressed lymphocyte responsiveness in autistic children. *J Autism Child Schizophr* 1977;7:49–55.

132. Warren RP, Margaretten NC, Pace NC, et al. Immune abnormalities in patients with autism. *J Autism Dev Disord* 1986;16:189–197.

133. Yonk LJ, Warren RP, Burger RA, et al. CD4+ helper T cell depression in autism. *Immunol Lett* 1990;25:341–345.

134. Warren RP, Yonk LJ, Burger RA, et al. Deficiency of suppressor-inducer (CD4+CD45RA+) T cells in autism. *Immunol Invest* 1990;19:245–251.

135. Plioplys AV. Autism: electroencephalogram abnormalities and clinical improvement with valproic acid. *Arch Pediatr Adolesc Med* 1994;148:220–222.

136. Warren RP, Yonk J, Burger RW, et al. DR-positive T cells in autism: association with decreased plasma levels of the complement C4B protein. *Neuropsychobiology* 1995;31:53–57.

137. Warren RP, Singh VK, Cole P, et al. Increased frequency of the null allele at the complement C4b locus in autism. *Clin Exp Immunol* 1991;83:438–440.

138. Warren RP, Singh VK, Cole P, et al. Possible association of the extended MHC haplotype B44-SC30-DR4 with autism. *Immunogenetics* 1992;36:203–207.

139. Daniels WW, Warren RP, Odell JD, et al. Increased frequency of the extended or ancestral haplotype B44-SC30-DR4 in autism. *Neuropsychobiology* 1995;32:120–123.

140. Warren RP, Singh VK, Averett RE, et al. Immunogenetic studies in autism and related disorders. *Mol Chem Neuropathol* 1996;28:77–81.

141. Xu X, Rogus JJ, Terwedow HA, et al. An extreme-sib-pair genome scan for genes regulating blood pressure. *Am J Hum Genet* 1999;64:1694–1701.

142. Risch N, Zhang H. Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science* 1995;268:1584–1589.

143. Collins FS, Patrinos A, Jordan E, et al. New goals for the U.S. Human Genome Project: 1998–2003. *Science* 1998;282:682–689.

144. Kwok PY, Gu Z. Single nucleotide polymorphism libraries: why and how are we building them? *Mol Med Today* 1999;5:538–543.

145. Ramsay G. DNA chips: state-of-the art. *Nat Biotechnol* 1998;16:40–44.

146. Khan J, Saal LH, Bittner ML, et al. Expression profiling in cancer using cDNA microarrays. *Electrophoresis* 1999;20:223–229.

147. Zhang MQ. Large-scale gene expression data analysis: a new challenge to computational biologists. *Genome Res* 1999;9:681–688.

148. Martin ER, Menold MM, Wolpert CM, et al. Analysis of linkage disequilibrium in γ-aminobutyric acid receptor subunit genes in autistic disorder. *Am J Med Genet* 2000;96:43–48.

149. Salmon B, Hallmayer J, Rogers T, et al. Absence of linkage and linkage disequilibrium to chromosome 15q11-q13 markers in 139 multiplex families with autism. *Am J Med Genet* 1999;88:551–556.

150. Rogers T, Kalaydjieva L, Hallmayer J, et al. Exclusion of linkage to the HLA region in ninety multiplex sibships with autism. *J Autism Dev Disord* 1999;29:195–201.

151. Cook EH, Jr., Courchesne RY, Cox NJ, et al. Linkage-disequilibrium mapping of autistic disorder with 15q11-13 markers. *Am J Hum Genet* 1998;62:1077–1083.

152. Lassig JP, Vachirasomtoon K, Hartzell K, et al. Physical mapping of the serotonin 5-HT(7) receptor gene (HTR7) to chromosome 10 and pseudogene (HTR7P) to chromosome 12, and testing of linkage disequilibrium between HTR7 and autistic disorder. *Am J Med Genet* 1999;88:472–475.

153. Hallmayer J, Pintado E, Lotspeich L, et al. Molecular analysis and test of linkage between the FMR-1 gene and infantile autism in multiplex families. *Am J Hum Genet* 1994;55:951–959.

154. Mbarek O, Marouillat S, Martineau J, et al. Association study of the NF1 gene and autistic disorder. *Am J Med Genet* 1999;88:729–732.

155. Veenstra-VanderWeele J, Gonen D, Leventhal BL, et al. Mutation screening of the UBE3A/E6-AP gene in autistic disorder. *Mol Psychiatry* 1999;4:64–67.

156. Herault J, Perrot A, Barthelemy C, et al. Possible association of c-Harvey-Ras-1 (HRAS-1) marker with autism. *Psychiatry Res* 1993;46:261–267.

157. Herault J, Petit E, Martineau J, et al. Autism and genetics: clinical approach and association study with two markers of HRAS gene. *Am J Med Genet* 1995;60:276–281.

158. Comings DE, Wu S, Chiu C, et al. Studies of the c-Harvey-Ras gene in psychiatric disorders. *Psychiatry Res* 1996;63:25–32.

159. Comings DE, Gade R, Muhleman D, et al. No association of a tyrosine hydroxylase gene tetranucleotide repeat polymorphism in autism, Tourette syndrome, or ADHD. *Biol Psychiatry* 1995;37:484–486.

160. Petit E, Herault J, Martineau J, et al. Association study with two markers of a human homeogene in infantile autism. *J Med Genet* 1995;32:269–274.

161. Herault J, Petit E, Buchler M, et al. Lack of association between three genetic markers of brain growth factors and infantile autism. *Biol Psychiatry* 1994;35:281–283.

42

Current and Emerging Therapeutics of Autistic Disorder and Related Pervasive Developmental Disorders

Christopher J. McDougle

Christopher J. McDougle: Department of Psychiatry, Section of Child and Adolescent Psychiatry, Indiana University School of Medicine; James Whitcomb Riley Hospital for Children, Indianapolis, Indiana.

DIAGNOSIS OF PERVASIVE DEVELOPMENTAL DISORDERS

Part of "42 - Current and Emerging Therapeutics of Autistic Disorder and Related Pervasive Developmental Disorders "

Pervasive developmental disorders (PDDs) are characterized by severe and pervasive impairment in several areas of development, including reciprocal social interaction skills, communication skills, and the presence of stereotyped behavior, interests, and activities (1). The qualitative impairments that define these disorders are abnormal relative to the individual's developmental level or mental age. These conditions are typically evident in the first 1 to 3 years of life and are usually associated with some degree of mental retardation. The PDDs are sometimes observed among a diverse group of identifiable biological abnormalities (e.g., chromosomal abnormalities, congenital infections, structural abnormalities of the brain). In the majority of cases, however, a specific etiologic factor is not found. Previously, terms like *psychosis* and *childhood schizophrenia* were used to refer to individuals with these conditions. However, there is now considerable evidence to demonstrate that PDDs are distinct from schizophrenia. There are five subtypes of PDD in the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV) (1). They include autistic disorder, Rett syndrome, childhood disintegrative disorder, Asperger's syndrome, and PDD not otherwise specified (NOS) (Table 42.1).

Autistic disorder
Rett syndrome
Childhood disintegrative disorder
Asperger's syndrome
Pervasive developmental disorder not otherwise specified

DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders,* fourth ed.

TABLE 42.1. DSM-IV PERVASIVE DEVELOPMENTAL DISORDERS (1)

Autistic Disorder

The essential features of autistic disorder are the presence of markedly abnormal or impaired development in social interaction and communication and a markedly restricted repertoire of activity and interests. The clinical presentation of the disorder varies significantly depending on the developmental level and chronologic age of the individual. In the past, autistic disorder was referred to as early infantile autism, childhood autism, or Kanner's syndrome.

The clinical features must include delays or abnormal functioning in social interaction, in language as used in social communication, or in symbolic or imaginative play prior to the age of 3 years. There is usually no period of clearly normal development, although 1 or 2 years of relatively normal progression can occur. In some instances, regression in language development, typically manifest as a complete loss of speech after a child has acquired from five to ten words, has been reported. If there is a period of normal development, it cannot extend past the age of 3 years.

In approximately 75% of cases, there is an associated diagnosis of mental retardation, commonly in the moderate range (IQ 35 to 50). A number of behavioral symptoms, including hyperactivity; inattention; impulsivity; aggression toward self, others, or property; and interfering repetitive thoughts and behavior are often present. The disorder is sometimes observed in association with an identifiable medical condition (e.g., herpes encephalitis, phenylketonuria, tuberous sclerosis, fragile X syndrome, anoxia during birth, maternal rubella). Seizures may develop, particularly in adolescence, in up to 25% to 33% of cases. The disorder is four to five times more common in males than in females, although females often have a more severe cognitive disability. Epidemiologic studies have identified rates of autistic disorder of two to five cases per 10,000. Language skills and IQ are the strongest predictors of eventual outcome.

Rett Syndrome

Rett syndrome differs from autistic disorder in its characteristic sex ratio and distinctive pattern of abnormal developmental

progression. The syndrome is much less common than autistic disorder and has been diagnosed almost exclusively in females. Following apparently normal prenatal and perinatal development, and a period of normal psychomotor development through the first 5 months of life, there is a characteristic pattern of head growth deceleration, loss of previously acquired purposeful hand skills, intermittent hyperventilation, and the appearance of ataxic gait or trunk movements. Individuals with Rett syndrome may exhibit difficulties in social interaction, particularly during the preschool years, but these may improve somewhat over time. Severe or profound mental retardation, seizures, and significant expressive and receptive language impairment are typical. Recently, a mutation in the gene (*MECP2*) encoding X-linked methyl-CpG-binding protein 2 (MeCP2) has been identified as the cause of some cases of Rett syndrome (2).

Childhood Disintegrative Disorder

Childhood disintegrative disorder contrasts with autistic disorder in that there is a distinctive pattern of developmental regression following at least 2 years of normal development. In autistic disorder, some developmental abnormalities are usually noted within the first year of life. After the first 2 years of life (but before the age of 10 years), the child with childhood disintegrative disorder has a clinically significant loss of previously acquired skills in at least two of the following areas: expressive or receptive language, social skills or adaptive behavior, bowel or bladder control, play, or motor skills. The onset, in most cases, is between the ages of 3 and 4 years and may be insidious or abrupt. To date, an underlying pathologic mechanism has not been identified. The disorder has been reported in association with metachromatic leukodystrophy and Schilder's disease, in some cases. Childhood disintegrative disorder is usually associated with severe mental retardation. It appears to be very rare, much less frequent than autistic disorder, and more common among males. The disorder has also been termed Heller's syndrome, dementia infantilis, or disintegrative psychosis.

Asperger's Syndrome

Asperger's syndrome can be distinguished from autistic disorder by the lack of delay in language and cognitive development, in addition to no significant abnormality in the development of age-appropriate self-help skills, adaptive behavior (other than in social interaction), and curiosity about the environment in childhood. Motor milestones may be delayed, and motor clumsiness is often observed. The syndrome appears to be more common in males. Asperger's syndrome is usually recognized somewhat later than autistic disorder, frequently in the context of school. All-encompassing preoccupations or circumscribed interests are typically present and can contribute to significant functional impairment.

Pervasive Developmental Disorder Not Otherwise Specified

PDD NOS is diagnosed when there is a severe and pervasive impairment in the development of reciprocal social interaction or verbal and nonverbal communication skills, but the criteria are not met for a specific PDD, schizophrenia, schizotypal personality disorder, or avoidant personality disorder. Stereotyped behavior, interests, and activities are often present. This category includes presentations associated with late age at onset, atypical or subthreshold symptoms, or both.

PHARMACOTHERAPY OF PERVASIVE DEVELOPMENTAL DISORDERS

Part of "42 - Current and Emerging Therapeutics of Autistic Disorder and Related Pervasive Developmental Disorders "

The treatment of autistic disorder and related PDDs is multimodal and largely based on educational interventions and behavior management principles. Speech therapy is usually indicated, and physical and occupational therapy are often needed as well. Despite educational and behavioral strategies, many children, adolescents, and adults with PDDs remain significantly impaired. Under these circumstances, pharmacologic treatment is often appropriate and warranted.

Adequate drug treatment studies that have been focused on subjects with particular subtypes of PDD, other than autistic disorder, have not been conducted. Many investigations have included mixed samples of subjects with autistic disorder, Asperger's syndrome, and PDD NOS. Because of the extreme rarity of Rett syndrome and childhood disintegrative disorder, essentially no systematic psychopharmacologic treatment research has occurred in subjects with these subtypes of PDD. More recently, researchers have been conducting drug studies in adults with PDDs, in addition to those in children and adolescents with these disorders. The results from these investigations have allowed for some assessment of the impact of developmental factors on drug efficacy and tolerability.

Drugs that have consistent, primary effects on the core social disability of autistic disorder have not yet been developed. Studies in laboratory animals have identified particular

neurochemical systems that mediate some elements of affiliative behavior (3 ,4). The translation of these findings into investigational applications in humans, however, has not yet occurred. The pharmacotherapy of autistic disorder currently involves the identification and treatment of symptoms including motor hyperactivity (primarily in prepubertal autistic individuals); inattention; aggression directed toward self, others, or the environment; and interfering repetitive thoughts and behavior. With reduction in these associated target symptoms, improvement in some aspects of social behavior can result secondarily.

Following a brief review of earlier drug studies, results from more current investigations, including those of atypical antipsychotics and serotonin reuptake inhibitors (SRIs), will be presented in some detail. For a more comprehensive review of drug treatment of PDDs, the reader is referred to other sources (5 ,6).

Early Drug Treatment Studies

Beginning in the 1960s, numerous agents, including lysergic acid diethylamide, methysergide, levodopa, triiodothyronine, imipramine, and 5-hydroxytryptophan were studied in autistic disorder. Many of these investigations were limited by a lack of diagnostic precision and inadequate methodologic design. In general, these initial studies identified no drug that resulted in consistent target symptom reduction.

Elevated levels of whole blood serotonin (5-hydroxytryptamine, 5-HT) have long been associated with autistic disorder in a large minority of subjects (7). Following reports that fenfluramine, an indirect 5-HT agonist, decreased blood and brain 5-HT in animals, this drug received extensive investigation. Despite early enthusiasm generated by small open-label reports, most controlled studies found no consistent efficacy for fenfluramine as a treatment for autistic disorder (8). Furthermore, increasing evidence of possible neurotoxic effects of the drug on 5-HT neurons in animals and the association of fenfluramine with primary pulmonary hypertension and (in combination with phentermine) valvular heart disease have eliminated its use as a safe agent.

Most of the available typical antipsychotic drugs were studied in heterogeneous groups of children that included autistic subjects. Many of these early investigations suffered from the lack of adequate diagnostic methods and nonstandardized outcome measures. Most of these trials were direct comparisons of two drugs, usually low-potency antipsychotics, and did not include a placebo control. A number of these agents were found to be effective for behavioral symptoms including motor hyperactivity, agitation, and stereotypies. Due to significant sedation and adverse cognitive effects secondary to the low-potency drugs, however, studies of higher potency conventional antipsychotics were next pursued.

Campbell and co-workers (9 ,10 ,11 and 12) conducted several well-designed controlled studies of haloperidol in autistic children. In doses of 1.0 to 2.0 mg per day, haloperidol was found to be more efficacious than placebo for withdrawal, stereotypy, hyperactivity, affective lability, anger, and temper outbursts. However, acute dystonic reactions along with withdrawal and tardive dyskinesias were not infrequent.

Beginning in the late 1980s, the opioid antagonist naltrexone was investigated as a potential treatment for the associated behavioral symptoms of autistic disorder, as well as the core social deficits. Again, results from initial open-label reports and small controlled studies suggested possible effectiveness for naltrexone. More recent large well-designed controlled investigations involving children, adolescents, and adults with autistic disorder, however, have failed to demonstrate improvement in the majority of target symptoms or social behavior (13 ,14). The most consistent findings from these controlled studies were that naltrexone is well tolerated and may be effective for reducing motor hyperactivity.

A number of other drugs have been studied in autistic disorder, although most of the trials were either uncontrolled or contained a small number of subjects (5 ,6). For example, ß-adrenergic antagonists have been reported to reduce aggression and self-injury in some small open-label pilot trials in adults with autistic disorder. Hypotension and bradycardia were common dose-related adverse effects. Case reports and small open-label studies have described mixed results with the $5\text{-}HT_{1A}$ partial agonist buspirone. Controlled investigations of mood stabilizers, including lithium, valproic acid, carbamazepine, and gabapentin, have not been reported in well-defined groups of autistic subjects.

The pharmacologic management of motor hyperactivity and impaired attentional mechanisms in individuals with PDDs has proven particularly challenging to clinicians and researchers. These symptoms are most prominent in younger-aged autistic children. Thus, these symptoms are largely present during a time when educational programming and interventions are most critical. The psychostimulants, such as methylphenidate and dextroamphetamine, are effective treatments for these symptoms in individuals with attention-deficit/hyperactivity disorder (ADHD). Early controlled studies of these agents in autistic children, however, produced mixed results at best (15 ,16). In a more recent double-blind crossover study of methylphenidate and placebo, ten autistic children, ages 7 to 11 years, received doses of 10 or 20 mg twice daily for 2 weeks (17). Statistically significant improvement was seen on the Conners Teacher Questionnaire (18) and on the hyperactivity factor, irritability factor, and total score of the Aberrant Behavior Checklist (19). Adverse effects were minimal. The authors' impression was that the effects of methylphenidate were modest. Following completion of the study, it was necessary to add haloperidol to the treatment regimen of two of the ten children due to continued symptoms of aggression. Anecdotal reports from physicians in clinical practice and in

academic centers commonly describe the onset or exacerbation of irritability, insomnia, and aggression in individuals with PDDs with psychostimulant treatment.

The α_2-adrenergic agonist clonidine has been shown to be an effective treatment for some individuals with ADHD. In a small double-blind, placebo-controlled study of clonidine (4 to 10 µg per kg daily) in eight children with autistic disorder, statistically significant improvement was recorded in hyperactivity and irritability on some teacher and parent ratings (20). No significant drug–placebo differences were identified on clinician ratings of videotaped observations, however. Adverse effects included hypotension, sedation, and irritability. In contrast, transdermal clonidine (5 µg per kg daily) was reported to be effective in a double-blind, placebo-controlled crossover study (4 weeks in each treatment phase) involving nine males (ages 5 to 33 years) with autistic disorder (21). Significant improvement was seen on the Clinical Global Impression Scale (CGI) (22), and hyperactivity and anxiety were also reduced. The most common adverse effects were sedation and fatigue.

Guanfacine is an α_2-adrenergic agonist with a longer half-life than clonidine that may be less sedating and cause less pronounced hypotension (23). No open-label or controlled studies have been published on the use of guanfacine in PDDs.

To more rigorously address the pharmacotherapy of symptoms of hyperactivity and inattention in PDD, the National Institute of Mental Health (NIMH)-sponsored Research Units on Pediatric Psychopharmacology (RUPP) Autism Network is planning a controlled investigation of a methylphenidate vs. placebo in children with PDDs.

Current Drug Treatment Studies

Atypical Antipsychotics

Over the past 5 to 10 years, considerable interest has been generated by the introduction of the atypical antipsychotics (24). These drugs appeared to have potential as a treatment for autistic disorder for a number of reasons. Initial studies in schizophrenia indicated that these agents were better tolerated and had a lower risk of acute and tardive dyskinesias compared with conventional antipsychotics. In addition, these drugs were shown to improve both the “positive” (hallucinations and delusions) and “negative” symptoms of schizophrenia. The negative symptoms include blunted affect, emotional and social withdrawal, lack of interest in interpersonal relationships, difficulty in abstract thinking, lack of spontaneity and flow of conversation, and stereotyped thinking (25). A number of investigators suggested that the negative symptoms of schizophrenia were comparable to those that characterize the social impairment of autistic disorder. To date, reports have appeared in which clozapine, risperidone, olanzapine, or quetiapine was used in the treatment of autistic disorder and other PDDs. The reader is referred to a recent publication that provides a comprehensive review of atypical antipsychotics in autistic disorder (26).

Clozapine

Clozapine has been shown to be effective for treatment-refractory schizophrenia (27). The drug’s ability to block 5-HT_{2A}, 5-HT_{2C}, 5-HT_3, and dopamine D_1-D_4 receptors has been proposed as its mechanism of action. There has been only one report to date describing the use of clozapine in autistic disorder (28). Three children with significant hyperactivity or aggression were given clozapine after they had not responded to typical antipsychotics. Improvement was observed in the three subjects after 3 months’ treatment at dosages up to 200 mg per day. The scarcity of reports describing the use of clozapine in autistic disorder might reflect concern regarding the risk of agranulocytosis or seizures in children or adolescents that are associated with the drug. Because autistic individuals typically have an impaired ability to communicate effectively and often a high pain threshold, infections secondary to a decreased white blood cell count may not be identified in a timely manner. Additionally, as mentioned above, many individuals with autistic disorder have comorbid seizure disorders. Furthermore, the necessary frequent blood draws are not ideal for children, particularly those with autistic disorder.

Risperidone

Risperidone is a highly potent 5-HT_{2A}/D_2 antagonist that has been shown in controlled trials to improve both the positive and negative symptoms of schizophrenia (29). A number of open-label reports describing improvement in aggression, self-injury, ritualistic behavior, irritability, impulsivity, hyperactivity, and social relatedness in children, adolescents, and adults with autistic disorder have appeared (26). Only one controlled study of risperidone, or any atypical antipsychotic for that matter, has been published in individuals with autistic disorder and related PDDs (30) (Table 42.2). In that study, 31 adults (mean age, 28.1 years) with autistic disorder (*n* = 17) or PDD NOS (*n* = 14) entered the 12-week trial. For subjects completing the study, eight (57%) of 14 treated with risperidone [mean ± standard deviation (SD) dose, 2.9 ± 1.4 mg daily; range 1 to 6 mg per day] were categorized as responders compared with none of 16 in the placebo group based on the CGI. Nine (60%) of 15 subjects who received open-label risperidone following the double-blind placebo phase responded. Specifically, risperidone was effective for reducing interfering repetitive behavior, as well as aggression toward self, others, and property.

TABLE 42.2. CONTROLLED TRIALS OF CURRENT DRUG TREATMENTS FOR AUTISTIC DISORDER

Drug	Reference	Sample and Design	Results
Atypical antipsychotics			
Risperidone	McDougle et al. (30)	31 adults, autistic disorder (*n* = 17), pervasive developmental disorder not otherwise specified (PDD NOS) (*n* = 14), 12-week double-blind, placebo-controlled	Risperidone—8 of 14 (57%) responders Placebo—0 of 16 responders
SRIs			
Clomipramine vs. Desipramine	Gordon et al. (40)	7 children and adolescents with autistic disorder, 10-week double-blind crossover following 2-week single-blind placebo phase	Clomipramine > desipramine for obsessive-compulsive symptoms and autistic symptoms Clomipramine = desipramine for hyperactivity
Clomipramine vs. Placebo	Gordon et al. (41)	12 children and adolescents with autistic disorder, 10-week double-blind crossover following 2-week single-blind placebo phase	Clomipramine > placebo for obsessive-compulsive symptoms, autistic symptoms and hyperactivity
and			
Clomipramine vs. Desipramine		12 children and adolescents with autistic disorder, 10-week double-blind crossover following 2-week single-blind placebo phase	Clomipramine > desipramine for obsessive-compulsive symptoms and autistic symptoms Clomipramine = desipramine for hyperactivity
Fluvoxamine	McDougle et al. (48)	30 adults with autistic disorder, 12 week double-blind, placebo-controlled	Fluvoxamine—8 of 15 (53%) responders Placebo—0 of 15 responders

SRI, serotonin reuptake inhibitor.

Although risperidone was better than placebo for decreasing the overall maladaptive behaviors of autistic disorder, as measured by the Ritvo-Freeman Real Life Rating Scale overall score (31), this finding was largely accounted

for by significant changes in sensory motor behaviors (e.g., rocking, flapping), affectual reactions (e.g., temper outbursts), and to some extent sensory responses (e.g., spinning objects, sniffing self or objects). Significant differences between risperidone and placebo were not captured on subscales of the Ritvo-Freeman Scale that measure social relationships to people and language. For many subjects, however, clinicians, parents, and other members of the treatment teams had the impression that anxiety associated with social interactions was reduced, allowing for enhanced social function. It may be that the rating scales used to assess social relatedness in this study were not sensitive enough to detect changes in this complex aspect of behavior.

In general, risperidone was well tolerated. Thirteen (87%) of 15 subjects randomized to risperidone had at least one adverse effect, although this included only mild, transient sedation in five subjects, compared with five (31%) of 16 subjects given placebo (agitation in all five cases). Interestingly, the weight gain that has been observed with risperidone in the treatment of some children and adolescents with PDDs did not occur to the same degree in this study of adults.

Based on these results and other clinical, preclinical, and safety data, the RUPP Autism Network chose risperidone as the first drug to study in children and adolescents with autistic disorder (26). When completed, this investigation will be the largest controlled drug trial conducted to date in autistic disorder.

Olanzapine

Three case reports and an open-label case series have described positive responses to the atypical antipsychotic olanzapine in subjects with PDDs. In the case series, six of seven children, adolescents, and adults with autistic disorder and other PDDs (mean ± SD age, 20.9 ± 11.7 years; range 5 to 42 years) who completed the 12-week open-label trial were responders based on the CGI (32). Significant improvement in overall symptoms of autistic disorder, motor restlessness or hyperactivity, social relatedness, affectual reactions, sensory responses, language usage, self-injurious behavior, aggression, irritability or anger, anxiety, and depression was observed. Significant changes in repetitive behavior did not occur for the group. The mean ± SD dose of olanzapine was 7.8 ± 4.7 mg daily (range 5 to 20 mg per day). The drug was well tolerated, with the most significant adverse effects being increased appetite and weight gain in six subjects and sedation in three.

Quetiapine

Only one report of quetiapine in the treatment of autistic disorder has been published (33). Six males with autistic disorder, 6.2 to 15.3 years of age (mean ± SD age, 10.9

± 3.3 years), entered a 16-week open-label trial of quetiapine. The mean ± SD daily dose of quetiapine was 225 ± 108 mg (range 100 to 350 mg per day). Overall, there was no statistically significant improvement for the group as a whole on various behavioral rating scales. Two subjects completed the entire 16-week trial; both were considered responders based on CGI scores. However, only one of these two subjects continued to benefit from longer term treatment with the drug. The other four subjects dropped out due to lack of response and sedation (three subjects), and a possible seizure during the fourth week of treatment (one subject). Other significant side effects included behavioral activation, increased appetite, and weight gain (range 0.9 to 8.2 kg). The authors concluded that quetiapine was poorly tolerated and ineffective in most subjects, although the sample size was small. More studies of quetiapine in autistic disorder and related PDDs are needed before definitive conclusions about its effectiveness and safety can be made.

Summary

The atypical antipsychotics have potent effects on 5-HT and dopamine neuronal systems, both of which have been implicated in the pathophysiology of PDDs. Unlike the typical antipsychotic, haloperidol, which has been shown to be effective for reducing many of the maladaptive behaviors associated with PDDs, the atypical agents' $5\text{-}HT_{2A}/D_2$ ratio of receptor blockade appears to produce a lower risk of acute and chronic extrapyramidal side effects, as well as enhanced efficacy for the "negative" symptoms of PDD. Because most of the studies of this class of drugs have been short-term open-label trials in small samples, larger scale controlled investigations are needed to confirm these preliminary results. Due to the possibility that many children and adolescents who demonstrate short-term benefit from these drugs will remain on them indefinitely, the longer term safety and efficacy of atypical antipsychotics in this population also needs to be determined.

Serotonin Reuptake Inhibitors

In Kanner's 1943 landmark description of 11 autistic children, the repetitive natures of behavior, speech, and modes of social interaction were designated as core clinical elements of the syndrome (34). Verbal and motor rituals, obsessive questioning, a rigid adherence to routine, a preoccupation with details, and an anxiously obsessive desire for the maintenance of sameness and completeness were all noted. Religious and somatic obsessions, repetitive hand washing, and tics were identified in family members.

Results from a more recent study indicated that adults with autistic disorder and obsessive-compulsive disorder (OCD) can be distinguished on the basis of their current types of repetitive thoughts and behavior (35). Compared with the OCD group, the autistic subjects were significantly less likely to have aggressive, contamination, sexual, religious, symmetry, and somatic thoughts. Repetitive ordering, hoarding, telling or asking (trend), touching, tapping, or rubbing, and self-damaging or self-mutilating behaviors were reported significantly more frequently in the autistic subjects, whereas cleaning, checking, and counting behaviors were less common in the autistic group compared with the OCD subjects.

The clear efficacy of potent SRIs in OCD and their differential effects when directly compared to drugs that potently inhibit norepinephrine (NE) uptake support the hypothesized importance of 5-HT in the treatment of obsessive-compulsive symptoms (36). Consistent with these drug response data is the hypothesis that a dysregulation of 5-HT function might contribute to the pathophysiology of at least some individuals with OCD (37). Abnormalities in 5-HT function have also been identified in subjects with autistic disorder and other PDDs (38). Based on the efficacy of SRIs in the treatment of OCD, the high prevalence of interfering repetitive thoughts and behavior in subjects with PDD, and evidence indicating that a dysregulation in 5-HT neurotransmission may contribute to the pathophysiology of some individuals with autistic disorder, researchers have been studying the clinical response and side effect profile of SRIs in children, adolescents, and adults with PDDs.

Clomipramine

Clomipramine, a tricyclic antidepressant (TCA) and potent, but nonselective, inhibitor of 5-HT uptake, has been shown to be more efficacious than the relatively selective NE uptake inhibiting TCA desipramine in the treatment of children and adolescents with OCD (39). In the first controlled investigation of clomipramine in autistic disorder, the drug was found to be more efficacious than desipramine and placebo on standardized ratings of autistic disorder and anger, as well as ratings of repetitive and compulsive behaviors (40). Seven subjects with autistic disorder, ages 6 to 18 years (means age, 9.6 years), completed the 10-week double-blind crossover trial of clomipramine (mean dose, 129 daily) and desipramine (mean dose, 111 mg daily) following a 2-week single-blind placebo phase. In general, the side effects were relatively minor and did not differ between the two drugs. Mild sleep disturbance, dry mouth, and constipation were observed, and one patient developed a minor tremor on clomipramine. Two subjects taking desipramine developed uncharacteristic and severe irritability and temper outbursts. The parents of all seven subjects chose to have their children continue on clomipramine after completion of the study.

As a follow-up to this pilot study, a larger double-blind comparison of clomipramine, desipramine and placebo was conducted in children and adolescents with autistic disorder (41). Following a 2-week single-blind placebo phase, 12 subjects completed a 10-week double-blind crossover comparison of clomipramine and placebo, and 12 different subjects completed a similar comparison of clomipramine and

desipramine. The latter study included data from the seven subjects who participated in the original pilot study described above. Clomipramine (mean dose, 152 mg daily) was superior to both placebo and desipramine (mean dose, 127 mg daily) on ratings of autistic symptoms, including stereotypies, anger, and ritualized behaviors, with no difference between desipramine and placebo. Clomipramine was equal to desipramine and both drugs were superior to placebo for reducing motor hyperactivity. One child developed prolongation of the corrected QT interval (0.45 seconds) and another developed severe tachycardia (resting heart rate, 160 to 170 beats per minute) during clomipramine treatment. A third child experienced a grand mal seizure.

Subsequent open-label studies of clomipramine have been published with mixed results and increased recognition of adverse effects. Clomipramine treatment of five young adults (ages 13 to 33 years) with autistic disorder led to ratings of “much improved” on the CGI in four patients, with improvement seen in social relatedness, obsessive-compulsive symptoms, aggression, and self-injurious behavior (42). In another study, 11 consecutively referred children and adolescents with developmental disabilities and chronic stereotypies or self-injurious behavior were treated with clomipramine (43). Four of the subjects (ages 13 to 20 years) had been diagnosed with autistic disorder and of them, three had a significant reduction in stereotypic, self-injurious behavior with clomipramine at doses of 50 to 125 mg daily. Adverse effects included constipation, aggression, rash, and enuresis. In another open-label study, clomipramine 200 mg daily, was associated with decreased abnormal motor movements and compulsions in five autistic boys ages 6 to 12 years (44).

A large prospective open-label study of clomipramine (mean dose, 139 mg daily) treatment of 35 adults diagnosed with different subtypes of PDD was described (45). Of the 33 subjects who completed the 12-week study, 18 (55%) were judged responders on the CGI with improvement seen in aggression, self-injury, interfering repetitive thoughts and behavior, and social relatedness. Thirteen of the 33 subjects had significant adverse effects including seizures (in three patients, including two who had preexisting seizure disorders stabilized on anticonvulsants), weight gain, constipation, sedation, agitation, and anorgasmia.

A number of studies have suggested that younger children may tolerate clomipramine less well and show a decreased response compared to adolescents and adults with PDDs. In one report, eight children (ages 3.5 to 8.7 years) were given clomipramine (mean dose, 103 mg daily) for 5 weeks in a prospective open-label manner (46). Among the seven children who completed the trial, only one child was rated as moderately improved on a clinical global consensus measure. Adverse effects were frequent and included urinary retention requiring catheterization, constipation, drowsiness, and increased aggression and irritability. In a follow-up report to the study described above, in which five autistic children had an initial positive response to clomipramine (44), it was noted that the drug was eventually discontinued in all cases due to adverse effects or continued maladaptive behavior (47). Adverse effects included the serotonin syndrome, increased seizure frequency, and exacerbation of agitation and aggressiveness that required hospitalization.

Because of their better side effect profile compared with clomipramine, including their lower propensity to decrease the seizure threshold, selective SRIs (SSRIs) have been receiving increasing attention as a potential treatment for the interfering symptoms associated with autistic disorder and other PDDs.

Fluvoxamine

To date, only one double-blind, placebo-controlled study of an SSRI in subjects with autistic disorder has been published (48). Fluvoxamine (mean dose, 276.7 mg daily) or placebo was given to 30 autistic adults for 12 weeks. Eight of 15 subjects who received fluvoxamine vs. none who received placebo were categorized as “much improved” or “very much improved” on the CGI. Fluvoxamine was significantly more effective than placebo for reducing repetitive thoughts and behavior, maladaptive behavior, and aggression. In addition, fluvoxamine reduced inappropriate repetitive language usage. Adverse effects included nausea and sedation, which were transient and of minor severity.

In contrast to the encouraging results from this study of fluvoxamine in autistic adults, a 12-week double-blind, placebo-controlled study in children and adolescents with autistic disorder and other PDDs found the drug to be poorly tolerated with limited efficacy at best (McDougle and co-workers, unpublished data). Thirty-four patients (five female, 29 male; age range 5 to 18 years, mean age, 9.5 years), 12 of whom met criteria for autistic disorder, eight for Asperger’s syndrome, and 14 for PDD NOS, participated. Of the 16 subjects randomized to placebo, none demonstrated any significant change in target symptoms. Adverse events that occurred in the placebo-treated subjects included increased motor hyperactivity (*n* = 2), insomnia (*n* = 2), dizziness and/or vertigo (*n* = 1), agitation (*n* = 1), diarrhea (*n* = 1), decreased concentration (*n* = 1), and increased self-stimulation (*n* = 1). Eighteen of the subjects were randomized to fluvoxamine (range 25 to 250 mg daily; mean dose, 106.9 mg per day). The drug was begun at 25 mg every other day and increased by 25 mg every 3 to 7 days as tolerated. Only one of the fluvoxamine-treated children demonstrated a significant improvement with the drug. Fourteen of the children randomized to fluvoxamine demonstrated adverse effects [insomnia (*n* = 9), motor hyperactivity (*n* = 5), agitation (*n* = 5), aggression (*n* = 5), increased rituals (*n* = 2), anxiety (*n* = 3), anorexia (*n* = 3), increased appetite (*n* = 1), irritability (*n* = 1), decreased concentration (*n* = 1), and increased impulsivity (*n* = 1)].

The marked difference in efficacy and tolerability of fluvoxamine in children and adolescents with autistic disorder

and other PDDs in this study, compared with that of autistic adults, underscores the importance of developmental factors in the pharmacotherapy of these subjects. This differential drug response is consistent with the hypothesis that ongoing brain development has a significant impact on the subjects' ability to tolerate and respond to a drug, at least with respect to fluvoxamine and possibly other SSRIs. Developmental changes in brain 5-HT function may contribute to these widely varying clinical responses between subjects with autistic disorder and other PDDs of different age groups.

Fluoxetine

Several case reports have described fluoxetine treatment of autistic subjects although, to date, no controlled studies have appeared.

In a large case series, Cook and associates (49) found fluoxetine (10 to 80 mg daily), given in an open-label manner, effective in 15 of 23 subjects (ages 7 to 28 years) with autistic disorder as determined by the CGI (49). Intolerable side effects, including restlessness, hyperactivity, agitation, elated affect, decreased appetite, and insomnia, occurred in six of 23 subjects.

In a retrospective investigation, fluoxetine (20 to 80 mg daily) and paroxetine (20 to 40 mg daily) were found to be effective in approximately one-fourth of adults (mean age, 39 years) with "intellectual disability" and autistic traits (50). The sample included all intellectually disabled subjects who had been treated with an SSRI over a 5-year period within a health care service in Great Britain. The mean duration of treatment was 13 months. Target symptoms were perseverative behaviors, aggression, and self-injury. Six of 25 subjects treated with fluoxetine and three of 12 subjects given paroxetine were rated as "much improved" or "very much improved" on the CGI.

In another study, 37 children (ages 2.25 to 7.75 years) with autistic disorder were treated with fluoxetine in an open-label fashion at doses ranging from 0.2 to 1.4 mg per kg daily (51). Eleven of the children had an "excellent" clinical response and 11 others had a "good" response. Improvement was seen in behavioral, cognitive, affective, and social areas. Interestingly, language acquistion seemed to improve with fluoxetine treatment. Drug-induced hyperactivity, agitation, and aggression were frequent causes of discontinuation of fluoxetine.

Sertraline

To our knowledge, no controlled studies of sertraline in subjects with autistic disorder or other PDDs have been published, although a number of open-label reports have appeared. In a 28-day trial of sertraline (at doses of 25 to 150 mg daily) in nine adults with mental retardation (five of whom had autistic disorder), significant decreases in aggression and self-injurious behavior occurred in eight as rated on the CGI severity rating (52). In a case series of nine autistic children (ages 6 to 12 years) treated with sertraline (25 to 50 mg daily), eight showed significant improvement in anxiety, irritability, and "transition-induced behavioral deterioration" or "need for sameness" (53). In three of the responders, a return of symptoms occurred after 3 to 7 months. Two children demonstrated agitation when the dose was raised to 75 mg daily.

A large prospective open-label study of 42 adults with PDDs (including subjects with autistic disorder, Asperger's syndrome, and PDD NOS) found sertraline (mean dose, 122 mg per day) effective for improving aggression and interfering repetitive behavior, but not impaired social relatedness as assessed by various rating scales over the course of the 12-week study (54). As determined by a CGI global improvement item score of "much improved" or "very much improved," 15 of 22 subjects with autistic disorder, none of six with Asperger's syndrome, and nine of 14 with PDD NOS were judged responders. Those subjects with autistic disorder and PDD NOS showed significantly more benefit from sertraline than those with Asperger's syndrome; the authors hypothesized that this might have been because those diagnosed with Asperger's syndrome were less symptomatic at baseline. Three of the 42 subjects dropped out of the study due to intolerable agitation and anxiety.

Paroxetine

Only a few reports, none of them controlled, have appeared on the use of paroxetine in autistic disorder. Paroxetine 20 mg per day decreased self-injury in a 15-year-old boy with "high-functioning" autistic disorder (55). In another report, paroxetine resulted in a reduction of irritability, temper tantrums, and interfering preoccupations in a 7-year-old boy with autistic disorder (56). The optimal dose of paroxetine was 10 mg daily; an increase of paroxetine to 15 mg per day led to agitation and insomnia. As described earlier, a retrospective case analysis found paroxetine to be effective in approximately 25% of adults with PDD NOS (50).

In a 4-month open-label study of 15 adults with severe and profound mental retardation (seven with PDD), paroxetine at doses of 20 to 50 mg daily was effective for symptoms of aggression at 1 month, but not at 4-month follow-up (57). The investigators hypothesized that adaptive changes may have occurred in 5-HT receptor density, availability of 5-HT, or in 5-HT transporter sensitivity.

Citalopram

To date, there have been no published reports on the effects of citalopram, an SSRI that has been recently introduced in the United States, in patients with autistic disorder or other PDDs.

Summary

Recent work has determined that the types of repetitive phenomena associated with autistic disorder are different from those that characterize OCD. Nevertheless, these signs

and symptoms can interfere with the autistic individual's quality of life. Studies of SRIs, the mainstay of treatment for the obsessions and compulsions of OCD, have yielded mixed results in autistic disorder. To date, only three controlled studies of SRIs in autistic disorder have been published, as reviewed above (Table 42.2). All three of these studies found the SRI to be helpful for the interfering repetitive phenomena associated with autistic disorder, as well as for aspects of aggression, self-injury, and impaired social relatedness. On the other hand, results from an unpublished controlled study of fluvoxamine in children and adolescents with autistic disorder and other PDDs indicated that the drug was poorly tolerated and of limited efficacy. The results from that study are consistent with those from a number of open-label reports suggesting that SRIs may be less well tolerated and less effective in younger (prepubertal) autistic subjects compared with autistic adolescents and adults (postpubertal). Although this developmental difference in tolerability and response to SRIs may be a dose-related phenomena, other factors need to be considered. Recent data indicate that significant changes in measures of 5-HT function occur during puberty in autistic individuals. For example, McBride and co-workers (58) found that mean platelet 5-HT levels were significantly higher in prepubertal autistic children than prepubertal normal controls, but no significant difference was found between postpubertal male autistic subjects and postpubertal normal controls (58). Furthermore, Chugani and associates (59) reported results from a positron emission tomography brain imaging study showing that changes in brain 5-HT synthesis capacity that normally occur in developing humans are disrupted in autistic children (59). Thus, pre- and postpubertal autistic subjects may have significant differences in brain 5-HT function that influence their ability to tolerate and respond to SRIs. Pharmacogenetic differences among autistic individuals, which may affect SRI tolerability and responsivity, will also require more investigation (60).

Novel Therapeutic Strategies

Secretin

Secretin is a polypeptide hormone secreted primarily by the endocrine cells in the upper gastrointestinal (GI) tract that is involved in regulating pancreatic exocrine secretion. A synthetic form of secretin is Food and Drug Administration (FDA) approved for use in the diagnosis of particular GI diseases. In 1998, Horvath and co-workers (61) published a report that described marked improvement in language and social behavior in three children with autistic disorder who received secretin as part of a routine diagnostic workup for GI problems.

These encouraging yet preliminary results, coupled with enthusiastic media reports, led to widespread excitement and optimism among many family members of autistic individuals about a potential "cure" for the disorder. In response, researchers began conducting controlled studies of secretin in autistic children. A double-blind, controlled study of single-dose intravenous secretin (0.4 μg per kg) or placebo administration was conducted in 60 children, ages 3 to 14 years, of whom 40 had autistic disorder and 20 had PDD NOS (62). No significant differences were found between secretin and placebo on primary outcome measures that assessed changes in autistic behaviors and adaptive functioning at days 1 and 2 and weeks 1, 2, and 4 following the infusion. No significant difference was found in adverse effects between secretin and placebo. Two additional controlled studies have reported similar findings (63 ,64). Based on the results of these systematic investigations, secretin cannot be recommended as a treatment for the target symptoms associated with autistic disorder.

Glutamatergic Agents

The *N*-methyl-D-aspartate (NMDA) subtype of glutamate receptor is central to developmental processes including neuronal migration, differentiation, and plasticity (65). Disturbances in glutamatergic function, via reduced neurotropic actions of glutamate or excessive neurotoxic effects, could alter neurodevelopment substantially (66). During the past 5 to 10 years, significant advances have been made in the identification of potential pharmacotherapies affecting glutamatergic function for a number of neuropsychiatric disorders (67).

Hypotheses regarding glutamatergic dysfunction in autistic disorder have been proposed (68). In addition, preliminary results from studies of drugs that modulate glutamate neurotransmission in autistic disorder have been published. Lamotrigine is a drug that attenuates some forms of cortical glutamate release via inhibition of sodium channels, P- and N-type calcium channels, and potassium channels. In one study, eight of 13 children and adolescents with autistic disorder given lamotrigine for intractable epilepsy showed a decrease in "autistic symptoms" (69). Another report described improvement in self-injurious behavior, irritability, and disturbed sleep in an 18-year-old woman with profound mental retardation and a generalized seizure disorder who was given open-label lamotrigine (70). Interestingly, the subject showed improvement in measures of "fixed facial expression, lacks emotional responsivity," "resists any form of physical contact," and "inactive, never moves spontaneously." The authors suggested that these changes might represent a "prosocial" effect of the drug. In a double-blind, placebo-controlled study, 39 subjects with autistic disorder, ages 5 to 19 years old, were given placebo or the NMDA receptor antagonist amantadine (71). The design included a single-blind placebo lead-in, followed by a single daily dose of amantadine (2.5 mg per kg) or placebo for the next week, and then twice daily dosing for the subsequent 3 weeks. No significant difference was found between drug

and placebo on parent ratings, although clinician-rated measures of hyperactivity and inappropriate speech showed statistically significant improvement. A trend toward greater response in the amantadine group, based on CGI ratings, occurred. Amantadine was well tolerated.

Based on these preliminary results, and reports that the “negative” symptoms of schizophrenia can be improved with drugs active at the NMDA receptor (72), additional research with these and other agents affecting the glutamatergic systems appears warranted. The group II/III metabotropic-glutamate receptor (mGluR II/III) agonists (73) and the positive allosteric modulator of α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptors, CX516 (Ampakine) (74) may hold promise in this regard. Interestingly, one mechanism of action underlying the relative efficacy of atypical antipsychotics, such as risperidone, for autistic disorder (30) may be the suppression of glutamate release via 5-HT_{2A} antagonism (75).

Neuroimmune Modulation

Neuroimmunologic dysfunction has been implicated in the pathophysiology of autistic disorder (76) and other neuropsychiatric conditions (77). To date, no consistent immunologic abnormalities have been found in autistic disorder, although viral and autoimmune hypotheses, among others, have been posited (76). Neurovirologic disease and other insults to the immune system can lead to increased production of catabolites of tryptophan, including quinolinate and kynurenate, which can cause significant neurotoxicity via activity at the NMDA receptor complex (78). Thus, neuroimmune dysregulation in autistic disorder would not be inconsistent with altered glutamatergic function, as described above. Results from small open-label studies of intravenous immunoglobulin have suggested that this intervention may be helpful in only a limited minority of subjects, if at all (79). Controlled studies of agents that have direct effects on immune function, however, have not been conducted in autistic disorder. Such research on neuroimmune interactions may yield important data on pathophysiology, if not etiology, in a subset of autistic subjects.

CONCLUSION

Part of "42 - Current and Emerging Therapeutics of Autistic Disorder and Related Pervasive Developmental Disorders "

Significant progress in the neuropsychopharmacology of autistic disorder and related PDDs has been made since the fourth edition of this text (80). Future research in this area should include controlled studies of atypical antipsychotics in children and adolescents with autistic disorder and other PDDs, such as that being conducted by the NIMH-sponsored RUPP Autism Network (26). Longitudinal efficacy and safety data will need to be gathered on atypical antipsychotics in this population, as well. Larger double-blind, placebo-controlled trials of SRIs in pre- vs. postpubertal individuals with autistic disorder, as well as studies designed to determine the effects of these drugs on the target symptoms of subjects with different subtypes of PDD, including Asperger’s syndrome, are also needed. In these studies, the optimal dosage for age and developmental level and the duration of adequate treatment should be determined. In addition, genetic predictors of treatment response, such as 5-HT transporter protein genotype, should be sought (60). The scientific community needs to continue to respond to reports of putative “cures” for autistic disorder, such as those that surrounded secretin, by conducting controlled studies of such agents. Accepting this responsibility will contribute to ensuring the continued safety of autistic subjects and provide family members with data on which to base informed decisions regarding their child’s medical care. Finally, exploration of promising novel therapeutic strategies, such as those affecting glutamatergic and neuroimmune function, may provide new insights into the neurobiology and treatment of this devastating group of disorders.

ACKNOWLEDGMENTS

Part of "42 - Current and Emerging Therapeutics of Autistic Disorder and Related Pervasive Developmental Disorders "

This work was supported by an independent investigator award (Seaver Foundation Investigator) from the National Alliance for Research in Schizophrenia and Depression, the Theodore and Vada Stanley Research Foundation, a research unit on Pediatric Psychopharmacology Contract, no. N01MH70001 from the National Institute of Mental Health, and the State of Indiana Division of Mental Health.

Dr. McDougle has received research support from Pfizer, Eli Lilly, and Janssen Pharmaceutica, and has served on speakers’ bureaus and/or as a consultant for Pfizer, Eli Lilly, Janssen, and Solvay Pharmaceuticals.

REFERENCES

1. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington, DC: American Psychiatric Association, 1994.

2. Amir RE, Van den Veyver IB, Wan M, et al. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genet* 1999;23:185–188.

3. Insel TR. Oxytocin-a neuropeptide for affiliation: evidence from behavioral, receptor autoradiographic, and comparative studies. *Psychoneuroendocrinology* 1992;17:3–35.

4. Young LJ, Nilsen R, Waymire KG, et al. Increased affiliative response to vasopressin in mice expressing the V1a receptor from a monogamous vole. *Nature* 1999;400:766–768.

5. McDougle CJ. Psychopharmacology. In: Cohen DJ, Volkmar FR, eds. *Handbook of autism and developmental disorders second ed*. New York: Wiley, 1997:707–729.

6. Posey DJ, McDougle CJ. The pharmacotherapy of target symptoms associated with autistic disorder and other pervasive developmental disorders. *Harvard Rev Psychiatry* 2000;8:45–63.

7. Schain RJ, Freedman DX. Studies on 5-hydroxyindole metabolism in autistic and other mentally retarded children. *J Pediatr* 1961;58:315–320.

8. Campbell M, Adams P, Small AM, et al. Efficacy and safety of fenfluramine in autistic children. *J Am Acad Child Adolesc Psychiatry* 1988;27:434–439.

9. Campbell M, Anderson LT, Meier M, et al. A comparison of haloperidol and behavior therapy and their interaction in autistic children. *J Am Acad Child Psychiatry* 1978;17:640–655.

10. Cohen IL, Campbell M, Posner D, et al. Behavioral effects of haloperidol in young autistic children. *J Am Acad Child Psychiatry* 1980;19:665–677.

11. Anderson LT, Campbell M, Grega DM, et al. Haloperidol in the treatment of infantile autism: effects on learning and behavioral symptoms. *Am J Psychiatry* 1984;141:1195–1202.

12. Anderson LT, Campbell M, Adams P, et al. The effects of haloperidol on discrimination learning and behavioral symptoms in autistic children. *J Autism Dev Disord* 1989;19:227–239.

13. Willemsen-Swinkels SHN, Buitelaar JK, van Engeland H. The effects of chronic naltrexone treatment in young autistic children: a double-blind placebo-controlled crossover study. *Biol Psychiatry* 1996;39:1023–1031.

14. Willemsen-Swinkels SHN, Buitelaar JK, Nijhof GJ, et al. Failure of naltrexone hydrochloride to reduce self-injurious and autistic behavior in mentally retarded adults. *Arch Gen Psychiatry* 1995;52:766–773.

15. Campbell M, Fish B, David R, et al. Response to triiodothyronine and dextroamphetamine: a study of preschool schizophrenic children. *J Autism Child Schizophr* 1972;2:343–358.

16. Campbell M, Small AM, Collins PJ, et al. Levodopa and levoamphetamine: a crossover study in young schizophrenic children. *Curr Ther Res* 1976;19:70–86.

17. Quintana H, Birmaher B, Stedge D, et al. Use of methylphenidate in the treatment of children with autistic disorder. *J Autism Dev Disord* 1995;25:283–294.

18. Conners CK. *Conners Rating Scales manual*. North Tonowanda, NY: Multi-Health Systems, 1989.

19. Aman MG, Singh NN, Stewart AW, et al. The Aberrant Behavior Checklist: a behavior rating scale for the assessment of treatment effects. *Am J Ment Defic* 1985;89:485–491.

20. Jaselskis CA, Cook EH Jr, Fletcher KE, et al. Clonidine treatment of hyperactive and impulsive children with autistic disorder. *J Clin Psychopharmacol* 1992;12:322–327.

21. Fankhauser MP, Karumanchi VC, German ML, et al. A double-blind, placebo-controlled study of the efficacy of transdermal clonidine in autism. *J Clin Psychiatry* 1992;53:77–82.

22. Guy W. *ECDEU assessment manual for psychopharmacology* (NIMH publication no. 76-338). Washington, DC: U.S. DHEW, NIMH, 1976.

23. Arnsten AFT, Cai JX, Goldman-Rakic PS. The alpha-2 adrenergic agonist guanfacine improves memory in aged monkeys without sedative or hypotensive side effects: evidence for alpha-2 receptor subtypes. *J Neurosci* 1988;8:4287–4298.

24. Potenza MN, McDougle CJ. Potential of atypical antipsychotics in the treatment of nonpsychotic disorders. *CNS Drugs* 1998;9:213–232.

25. Kay SR, Opler LA, Lindemayer JP. Reliability and validity of the positive and negative symptom scale for schizophrenia. *Psychiatry Res* 1987;23:99–110.

26. McDougle CJ, Scahill L, McCracken JT, et al. Research Units on Pediatric Psychopharmacology (RUPP) Autism Network: background and rationale for an initial controlled study of risperidone. *Child Adolesc Psychiatry Clin North Am* 2000;9:201–224.

27. Kane J, Honigfeld G, Singer J, et al. Clozapine for the treatment-resistant schizophrenic: a double-blind comparison with chlorpromazine. *Arch Gen Psychiatry* 1988;45:789–796.

28. Zuddas A, Ledda MG, Fratta A, et al. Clinical effects of clozapine on autistic disorder. *Am J Psychiatry* 1996;153:738.

29. Chouinard G, Jones B, Remington G, et al. A Canadian multicenter placebo-controlled study of fixed doses of risperidone and haloperidol in the treatment of chronic schizophrenic patients. *J Clin Psychopharmacol* 1993;13:25–40.

30. McDougle CJ, Holmes JP, Carlson DC, et al. A double-blind placebo-controlled study of risperidone in adults with autistic disorder and other pervasive developmental disorders. *Arch Gen Psychiatry* 1998;55:633–641.

31. Freeman BJ, Ritvo ER, Yokota A, et al. A scale for rating symptoms of patients with the syndrome of autism in real life settings. *J Am Acad Child Psychiatry* 1986;25:130–136.

32. Potenza MN, Holmes JP, Kanes SJ, et al. Olanzapine treatment of children, adolescents, and adults with pervasive developmental disorders: an open-label pilot study. *J Clin Psychopharmacol* 1999;19:37–44.

33. Martin A, Koenig K, Scahill L, et al. Open-label quetiapine in the treatment of children and adolescents with autistic disorder. *J Child Adolesc Psychopharmacol* 1999;9:99–107.

34. Kanner L. Autistic disturbances of affective contact. *Nerv Child* 1943;2:217–250.

35. McDougle CJ, Kresch LE, Goodman WK, et al. A case-controlled study of repetitive thoughts and behavior in adults with autistic disorder and obsessive-compulsive disorder. *Am J Psychiatry* 1995;152:772–777.

36. Goodman WK, Price LH, Delgado PL, et al. Specificity of serotonin reuptake inhibitors in the treatment of obsessive-compulsive disorder: comparison of fluvoxamine and desipramine. *Arch Gen Psychiatry* 1990;47:577–585.

37. Insel TR, Mueller EA, Alterman I, et al. Obsessive-compulsive disorder and serotonin: is there a connection? *Biol Psychiatry* 1985;20:1174–1188.

38. Potenza MN, McDougle CJ. The role of serotonin in autism-spectrum disorders. *CNS Spectrums* 1997;2:25–42.

39. Leonard HL, Swedo SE, Rapoport JL, et al. Treatment of obsessive-compulsive disorder with clomipramine and desipramine in children and adolescents. *Arch Gen Psychiatry* 1989;46:1088–1092.

40. Gordon CT, Rapoport JL, Hamburger SD, et al. Differential response of seven subjects with autistic disorder to clomipramine and desipramine. *Am J Psychiatry* 1992;149:363–366.

41. Gordon CT, State RC, Nelson JE, et al. A double-blind comparison of clomipramine, desipramine, and placebo in the treatment of autistic disorder. *Arch Gen Psychiatry* 1993;50:441–447.

42. McDougle CJ, Price LH, Volkmar FR, et al. Clomipramine in autism: preliminary evidence of efficacy. *J Am Acad Child Adolesc Psychiatry* 1992;31:746–750.

43. Garber HJ, McGonigle JJ, Slomka GT, et al. Clomipramine treatment of stereotypic behaviors and self-injury in patients with developmental disabilities. *J Am Acad Child Adolesc Psychiatry* 1992;31:1157–1160.

44. Brasic JR, Barnett JY, Kaplan D, et al. Clomipramine ameliorates adventitious movements and compulsions in prepubertal boys with autistic disorder and severe mental retardation. *Neurology* 1994;44:1309–1312.

45. Brodkin ES, McDougle CJ, Naylor ST, et al. Clomipramine in adults with pervasive developmental disorders: a prospective open-label investigation. *J Child Adolesc Psychopharmacol* 1997;7:109–121.

46. Sanchez LE, Campbell M, Small AM, et al. A pilot study of clomipramine in young autistic children. *J Am Acad Child Adolesc Psychiatry* 1996;35:537–544.

47. Brasic JR, Barnett JY, Sheitman BB, et al. Behavioral effects of clomipramine on prepubertal boys with autistic disorder and severe mental retardation. *CNS Spectrums* 1998;3:39–46.

48. McDougle CJ, Naylor ST, Cohen DJ, et al. A double-blind, placebo-controlled study of fluvoxamine in adults with autistic disorder. *Arch Gen Psychiatry* 1996;53:1001–1008.

49. Cook EH, Rowlett R, Jaselskis C, et al. Fluoxetine treatment of children and adults with autistic disorder and mental retardation. *J Am Acad Child Adolesc Psychiatry* 1992;31:739–745.

50. Branford D, Bhaumik S, Naik B. Selective serotonin re-uptake inhibitors for the treatment of perseverative and maladaptive behaviours of people with intellectual disability. *J Intellect Disabil Res* 1998;42:301–306.

51. DeLong GR, Teague LA, Kamran MM. Effects of fluoxetine treatment in young children with idiopathic autism. *Dev Med Child Neurol* 1998;40:551–562.

52. Hellings JA, Kelley LA, Gabrielli WF, et al. Sertraline response in adults with mental retardation and autistic disorder. *J Clin Psychiatry* 1996;57:333–336.

53. Steingard RJ, Zimnitzky B, DeMaso DR, et al. Sertraline treatment of transition-associated anxiety and agitation in children with autistic disorder. *J Child Adolesc Psychopharmacol* 1997;7:9–15.

54. McDougle CJ, Brodkin ES, Naylor ST, et al. Sertraline in adults with pervasive developmental disorders: a prospective open-label investigation. *J Clin Psychopharmacol* 1998;18:62–66.

55. Snead RW, Boon F, Presberg J. Paroxetine for self-injurious behavior. *J Am Acad Child Adolesc Psychiatry* 1994;33:909–910.

56. Posey DJ, Litwiller M, Koburn A, et al. Paroxetine in autism. *J Am Acad Child Adolesc Psychiatry* 1999;38:111–112.

57. Davanzo PA, Belin TR, Widawski MH, et al. Paroxetine treatment of aggression and self-injury in persons with mental retardation. *Am J Ment Retard* 1998;102:427–437.

58. McBride PA, Anderson GM, Hertzig ME, et al. Effects of diagnosis, race, and puberty on platelet serotonin levels in autism and mental retardation. *J Am Acad Child Adolesc Psychiatry* 1998;37:767–776.

59. Chugani DC, Muzik O, Behen M, et al. Developmental changes in brain serotonin synthesis capacity in autistic and nonautistic children. *Ann Neurol* 1999;45:287–295.

60. Cook EH, Courchesne R, Lord C, et al. Evidence of linkage between the serotonin transporter and autistic disorder. *Mol Psychiatry* 1997;2:247–250.

61. Horvath K, Stefanatos G, Sokolski KN, et al. Improved social and language skills after secretin administration in patients with autistic spectrum disorders. *J Assoc Acad Minor Phys* 1998;9:9–15.

62. Sandler AD, Sutton KA, DeWeese J, et al. Lack of benefit of a single dose of synthetic human secretin in the treatment of autism and pervasive developmental disorder. *N Engl J Med* 1999;341:1801–1806.

63. Owley T, Steele E, Corsello C, et al. A double-blind, placebo-controlled trial of secretin for the treatment of autistic disorder. *Medscape Gen Med* 1999;1(10) [Available at: http://www.medscape.com/medscape/GeneralMedicine/journal/1999/v01.n10/mgm1006.owle/mgm1006.owle-01.html].

64. Chez MG, Buchanan CP, Bagan BT, et al. Secretin and autism: a two-part clinical investigation. *J Autism Dev Disord* 2000;30(2):87–94.

65. Coyle JT. The glutamatergic dysfunction hypothesis for schizophrenia. *Harvard Rev Psychiatry* 1996;3:241–253.

66. Krystal JH, Belger A, D'Souza DC, et al. Therapeutic implications of the hyperglutamatergic effects of NMDA antagonists. *Neuropsychopharmacology* 1999;22:S143–S157.

67. Krystal JH, D'Souza DC, Petrakis IL, et al. NMDA agonists and antagonists as probes of glutamatergic dysfunction and pharmacotherapies in neuropsychiatric disorders. *Harvard Rev Psychiatry* 1999;7:125–143.

68. Carlsson ML. Hypothesis: Is infantile autism a hypoglutamatergic disorder? Relevance of glutamate-serotonin interactions for pharmacotherapy. *J Neural Transm* 1998;105:525–535.

69. Uvebrant P, Bauziene R. Intractable epilepsy in children: the efficacy of lamotrigine treatment, including non-seizure-related benefits. *Neuropediatrics* 1994;25:284–289.

70. Davanzo PA, King BH. Open trial of lamotrigine in the treatment of self-injurious behavior in an adolescent with profound mental retardation. *J Child Adolesc Psychopharmacol* 1996; 6:273–279.

71. King BH, Wright DM, Handen BL, et al. Double-blind, placebo-controlled study of amantadine hydrochloride in the treatment of children with autistic disorder. *J Am Acad Adolesc Psychiatry* 2001;40:658–665.

72. Goff DC, Tsai G, Levitt J, et al. A placebo-controlled trial of D-cycloserine added to conventional neuroleptics in patients with schizophrenia. *Arch Gen Psychiatry* 1999;56:21–27.

73. Moghaddam B, Adams BW. Reversal of phencyclidine effects by a group II metabotropic glutamate receptor agonist in rats. *Science* 1998;281:1349–1352.

74. Goff DC, Leahy L, Berman I, et al. A placebo-controlled pilot study of the ampakine, CX516, added to clozapine in schizophrenia. *Am J Psychiatry* (*in press*).

75. Aghajanian GK, Marek GJ. Serotonin-glutamate interactions: a new target for antipsychotic drugs. *Neuropsychopharmacology* 1999;21:S122–S133.

76. van Gent T. Autism and the immune system. *J Child Psychol Psychiat* 1997;38:337–349.

77. Kronfol Z, Remick DG. Cytokines and the brain: implications for clinical psychiatry. *Am J Psychiatry* 2000;157:683–694.

78. Heyes MP, Saito H, Crowley JS, et al. Quinolinic acid and kynurenine pathway metabolism in inflammatory and non-inflammatory neurologic disease. *Brain* 1992;115:1249–1273.

79. DelGiudice-Asch G, Simon L, Schmeidler J, et al. Brief report: a pilot open clinical trial of intravenous immunoglobulin in childhood autism. *J Autism Dev Disord* 1999;29:157–160.

80. Lotspeich LJ. Autism and pervasive developmental disorders. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:1653–1662.

43

Pathophysiology of Attention-Deficit/Hyperactivity Disorder

Stephen V. Farone

Joseph Biederman

Stephen V. Farone: Pediatric Psychopharmacology Unit, Child Psychiatry Service, Massachusetts General Hospital; Harvard Medical School; Massachusetts Mental Health Center; Harvard Institute of Psychiatric Epidemiology and Genetics, Boston, Massachusetts.

Joseph Biederman: Pediatric Psychopharmacology Unit, Child Psychiatry Service, Massachusetts General Hospital; Harvard Medical School, Boston, Massachusetts.

Attention-deficit/hyperactivity disorder (ADHD) is a childhood-onset, clinically heterogeneous disorder of inattention, hyperactivity, and impulsivity. Its impact on society is enormous in terms of its financial cost, stress to families, adverse academic and vocational outcomes, and negative effects on self-esteem (1). Children with ADHD are easily recognized in clinics, in schools, and in the home. Their inattention leads to daydreaming, distractibility, and difficulties in sustaining effort on a single task for a prolonged period. Their impulsivity makes them accident prone, creates problems with peers, and disrupts classrooms. Their hyperactivity, often manifest as fidgeting and excessive talking, is poorly tolerated in schools and is frustrating to parents, who can easily lose them in crowds and cannot get them to sleep at a reasonable hour. In their teenage years, symptoms of hyperactivity and impulsivity diminish, but in most cases the symptoms and impairments of ADHD persist. The teen with ADHD is at high risk of low self-esteem, poor peer relationships, conflict with parents, delinquency, smoking, and substance abuse (1).

The validity of diagnosing ADHD in adults has been a source of much controversy (2). Some investigators argue that most cases of ADHD remit by adulthood (3), a view that questions the validity of the diagnosis in adulthood. Others argue that the diagnosis of ADHD in adults is both reliable and valid (2). These investigators point to longitudinal studies of children with ADHD, studies of clinically referred adults, family-genetic studies, and psychopharmacologic studies. Longitudinal studies have found that as many as two thirds of children with ADHD have impairing ADHD symptoms as adults. Studies of clinically referred adults with retrospectively defined childhood-onset ADHD show them to have a pattern of psychosocial disability, psychiatric comorbidity, neuropsychological dysfunction, familial illness, and school failure that resemble the well known features of children with ADHD.

Throughout the life cycle, a key clinical feature observed in patients with ADHD is comorbidity with conduct, depressive, bipolar, and anxiety disorders (4 ,5). Although spurious comorbidity can result from referral and screening artifacts (5), these artifacts cannot explain the high levels of psychiatric comorbidity observed for ADHD (4). Notably, epidemiologic investigators find comorbidity in unselected general population samples (6 ,7), a finding that cannot be caused by the biases that inhere in clinical samples. Moreover, as we discuss later, family studies of comorbidity dispute the notion that artifacts cause comorbidity; instead, they assign a causal role to etiologic relationships among disorders.

NEUROPSYCHOPHARMACOLOGY

Part of "43 - Pathophysiology of Attention-Deficit/Hyperactivity Disorder "

Pharmacotherapy

Any pathophysiologic theory about ADHD must address the large pharmacotherapy literature about the disorder. The mainline treatments for ADHD are the stimulant medications methylphenidate, pemoline, and dextroamphetamine. These compounds are safe and effective for treating ADHD in children, adolescents, and adults (8 ,9). In addition, to improving ADHD's core symptoms of inattentiveness, hyperactivity, and impulsivity, stimulants also improve associated behaviors, including on-task behavior, academic performance, and social functioning in the home and at school. In adults, occupational and marital dysfunction tend to improve with stimulant treatment. There is little evidence of a differential response to methylphenidate, pemoline, and dextroamphetamine. The average response rate for each is 70%.

Stimulants enhance social skills at home and in school.

They also improve maternal-child and sibling interactions. Children with ADHD who are treated with stimulants have increased abilities to perceive peer communications and situational cues and to modulate the intensity of their behavior. They also show improved communication, greater responsiveness, and fewer negative interactions. Neuropsychological studies show that stimulants improve vigilance, cognitive impulsivity, reaction time, short-term memory, and learning of verbal and nonverbal material in children with ADHD.

Although stimulants are the mainstay of anti-ADHD pharmacotherapy, tricyclic antidepressants (TCAs) also are effective anti-ADHD agents. TCAs include secondary and tertiary amines with a wide range of receptor actions, efficacy, and side effects. Secondary amines are more selective (noradrenergic) with fewer side effects. Most studies of TCAs have found either a moderate or robust response rate of ADHD symptoms (8 ,9 and 10). These studies show anti-ADHD efficacy for imipramine, desipramine, amitriptyline, nortriptyline, and clomipramine. Both short- and long-term studies show that TCAs produce moderate to strong effects on ADHD symptoms. In contrast, neurocognitive symptoms are do not respond well to TCA treatment. Because of rare reports of sudden death among TCA-treated children, these drugs are not a first-line treatment for ADHD and are only used after carefully weighing the risks and benefits of treating or not treating a child who does not respond to other agents.

Other noradrenergic agents help to control ADHD symptoms. Bupropion hydrochloride, which has both dopaminergic and noradrenergic effects, is effective for ADHD in children (11 ,12)as well as in adults (13). Although they are rarely used because of their potential for hypertensive crisis, several studies suggested that monoamine oxidase inhibitors may be effective in juvenile and adult ADHD (14). The experimental noradrenergic compound tomoxetine showed efficacy in a controlled study of adults with ADHD (15) and in an open study of children with ADHD (16).

In contrast to the beneficial effects of stimulants and TCAs, there is only weak evidence that either α_2-noradrenergic agonists or serotonin reuptake inhibitors effectively combat ADHD (17). A controlled clinical trial showed that transdermal nicotine improved ADHD symptoms and neuropsychological functioning in adults with ADHD (18). Consistent with this finding, a controlled study found the experimental compound ABT-418 to treat adult ADHD effectively (19). ABT-418 is a potent and selective agonist for $\alpha_4\beta_2$-subtype central nervous system neuronal nicotinic receptors.

Catecholamine Hypothesis

As the foregoing review shows, effective medications for ADHD act in noradrenergic and dopaminergic systems. Stimulants block the reuptake of dopamine and norepinephrine into the presynaptic neuron and increase the release of these monoamines into the extraneuronal space (20). Solanto suggested that stimulants may also activate presynaptic inhibitory autoreceptors and may lead to reduced dopaminergic and noradrenergic activity (21). The maximal therapeutic effects of stimulants occur during the absorption phase of the kinetic curve, within 2 hours after ingestion. The absorption phase parallels the acute release of neurotransmitters into synaptic clefts, a finding providing support for the hypothesis that alteration of monoaminergic transmission in critical brain regions may be the basis for stimulant action in ADHD (22). A plausible model for the effects of stimulants in ADHD is that, through dopaminergic or noradrenergic pathways, these drugs increase the inhibitory influences of frontal cortical activity on subcortical structures (22).

Human studies of the catecholamine hypothesis of ADHD that focused on catecholamine metabolites and enzymes in serum and cerebrospinal fluid produced conflicting results (23 ,24). Perhaps the best summary of this literature is that aberrations in no single neurotransmitter system can account for the available data. Of course, because studies of neurotransmitter systems rely on peripheral measures, which may not reflect brain concentrations, we cannot expect such studies to be completely informative. Nevertheless, although such studies do not provide a clear profile of neurotransmitter dysfunction in ADHD, on balance, they are consistent with the idea that catecholaminergic dysregulation plays a role in the origin of at least some cases of ADHD.

The catecholamine hypothesis of ADHD finds further support from animal studies. One approach has been the use of 6-hydroxydopamine to create lesions in dopamine pathways in developing rats. Because these lesions created hyperactivity, they were thought to provide an animal model of ADHD (25). Disruption of catecholaminergic transmission with chronic low-dose *N*-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), a neurotoxin, creates an animal model of ADHD in monkeys. In this latter work, MPTP administration to monkeys caused cognitive impairments on tasks thought to require efficient frontal-striatal neural networks. These cognitive impairments mirrored those seen in monkeys with frontal lesions (26 ,27). Like children with ADHD, MPTP-treated monkeys show attentional deficits and task impersistence. Methylphenidate and the dopamine D2 receptor agonist LY-171555 reversed the behavioral deficits but not the cognitive dysfunction (28 ,29).

Several investigators used the spontaneously hypertensive rat (SHR) as an animal model of ADHD because of the animal's locomotor hyperactivity and impaired discriminative performance. Studies using the SHR have implicated dopaminergic and noradrenergic systems. For example, the dopamine D2 receptor agonist, quinpirole, caused significantly greater inhibition of dopamine release from caudate-putamen

but not from nucleus accumbens or prefrontal cortex slices in SHR compared with control mice (30). In another study, dopamine release secondary to electrical stimulation was significantly lower in caudate-putamen and prefrontal cortex slices of SHR compared with control mice. These findings were attributed to increased autoreceptor-mediated inhibition of dopamine release in caudate-putamen slices but not in the prefrontal cortex. Another study showed that the altered presynaptic regulation of dopamine in SHR led to the down-regulation of the dopamine system (31). The authors hypothesized that this may have occurred early in development as a compensatory response to abnormally high dopamine concentrations.

Other SHR studies implicated an interaction between the noradrenergic and dopaminergic system in the nucleus accumbens, but they ruled out the idea that a dysfunctional locus ceruleus and A2 nucleus impairs dopaminergic transmission in the nucleus accumbens through α_2-adrenoceptor–mediated inhibition of dopamine release (32). Papa et al. used molecular imaging techniques to assess the neural substrates of ADHD-like behaviors in the SHR rat (33). Their data showed the corticostriatopallidal system to mediate these behaviors. King et al. showed that exposure to excess androgen levels early in development led to decreased catecholamine innervation in frontal cortex and enhanced expression of ADHD-like behaviors (34). Carey et al. used quantitative receptor autoradiography and computer-assisted image analysis to show a higher density of low-affinity D1 and D5 dopamine receptors in the caudate-putamen, the nucleus accumbens, and the olfactory tubercle of SHR (35). Stimulant treatment normalized these receptors by decreasing the number of binding sites and increasing affinity to the control level.

In contrast to the large body of evidence implicating dopaminergic and noradrenergic systems in ADHD, evidence implicating serotonergic systems is mixed. Although the tertiary amines (imipramine and amitriptyline) are more selective for the serotonin transporter than the norepinephrine transporter (36), the secondary amines (desipramine, nortriptyline, and protriptyline) are more selective for the norepinephrine transporter (36). Moreover, measures of serotonin metabolism appear minimally related to the clinical efficacy of the stimulants (22), a finding consistent with the lack of efficacy of serotonergic drugs for treating ADHD. This suggests that the anti-ADHD efficacy of the TCAs stems from their actions on catecholamine reuptake, particularly that of norepinephrine.

Despite these equivocal findings, work by Gainetdinov et al. suggests that we cannot rule out a role for serotonergic systems in the pathophysiology of ADHD (37). These authors studied knockout mice lacking the gene encoding the dopamine transporter (DAT). These mice have elevated dopaminergic tone, are hyperactive, and show decreased locomotion in response to stimulants. Gainetdinov et al. showed that the effects of stimulants were mediated by serotonergic neurotransmission (37).

The anti-ADHD efficacy of nicotine and ABT-418 suggests that nicotinic dysregulation may also play a role in the pathophysiology of ADHD. Patients with ADHD are more likely to smoke and have an earlier age of onset of smoking than persons who do not have ADHD (38 ,39 and 40). In addition, maternal smoking during pregnancy appears to increase the risk of ADHD in the children (41), and *in utero* exposure to nicotine in animals confers a heightened risk of an ADHD-like syndrome in the newborn (42 ,43). That nicotine dysregulation could play an important role in the pathophysiology of ADHD is not surprising considering that nicotinic activation enhances dopaminergic neurotransmission (44 ,45).

BRAIN ABNORMALITIES

Part of "43 - Pathophysiology of Attention-Deficit/Hyperactivity Disorder "

Satterfield and Dawson were among the first to propose that ADHD symptoms were caused by frontolimbic dysfunction (46). These investigators suggested that weak frontal cortical inhibitory control over limbic functions could lead to ADHD. A review of the neurologic literature showing similarities in disinhibited behavior between adult patients with frontal lobe damage and children with ADHD provided further evidence that the frontal lobes could be involved in the pathophysiology of the disorder (47). Two sources of data have tested the frontolimbic hypothesis of ADHD: neuropsychological studies and neuroimaging studies.

Neuropsychological Studies

Neuropsychological tests indirectly assess brain functioning by assessing features of human perception, cognition, or behavior that have been clinically or experimentally linked to specific brain functions (48). Although limited in their ability to localize brain dysfunction, these tests have several advantages. Many of these tests have been standardized on large populations, thus making it straightforward to define deviant performance. Because of the extensive use of these tests in brain-damaged populations, performance on many of these tests can lead to hypotheses, albeit weak, about the locus of brain dysfunction. Being noninvasive and inexpensive, neuropsychological tests are frequently used to generate hypotheses about brain dysfunction.

Given that inattention is a one of the defining clinical features of ADHD, many neuropsychological studies of the disorder have assessed the attention of children with ADHD. The most commonly used measure of attention is the *continuous performance test*, which requires subjects to sustain their attention to subtle sensory signals, to avoid being distracted by irrelevant stimuli, and to maintain alertness for the duration of the session. Most of these studies

find children with ADHD to be impaired on this measure (1).

Children with ADHD also perform poorly on tasks requiring inhibition of motor responses, organization of cognitive information, planning, complex problem solving, and the learning and recall of verbal material (49). Examples of tests that measure these functions are the Stroop Test, the Wisconsin Card Sorting Test, the Rey-Osterrieth Test, the Freedom from Distractibility factor from Wechsler's Tests of Intelligence, and the California Verbal Learning Test.

Some studies suggest that the impairments found in children with ADHD cannot be accounted for by psychiatric comorbidity (50). Moreover, having a family history of ADHD may predict a greater degree of neuropsychological impairment. This latter finding suggests that familial ADHD and neuropsychological impairment identify a more biologically based type of ADHD. In contrast, nonfamilial cases of ADHD with lesser neuropsychological impairments may have other etiologic factors. Children with ADHD do not appear to be impaired on simple motor speed, verbal fluency, or visual spatial accuracy, findings that suggest that observed neuropsychological impairments are caused by specific, not generalized, deficits (51).

Notably, neuropsychological studies have consistently found adults with ADHD to be impaired on measures of vigilance using the continuous performance test (52 ,53). These studies have also shown adults with ADHD to be impaired in other functions known to affect children with ADHD. These include the following: perceptual-motor speed as assessed by the digit symbol/coding tests (54 ,55); working memory as assessed by digit span tests (53 ,56); verbal learning, especially semantic clustering (52 ,56); and response inhibition as assessed by the Stroop Color-Word Test (57 ,58). Because neuropsychological tests are free of the potential biases of self-reported symptoms, the finding that the neurocognitive profiles of adults with ADHD are similar to those of children with ADHD suggests that the diagnosis of ADHD is valid as applied in adulthood.

Our description of neuropsychological dysfunction in ADHD describes trends that have emerged in the research literature, not findings that have been consistently replicated. Although there are inconsistencies among studies, it is notable that the pattern of deficits that has emerged is similar to what has been found among adults with frontal lobe damage. Thus, the neuropsychological data tend to support the hypothesis that the frontal cortex or regions projecting to the frontal cortex are dysfunctional in at least some children with ADHD.

Because neuropsychological tests provide indirect measures of brain function, we must be cautious in using them to make inferences about the locus of brain impairment in ADHD. Yet because many of these tests have been standardized on normative populations and administered extensively to brain-damaged populations, observed deficits tests can stimulate hypotheses about the role of specific brain regions in the pathophysiology of ADHD.

With this considerations in mind, we view the pattern of neuropsychological impairment in children with ADHD as consistent with Satterfield and Dawson's (46) idea that symptoms of ADHD derive from abnormalities of prefrontal cortex or its neural connections to subcortical structures. This inference derives from the clinical and behavioral features that have been linked to regions of the prefrontal cortex (59). Notably, orbital frontal lesions predict social disinhibition and impulsivity, and dorsolateral lesions affect organizational abilities, planning, working memory, and attention. Studies of children with ADHD find impairment in all these neuropsychological domains. Thus, the neuropsychological test data—along with the clinical features of the disorder—implicate both orbitofrontal and dorsolateral prefrontal dysfunction in ADHD. In contrast, the mesial prefrontal region, where lesions predict dysfluency and the slowing of spontaneous behavior, is not implicated in ADHD.

Given the complexity of prefrontal circuitry (60), along with the limitations of neuropsychological inference, we cannot endorse a simple lesion model of ADHD. The "prefrontal" abnormalities in ADHD may result from abnormalities of prefrontal cortex, but they may also reflect the dysfunction of brain areas with projections to prefrontal cortex. Given the known role of subcortical networks as modulators of prefrontal functioning, the term *frontosubcortical* seems appropriate for ADHD. This term denotes a behavioral or cognitive dysfunction that looks "frontal" but may be influenced by subcortical projections.

The neuropsychological findings in ADHD provide a fertile resource for speculations about the role of subcortical structures. For example, the cingulate cortex influences motivational aspects of attention and in response selection and inhibition. The brainstem reticular activating system regulates attentional tone and reticular thalamic nuclei filter interference. Working memory deficits implicate a distributed network including anterior hippocampus, ventral anterior and dorsolateral thalamus, anterior cingulate, parietal cortex, and dorsolateral prefrontal cortex. Moreover, the attentional problems of children with ADHD may implicate a wider distribution of neural networks. A system mainly involving right prefrontal and parietal cortex is activated during sustained and directed attention across sensory modalities. The inferior parietal lobule and superior temporal sulcus are polymodal sensory convergence areas that provide a representation of extrapersonal space and play an important role in focusing on and selecting a target stimulus.

Neuroimaging Studies

Fortunately, hypotheses based on neuropsychological inference can be tested with neuroimaging paradigms. Because neuroimaging studies provide direct assessments of brain

structure and function, they are ideal for testing hypotheses about the locus of brain dysfunction. Table 43.1 reviews 18 structural neuroimaging studies of children, adolescents, and adults with ADHD that used computed tomography or magnetic resonance imaging. Among these studies, the most consistent findings implicated frontal cortex, usually limited to the right side, cerebellum, globus pallidus, caudate, and corpus callosum. Several other regions were less consistently implicated. Consistent with these findings, the I/LnJ mouse strain shows total callosal agenesis along with behavioral features that resemble ADHD (61). These mice show learning impairments, impulsiveness, and hyperactivity. Metabolic mapping studies suggest that their behavioral deficits are associated with lower 2-deoxyglucose uptake in the left striatum and the frontal and parietal cortex (61).

Study	Diagnosis	Method	Findings
Shaywitz et al. (199)	ADD	CT	No abnormalities found
Nasrallah et al. (200)	HYP	CT	Sulcal widening, cerebellar atrophy
Lou et al. (201)	ADD	CT	Slight frontal cortex atrophy
Hynd et al. (202)	ADD/H	MRI	Smaller frontal cortex Loss of normal asymmetry in frontal cortex
Hynd et al. (203)	ADHD	MRI	Smaller corpus callosum
Aylward et al. (204)	ADHD	MRI	Smaller left globus pallidus
Singer et al. (205)	ADHD+TS	MRI	Smaller left globus pallidus
Baumgardner (206)	ADHD	MRI	Small corpus callosum
Semrud-Clikeman et al. (207)	ADHD	MRI	Small corpus callosum
Castellanos et al. (208)	ADHD	MRI	Smaller right prefrontal cortex, right caudate, and globus pallidus
Mostofsky et al. (209)	ADHD	MRI	Smaller inferior posterior vermis of cerebellum
Nopoulos et al. (70)	ADHD	MRI	Neural migration anomalies and excess cerebrospinal fluid in the posterior fossa but no differences in cavum septi pellucidi
Overmeyer et al. (210)	ADHD	MRI	No corpus callosum abnormalities
Mataro et al. (211)	ADHD	MRI	Larger right caudate nucleus
Kayl et al. (212)	ADHD[a]	MRI	Increased severity of attention problems was associated with small total callosal areas
Berquin et al. (213)	ADHD	MRI	Smaller inferior posterior vermis of cerebellum
Casey et al. (214)	ADHD	MRI	Poor response inhibition associated with right sided abnormalities prefrontal cortex, caudate, and globus pallidus, but not putamen
Filipek et al. (215)	ADHD	MRI	Smaller left caudate, right frontal cortex, and bilateral peribasal ganglia and parietal-occipital regions

ADD, DSM-III attention-deficit disorder; ADD/H, DSM-III ADD with hyperactivity; ADHD, DSM-III-R attention-deficit hyperactivity disorder; CT, computed tomography; HYP, DSM-II hyperkinesis; MRI, magnetic resonance imaging; TS, Tourette syndrome.
[a]In this study, ADHD was secondary to neurofibromatosis.

TABLE 43.1. STRUCTURAL NEUROIMAGING STUDIES OF ADHD

Table 43.2 reviews 14 functional neuroimaging studies of ADHD using regional cerebral blood flow, positron emission tomography, single photon emission tomography, functional magnetic resonance imaging, or electroencephalography. The most consistent findings were hypoactivity of frontal cortex and subcortical structures, usually on the right side. Because Ernst et al. found significant brain dysfunction for girls, but not boys, with ADHD (62), and Baving et al. found gender differences in lateralization (63), future studies will need to assess gender differences and to determine how they may be related to the male predominance of the disorder.

Study	Diagnosis	Method	Findings
Lou et al. (201)	ADD	rCBF	Hypoperfusion in frontal cortex and caudate, hyperperfusion in occipital cortex
Lou et al. (216)	ADHD	rCBF	Hypoperfusion in right striatal region, hyperperfusion in occipital cortex, left sensorimotor, and primary auditory regions
Lou et al. (217)	ADHD	rCBF	Hypoperfusion in striatal and posterior periventricular regions; hyperperfusion in occipital cortex, left sensorimoter, and primary auditory regions
Zametkin et al. (66)	ADHD	PET	Lower glucose metabolism in premotor and superior prefrontal cortex, right thalamus, right caudate, right hippocampus, and right cingulate
Ernst et al. (62)	ADHD	PET	ADHD girls (but not boys) show lower glucose metabolism in right prefrontal cortex, right temporal cortex, right and left posterior putamen, and middle cingulate
Amen et al. (218)	ADHD	SPECT	Decreased perfusion in prefrontal cortex
Rubia et al. (219)	ADHD	fMRI	Lower activation in right mesial prefrontal cortex, right inferior prefrontal cortex, and left caudate
Baving et al. (63)	ADHD	EEG	Boys show a less right-lateralized frontal activation pattern; girls show a more right-lateralized frontal activation pattern than healthy control girls
Schweitzer et al. (220)	ADHD	rCBF	Task-related changes in rCBF in non-ADHD men without ADHD were prominent in frontal and temporal regions; changes in ADHD men were more widespread, suggesting the use of compensatory mental and neural strategies
Silberstein et al. (221)	ADHD	EEG	Increased speed of prefrontal processing in non-ADHD children, ADHD following priming stimulus, and a deficit in such processes in ADHD children
Vaidya et al. (222)	ADHD	fMRI	ADHD is characterized by atypical frontal-striatal function, and methylphenidate affects striatal activation differently in ADHD than in healthy children
Ernst et al. (223)	ADHD	PET	More accumulation of [^{18}F]DOPA in the right midbrain correlated with symptom severity
Bush et al. (65)	ADHD	fMRI	ADHD adults show weak activation of anterior cingulate cognitive division during counting Stroop task
Dougherty et al. (67)	ADHD	SPECT	Dopamine transporter density in striatum greater in ADHD adults

ADD, DSM-III attention-deficit disorder; ADHD, DSM-III-R attention-deficit hyperactivity disorder; EEG, Electroencephalogram; fMRI, functional magnetic resonance imaging; PET, position emission tomography, rCBF, regional cerebral blood flow; SPECT, photon emission computed tomography.

TABLE 43.2. FUNCTIONAL NEUROIMAGING STUDIES OF ADHD

Ernst et al. noted that findings of frontal hypoactivity are stronger in adult ADHD compared with adolescent ADHD (64). They offered two explanations for this finding. First, the adolescent samples studied may have been more heterogeneous than the adult samples. Although all the adults had persistent ADHD, some of the adolescent cases may have remitted by adulthood. Thus, frontal dopaminergic hypoactivity may be associated with persistent ADHD only. Alternatively, Ernst et al. speculated that, because of brain maturation, the locus of ADHD's dopamine abnormality may shift from the midbrain in childhood to the prefrontal cortex in adults.

Anterior cingulate cortex, lying on the medial surface of the frontal lobe, has strong connections to dorsolateral prefrontal cortex. Bush et al. used a Stroop task to compare anterior cingulate cortex activation in adults with ADHD and those who did not have ADHD (65). In contrast to controls, the adults with ADHD failed to activate the anterior cingulate cortex. Notably, in the prior study by Zametkin et al. (66), cingulate cortex was one of only four (of 60) regions evaluated that still showed regional hypoactivity after global normalization.

The neurochemical basis of brain dysfunction in ADHD was studied by Dougherty et al. (67). They measured DAT density by single photon emission computed tomography with the radiopharmaceutical iodine 123–labeled altropane. Their findings were consistent with the catecholamine hypothesis

of ADHD in showing the DAT to be elevated by about 70% in adults with ADHD.

The functional studies are consistent with the structural studies in implicating frontosubcortical system in the pathophysiology of ADHD. Taken together, the brain imaging studies fit well with the idea that dysfunction in frontosubcortical pathways occurs in ADHD. They are also consistent with the report of a father and son, both having methylphenidate-responsive ADHD secondary to frontal lobe epilepsy (68). Notably, the frontosubcortical systems that control attention and motor behavior are rich in catecholamines, which have been implicated in ADHD by the mechanism of action of stimulants.

In a novel approach to assessing brain regions implicated in ADHD, Herskovits et al. used magnetic resonance imaging to assess the spatial distribution of lesions in children who developed ADHD after closed-head injuries (69). Compared with head-injured children who did not develop ADHD, the children with ADHD had more lesions in the right putamen and a trend for more lesions in the right caudate nucleus and right globus pallidus.

Very little is known about when ADHD-related brain abnormalities emerge. To address this issue, Nopoulos et al. assayed four brain abnormalities believed to occur before birth: neural migration anomalies, corpus callosum agenesis or partial agenesis, enlarged cavum septi pellucidi, and malformations of the posterior fossa (70). Neural migration anomalies and malformations of the posterior fossa were more common among patients with ADHD compared with control subjects. Both these abnormalities were rare. However, given that several other studies showed partial agenesis of the corpus callosum or anomalies of the cerebellar vermis (also formed before birth), it seems reasonable to conclude that at least some children with ADHD have a very early onset of brain abnormalities.

GENETICS

Part of "43 - Pathophysiology of Attention-Deficit/Hyperactivity Disorder "

Family Studies

Figure 43.1A shows rates of hyperactivity among the siblings of hyperactive probands (71 ,72 ,73 ,74 and 75). Figure 43.1B shows an elevated prevalence of ADHD among mothers and fathers of children with ADHD that provides further support for the familiality of the disorder and evidence that the adult diagnosis is valid. These studies leave no doubt that ADHD

is familial. Moreover, studies of more distant relatives are consistent with this idea as well (76).

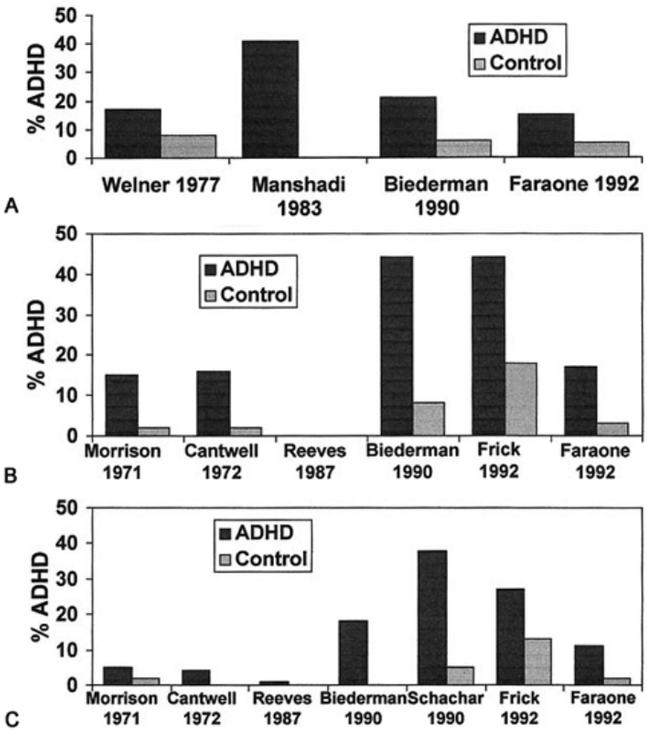


FIGURE 43.1. ADHD in relatives of ADHD and controls children. **A:** ADHD in siblings. **B:** ADHD in fathers. **C:** ADHD in mothers.

Family studies of ADHD suggest that its psychiatric comorbidities may help to clarify its genetic heterogeneity. The Harvard/Massachusetts General Hospital (Boston) ADHD family project studied two independent samples of children with attention-deficit disorder (ADD) as defined by the DSM-III (74) and ADHD as defined by the DSM-III-R (77). These data show that (a) ADHD and major depression share common familial vulnerabilities (78 ,79), (b) children with ADHD who have conduct (80 ,81) and bipolar (82 ,83) disorders may comprise a distinct familial subtype of ADHD, and (c) ADHD is familially independent of anxiety disorders (84) and learning disabilities (85). Thus, stratification by conduct and bipolar disorders may cleave the universe of children with ADHD into more familially homogeneous subgroups. In contrast, major depression may be a nonspecific manifestation of different ADHD subforms. In a sample of 132 ADHD sib-pair families, Smalley et al. reported further evidence that ADHD with conduct disorder is a distinct subtype (86). These investigators also examined comorbidity with learning disability, but these data produced equivocal results.

Faraone et al. proposed that stable or persistent ADHD may be a useful subtype of ADHD for genetic studies (87). These investigators reasoned that cases that remit before adolescence could have a smaller genetic component to their disorder than persistent cases. Evidence supporting this hypothesis derives from several studies. In a prospective follow-up study, Biederman et al. showed that by midadolescence, 85% of boys with ADHD continued to have ADHD; 15% remitted (88). The prevalence of ADHD among parents was 16.3% for the persistent ADHD probands and 10.8% for the remitted ADHD probands. For sibs, the respective prevalences were 24.4% and 4.6%. Thus, these data suggest that children with persistent ADHD have a more familial form of ADHD than those whose ADHD remits by adolescence.

Consistent with this finding, Biederman et al.(188) showed that children of parents with clinically referred, childhood-onset, ADHD were at high risk of meeting diagnostic criteria for ADHD: 84% of the adults with ADHD who had children had at least one child with ADHD, and 52% had two or more children with ADHD (89). The 57% rate of ADHD among children of adults with ADHD was much higher than the more modest 15% risk for ADHD in siblings of referred children with this disorder. These findings were consistent with a prior study by Manshadi et al. (72). They studied the siblings of 22 alcoholic adult psychiatric patients who met DSM-III criteria for ADD, residual type. The authors compared these patients with 20 patients matched for age and comorbid psychiatric diagnoses. Forty-one percent of the siblings of the adult ADD probands were diagnosed with ADHD compared with 0% of the non-ADHD comparison siblings.

In another retrospective study, Biederman et al.(188) compared adolescents with ADHD having retrospectively reported childhood-onset ADHD with children with ADHD (90). These investigators found that the relatives of adolescent probands had higher rates of ADHD compared with the relatives of child probands. Thus, a prospective study of children and retrospective studies of adolescents and adults suggested that, when ADHD persists into adolescence and adulthood, it is highly familial. This idea is consistent with one of Ernst's explanations for the finding that frontal dopaminergic hypoactivity is stronger in adult ADHD compared with adolescent ADHD; that is, frontal dopaminergic hypoactivity may be associated with persistent ADHD.

Twin and Adoption Studies

Although family studies provide much useful information, they cannot disentangle genetic from environmental sources of transmission. To do so, we must turn to twin and adoption studies. There are two types of twins: identical or monozygotic twins share 100% of their genes in common. In contrast, fraternal or dizygotic twins are no more genetically alike than siblings and therefore share only 50% of

their genes. Thus, the occurrence of twinning creates a natural experiment in psychiatric genetics (91). If a disorder is strongly influenced by genetic factors, then the risk to co-twins of ill probands should be greatest when the twins are monozygotic. The risk to dizygotic twins should exceed the risk to controls but should not be greater than the risk to siblings.

Twin data are used to estimate *heritability*, which measures the degree to which a disorder is influenced by genetic factors. Heritability ranges from zero to one, with higher levels indicating a greater degree of genetic determination. Figure 43.2 presents heritability data from 11 twin studies of ADHD. These data attribute about 80% of the origin of ADHD to genetic factors.

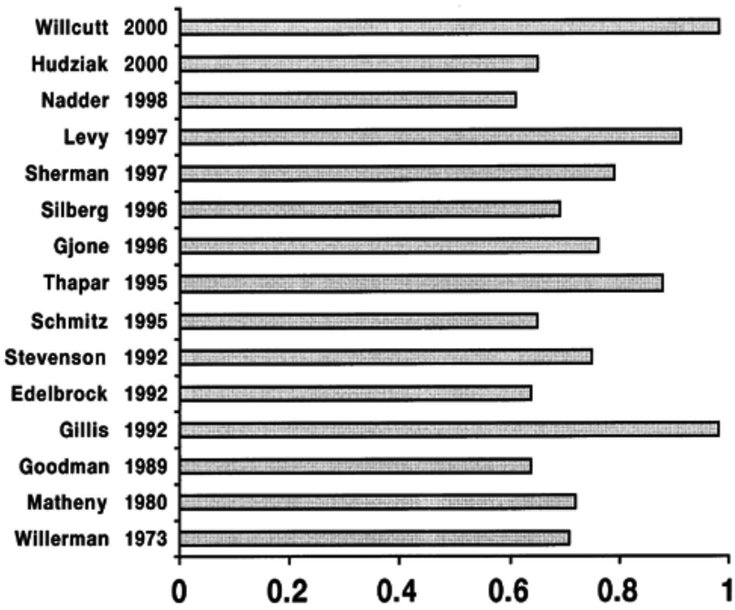


FIGURE 43.2. Heritability of ADHD.

Goodman and Stevenson found the heritability of hyperactivity to be 64% (92 ,93). In a repeat analysis of these data, Stevenson reported that the heritability of mother-reported activity levels was 75%, and the heritability of a psychometric measure of attention was 76% (94). In a study of ADHD in twins who also had reading disability, Gilger et al. estimated the heritability of attention-related behaviors as 98% (95). In a study of 288 male twin pairs, Sherman et al. examined inattentive and impulsive-hyperactive symptoms using both mother and teacher reports (96). Within both raters, the heritability of the impulsivity-hyperactivity dimension exceeded that of the inattention dimension; however, mothers' ratings showed higher heritability than did teachers' ratings. Specifically, mothers' ratings produced a heritability of 91% for impulsivity and hyperactivity and 69% for inattention. Teachers' ratings yielded a heritability of 69% for impulsivity and hyperactivity and 39% for inattention. Using the Child Behavioral Checklist as a dimensional measure, Hudziak and colleagues found a similar heritability (60% to 68%) for mother-reported attention problems (97).

Other studies of inattentive and hyperactive symptoms found a high heritability and minimal impact of the shared environment (98 ,99). Rhee et al. examined gender differences in heritability using twin and sibling pairs from Australia (99). Specific genetic and environmental influences were highly similar for boys and girls. Slight differences that emerged were related to more influence of the shared environment in girls and some evidence genetic dominance in boys.

Several twin studies examined the genetic contribution to the comorbidity of ADHD and other disorders. Data from Gilger et al. (95) were consistent with a prior family study (85) in suggesting that ADHD and reading disability were genetically independent; however, the existence of a genetically mediated subtype of both disorders could not be excluded. In contrast, two twin studies suggested that ADHD and reading disability share some genes in common (100 ,101). That this relationship may be complex is suggested by the report by Willicutt et al. of genetic overlap between reading disability and inattention but not between reading disability and hyperactive impulsive symptoms (102).

Nadder et al. examined whether ADHD and comorbid conduct and oppositional defiant disorder symptoms shared genetic risk factors (98). These investigators found that 50% of the correlation between the ADHD and comorbid conduct was the result of shared genes. Similarly, the twin study of Silberg et al. found that genes influencing variation in hyperactivity scores were also responsible for variation in conduct problems (103). Between 76% and 88% of the correlation between hyperactivity and conduct scores were attributed to genes. These investigators concluded that the results were consistent with the existence of a biologically based group of children who manifest both hyperactivity and conduct disturbances. Further evidence that the ADHD plus comorbid conduct subgroup may be etiologically meaningful comes from a study showing differences in serotonergic functioning between aggressive and nonaggressive children with ADHD (104).

Like twinning, adoption provides another useful experiment for psychiatric genetics (91). Whereas parents can confer a disease risk to their biological children by both biological and environmental pathways, to adoptive children they can confer risk only by an environmental pathway. Thus, by examining both the adoptive and the biological relatives of ill probands, we can disentangle genetic and environmental sources of familial transmission.

Adoption studies of ADHD also implicate genes in its origin. The adoptive relatives of children with ADHD are less likely to have ADHD or associated disorders than are the biological relatives of children with ADHD (105 ,106). Biological relatives of children with ADHD also do more

poorly on standardized measures of attention than do adoptive relatives of children with ADHD (107).

Segregation Analysis Studies

Segregation analysis provides evidence of genetic transmission by demonstrating that the pattern of illness in families is consistent with known genetic mechanisms. An early approach to this was reported by Morrison and Stewart, who concluded that polygenic inheritance was a likely mode of transmission for ADHD (108). Contrasting data were presented by Deutsch et al. (109). They found preliminary evidence for a single dominant gene regulating the transmission of ADHD and minor physical anomalies in 48 families. Similarly, Faraone et al. reported that the familial distribution of ADHD was consistent with the effects of a single major gene (75). Similar results were since reported in a twin study by Eaves et al. (110) and in a pedigree study by Hess et al. (111). Consistent findings also emerged from South America (112). Based on a sample of families from Colombia, the only models of inheritance that could not be rejected were those of dominant and codominant major gene effects. Finally, when families of ADHD probands were ascertained by the father's diagnosis of substance abuse, Maher et al. found that a sex-dependent mendelian codominant model was the best explanation for the pattern of transmission of ADHD (113).

Although the segregation analyses of ADHD suggest that a single gene of major effect is involved in the origin of ADHD, the differences in fit among genetic models was modest. This was especially true for the comparison of multifactorial and single gene inheritance. Several interpretations of these results are possible. If ADHD had more than one genetic cause, then the evidence of any single mode of transmission could be relatively weak. Alternatively, ADHD may be caused by several interacting genes of modest effect. This latter idea is consistent with ADHD's high population prevalence (2% to 7% for ADHD) and high concordance in monozygotic twins but modest recurrence risks in first-degree relatives.

The studies by Deutsch et al. and Faraone et al. predicted that only about 40% of children carrying the putative ADHD gene would develop ADHD. This finding and other features of the genetic epidemiology of ADHD suggest that such a gene likely interacts with other genes and environmental factors to produce ADHD. Moreover, the segregation studies indicated that about 2% of people without the ADHD gene would develop ADHD, a finding suggesting that nongenetic forms of ADHD may exist.

Chromosomal Anomalies and Molecular Genetic Studies

Anomalies in the number or gross structure of chromosomes usually lead to very early-onset disorders having severe clinical manifestations (e.g., mental retardation, gross physical anomalies). No systematic studies of gross chromosomal anomalies in ADHD have been conducted, but there are several reports that such anomalies cause hyperactivity and inattention. Examples include the fragile X syndrome, duplication of the Y chromosome in boys, and loss of an X chromosome in girls. These associations are intriguing but rare. Thus, they can account for only a very small proportion of cases of ADHD.

Molecular genetic studies use the methods of linkage and association to search for aberrant genes that cause disease. Such studies of ADHD are relatively new and far from definitive. Hauser et al. demonstrated that a rare familial form of ADHD is associated with generalized resistance to thyroid hormone, a disease caused by mutations in the thyroid receptor-ß gene (114). The thyroid receptor-ß gene cannot, however, account for many cases of ADHD because the prevalence of generalized resistance to thyroid hormone is very low among patients with ADHD (1 in 2,500) (115), and, among pedigrees with generalized resistance to thyroid hormone, the association between ADHD and the thyroid receptor-ß gene has not been consistently found (116).

Several research teams have examined candidate genes in dopamine pathways because, as discussed earlier, animal models, theoretic considerations, and the effectiveness of stimulant treatment implicate dopaminergic dysfunction in the pathophysiology of this disorder. Several groups have reported an association between ADHD and dopamine D4 receptor gene *(DRD4)* gene (117 ,118 ,119 ,120 ,121 ,122 and 123). Notably, each study showed the 7-repeat allele of *DRD4* to be associated with ADHD despite the use of different diagnostic systems (DSM-IIIR and DSM-IV) and measures of ADHD (rating scales and structured interviews). However, like many findings in psychiatric genetics (91), these positive findings are offset by some negative studies (124 ,125 ,126 ,127 and 128).

The positive *DRD4* findings could be caused by another gene in linkage disequilibrium with *DRD4* or another variant within *DRD4*. However, because the *DRD4* 7-repeat allele mediates a blunted response to dopamine, it is a biologically reasonable risk factor for ADHD (129). The 7-repeat allele has also been implicated in novelty seeking, a personality trait related to ADHD (130 ,131). Moreover, both norepinephrine and dopamine are potent agonists of *DRD4* (132).

When the D4 gene is disabled in a knockout mouse model, dopamine synthesis increases in the dorsal striatum, and the mice show locomotor supersensitivity to ethanol, cocaine, and methamphetamine. (133). D4 knockout mice also show reduced novelty-related exploration (134), a finding consistent with human data suggesting a role for D4 in novelty-seeking behaviors.

Cook et al. reported an association between ADHD and the 480-bp allele of the *DAT* gene using a family-based association study (135). This finding was replicated by Gill et al. (136), Daly et al. (126), and Waldman et al. (137), but

not in other studies (124 ,138). In the study by Waldman et al. (137), hyperactive-impulsive symptoms but not inattentive symptoms were related to the number of DAT high-risk alleles. Further support for a link between the *DAT* gene and ADHD comes from a study that relates this gene to poor methylphenidate response in children with ADHD (139) and from the neuroimaging study (Table 43.2) showing that DAT activity in the striatum is elevated by 70% in adults with ADHD (67).

In mice, eliminating *DAT* gene function leads to several features suggestive of ADHD: hyperactivity, deficits in inhibitory behavior, and a paradoxical response to stimulants (i.e., stimulants reduce hyperactivity) (37 ,140). Studies of this knockout mouse model show the potential complexities of gene–disease associations. The loss of the *DAT* gene has many biological effects: increased extracellular dopamine, a doubling of the rate of dopamine synthesis (141), decreased dopamine and tyrosine hydoxylase in striatum (142), and a nearly complete loss of functioning of dopamine autoreceptors (143). Because ADHD is believed to be a hypodopaminergic disorder, the decreased striatal dopamine may be most relevant to the disorder.

Gainetdinov et al. showed that enhancement of serotonergic transmission mediates the mouse's paradoxical response to stimulants (37). These researchers attributed this to the effects of stimulants on the serotonin transporter. To complicate matters further, Bezard et al. showed that *DAT* knockout mice did not experience MPTP-induced dopaminergic cell death (144), and another study found a gradient effect such that mice with zero, one, and two functional *DAT* genes showed increasing susceptibility to MPTP (145). These latter findings suggest that individual differences in the *DAT* gene may mediate susceptibility to neurotoxins having an affinity for the DAT.

A population-based association study has also implicated the A1 allele of the dopamine D2 receptor gene in ADHD (146). Absence of the D2 gene in mice leads to significantly reduced spontaneous movements, a finding suggesting that D2 plays a role in the regulation of activity levels (147 ,148). Mice without D2 genes also show decreased striatal DAT functioning (149), a finding that illustrates the potential effects of gene–gene interaction on simple phenotypes such as locomotion in mice. In addition, Calabresi et al. used the D2 knockout mouse to study the role of the D2 receptor in striatal synaptic plasticity (150). In these mice, these researchers found abnormal synaptic plasticity at corticostriatal synapses and long-term changes in synaptic efficacy in the striatum.

The only human study of the D3 receptor gene found no evidence of an association with ADHD (151). However, homozygous mice lacking D3 receptors displayed increased locomotor activity, and heterozygous mice showed less pronounced hyperactivity. These results led Accili et al. to conclude that D3 receptors play an inhibitory role in the control of certain behaviors (152).

Four human studies of ADHD have examined the catechol-O-methyltransferase (COMT) gene, the product of which is involved in the breakdown of dopamine and norepinephrine. Although one study found that ADHD was associated with the Val allele (153), others have found no association between the *COMT* polymorphism and ADHD in Irish (154), Turkish (155), and Canadian (156) samples. Despite the negative finding, the positive finding is intriguing because the Val allele leads to high COMT activity and an increased breakdown of catecholamines.

Another study found an association with the DXS7 locus of the X chromosome, a marker for monoamine oxidase that encode enzymes that metabolize dopamine and other neurotransmitters (157). Finally, Comings and colleagues found associations and additive effects of polymorphisms at three noradrenergic genes (the adrenergic α_{2A}, adrenergic α_{2C}, and dopamine-ß-hydroxylas) on ADHD symptoms in a sample of patients with Tourette syndrome (158), but they found no association between the tyrosine hydroxylase gene and ADHD in this sample (159).

Some investigators have used the coloboma mouse model to investigate the genetics of ADHD. These mice have the coloboma mutation, a hemizygous, 2-centimorgan deletion of a segment on chromosome 2q. The mutation leads to spontaneous hyperactivity (which is reversed by stimulants), delays in achieving complex neonatal motor abilities, deficits in hippocampal physiology that may contribute to learning deficiencies, and deficits in Ca^{2+}-dependent dopamine release in dorsal striatum (160).

The coloboma deletion region includes the gene encoding SNAP-25, a neuron-specific protein implicated in exocytotic neurotransmitter release. Hess et al. suggested that interference with SNAP-25 may mediate the mouse's hyperactivity (161). As predicted by this hypothesis, when these investigators bred a SNAP-25 transgene into coloboma mice, the animals' hyperactivity was reduced. Moreover, other work suggested that reduced SNAP-25 expression leads to striatal dopamine and serotonin deficiencies, which may be involved in hyperactivity (162).

Hess et al. tested the idea that the human homologue of the mouse coloboma gene could be responsible for ADHD by completing linkage studies of families with ADHD by using markers on human chromosome 20p11-p12, which is syntenic to the coloboma deletion region (111). These investigators used five families for which segregation analysis suggested that ADHD was the result of a sex-influenced, single gene. However, no significant linkage was detected between ADHD and markers on chromosome 20p11-p12.

ENVIRONMENTAL RISK FACTORS

Although genetic studies of ADHD unequivocally show that genes are risk factors for the disorder, they also show

that the environment has a strong influence on the emergence of the disorder. This conclusion follows from studies of identical twins, which show that when one twin has ADHD, the probability of the other, genetically identical, twin's having ADHD is only about 60%. This less than perfect identical twin concordance implicates environmental risk factors. The nature of these risk factors has emerged from studies assessing features of the biological and psychosocial environment that may increase the risk of ADHD.

Biological Adversity

The idea that certain foods could cause ADHD received much attention in the popular press after claims were made that ADHD could be cured by eliminating food additives from the diet. The Feingold diet for ADHD was popularized by the media and was accepted by many parents of ill children. Systematic studies, however, showed the diet was not effective and concluded that food additives do not cause ADHD (163). Another popular theory posited that excessive sugar intake would lead to ADHD symptoms. Although some positive studies supported this idea, the bulk of systematic, controlled research did not (164).

In contrast to the mostly negative studies of dietary factors, some toxins have been implicated in the origin of at least some cases of ADHD. Several groups have shown that lead contamination leads to distractibility, hyperactivity, restlessness, and lower intellectual functioning (165). However, many children with ADHD do not show lead contamination, and many children with high lead exposure do not develop ADHD. Thus, lead exposure cannot account for the bulk of cases of ADHD.

The literature examining the association of ADHD with pregnancy and delivery complications (PDCs) presents conflicting results; it tends to support the idea that PDCs can predispose children to ADHD (166 ,167 and 168), although some investigators do not (169). The PDCs implicated in ADHD frequently lead to hypoxia and tend to involve *chronic* exposures to the fetus, such as toxemia, rather than *acute*, traumatic events, such as delivery complications.

For example, Conners reported that mothers of children with ADHD had high rates of toxemia during pregnancy (166). Hartsough and Lambert described eight PDCs associated with ADHD: maternal illness, toxemia, eclampsia, older maternal age, parity of child, fetal postmaturity, duration of labor, and fetal distress during labor or birth (170). Nichols and Chen found that hyperactivity was significantly associated with low birth weight (171), and Chandola et al. reported antepartum hemorrhage, maternal age, length of labor, sex, and 1-minute Apgar scores to be significant prenatal and perinatal risk factors for subsequent referral for hyperactivity (172).

Sprich-Buckminster et al. showed that the association between ADHD and PDCs was strongest for children with ADHD who had psychiatric comorbidity (168). PDCs were also elevated among children with ADHD who had no family history of ADHD. These investigators concluded that PDCs may be more common among those children with ADHD having a weaker genetic predisposition, but this hypothesis was not confirmed in another study by the same group (167). The latter study found that children with ADHD and a history of PDCs showed more school failure and psychometric evidence of cognitive impairment than other children with ADHD. In addition to confirming the etiologic role of medical complications, this study showed that psychosocial stress during pregnancy predicted subsequent ADHD and poor cognitive performance in children. Notably, catecholamines are secreted in response to stress, and mouse studies showed that catecholamine administration produces uterine vasoconstriction and fetal hypoxia (173).

One extensively studied risk factor has been maternal smoking during pregnancy. By exposing the fetus to nicotine, maternal smoking can damage the brain at critical times in the developmental process. The smoking mother is at increased risk of antepartum hemorrhage, low maternal weight, and abruptio placentae (173). Her fetus is at risk of low birth weight (173 ,174), and because smoking increases carboxyhemoglobin levels in both maternal and fetal blood, it places the fetus at risk of hypoxia (175). Consistent with these effects, maternal smoking during pregnancy predicts behavioral and cognitive impairment in children and ADHD (41 ,176).

Animal studies in pregnant mice and rats have shown a positive association between chronic exposure to nicotine and hyperactivity in offspring (42). Neonatal nicotine exposure prevents the development of low-affinity nicotine receptors (177), and chronic exposure results in tolerance to the drug and an increase in brain nicotinic receptors (178 ,179 ,180 and 181). Because nicotinic receptors modulate dopaminergic activity and dopaminergic dysregulation may be involved in the pathophysiology of ADHD, it is theoretically compelling to consider maternal smoking as a risk factor for ADHD.

Little is known about the potential role of *in utero* exposure to viral infections. Because maternal viral infections can affect the fetus and can have an adverse impact on the developing brain, viral infections could be associated with later psychopathology. Because viral infections occur more commonly in winter than in other seasons, season-of-birth data have been used to implicate *in utero* viral infection for several disorders including schizophrenia (182), autism (183), and dyslexia (184)

Although Mick et al. found no evidence of a strong seasonal pattern of birth in children with ADHD (185), they did find statistically significant peaks for September births in children with ADHD who had comorbid learning disabilities and in children with ADHD who had no additional psychiatric comorbidity. Thus, it is possible that winter infections during the first trimester of pregnancy may account

for some subtypes of ADHD. Mick et al. found no evidence favoring the idea that putative viral exposure led to a nonfamilial form of ADHD. In contrast, they found a weak trend toward an increase in winter births for children with ADHD who have a positive family history of ADHD. If replicated, this finding suggests that a seasonally mediated infection at birth may be an environmental "trigger" for the genetic predisposition to the disorder.

Psychosocial Adversity

The delineation of psychosocial features in the child's environment associated with more impaired outcome in children with ADHD has potentially important clinical, scientific, and public health implications. Such efforts can help to identify etiologic risk factors associated with more impaired outcome in ADHD and can characterize early predictors of persistence and morbidity of this disorder. Moreover, finding environmental risk factors for ADHD could help to design improved preventive and therapeutic intervention programs.

The classic studies by Rutter et al. of the Isle of Wight and the inner borough of London provide a compelling example of how psychosocial risk factors influence child psychopathology (186). Compelling examples of how psychosocial risk factors affect child psychopathology, these studies examined the prevalence of mental disorders in children living in two very different geographic areas. This research revealed six risk factors within the family environment that correlated significantly with childhood mental disturbances: (a) severe marital discord, (b) low social class, (c) large family size, (d) paternal criminality, (e) maternal mental disorder, and (f) foster placement. This work found that it was the aggregate of adversity factors, rather than the presence of any single one, that impaired development. Other studies also found that as the number of adverse conditions accumulated, the risk of impaired outcome in the child increased proportionally (187). Biederman et al. found a positive association between Rutter's index of adversity and ADHD, measures of ADHD-associated psychopathology, impaired cognition, and psychosocial dysfunction (188).

Other cross-sectional and longitudinal studies have identified variables such as marital distress, family dysfunction, and low social class as risk factors for psychopathology and dysfunction in children. For example, the Ontario Child Health Study in Canada showed that family dysfunction and low income predicted persistence and onset of one or more psychiatric disorders over a 4-year follow-up period (189). Other work implicated low maternal education, low social class, and single parenthood as important adversity factors for ADHD (171 ,190). These studies suggested that the mothers of children with ADHD had more negative communication patterns, more conflict with their children, and a greater intensity of anger than did control mothers.

Biederman et al. showed that long-term conflict, decreased family cohesion, and exposure to parental psychopathology, particularly maternal psychopathology, were more common in ADHD-affected families compared with control families (191). The differences between children with ADHD and control children could not be accounted for by either socioeconomic status or parental history of major psychopathology. Moreover, increased levels of family-environment adversity predicted impaired psychosocial functioning. Measures indexing long-term family conflict showed a more pernicious impact on the exposed child than those indexing exposure to parental psychopathology. Indeed, marital discord in families has consistently predicted disruptive behaviors in boys (192). This research shows that the extent of discord and overt conflict, regardless of whether the parents are separated, predicts the child's risks of psychopathology and dysfunction (193).

Thus, dysfunctional family environments appear to be a nonspecific risk factor for psychiatric disorders and psychological distress. Reid and Crisafulli reported a metaanalysis of the impact of marital discord on the psychological adjustment of children and found that parental conflict significantly predicted a variety of child behavior problems (194). The Ontario Child Health Study provided a prospective example of the impact of parental conflict on children's mental health: family dysfunction (and low income) predicted persistence and onset of one or more psychiatric disorders over a 4-year period (189).

Low maternal warmth and high maternal malaise and criticism were previously associated with ADHD in children (195), and an epidemiologic study examining family attributes in children who had undergone stressful experiences found that children's perceptions of mothers, but not fathers, differentiated stress-resilient and stress-affected children (196).

An extensive literature documents maternal depression as a risk factor for psychological maladjustment and psychiatric disorder in children (197). This is consistent with the known familial link between ADHD and depression (79). Some investigators have suggested that depressed mood may lead mothers to perceive their children as more deviant than warranted by the child's behavior. Richters, however, reviewed 22 studies of this issue and concluded that, owing to methodologic problems with research in the area, there was no empiric foundation for this claim (198).

Other data revealed a link between maternal depression and child functioning that was independent of the mother's perceptions. These data suggested that depressed mothers accurately perceive symptomatic behavior but react to it in a negative manner that worsens the condition of the child. This conclusion was echoed by Gelfand and Teti (197). Their comprehensive review of relevant literature found many studies to document the assertion that depressed mothers have attitudes of insensitivity, disengagement, disapproval, and hostility toward their children. They also

found maternal depression to be associated with undesirable parenting practices such as intrusiveness, unresponsiveness, and inept discipline. In addition, their review supported the idea that depressed mothers had negative perceptions of their children.

Other work shows that ADHD in children predicts depression in mothers, but maternal depression provides no additional information for predicting ADHD in siblings of ADHD probands. This finding suggests that maternal depression is a heterogeneous disorder. It may be that some mothers have a disorder that is genetically linked to ADHD, whereas others may experience depression resulting from the stress of raising a child with ADHD (and perhaps living with an ADHD-affected or antisocial husband). Furthermore, it is possible that maternal depression exacerbates family conflict and poor parenting, both of which could exacerbate ADHD symptoms.

Notably, although many studies provide strong evidence of the importance of psychosocial adversity for ADHD, these factors tend to emerge as universal predictors of children's adaptive functioning and emotional health, not predictors that are specific to ADHD. Thus, they can be conceptualized as nonspecific triggers of an underlying predisposition or as modifiers of the course of illness.

SUMMARY AND CONCLUSIONS

Part of "43 - Pathophysiology of Attention-Deficit/Hyperactivity Disorder "

It is not yet possible to describe the origin and pathophysiology of ADHD completely. Nevertheless, converging evidence from the studies reviewed in this chapter supports several empiric generalizations, which should be useful in guiding future research and theory.

Catecholamine Hypothesis

Much research supports the idea that catecholaminergic systems mediate the onset and expression of ADHD symptoms. The key data supporting this idea are as follows: (a) anti-ADHD medications have noradrenergic and dopaminergic effects; (b) lesion studies in mouse and monkey models implicate dopaminergic pathways; (c) the SHR rat shows deficits in catecholaminergic systems; (d) D2, D3, and D4 knockout mice studies show that these genes regulate locomotor activity; and (e) human studies implicate the *DRD4* and *DAT* genes in the origin of ADHD.

Although the role of catecholamine systems cannot be disputed, future work must also consider other neurotransmitter systems that exert upstream effects on catecholamines. Two prime candidates are nicotinic and serotonergic systems. Nicotinic agonists help to control the symptoms of ADHD, and nicotinic activation enhances dopaminergic neurotransmission. Serotonergic drugs have not been shown to be effective anti-ADHD agents, but knockout mice studies suggest that the paradoxical effects of stimulants on hyperactivity are mediated by serotonergic neurotransmission. Moreover, SNAP-25, which has been implicated in studies of the coloboma mouse, leads to striatal dopamine and serotonin deficiencies. These data call for further studies of serotonergic and nicotinic systems.

Brain Systems

Several types of study provide information about the locus of ADHD's pathophysiology in the brain: neuropsychological studies, neuroimaging studies, and animal models. Taken together, these studies support the idea that ADHD arises from the dysregulation of frontal cortex, subcortical structures, and networks connecting them. This idea fits with the pharmacotherapy of ADHD because a plausible model for the effects of stimulants is that, through dopaminergic or noradrenergic pathways, these drugs increase the inhibitory influences of frontal cortical activity on subcortical structures.

Additional data supporting frontal-subcortical involvement in ADHD are as follows: (a) neuropsychological studies implicate orbitofrontal and dorsolateral prefrontal cortex or regions projecting to these regions; (b) the monkey model of ADHD implicates frontal-striatal neural networks; (c) studies of the SHR rat implicate caudate, putamen, nucleus accumbens, and frontal cortex; patients with frontal lobe damage show ADHD-like behaviors; (d) structural neuroimaging implicates frontal cortex, usually limited to the right side, cerebellum, globus pallidus, caudate, and corpus callosum; (e) the I/LnJ mouse strain shows total callosal agenesis along with behavioral features that resemble ADHD; (f) functional neuroimaging finds hypoactivity of frontal cortex, anterior cingulate cortex, and subcortical structures, usually on the right side; (g) ADHD secondary to brain injury shows lesions in right putamen, right caudate nucleus, and right globus pallidus; (h) disabling the D4 gene in mice leads to increased dopamine synthesis in dorsal striatum; (i) mice without D2 genes also show decreased striatal DAT functioning, abnormal synaptic plasticity at corticostriatal synapses, and long-term changes in synaptic efficacy in the striatum; and (j) the coloboma mouse shows deficient dopamine release in dorsal striatum.

Etiologic Factors

In a word, the origin of ADHD is complex. Although rare cases may have a single cause such as lead exposure, generalized resistance to thyroid hormone, head injury, and frontal lobe epilepsy, most cases of ADHD are probably caused by a complex combination of risk factors.

From the many twin studies of ADHD, we know for certain that genes mediate susceptibility to ADHD. Molecular genetic studies suggest that two of these genes may be the *DRD4* gene and the *DAT* gene. To confirm these findings, we need much more work because, even if the positive

studies are correct, they may implicate neighboring genes instead of those targeted by the studies. It seems unlikely that a single "ADHD gene" causes ADHD with certainty. Instead, it seems likely that several genes act together to form the genetic substrate of the disorder.

When the ADHD-related variants of these genes are discovered, they will probably be "normal" variants and will most certainly not have the devastating effects seen in knockout mouse models. For example, suppose future work confirms that the 7-repeat allele is a risk factor for ADHD. We would consider this a normal variant because about 20% of people who do not have ADHD carry this version of the *DRD4* gene. Most of these people do not develop ADHD despite the blunted dopaminergic transmission associated with that allele, and many patients with ADHD do not carry the allele. Thus, the 7-repeat allele cannot be a necessary or sufficient cause of the disorder. Instead, it acts in concert with other genes and environmental risk factors to bring forth ADHD.

Like genetic studies, studies of environmental risk factors suggest that most of these risks exert small but significant influences on the origin of ADHD. For example, most children with a history of PDCs do not develop ADHD, and most children with ADHD do not have a history of ADHD. Nevertheless, research suggests that such complications are more common among children with ADHD.

These considerations lead us to conclude that the origin of ADHD is multifactorial. A simple multifactorial model posits ADHD to arise a pool of genetic and environmental variables—each of small effect—that act in concert to produce vulnerability to ADHD. If a person's cumulative vulnerability exceeds a certain threshold, he or she will manifest the signs and symptoms of ADHD. According to the multifactorial model, no single factor is a necessary or sufficient cause for ADHD, and each of the etiologic factors is interchangeable (i.e., it does not matter which factors one has; only the total number is important). Whether risk factors combine in an additive or interactive manner is unknown.

The mouse models of ADHD we described provide examples of multifactorial causation in a simple system. One model showed that individual differences in the *DAT* gene could directly produce a hypodopaminergic state; these studies showed that dopamine transporter variants differ in their affinity for neurotoxins. Thus, dopamine transporter abnormalities could interact with environmental toxins to produce hyperactivity. Another line of work shows that catecholamines are secreted in response to stress, and catecholamine administration produces fetal hypoxia. Human studies implicate both stress during pregnancy and fetal hypoxia as risk factors for ADHD.

These simple examples suggest that unraveling the complexities of multifactorial causation will be a difficult task for ADHD researchers. However, because technological developments in neuroscience and molecular genetics are moving at a rapid pace, the next decade of work should provide us with more accurate assessments of the brain along with a complete sequence of the human genome. These advances should set the stage for breakthroughs in our understanding of the neurobiology of ADHD and in our ability to treat affected persons.

DISCLAIMERS

Part of "43 - Pathophysiology of Attention-Deficit/Hyperactivity Disorder "

Dr. Biederman receives research support from Shire Laboratories, Gliatec, Cephalon, Novartis Pharmaceuticals, and Eli Lilly & Company. In addition, he serves on speaking bureaus for SmithKline Beecham, Eli Lilly & Company, and Pfizer Pharmaceuticals.

REFERENCES

1. Barkley RA. *Attention deficit hyperactivity disorder: a handbook for diagnosis and treatment*. New York: Guilford, 1998.

2. Spencer T, Biederman J, Wilens T, et al. Is attention deficit hyperactivity disorder in adults a valid disorder? *Harvard Rev Psychiatry* 1994;1:326–335.

3. Hill J, Schoener E. Age-dependent decline of attention deficit hyperactivity disorder. *Am J Psychiatry* 1996;153:1143–1146.

4. Biederman J, Newcorn J, Sprich S. Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders. *Am J Psychiatry* 1991;148:564–577.

5. Caron C, Rutter M. Comorbidity in child psychopathology: concepts, issues and research strategies. *J Child Psychol Psychiatry* 1991;32:1063–1080.

6. Bird HR, Canino G, Rubio-Stipec M, et al. Estimates of the prevalence of childhood maladjustment in a community survey in Puerto Rico: the use of combined measures. *Arch Gen Psychiatry* 1988;45:1120–1126.

7. Anderson JC, Williams S, McGee R, et al. DSM-III disorders in preadolescent children: prevalence in a large sample from the general population. *Arch Gen Psychiatry* 1987;44:69–76.

8. Spencer TJ, Biederman J, Wilens T, et al. Pharmacotherapy of attention deficit hyperactivity disorder across the lifecycle: a literature review. *J Am Acad Child Adolesc Psychiatry* 1996;35:409–432.

9. Wilens T, Biederman J, Spencer T, et al. Pharmacotherapy of adult attention deficit/hyperactivity disorder: a review. *J Clin Psychopharmacol* 1995;15:270–279.

10. Prince J, Wilens T, Biederman J, et al. A controlled study of nortriptyline in children and adolescents with attention deficit hyperactivity disorder. In: Scientific proceedings of the annual meeting of the American Academy of Child and Adolescent Psychiatrists XV, Chicago, 1999.

11. Casat CD, Pleasants DZ, Van Wyck Fleet J. A double-blind trial of bupropion in children with attention deficit disorder. *Psychopharmacol Bull* 1987;23:120–122.

12. Casat CD, Pleasants DZ, Schroeder DH, et al. Bupropion in children with attention deficit disorder. *Psychopharmacol Bull* 1989;25:198–201.

13. Wender PH, Reimherr FW. Bupropion treatment of attention-deficit hyperactivity disorder in adults. *Am J Psychiatry* 1990;147:1018–1020.

14. Ernst M, Liebenauer LL, Jons PH, et al. Selegiline in adults with attention deficit hyperactivity disorder: clinical efficacy and safety. *Psychopharmacol Bull* 1996;32:327–334.

15. Spencer T, Wilens TE, Biederman J. A double-blind, crossover comparison of tomoxetine and placebo in adults with ADHD. In: Scientific proceedings of the annual meeting of the American Academy of Child and Adolescent Psychiatrists XII, New Orleans, 1995.

16. Spencer T, Biederman J, Wilens T, et al. An open, dose ranging study of tomoxetine in children with ADHD. In: Scientific Proceedings of the annual meeting of the American Academy of Child and Adolescent Psychiatry XV, Chicago, 1999.

17. Biederman J, Spencer T, Wilens T. Psychopharmacology in children and adolescents. In: Wiener J, ed. *Textbook of child and adolescent psychiatry*. Washington, DC: American Psychiatric Press, 1997:779–813.

18. Levin ED, Conners CK, Sparrow E, et al. Nicotine effects on adults with attention-deficit/hyperactivity disorder. *Psychopharmacology* 1996;123:55–63.

19. Wilens TE, Biederman J, Spencer TJ, et al. A pilot controlled clinical trial of ABT-418, a cholinergic agonist, in the treatment of adults with attention deficit hyperactivity disorder. *Am J Psychiatry* 1999;156:1931–1937.

20. Elia J, Borcherding BG, Potter WZ, et al. Stimulant drug treatment of hyperactivity: biochemical correlates. *Clin Pharmacol Ther* 1990;48:57–66.

21. Solanto M. Neuropsychopharmacological mechanisms of stimulant drug action in attention-deficit hyperactivity disorder: a review and integration. *Behav Brain Res* 1998;94:127–152.

22. Zametkin AJ, Rapoport JL. Noradrenergic hypothesis of attention deficit disorder with hyperactivity: a critical review. In: Meltzer HY, ed. *Psychopharmacology: the third generation of progress*. New York: Raven, 1987:837–842.

23. Zametkin AJ, Rapoport JL.Neurobiology of attention deficit disorder with hyperactivity: where have we come in 50 years? *J Am Acad Child Adolesc Psychiatry* 1987;26:676–686.

24. Pliszka S, McCracken J, Maas J. Catecholamines in attention-deficity hyperactivity disorder: current perspectives. *J Am Acad Child Adolesc Psychiatry* 1996;35:264–272.

25. Shaywitz SE, Cohen DJ, Shaywitz BA. The biochemical basis of minimal brain dysfunction. *J Pediatr* 1978;92:179–187.

26. Schneider JS, Roeltgen DP. Delayed matching-to-sample, object retrieval, and discrimination reversal deficits in chronic low dose MPTP-treated monkeys. *Brain Res* 1993;615:351–354.

27. Schneider JS, Kovelowski CJD. Chronic exposure to low doses of MPTP. I. Cognitive deficits in motor asymptomatic monkeys. *Brain Res* 1990;519:122–128.

28. Schneider JS, Sun ZQ, Roeltgen DP. Effects of dopamine agonists on delayed response performance in chronic low-dose MPTP-treated monkeys. *Pharmacol Biochem Behav* 1994;48:235–240.

29. Roeltgen DP, Schneider JS. Task persistence and learning ability in normal and chronic low dose MPTP-treated monkeys. *Behav Brain Res* 1994;60:115–124.

30. Russell V, de Villiers A, Sagvolden T, et al. Altered dopaminergic function in the prefrontal cortex, nucleus accumbens and caudate-putamen of an animal model of attention-deficit hyperactivity disorder: the spontaneously hypertensive rat. *Brain Res* 1995;676:343–351.

31. Russell VA. The nucleus accumbens motor-limbic interface of the spontaneously hypertensive rat as studied *in vitro* by the superfusion slice technique. *Neurosci Biobehav Rev* 2000;24:133–136.

32. de Villiers AS, Russell VA, Sagvolden T, et al. Alpha 2-adrenoceptor mediated inhibition of [^{3}H]dopamine release from nucleus accumbens slices and monoamine levels in a rat model for attention-deficit hyperactivity disorder. *Neurochem Res* 1995;20:427–433.

33. Papa M, Berger DF, Sagvolden T, et al. A quantitative cytochrome oxidase mapping study, cross-regional and neurobehavioural correlations in the anterior forebrain of an animal model of attention deficit hyperactivity disorder. *Behav Brain Res* 1998;94:197–211.

34. King JA, Barkley RA, Delville Y, et al. Early androgen treatment decreases cognitive function and catecholamine innervation in an animal model of ADHD. *Behav Brain Res* 2000;107:35–43.

35. Carey MP, Diewald LM, Esposito FJ, et al. Differential distribution, affinity and plasticity of dopamine D-1 and D-2 receptors in the target sites of the mesolimbic system in an animal model of ADHD. *Behav Brain Res* 1998;94:173–185.

36. Tatsumi M, Groshan K, Bakely R, et al. Pharmacological profile of antidepressants and related compounds at human monoamine transporters. *Eur J Pharmacol* 1997;340:249–258.

37. Gainetdinov RR, Wetsel WC, Jones SR, et al. Role of serotonin in the paradoxical calming effect of psychostimulants on hyperactivity. *Science* 1999;283:397–402.

38. Milberger S, Biederman J, Faraone S, et al. Further evidence of an association between attention-deficit/hyperactivity disorder and cigarette smoking: findings from a high-risk sample of siblings. *Am J Addict* 1997;6:205–217.

39. Milberger S, Biederman J, Faraone SV, et al. Attention deficit hyperactivity disorder is associated with early initiation of cigarette smoking in children and adolescents. *J Am Acad Child Adolesc Psychiatry* 1997;36:37–44.

40. Riggs PD, Mikulich SK, Whitmore EA, et al. Relationship of ADHD, depression, and non-tobacco substance use disorders to nicotine dependence in substance-dependent delinquents. *Drug Alcohol Depend* 1999;54:195–205.

41. Milberger S, Biederman J, Faraone S, et al. Is maternal smoking during pregnancy a risk factor for attention deficit hyperactivity disorder in children? *Am J Psychiatry* 1996;153:1138–1142.

42. Johns JM, Louis TM, Becker RF, et al. Behavioral effects of prenatal exposure to nicotine in guinea pigs. *Neurobehav Toxicol Teratol* 1982;4:365–369.

43. Fung YK, Lau YS. Effects of prenatal nicotine exposure on rat striatal dopaminergic and nicotinic systems. *Pharmacol Biochem Behav* 1989;33:1–6.

44. Westfall TC, Grant H, Perry H. Release of dopamine and 5-hydroxytryptamine from rat striatal slices following activation of nicotinic cholinergic receptors. *Gen Pharmacol* 1983;14:321–325.

45. Mereu G, Yoon K, Gessa G, et al. Preferential stimulation of ventral tegmental area dopaminergic neurons by nicotine. *Eur J Pharmacol* 1987;141:395-399.

46. Satterfield JH, Dawson ME. Electrodermal correlates of hyperactivity in children. *Psychophysiology* 1971;8:191–197.

47. Mattes JA. The role of frontal lobe dysfunction in childhood hyperkinesis. *Comp Psychiatry* 1980;21:358–369.

48. Weiss JL, Seidman LJ. The clinical use of psychological and neuropsychological tests. In: Nicholi A, ed. *The new Harvard guide to psychiatry*. Cambridge, MA: Harvard University Press, 1988:46–69.

49. Barkley RA, Grodzinsky G, DuPaul GJ. Frontal lobe functions in attention deficit disorder with and without hyperactivity: a review and research report. *J Abnorm Child Psychol* 1992;20:163–188.

50. Seidman LJ, Biederman J, Faraone SV, et al. Effects of family history and comorbidity on the neuropsychological performance of children with ADHD: preliminary findings. *J Am Acad Child Adolesc Psychiatry* 1995;34:1015–1024.

51. Seidman LJ, Biederman J, Faraone SV, et al. Toward defining a neuropsychology of attention deficit-hyperactivity disorder: performance of children and adolescents from a large clinically referred sample. *J Consult Clin Psychol* 1997;65:150–160.

52. Seidman LJ, Biederman J, Weber W, et al. Neuropsychological function in adults with attention-deficit hyperactivity disorder. *Biol Psychiatry* 1998;44:260–268.

53. Barkley R, Murphy K, Kwasnik D. Psychological adjustment and adaptive impairments in young adults with ADHD. *J Atten Disord* 1996;1:41–54.

54. Buchsbaum MS, Haier RJ, Sostek AJ, et al. Attention dysfunction and psychopathology in college men. *Arch Gen Psychiatry* 1985;42:354–360.

55. Gualtieri CT, Ondrusek MG, Finley C. Attention deficit disorders in adults. *Clin Neuropharmacol* 1985;8:343–356.

56. Holdnack JA, Moberg PJ, Arnold SE, et al. Speed of processing and verbal learning deficits in adults diagnosed with attention deficit disorder. *Neuropsychiatry Neuropsychol Behav Neurol* 1995;8:282–292.

57. Lovejoy DW, Ball JD, Keats M, et al. Neuropsychological performance of adults with attention deficit hyperactivity disorder (ADHD): diagnostic classification estimates for measures of frontal lobe/executive functioning. *J Int Neuropsychol Soc* 1999;5:222–233.

58. Taylor CJ, Miller DC. Neuropsychological assessment of attention in ADHD adults. *J Atten Disord* 1997;2:77–88.

59. Fuster J. *The prefrontal cortex*. New York: Raven, 1989.

60. Cummings JL. Frontal-subcortical circuits and human behavior. *Arch Neurol* 1993;50:873–880.

61. Magara F, Ricceri L, Wolfer DP, et al. The acallosal mouse strain I/LnJ: a putative model of ADHD? *Neurosci Biobehav Rev* 2000;24:45–50.

62. Ernst M, Liebenauer L, King A, et al. Reduced brain metabolism in hyperactive girls. *J Am Acad Child Adolesc Psychiatry* 1994;33:858–868.

63. Baving L, Laucht M, Schmidt MH. Atypical frontal brain activation in ADHD: preschool and elementary school boys and girls. *J Am Acad Child Adolesc Psychiatry* 1999;38:1363–1371.

64. Ernst M, Zametkin A, Matochik J, et al. DOPA decarboxylase activity in attention deficit hyperactivity disorder adults: a [fluorine-18]fluorodopa positron emission tomographic study. *J Neurosci* 1998;18:5901–5907.

65. Bush G, Frazier JA, Rauch SL, et al. Anterior cingulate cortex dysfunction in attention deficit/hyperactivity disorder revealed by fMRI and the counting stroop. *Biol Psychiatry* 1999;45:1542–1552.

66. Zametkin AJ, Nordahl TE, Gross M, et al. Cerebral glucose metabolism in adults with hyperactivity of childhood onset. *N Engl J Med* 1990;323:1361–1366.

67. Dougherty DD, Bonab AA, Spencer TJ, et al. Dopamine transporter density is elevated in patients with attention deficit hyperactivity disorder. *Lancet* 1999;354:2132–2133.

68. Powell AL, Yudd A, Zee P, et al. Attention deficit hyperactivity disorder associated with orbitofrontal epilepsy in a father and a son. *Neuropsychiatry Neuropsychol Behav Neurol* 1997;10:151–154.

69. Herskovits EH, Megalooikonomou V, Davatzikos C, et al. Is the spatial distribution of brain lesions associated with closed-head injury predictive of subsequent development of attention-deficit/hyperactivity disorder? Analysis with brain-image database. *Radiology* 1999;213:389–394.

70. Nopoulos P, Berg S, Castellenos FX, et al. Developmental brain anomalies in children with attention-deficit hyperactivity disorder. *J Child Neurol* 2000;15:102–108.

71. Welner Z, Welner A, Stewart M, et al. A controlled study of siblings of hyperactive children. *J Nerv Ment Dis* 1977;165:110–117.

72. Manshadi M, Lippmann S, O'Daniel R, et al. Alcohol abuse and attention deficit disorder. *J Clin Psychiatry* 1983;44:379–380.

73. Pauls DL, Shaywitz SE, Kramer PL, et al. Demonstration of vertical transmission of attention deficit disorder. *Ann Neurol* 1983;14:363.

74. Biederman J, Faraone SV, Keenan K, et al. Family-genetic and psychosocial risk factors in DSM-III attention deficit disorder. *J Am Acad Child Adolesc Psychiatry* 1990;29:526–533.

75. Faraone S, Biederman J, Chen WJ, et al. Segregation analysis of attention deficit hyperactivity disorder: evidence for single gene transmission. *Psychiatr Genet* 1992;2:257–275.

76. Faraone SV, Tsuang MT. Methods in psychiatric genetics. In: Tohen M, Tsuang MT, Zahner GEP, eds. *Textbook in psychiatric epidemiology*. New York: John Wiley, 1995:81–134.

77. Biederman J, Faraone SV, Keenan K, et al. Further evidence for family-genetic risk factors in attention deficit hyperactivity disorder: patterns of comorbidity in probands and relatives psychiatrically and pediatrically referred samples. *Arch Gen Psychiatry* 1992;49:728–738.

78. Biederman J, Faraone SV, Keenan K, et al. Evidence of familial association between attention deficit disorder and major affective disorders. *Arch Gen Psychiatry* 1991;48:633–642.

79. Faraone SV, Biederman J. Do attention deficit hyperactivity disorder and major depression share familial risk factors? *J Nerv Ment Dis*s 1997;185:533–541.

80. Faraone S, Biederman J, Monuteaux MC. Attention deficit disorder and conduct disorder in girls: evidence for a familial subtype. *Biol Psychiatry* 2000;48:21–29.

81. Faraone S, Biederman J, Garcia Jetton J, et al. Attention deficit disorder and conduct disorder: longitudinal evidence for a familial subtype. *Psychol Med* 1997;27:291–300.

82. Faraone SV, Biederman J, Mennin D, et al. Bipolar and antisocial disorders among relatives of ADHD children: parsing familial subtypes of illness. *Am J Med Genet* 1998;81:108–116.

83. Faraone SV, Biederman J, Mennin D, et al. Attention-deficit hyperactivity disorder with bipolar disorder: a familial subtype? *J Am Acad Child Adolesc Psychiatry* 1997;36:1378–1387;discussion 1387–1390.

84. Biederman J, Faraone SV, Keenan K, et al. Familial association between attention deficit disorder and anxiety disorders. *Am J Psychiatry* 1991;148:251–256.

85. Faraone S, Biederman J, Krifcher Lehman B, et al. Evidence for the independent familial transmission of attention deficit hyperactivity disorder and learning disabilities: results from a family genetic study. *Am J Psychiatry* 1993;150:891–895.

86. Smalley SL, McCracken J, McGough J. Refining the ADHD phenotype using affective sibling pair families. *Am J Med Genet* 2001;105:31–33.

87. Faraone SV, Biederman J, Monuteaux MC. Toward guidelines for pedigree selection in genetic studies of attention deficit hyperactivity disorder. *Genet Epidemiol* 2000;18:1–16.

88. Biederman J, Faraone SV, Milberger S, et al. Predictors of persistence and remission of ADHD: results from a four-year prospective follow-up study of ADHD children. *J Am Acad Child Adolesc Psychiatry* 1996;35:343–351.

89. Biederman J, Faraone SV, Mick E, et al. High risk for attention deficit hyperactivity disorder among children of parents with childhood onset of the disorder: a pilot study. *Am J Psychiatry* 1995;152:431–435.

90. Biederman J, Faraone SV, Taylor A, et al. Diagnostic continuity between child and adolescent ADHD: findings from a longitudinal clinical sample. *J Am Acad Child Adolesc Psychiatry* 1998;37:305–313.

91. Faraone SV, Tsuang D, Tsuang MT. *Genetics and mental disorders: a guide for students, clinicians, and researchers*. New York: Guilford, 1999.

92. Goodman R, Stevenson J. A twin study of hyperactivity. II. The aetiological role of genes, family relationships and perinatal adversity. *J Child Psychol Psychiatry* 1989;30:691–709.

93. Goodman R, Stevenson J. A twin study of hyperactivity. I. An examination of hyperactivity scores and categories derived from Rutter teacher and parent questionnaires. *J Child Psychol Psychiatry* 1989;30:671–689.

94. Stevenson J. Evidence for a genetic etiology in hyperactivity in children. *Behav Genet* 1992;22:337–344.

95. Gilger JW, Pennington BF, DeFries C. A twin study of the etiology of comorbidity: attention deficit hyperactivity disorder and dyslexia. *J Am Acad Child Adolesc Psychiatry* 1992;31:343–348.

96. Sherman D, Iacono W, McGue M. Attention deficit hyperactivity disorder dimensions: a twin study of inattention and impulsivity hyperactivity. *J Am Acad Child Adolesc Psychiatry* 1997;36:745–753.

97. Hudziak JJ, Rudiger LP, Neale MC, et al. A twin study of inattentive, aggressive, and anxious/depressed behaviors. *J Am Acad Child Adolesc Psychiatry* in press.

98. Nadder TS, Silberg JL, Eaves LJ, et al. Genetic effects on ADHD symptomatology in 7- to 13-year-old twins: results from a telephone survey. *Behav Genet* 1998;28:83–99.

99. Rhee SH, Waldman ID, Hay DA, et al. Sex differences in genetic and environmental influences on DSM-III-R attention-deficit/hyperactivity disorder. *J Abnorm Psychol* 1999;108:24–41.

100. Light JG, Pennington BF, Gilger J, et al. Reading disability and hyperactivity disorder: evidence for a common genetic etiology. *Dev Neuropsychol* 1995;11:323–335.

101. Stevenson J, Pennington BF, Gilger JW, et al. Hyperactivity and spelling disability: testing for shared genetic aetiology. *J Child Psychol Psychiatry* 1993;34:1137–1152.

102. Willicutt EG, Pennington BF, DeFries JC. A twin study of the etiology of comorbidity between reading disability and attention-deficit/hyperactivity disorder. *Am J Med Genet* 2000;96:293–301.

103. Silberg J, Rutter M, Meyer J, et al. Genetic and environmental influences on the covariation between hyperactivity and conduct disturbance in juvenile twins. *J Child Psychol Psychiatry* 1996;37:803–816.

104. Halperin J, Sharma V, Siever L, et al. Serotonergic function in aggressive and nonaggressive boys with attention deficit hyperactivity disorder. *Am J Psychiatry* 1994;151:243–248.

105. Cantwell DP. Genetics of hyperactivity. *J Child Psychol Psychiatry* 1975;16:261–264.

106. Morrison JR, Stewart MA. The psychiatric status of the legal families of adopted hyperactive children. *Arch Gen Psychiatry* 1973;28:888–891.

107. Alberts-Corush J, Firestone P, Goodman JT. Attention and impulsivity characteristics of the biological and adoptive parents of hyperactive and normal control children. *Am J Orthopsychiatry* 1986;56:413–423.

108. Morrison JR, Stewart MA. Bilateral inheritance as evidence for polygenicity in the hyperactive child syndrome. *J Nerv Ment Dis* 1974;158:226–228.

109. Deutsch CK, Matthysse S, Swanson JM, et al. Genetic latent structure analysis of dysmorphology in attention deficit disorder. *J Am Acad Child Adolesc Psychiatry* 1990;29:189–194.

110. Eaves L, Silberg J, Hewitt J, et al. Genes, personality, and psychopathology: a latent class analysis of liability to symptoms of attention-deficit hyperactivity disorder in twins. In: Plomin R, McLearn G, eds. *Nature, Nurture and Psychology*. Washington, DC: American Psychological Association, 1993:285–306.

111. Hess EJ, Rogan PK, Domoto M, et al. Absence of linkage of apparently single gene mediated ADHD with the human syntenic region of the mouse mutant coloboma. *Am J Med Genet* 1995;60:573–579.

112. Lopera F, Palacio LG, Jimenez I, et al. Discrimination between genetic factors in attention deficit. *Rev Neurol* 1999;28:660–664.

113. Maher BS, Marazita ML, Moss HB. Segregation analysis of attention deficit hyperactivity disorder. *Am J Med Gen* 1999;88:71–78.

114. Hauser P, Zametkin A, Martinez P, et al. Attention deficit-hyperactivity disorder in people with generalized resistance to thyroid hormone. *N Engl J Med* 1993;328:997–1001.

115. Weiss R, Stein M, Trommer B, et al. Attention-deficit hyperactivity disorder and thyroid function. *J Pediatr* 1993;123:539–545.

116. Weiss RE, Stein MA, Duck SC, et al. Low intelligence but not attention deficit hyperactivity disorder is associated with resistance to thyroid hormone caused by mutation R316H in the thyroid hormone receptor *B* gene. *J Clin Endocrinol Metab* 1994;78:1525–1528.

117. Barr CL, Wigg KG, Bloom S, et al. Further evidence from haplotype analysis for linkage of the dopamine D4 receptor gene and attention-deficit hyperactivity disorder. *Am J Med Genet* 2000;96:262–267.

118. Comings DE, Gonzalez N, Wu S, et al. Studies of the 48 bp repeat polymorphism of the DRD4 gene in impulsive, compulsive, addictive behaviors: Tourette syndrome, ADHD, pathological gambling, and substance abuse. *Am J Med Genet* 1999;88:358–368.

119. Faraone SV, Biederman J, Weiffenbach B, et al. Dopamine D4 gene 7-repeat allele and attention deficit hyperactivity disorder. *Am J Psychiatry* 1999;156:768–770.

120. LaHoste GJ, Swanson JM, Wigal SB, et al. Dopamine D4 receptor gene polymorphism is associated with attention deficit hyperactivity disorder. *Mol Psychiatry* 1996;1:121–124.

121. Rowe DC, Stever C, Giedinghagen LN, et al. Dopamine DRD4 receptor polymorphism and attention deficit hyperactivity disorder [see Comments]. *Mol Psychiatry* 1998;3:419–426.

122. Smalley SL, Bailey JN, Palmer CG, et al. Evidence that the dopamine D4 receptor is a susceptibility gene in attention deficit hyperactivity disorder [see Comments]. *Mol Psychiatry* 1998;3:427–430.

123. Swanson JM, Sunohara GA, Kennedy JL, et al. Association of the dopamine receptor D4 (DRD4) gene with a refined phenotype of attention deficit hyperactivity disorder (ADHD): a family-based approach. *Mol Psychiatry* 1998;3:38–41.

124. Asherson P, Virdee V, Curran S, et al. Association of DSM-IV attention deficit hyperactivity disorder and monoamine pathway genes. *Am J Med Genet* 1998;81:548.

125. Castellanos FX, Lau E, Tayebi N, et al. Lack of an association between a dopamine-4 receptor polymorphism and attention-deficit/hyperactivity disorder: genetic and brain morphometric analyses [see Comments]. *Mol Psychiatry* 1998;3:431–434.

126. Daly G, Hawi Z, Fitzgerald M, et al. Attention deficit hyperactivity disorder: association with the dopamine transporter (DAT1) but not with the dopamine D4 receptor (DRD4). *Am J Med Genet* 1998;81:501.

127. Eisenberg J, Zohar A, Mei-Tal G, et al. A halotype relative risk study of the dopamine D4 receptor (DRD4) exon III repeat polymorphism and attention deficit hyperactivity disorder (ADHD). *Am J Med Genet* 2000;96:258–261.

128. Hawi Z, McCarron M, Kirley A, et al. No association of dopamine DRD4 receptor (DRD4) gene polymorphism in attention deficit hyperactivity disorder (ADHD) in the Irish population. *Am J Med Genet* 2000;96:268–272.

129. Asghari V, Sanyal S, Buchwaldt S, et al. Modulation of intracellular cyclic AMP levels by different human dopamine D4 receptor variants. *J Neurochem* 1995;65:1157–1165.

130. Ebstein RP, Novick O, Umansky R, et al. Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of novelty seeking. *Nat Genet* 1996;12:78–80.

131. Benjamin J, Patterson C, Greenberg BD, et al. Population and familial association between the D4 dopamine receptor gene and measures of novelty seeking. *Nat Genet* 1996;12:81–84.

132. Lanau F, Zenner M, Civelli O, et al. Epinephrine and norepinephrine act as potent agonists at the recombinant human dopamine D4 receptor. *J Neurochem* 1997;68:804–812.

133. Rubinstein M, Phillips TJ, Bunzow JR, et al. Mice lacking dopamine D4 receptors are supersensitive to ethanol, cocaine, and methamphetamine. *Cell* 1997;90:991–1001.

134. Dulawa SC, Grandy DK, Low MJ, et al. Dopamine D4 receptor-knock-out mice exhibit reduced exploration of novel stimuli. *J Neurosci* 1999;19:9550–9556.

135. Cook EH, Stein MA, Krasowski MD, et al. Association of attention deficit disorder and the dopamine transporter gene. *Am J Med Genet* 1995;56:993–998.

136. Gill M, Daly G, Heron S, et al. Confirmation of assocation between attention deficit hyperactivity disorder and a dopamine transporter polymorphism. *Mol Psychiatry* 1997;2:311–313.

137. Waldman ID, Rowe DC, Abramowitz A, et al. Association and linkage of the dopamine transporter gene and attention-deficit hyperactivity disorder in children: heterogeneity owing to diagnostic subtype and severity. *Am J Med Genet* 1998;63:1767–1776.

138. Poulton K, Holmes J, Hever T, et al. A molecular genetic study of hyperkinetic disorder/attention deficit hyperactivity disorder. *Am J Med Genet* 1998;81:458.

139. Winsberg BG, Comings DE. Association of the dopamine transporter gene (DAT1) with poor methylphenidate response [see Comments]. *J Am Acad Child Adolesc Psychiatry* 1999;38:1474–1477.

140. Giros B, Jaber M, Jones SR, et al. Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. *Nature* 1996;379:606–612.

141. Gainetdinov RR, Jones SR, Fumagalli F, et al. Re-evaluation of the role of the dopamine transporter in dopamine system homeostasis. *Brain Res Brain Res Rev* 1998;26:148–153.

142. Jaber M, Dumartin B, Sagne C, et al. Differential regulation of tyrosine hydroxylase in the basal ganglia of mice lacking the dopamine transporter. *Eur J Neurosci* 1999;11:3499–3511.

143. Jones SR, Gainetdinov RR, Hu XT, et al. Loss of autoreceptor functions in mice lacking the dopamine transporter. *Natl Neurosci* 1999;2:649–655.

144. Bezard E, Gross CE, Fournier MC, et al. Absence of MPTP-induced neuronal death in mice lacking the dopamine transporter. *Exp Neurol* 1999;155:268–273.

145. Gainetdinov RR, Fumagalli F, Jones SR, et al. Dopamine transporter is required for *in vivo* MPTP neurotoxicity: evidence from mice lacking the transporter. *J Neurochem* 1997;69:1322–1325.

146. Comings DE, Comings BG, Muhleman D, et al. The dopamine D2 receptor locus as a modifying gene in neuropsychiatric disorders. *JAMA* 1991;266:1793–1800.

147. Baik JH, Picetti R, Saiardi A, et al. Parkinsonian-like locomotor impairment in mice lacking dopamine D2 receptors. *Nature* 1995;377:424–428.

148. Kelly MA, Rubinstein M, Phillips TJ, et al. Locomotor activity in D2 dopamine receptor-deficient mice is determined by gene dosage, genetic background, and developmental adaptations. *J Neurosci* 1998;18:3470–3479.

149. Dickinson SD, Sabeti J, Larson GA, et al. Dopamine D2 receptor-deficient mice exhibit decreased dopamine transporter function but no changes in dopamine release in dorsal striatum. *J Neurochem* 1999;72:148–156.

150. Calabresi P, Saiardi A, Pisani A, et al. Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. *J Neurosci* 1997;17:4536–4544.

151. Barr CL, Wigg KG, Wu J, et al. Linkage study of two polymorphisms at the dopamine D3 receptor gene and attention-deficit hyperactivity disorder. *Am J Med Genet* 2000;96:114–117.

152. Accili D, Fishburn CS, Drago J, et al. A targeted mutation of the D3 dopamine receptor gene is associated with hyperactivity in mice. *Proc Natl Acad Sci USA* 1996;93:1945–1949.

153. Eisenberg J, Mei-Tal G, Steinberg A, et al. Haplotype relative risk study of catechol-O-methyltransferase (COMT) and attention deficit hyperactivity disorder (ADHD): association of the high-enzyme activity Val allele with ADHD impulsive-hyperactive phenotype. *Am J Med Genet* 1999;88:497–502.

154. Hawi Z, Millar N, Daly G, et al. No association between catechol-O-methyltransferase (COMT) gene polymorphism and attention deficit hyperactivity disorder (ADHD) in an Irish sample. *Am J Med Genet* 2000;96:282–284.

155. Tahir E, Curran S, Yazgan Y, et al. No association between low and high activity catecholamine-methyl-transferase (COMT) and attention deficit hyperactivity disorder (ADHD) in a sample of Turkish children. *Am J Med Genet* 2000;96:285–288.

156. Barr CL, Wigg K, Malone M, et al. Linkage study of catechol-O-methyltransferase and attention-deficit hyperactivity disorder. *Am J Med Genet* 1999;88:710–713.

157. Jiang S, Xin R, Wu X, et al. Association between attention deficit disorder and the DXS7 locus. *Am J Med Genet* 2000;96:289–292.

158. Comings D, Gade-Andavolu R, Gonzalez N, et al. Additive effect of three noradenergic genes (ADRA2A, ADRA2C, DBH) on attention-defecit hyperactivity disorder and learning disabilities an Tourette syndrome subjects. *Clin Genet* 1999;55:160–172.

159. Comings D, Gade R, Muhleman D, et al. No association of a tyrosine hydroxylase gene tetranucleotide repeat polymorphism in autism, Tourette syndrome, or ADHD. *Biol Psychiatry* 1995;37:484–486.

160. Wilson MC. Coloboma mouse mutant as an animal model of hyperkinesis and attention deficit hyperactivity disorder. *Neurosci Biobehav Rev* 2000;24:51–57.

161. Hess EJ, Jinnah HA, Kozak CA, et al. Spontaneous locomotor hyperactivity in a mouse mutant with a deletion including the Snap gene on chromosome 2. *J Neurosci* 1992;12:2865–2874.

162. Raber J, Mehta PP, Kreifeldt M, et al. Coloboma hyperactive mutant mice exhibit regional and transmitter-specific deficits in neurotransmission. *J Neurochem* 1997;68:176–186.

163. Conners CK. *Food additives and hyperactive children*. New York: Plenum, 1980.

164. Wolraich M, Wilson D, White W. The effect of sugar on behavior or cognition in children. *JAMA* 1995;274:1617–1621.

165. Needleman HL. The neuropsychiatric implications of low level exposure to lead. *Psychol Med* 1982;12:461–463.

166. Conners CK. Controlled trial of methylphenidate in preschool children with minimal brain dysfunction. *Int J Ment Health* 1975;4:61–74.

167. Milberger S, Biederman J, Faraone S, et al. Pregnancy delivery and infancy complications and ADHD: issues of gene-environment interactions. *Biol Psychiatry* 1997;41:65–75.

168. Sprich-Buckminster S, Biederman J, Milberger S, et al. Are perinatal complications relevant to the manifestation of ADD? Issues of comorbidity and familiality. *J Am Acad Child Adolesc Psychiatry* 1993;32:1032–1037.

169. Schmidt MH, Esser G, Allehoff W, et al. Evaluating the significance of minimal brain dysfunction: results of an epidemiologic study. *J Child Psychol Psychiatry* 1987;28:803–821.

170. Hartsough CS, Lambert NM. Medical factors in hyperactive and normal children: prenatal, developmental, and health history findings. *Am J Orthopsychiatry* 1985;55:191–201.

171. Nichols PL, Chen TC. *Minimal brain dysfunction: a prospective study*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1981.

172. Chandola C, Robling M, Peters T, et al. Pre-and perinatal factors and the risk of subsequent referral for hyperactivity. *Child Psychol Psychiatry* 1992;33:1077–1090.

173. Kline J, Stein Z, Susser M. *Conception to birth: epidemiology of prenatal development*. New York: Oxford University Press, 1989.

174. Landesman-Dwyer S, Emmanuel I. Smoking during pregnancy. *Teratology* 1979;19:119.

175. Fielding JE. Smoking: health effects and control. *N Engl J Med* 1985;313:491–498.

176. Denson R, Nanson J, McWatters J. Hyperkinesis and maternal smoking. *Can Psychiatr Assoc J* 1975;20:183–187.

177. Eriksson P, Ankarberg E, Fredriksson A. Exposure to nicotine during a defined period in neonatal life induces permanent changes in brain nicotinic receptors and in behaviour of adult mice. *Brain Res* 2000;853:41–48.

178. Hagino N, Lee J. Effect of maternal nicotine on the development of sites for [^{3}H] nicotine binding in the fetal brain. *Int J Dev Neurosci* 1985;3:567–571.

179. Marks MJ, Pauly JR, Gross SD, et al. Nicotine binding and nicotinic receptor subunit RNA after chronic nicotine treatment. *J Neurosci* 1992;12:2765–2784.

180. Marks MJ, Grady SR, Collins AC. Downregulation of nicotinic receptor function after chronic nicotine infusion. *J Pharmacol Exp Ther* 1993;266:1268–1276.

181. Slotkin TA, Lappi SE, Seidler FJ. Impact of fetal nicotine exposure on development of rat brain regions: critical sensitive periods or effects of withdrawal? *Brain Res Bull* 1993;31:319–328.

182. Bradbury T, Miller GA. Season of birth in schizophrenia: a review of evidence, methodology, and etiology. *Psychol Bull* 1985;98:569–594.

183. Mouridsen SE, Nielsen S, Rich B, et al. Season of birth in infantile autism and other types of childhood psychoses. *Child Psychiatry Hum Dev* 1994;25:31–43.

184. Livingston R, Adam B, Bracha H. Season of birth and neurodevelopmental disorders: summer birth is associated with dyslexia. *J Am Acad Child Adolesc Psychiatry* 1993;32:612–616.

185. Mick E, Biederman J, Faraone SV. Is season of birth a risk factor for attention deficit hyperactivity disorder? *J Am Acad Child Adolesc Psychiatry* 1996;35:1470–1476.

186. Rutter M, Cox A, Tupling C, et al. Attainment and adjustment in two geographical areas. I. The prevalence of psychiatric disorders. *Br J Psychiatry* 1975;126:493–509.

187. Blanz B, Schmidt MH, Esser G. Familial adversities and child psychiatric disorders. *J Child Psychol Psychiatr Disord* 1991;32:939–950.

188. Biederman J, Milberger S, Faraone SV, et al. Family-environment risk factors for attention deficit hyperactivity disorder: a test of Rutter's indicators of adversity. *Arch Gen Psychiatry* 1995;52:464–470.

189. Offord DR, Boyle MH, Racine YA, et al. Outcome, prognosis and risk in a longitudinal follow-up study. *J Am Acad Child Adolesc Psychiatry* 1992;31:916–923.

190. Palfrey JS, Levine MD, Walker DK, et al. The emergence of attention deficits in early childhood: a prospective study. *Dev Behav Pediatr* 1985;6:339–348.

191. Biederman J, Milberger S, Faraone SV, et al. Impact of exposure to parental psychopathology and conflict on adaptive functioning and comorbidity in children with attention deficit hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 1995;34:1495–1503.

192. Institute of Medicine. *Research on children and adolescents with mental, behavioral and developmental disorders*. Washington, DC: National Academy Press, 1989.

193. Hetherington EM, Cox M, Cox R. Effects of divorce on parents and children. In: Lamb M, ed. *Non-traditional families*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1982:223–285.

194. Reid WJ, Crisafulli A. Marital discord and child behavior problems: a meta-analysis. *J Abnorm Child Psychol* 1990;18:105–117.

195. Barkley RA, Fischer M, Edlebrock C, et al. The adolescent outcome of hyperactive children diagnosed by research criteria. III. Mother-child interactions, family conflicts and maternal psychopathology. *J Child Psychol Psychiatry* 1991;32:233–255.

196. Wyman PA, Cowen EL, Work WC, et al. Interviews with children who experienced major life stress: family and child attributes that predict resilient outcomes. *J Am Acad Child Adolesc Psychiatry* 1992;31:904–911.

197. Gelfand DM, Teti DM. The effects of maternal depression on children. *Clin Psychol Rev* 1990;10:329–353.

198. Richters JE. Depressed mothers as informants about their children: a critical review of the evidence for distortion. *Psychol Bull* 1992;112:485–499.

199. Shaywitz BA, Shaywitz SE, Byrne T, et al. Attention deficit disorder: quantitative analysis of CT. *Neurology* 1983;33:1500–1503.

200. Nasrallah HA, Loney J, Olson SC, et al. Cortical atrophy in young adults with a history of hyperactivity in childhood. *Psychiatry Res* 1986;17:241–246.

201. Lou H, Henriksen L, Bruhn P. Focal cerebral hypoperfusion in children with dysphasia and/or attention defeicit disorder. *Arch Neurol* 1984;41:825–829.

202. Hynd GW, Semrud-Clikeman MS, Lorys AR, et al. Brain morphology in developmental dyslexia and attention deficit/hyperactivity. *Arch Neurol* 1990;47:919–926.

203. Hynd GW, Semrud-Clikeman M, Lorys AR, et al. Corpus callosum morphology in attention-deficit hyperactivity disorder: morphometric analysis of MRI. *J Learn Disabil* 1991;24:141–146.

204. Aylward EH, Reiss AL, Reader MJ, et al. Basal ganglia volumes in children with attention-deficit hyperactivity disorder. *J Child Neurol* 1996;11:112–115.

205. Singer HS, Reiss AL, Brown JE. Volumetric MRI changes in basal ganglia of children with Tourette's syndrome. *Neurology* 1993;43:950–956.

206. Baumgardner TL, Singer HS, Denckla MB, et al. Corpus callosum morphology in children with Tourette syndrome and attention deficit hyperactivity disorder. *Neurology* 1996;47:1–6.

207. Semrud-Clikeman MS, Filipek PA, Biederman J, et al. Attention-deficit hyperactivity disorder: magnetic resonance imaging morphometric analysis of the corpus callosum. *J Am Acad Child Adolesc Psychiatry* 1994;33:875–881.

208. Castellanos F, Giedd J, Marsh W, et al. Quantitative brain magnetic resonance imaging in attention deficit hyperactivity disorder. *Arch Gen Psychiatry* 1996;53(July):607–616.

209. Mostofsky S, Reiss AL, Lockhart P, et al. Evaluation of cerebellar size in attention-deficit hyperactivity disorder. *J Child Neurol* 1998;13:434–439.

210. Overmeyer S, Simmons A, Santosh J, et al. Corpus callosum may be similar in children with ADHD and siblings of children with ADHD. *Dev Med Child Neurol* 2000;42:8–13.

211. Mataro M, Garcia-Sanchez C, Junque C, et al. Magnetic resonance imaging measurement of the caudate nucleus in adolescents with attention-deficit hyperactivity disorder and its relationship with neuropsychological and behavioral measures. *Arch Neurol* 1997;54:963–968.

212. Kayl AE, Moore BD 3rd, Slopis JM, et al. Quantitative morphology of the corpus callosum in children with neurofibromatosis and attention-deficit hyperactivity disorder. *J Child Neurol* 2000;15:90–96.

213. Berquin PC, Giedd JN, Jacobsen LK, et al. Cerebellum in attention-deficit hyperactivity disorder: a morphometric MRI study. *Neurology* 1998;50:1087–1093.

214. Casey B, Castellanos X, Giedd J, et al. Implication of right frontostriatal circuitry in response inhibition and attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 1997;36:374–383.

215. Filipek PA, Semrud-Clikeman M, Steingrad R, et al. Volumetric MRI analysis: comparing subjects having attention-deficit hyperactivity disorder with normal controls. *Neurology* 1997;48:589–601.

216. Lou HC, Henriksen L, Bruhn P, et al. Striatal dysfunction in attention deficit and hyperkinetic disorder. *Arch Neurol* 1989;46:48–52.

217. Lou H, Henriksen L, Bruhn P. Focal cerebral dysfunction in developmental learning disabilities. *Lancet* 1990;335:8–11.

218. Amen D, Carmichael B. High-resolution brain SPECT imaging in ADHD. *Ann Clin Psychiatry* 1997;9:81–86.

219. Rubia K, Overmeyer S, Taylor E, et al. Hypofrontality in attention deficit hyperactivity disorder during higher-order motor control: a study with functional MRI. *Am J Psychiatry* 1999;156:891–896.

220. Schweitzer JB, Faber TL, Grafton ST, et al. Alterations in the functional anatomy of working memory in adult attention deficit hyperactivity disorder. *Am J Psychiatry* 2000;157:278–280.

221. Silberstein RB, Farrow M, Levy F, et al. Functional brain electrical activity mapping in boys with attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry* 1998;55:1105–1112.

222. Vaidya C, Austin G, Kirkorian G, et al. Selective effects of methylphenidate in attention deficit hyperactivity disorder: a functional magnetic resonance study. *Proc Natl Acad Sci USA* 1998;95:14494–14499.

223. Ernst M, Zametkin AJ, Matochik JA, et al. High midbrain [^{18}F]DOPA accumulation in children with attention deficit hyperactivity disorder. *Am J Psychiatry* 1999;156:1209–1215.

44

Learning Disorders

C. Keith Conners

Ann C. Schulte

C. Keith Conners: Behavioral Neurology Department, Durham, North Carolina.

Ann C. Schulte: Department of Psychology, North Carolina State University, Raleigh, North Carolina.

NOSOLOGY AND CLASSIFICATION

Part of "44 - Learning Disorders "

Current conceptualizations of *learning disorders* (LDs), formerly referred to as “academic skills disorders” (1), follow the traditional approach of classifying learning by specific academic skills. These skills include reading, mathematics, and written expression. In each case, the skills are measured by standardized tests whose scores must fall substantially below the level expected from chronologic age, intelligence, and age-appropriate education. The deficits must significantly interfere with academic or daily living activities requiring the skills. When LDs result from sensory, medical, or neurologic conditions, they are coded on Axis III (medical conditions) within the DSM-IV nomenclature.

Commonly associated features of LDs include low selfesteem and demoralization, social skills deficits, school dropout, and difficulties in employment or social adjustment. Patients with conduct disorder, oppositional disorder, attention-deficit/hyperactivity disorder (ADHD), major depression, dysthymic disorder, and Tourette syndrome all have substantially elevated rates of LD. Academic skills in pervasive developmental disorders are often not discrepant from the measured intelligence and language abilities associated with the pervasive development disorder. Communication disorders and motor skills disorders are also common in LDs, including expressive language disorders, phonologic disorder, and stuttering. Spelling disorders are usually not considered separate from other reading- and writing-related deficits.

Although this approach to classification of LDs is useful in a practical context and allows for an operational definition for detection and remediation, it has several drawbacks from a theoretic and scientific point of view. Reading, mathematics, and writing comprise many processing skills, giving rise to subtypes with different underlying mechanisms. Thus, reading at the word level may involve visual, lexical, or semantic processes (2 ,3), with correspondingly different neuroanatomic circuitry and computational mechanisms within the brain. There are subtypes characterized both by the pattern of skills deficits (e.g., reading and spelling, but no mathematics disorder) and by different patterns of neuropsychological function, such as the relative strength of verbal and nonverbal factors on intelligence tests (4). There are also important developmental changes in LD, such that variables characterizing the disorder at earlier ages may be different from those seen in older patients (5). Advances in the genetics and neuroimaging of LDs will depend on more homogeneous clinical definitions at the symptomatic level (6).

PREVALENCE

Part of "44 - Learning Disorders "

The DSM-IV reports prevalence estimates of 2% to 10% for LDs, depending on the nature of ascertainment and the definitions applied (1). In most prevalence studies, a diagnosis of LD has been made on the basis of a significant discrepancy between IQ and achievement in one or more areas (7), with studies varying in terms of the manner in which a discrepancy has been determined and the cutoff score for considering a discrepancy “severe.” One study of the prevalence of regression-based ability/achievement discrepancies using the co-norming sample from the Wechsler Intelligence Scale for Children III and the Wechsler Individual Achievement Scales found that 17% of the norming group had ability/achievement discrepancies at the .05 significance level in one or more areas of achievement (8). This figure can probably be considered the upper limit for LD prevalence estimates based on ability/achievement discrepancies, given that a diagnosis of LD would also require determining both that the discrepancy was not the result of poor instruction and that it interfered with daily functioning.

Several researchers have questioned the conceptual and empiric basis for the use of ability/achievement discrepancies in the diagnosis of LDs, as well as current operationalizations of the exclusionary criteria. Reasons for concern on

the use of ability/achievement discrepancies are (a) findings that the cognitive profiles of children with low achievement are similar regardless of whether they evidence an ability/achievement discrepancy (9), (b) findings that the same deficits that lead to poor achievement may also lower IQ (10), and (c) the finding that the use of such definitions prevents early identification and treatment because the underlying cognitive deficits that cause the disability must retard growth in academic skills before intervention can begin (11). Alternate proposals for identification include simply using a low achievement criterion (e.g., academic functioning 1 to 2 standard deviations [SD] below the mean), using a definition that combines the ability/achievement discrepancy and low achievement approaches, and use of domain-specific rather than general cognitive ability tests as predictors of academic achievement. In general, these alternate procedures are likely to raise prevalence rates.

There is some indication that more rigorous operationalization of the exclusionary criteria in the LD definition could substantially reduce LD prevalence rates. For example, when Vellutino and his colleagues used daily tutoring as a “first cut” diagnostic criterion to distinguish between children who had reading difficulties caused by cognitive deficits and those whose deficits were the result of poor instruction, they found that two thirds of their sample scored within the average range in reading (thirtieth percentile and higher) after one semester of one-to-one tutoring (12). This relatively stringent criterion for establishing an “adequate educational environment” resulted in a drop in the prevalence rate of reading disorders (RDs) from 9% to 3%. Geary used failure to respond to short-term intensive remedial instruction as a diagnostic criterion for mathematics disability (MD) and noted a marked drop in prevalence (13). Clearly, the definition of caseness in these studies has implications for how phenotypes are characterized in genetic and neurobiological investigations. The use of the more conservative methods of case definition are clearly more costly for selecting subjects, but they may prove more valid and useful in finding biological markers of LD.

COMORBIDITY

Part of "44 - Learning Disorders "

Many psychiatric and medical conditions include LD as an associated deficit. The most common childhood condition comorbid with LD is ADHD. Estimates of comorbidity range from 20% to 90%, with the lower figures appearing in epidemiologic samples and the higher figures appearing in clinically referred samples. The high degree of overlap in clinical samples suggests that common mechanisms may be at work in the neurologic basis for both disorders. LDs were once considered a necessary criterion for minimal brain dysfunction. Although some studies suggest that ADHD may simply be the result of an LD, most studies indicate that when both conditions are present, characteristics of each are found, whereas in LDs alone, only symptoms of LD, not those of ADHD, are present, and vice versa (14). The high degree of overlap has the practical implication that when one disorder is identified, it is always prudent to expect the presence of the other and to make appropriate diagnostic probes.

PHONOLOGIC PROCESSING

Part of "44 - Learning Disorders "

As noted earlier, the present classification approach to LD subdivides the disorder on the basis of impairment in specific academic areas (reading, math, written expression). However, given that performance in each of these area draws on numerous cognitive processes, it is likely that the present classification system will eventually be replaced by one that focuses on the specific cognitive deficits that underlie poor academic performance and their impact on the development and execution of specific subskills within and across academic areas.

The greatest progress in specifying the cognitive and neuropsychological dysfunctions underlying LDs has occurred in reading. Numerous investigations using longitudinal, intervention, genetic, and neuroimaging methods have produced strong and converging evidence that deficits in phonologic processing are the proximal cause of reading difficulties in a large proportion of children with RDs (see ref. 15 and ref. 16 for reviews). Deficits in phonologic processing also appear to affect spelling, written expression, and mathematics.

Phonologic processing refers to the ability to use and manipulate the sound structure of one’s oral language (17). Although conceptualizations of phonologic processing and its components vary, within the Wagner and Torgesen model of phonologic processing, it consists of three related abilities: phonologic awareness, phonologic memory, and rapid naming (18 ,19). *Phonologic awareness* refers to the understanding that words can be broken down into phonemes and the ability to identify phonemes and manipulate them in words (16). Phonemes are the smallest sound unit that changes the meaning of a word (e.g., tap and lap differ by one phoneme). Phonologic awareness is a critical ability in learning to read because it allows beginning readers to link letters and letter combinations in text to sound strings in oral language (20). Knowledge of these links allows readers to discover the regularities in written text so written words can be rapidly translated into their spoken equivalents. Such recoding allows the reader to access the semantic code (or meaning) for the letter string. The repeated pairing of the visual letter string and its spoken equivalent is thought eventually to allow the reader to develop direct visual word recognition strategies that bypass the phonologic code (10 ,21).

Phonologic memory is an individual’s ability to represent verbal information in working memory in terms of a sequence

of sounds or a phonetic code. When children have difficulty with phonologic coding, reading acquisition is impaired because of difficulty in performing the rapid comparison and blending needed to identify unfamiliar written words. Difficulties in verbal short-term memory are also hypothesized to be a major factor underlying MDs (22), and they may affect the acquisition of foreign languages (20).

Rapid naming is the ability to access phonologic information that is stored in long-term memory rapidly. It is typically assessed by asking children to name well-known items as rapidly as possible (e.g., presentation of a series of colored squares with the child naming the color of each square as fast as possible). Such tasks are thought to tap many of the same cognitive processes required in skilled reading, such as rapid scanning, sequencing and processing of serially presented visual stimuli, and rapid access to strings of phonemes (e.g., color names) (16). There is debate about whether the difficulty with rapid naming tasks observed in many children with RDs is a reflection of a core deficit in phonologic processing or whether it represents a deficit in a second set of processes that impairs reading. If this is the case, there may be “double-deficit” readers who are impaired in both phonologic and rapid naming processes (23). Such disabled readers would be less responsive to interventions that address phonologic processing and would require additional interventions targeted toward increasing language and reading fluency.

Efforts are also under way to understand more fully the core cognitive deficits underlying other types of LD. For example, Berninger et al. proposed a model of the cognitive processes underlying written language and writing disabilities (24). Geary proposed that there are three subtypes of MDs, with corresponding deficits in semantic memory, procedural knowledge of mathematics, and visuospatial processes (22).

GENETICS

Part of "44 - Learning Disorders "

It has been known for decades that LDs run in families. In the 1990s, family aggregation studies, twin studies, and genetic linkage analyses confirmed the strong hereditary influences on RD and MD (Table 44.1). The genetic studies also confirm the heterogeneity of the phenotype, with both orthographic and phonologic traits implicated but not having identical sources of genetic influence. A genetic link between RD and MD was confirmed in several studies. A strong link of Tourette syndrome, ADHD, and LD has been suggested by studies of patients who have Tourette syndrome with and without ADHD. Evidence has accumulated that locations on the short arm of chromosome 6

(6p21.3) and the short arm of chromosome 15 are involved. Odds for linkage to chromosome 15 are reported as being 1,000 to 1, with evidence that 30% of an extended series of families showed linkage to chromosome 15 polymorphisms. Some variations in results may reflect sampling methods or trait markers. The excess of affected males with LDs identified in clinic and referred samples disappears in research-based samples (25).

Study	Subjects	Method	Comment
Comings and Comings, 1987 (99)	47 normal controls, 246 TS	Comparison of TS and control	27% of TS had LD vs 4.2% of controls
Comings and Comings, 1990 (100)	130 TS probands with 1,851 relatives, 25 control probands with 541 relatives	Comparison of TS and control	Suggests LD/ADHD are integral part of the expression of the Gts gene(s)
Comings et al., 1999 (101)	274 TS and 62 normal controls	Tested associations and additive effects between polymorphisms at 3 nonadrenergic gene sites	Suggests additive effects of nonadrenergic genes related to presence of LD
DeFries et al., 1987 (102)	64 pairs identical, 55 pairs fraternal twins in which at least one member of pair is dyslexic	Multiple regression analysis	Significant genetic etiology for dyslexia
Fagerheim et al., 1999 (103)	80 Norwegian family members	Genome search for linkage and non-parametric multipoint GENEHUNTER analysis	Localization to 2p15–16 and to 6p21.3–23 give strong evidence of genetic heterogeneity in dyslexia
Field and Kaplan, 1998 (104)	79 families having at least 2 affected sibs with phonologic coding dyslexia (617 genotyped, 294 affected)	Tested for linkage	No evidence for linkage by LOD score analysis or affected-sib-pair methods; however, affected-pedigree-member (APM) method detects significant linkage; concludes APM may generate false-positive results
Fisher et al., 1999 (105)	181 sib pairs from 82 nuclear families with a dyslexic proband	Assessed linkage directly for several quantitative measures rather than a single composite measure or categoric definition	Pointwise analysis of sib-pair trait differences suggests presence in 6p21.3 of a QTL influencing multiple components of dyslexia: reading of irregular words and nonwords; shows that both orthographic and phonological skills are affected
Gayan and Olson, 1999 (106)	—	Review with focus on twin study design and sib-pair linkage techniques	DeFries-Fulker multiple regression analyses show significant estimates of heritability for group deficits on several reading and language measures, and presence of significant common and independent genetic effects on individual differences on reading skills; linkage techniques confirm a candidate locus for RD on chromosome 6
Gayan et al., 1999 (107)	126 sib pairs	Multipoint mapping method and 8 informative DNA markers on chromosome 6	Significant linkage across a distance of at least 5 cM for deficits in orthographic (LOD = 3.10) and phonological (LOD = 2.42) skills, confirming previous findings
Gillis et al., 1992 (108)	264 RD twin pairs and 182 matched control twin pairs	Multivariate behavior genetic analysis	Individual differences in both reading and math performance are highly heritable and appear to be caused by many of the same genetic influences
Knopik et al., 1997 (109)	102 identical and 77 same-sex fraternal twin pairs in which at least one member of each pair is reading disabled; and 42 identical and 23 same-sex fraternal twin pairs in which at least one member is math disabled	Multiple regression model for the analysis of selected twin data and its bivariate extension	The comorbidity between math and RD is due in part to genetic influences
Knopik and DeFries, 1999 (110)	526 twin pairs selected for RD (290 identical and 236 same-sex fraternal); and 355 control pairs (220 identical and 135 same sex fraternal)	Confirmatory factor analyses and heritability estimation	Heritability in proband and controls were 0.81 and 0.69; and those for math 0.88 and 0.67; genetic influences accounted for 83% of the covariation between reading and math factors in the proband group and 58% in the control group; shared environmental influences did not contribute to the relationship between reading and math factors, nor to their independent variation
Petryshen et al., 2000 (111)	79 families with at least 2 affected sibs	Two-point and multipoint quantitative-trait sib-pair linkage and variance-components analyses	No evidence for a locus in the 6p23-p21.3 region for several quantitative measures; speculates that perhaps families with subtypes of dyslexia linked to this region are underrepresented in the sample, either by chance or ascertainment criteria
Reynolds et al., 1996 (112)	Twins of the Virginia Twin Study	—	69% of variability in oral reading due to heredity vs 13% due to shared environmental effects; genetic and environmental influences were equivalent for males and females, but males showed greater phenotypic variability than females

ADHD, attention deficit hyperactivity disorder; LD, learning disorder; QTL, quantitative trait locus; RD, reading disorder; TS, Tourette syndrome.

TABLE 44.1. GENETICS OF LEARNING DISORDERS

NEUROIMAGING

Part of "44 - Learning Disorders "

The neuroanatomic and functional pathways in the brain involved in LDs were greatly clarified in the 1990s by a variety of neuroimaging techniques. Reviews of neuroimaging of LDs describe rapid progress in identifying the brain regions and functional pathways involved (20 ,26 ,27 ,28 and 29). However, these reviews also call attention to discrepancies in findings, possibly the result of small cohorts, variations in sampling, and heterogeneity of the LDs. Table 44.2 provides selected studies from several hundred investigations, mainly of RDs. Many studies confirm earlier findings of abnormalities of the microstructure of the planum temporale from autopsy studies, but conflicting data emerge, possibly related to the method employed or the sampling techniques and definition of the RD (30). Although most studies implicate abnormalities in left temporal-parietal anatomic areas, additional findings have identified white matter, right hemisphere anomalies, motor cortex, cingulate gyrus, and the splenium of the corpus callosum.

TABLE 44.2. NEUROIMAGING IN LEARNING DISORDERS

Study	Method	Subjects	Results
Klingberg et al., 2000 (113)	Diffusion tensor magnetic resonance	Adults with poor or normal reading ability	Subjects with reading difficulty exhibited decreased diffusion anisotropy bilaterally in temporoparietal white matter. White matter diffusion anisotropy in the temporoparietal region of the left hemisphere was significantly correlated with reading scores within the reading-impaired adults and within the control group. The anisotropy reflects microstructure of white matter tracts, which may contribute to reading ability by determining the strength of communication between cortical areas involved in visual, auditory, and language processing.
Fersten et al., 1999 (114)	Blood flow velocity in MCA in left and right hemisphere measured with the transcranial Doppler method	10 dysgraphic or dysorthographic students and 10 normal subjects	The dysgraphic persons had significantly higher blood flow velocity in the right hemisphere compared to the reference group.
Duncan et al., 1994 (115)	Event-related brain potentials	13 severely dyslexic men, 15 matched controls	As task demands increased, visual P300 was reduced in the dyslexic men as compared with the normal readers. Dyslexics with a history of many symptoms of ADHD in childhood (high ADHD) accounted for the group differences in P300; the dyslexics with a history of few or no such symptoms (low ADHD) were indistinguishable from the controls at all electrode sites. The results are interpreted as suggesting that a distinct brain organization may characterize dyslexic men with a history of concomitant deficits in attention.
Georgiewa et al., 1999 (116)	fMRI	17 phonologically impaired developmental dyslexics and 17 normal reading children	Significant differences in Broca area and the left inferior temporal region for both, nonword reading and the phonologic transformation task.
Rumsey et al., 1999 (117)	rCBF	17 right-handed dyslexic men, ages 18–40, and 14 matched controls	Correlations between reading skill and rCBF during a series of reading tasks; uniquely identified the left angular gyrus as the most probable site of a functional lesion in dyslexia: Here, higher rCBF was associated with better reading skill in controls ($p < .01$), but with worse reading skill in dyslexia ($p < .01$).
Helenius et al., 1999 (118)	Magnetoencephalography	10 dyslexic male adults and 10 normal controls	Early visual responses were similar in dyslexic and nonimpaired readers. In contrast, the letter-string–specific responses peaking around 150 ms predominantly in the left inferior occipitotemporal cortex in fluent readers were undetectable in dyslexic readers. Thus, while the early visual processing seems intact in dyslexic adults, the pattern of cortical activation starts to differ from that of fluent readers at the point where letter-string–specific signals first emerge during reading.
Best and Demb, 1999 (32)	Sagittal magnetic resonance images of PT and magnocellular visual pathway	Dyslexics with documented MC deficits and controls	Dyslexic subjects did not deviate from normal leftward PT asymmetry, but both groups became less left-lateralized with methods that excluded sulcul tissue. Results suggest that dyslexic subjects with a magnocellular deficit do not always have abnormal symmetry of the PT. PT symmetry may instead be related to a different subtype of dyslexia. In addition, PT asymmetry in any subject group depends on the measurement method.
Nicolson et al., 1999 (119)	PET	6 dyslexic adults and 6 matched controls	Brain activation was significantly lower ($P < .01$) for the dyslexic adults than for the controls in the right cerebellar cortex and the left cingulate gyrus when executing a prelearned motor sequence, and in the right cerebellar cortex when learning the new sequence.
Pennington et al., 1999 (120)	MRI	75 subjects with RD and 22 controls	Insula and anterior superior neocortex were smaller and the retrocallosal cortex was larger in the RD group. In contrast, no group main or interaction effects for the subcortical or callosal structures. Results were not due to ADHD.
Green et al., 1999 (121)	MRI-based surface reconstruction technique that models the curvature of the cerebral cortex in three dimensions to obtain whole-hemisphere and regional surface area estimates	8 male right-handed male dyslexics and matched controls	The caudal infrasylvian surface that encompasses the supratemporal plane and the inferior bank of the posterior ascending ramus of the sylvian fissure was significantly larger than that of control subjects, and this result was not attributable to a difference in whole-hemisphere surface area.

Richards et al., 1999 (122)	MR spectroscopic imaging technique called proton echo-planar spectroscopic imaging	6 dyslexic boys and 7 age- and IQ-matched right-handed good readers	Dyslexic boys showed a greater area of brain lactate elevation (2.33+/-SE 0.843 voxels) as compared with the control group during a phonological task in the left anterior quadrant. No significant differences were observed in the nonlanguage tasks.
Price et al., 1998 (31)	fMRI	2 boys with deep dyslexia	Activation patterns primarily reflect semantic and phonologic systems in spared regions of the left hemisphere. These results preclude an explanation of deep dyslexia in terms of purely right-hemisphere word processing.
Demb et al., 1998 (123)	fMRI	Group of dyslexic and normal readers	Dyslexics showed reduced brain activity compared with controls both in primary visual cortex (VI) and in several extrastriate areas, including area MT and adjacent motion-sensitive areas (MT+) that are believed to receive a predominant magnocellular pathway input.
McPherson et al., 1998 (124)	Event-related potentials	Adolescents who were good phonetic decoders or poor (dysphonetic)	Phonetics showed both orthographic and phonological priming but had a marked reduction in their CNV. These results support the separation of the reading disabled into a group that has difficulty translating orthography into phonology and a group that is slower functioning and has reduced capacity in preparing for a response.
Shaywitz et al., 1998 (125)	fMRI	Dyslexic and normal readers	Brain activation patterns differed significantly between the groups with dyslexic readers showing relative underactivation in posterior regions (Wernicke area, the angular gyrus, and striate cortex) and relative overactivation in an anterior region (inferior frontal gyrus). These results support a conclusion that the impairment in dyslexia is phonologic and that these brain activation patterns may provide a neural signature for this impairment.
Richardson et al., 1997 (126)	*In vivo* cerebral phosphorus-31 magnetic resonance spectroscopy	12 dyslexic and 10 nondyslexic adults	Membrane phospholipid metabolism is abnormal in dyslexia.
Halperin et al., 1997 (127)	Plasma levels of MHPG	ADHD children with and without RD	Plasma levels of MHPG were significantly lower in ADHD children without RD, compared with those with RD, replicating a published finding.

Rumsey et al., 1997 (30)	MRI	16 RH dyslexic men 18–40 and 14 matched controls	Results challenge the notion that anomalous asymmetry of the PT is strongly associated with developmental dyslexia. Given the heterogeneity of the dyslexic population, some subgroup of dyslexic individuals (i.e., those with developmental language disorders) may show unusual symmetry or reversed asymmetry in this region. However, anomalous asymmetry of the planum did not contribute to functional abnormalities demonstrated in these patients by positron emission tomography.
Rumsey et al., 1996 (128)	MRI	21 dyslexic men and 19 matched controls	As predicted, the area of the posterior third of the corpus callosum, roughly equivalent to the isthmus and splenium, was larger in dyslexic men than in controls. No differences were seen in the anterior or middle corpus callosum. The increased area of the posterior corpus callosum may reflect anatomical variation associated with deficient lateralization of function in posterior language regions of the cortex and their right-sided homologues, hypothesized to differ in patients with dyslexia.
Paulesu et al., 1996 (129)	PET	5 adult dyslexics with phonological processing deficits	Proposes that the defective phonologic system of these dyslexics is due to weak connectivity between anterior and posterior language areas. This could be due to a dysfunctional left insula which may normally act as an anatomic bridge among Broca area, superior temporal, and inferior parietal cortex. The independent activation of the posterior and anterior speech areas in dyslexics supports the notion that representations of unsegmented and segmented phonology are functionally and anatomically separate.
Eden et al., 1996 (130)	Review	—	The pathophysiology of developmental dyslexia is more complex than originally thought, extending beyond the classically defined language areas of the brain.
Shapleske et al., 1999 (131)	Review	PT studies	Overall, there is a significant leftward asymmetry in normals, which is reduced in left handers and females. The leftward asymmetry is much reduced in patients with schizophrenia due to a relatively larger right PT than normal controls.
Deb et al., 1997 (26)	Review	—	Brain abnormalities can be detected in cases of idiopathic and nonidiopathic learning disability, but their significance is not clear due to discrepancies in study findings and the small cohorts involved.

ADHD, attention deficit hyperactivity disorder; fMRI, functional magnetic resonance imaging; MCA, middle cerebral artery; MHPG, XXX; MT, XXX; PET, positron emission tomography; PT, planum temporale; rCBF, regional cerebral blood flow; RD, reading disorder.

One of the older controversies regarding the functional brain basis of dyslexia is whether dyslexia represents a visual (orthographic) disorder or a language-based (phonologic processing) disorder. Neuroimaging studies now appear to provide evidence that brain structures involving both the

striate visual magnocellular pathways and specific phonologic processing pathways in the left hemisphere are involved in dyslexia, a finding possibly reflecting different subtypes at the behavioral level. As noted earlier, cognitive behavioral analysis suggests that distinctive mechanisms for visual, lexical, and semantic processing are required to explain normal human reading (2 ,3). Pathologic studies indicate that each of these mechanisms can be affected separately in acquired dyslexias. For example, in *deep dyslexia*, it is primarily the semantic aspects of reading that are disturbed, whereas orthography and lexicality are preserved. Thus, a patient may read “spirit” as “whiskey,” or “church” as “priest.” Evidence suggests that, unlike the more typical left-hemisphere–based phonologic and visual deficits in dyslexia, deep dyslexia may reflect a right-hemisphere—based processing mechanism (31).

Whereas some investigators interpret functional magnetic resonance imaging studies as giving strong support to the hypothesis that dyslexia represents a disorder of the language system, involving the segmentation and synthesis of phonemes (20), others find evidence that magnocellular pathways without involvement of phonologic regions occur in dyslexia (32 ,33). As noted by Filipek, cognitive neuroscience identifies specific computational tasks that should be used to provide more homogeneous samples at the behavioral level for further advances in the neurobiology of developmental disorders (28). For example, rather than using classic clinical criteria for dyslexia, which leads to samples with diverse subtypes, neuroimaging studies may do better to select samples by visual, lexical, and semantic criteria first.

EDUCATIONAL MANAGEMENT

Part of "44 - Learning Disorders "

Various educational treatments have been developed for LD. In general, the most effective treatment approach is one that involves careful delineation of the specific academic deficits evidenced by the child and intensive instruction in the skill areas in which deficiencies are documented (34). Response to treatment varies by individuals, so it is important that careful monitoring take place throughout treatment to ensure that an intervention is effective for a particular child (35). In this section, we briefly summarize the educational treatment literature by academic area and then summarize research related to treatment monitoring or formative evaluation of interventions.

Reading

Considerable progress has been made in the development of preventive and early intervention approaches for beginning readers. Several studies have demonstrated that explicit instruction in phonologic awareness (generally combined with letter identification and reading instruction) in preschool and early elementary years can reduce the overall rate of RDs (36 ,37) and can improve outcomes for children who are at high risk of RD (38 ,39). One metaanalysis reported a combined effect size for phonologic awareness training of 1.16 for phonologic awareness skills and .40 for reading skills across studies that used samples of normal readers and .54 and .60 for studies that used samples of students who were either at risk of, or had shown evidence of, reading difficulty (40).

The difference in training effect on phonologic awareness between normal and impaired readers appears to reflect the difficulty many poor readers have in mastering phonologic processing, even when they are provided with intensive instruction to address these difficulties (40). Torgesen examined results from five large-scale early reading intervention studies and concluded that even with use of the best current methods of early reading remediation, 2% to 6% of children would still evidence inadequate reading skills in the early elementary grades (41). Such findings point to the need for the development of even more powerful intervention techniques to facilitate the acquisition of early reading skills.

Current models of reading skill acquisition characterize phonologic awareness as a necessary, but not sufficient condition for the development of skilled reading (15). Fluent reading requires the development of orthographic reading skills or the ability to recognize words by sight (41). Impaired readers generally show deficits in this area that persist into adulthood (41 ,42). Interventions to improve fluency are less well developed than interventions for the development of decoding skills (i.e., phonologic awareness interventions). The *repeated readings technique*, which involves multiple readings of the same passages, is the most researched approach to improving fluency (43), and it has shown limited but positive effects on fluency (44). The increased attention to issues of fluency in reading research has resulted in the development of new, comprehensive intervention approaches that ultimately may be more effective than existing techniques in addressing fluency deficits (23). At present, however, fluency deficits remain one of the most persistent and intransigent symptoms of RD (20).

Although most children with RDs show deficits in word recognition skills, comprehension deficits are also common. These may occur alone or in the presence of impaired word recognition skills (45). When impaired word recognition is the primary source of the comprehension deficit, decoding and fluency interventions such as those discussed earlier can improve reading comprehension (46). However, interventions have also been developed to address comprehension deficits directly. Two metaanalyses found substantial improvements for disabled readers who receive intensive instruction in reading comprehension (47 ,48). In both studies, metacognitive approaches (e.g., self-questioning, comprehension monitoring) produced the largest effect sizes.

Math

Geary characterized research in the area of MDs as “primitive” in comparison with studies of RDs (22). Nevertheless, effective remediation techniques for MDs have been developed. Mastropieri et al. presented a comprehensive review of mathematics instructional techniques that have been effective for students with LDs (49). However, Cawley et al. questioned the efficacy of available math computation instructional techniques (50). In a metaanalysis of math word problem interventions, Xin and Jitendra found that instruction in problem representation was an effective remedial strategy for addressing this type of difficulty in children with a range of mild disabilities (51). These investigators also found that problem representation instruction was most effective when it was presented in a computer-assisted format. Long-term interventions (i.e., more than 1 month) resulted in better maintenance and generalization of training.

Written Expression

Difficulties with composition and writing fluency are common in children with LDs. Several researchers have shown that cognitive strategy instruction is effective in improving the composition skills of children with written language deficits (52 ,53 and 54). Generally, such interventions provide students with explicit instruction in thinking and problem-solving strategies that allow them to break down the complex task of composing written text into manageable substeps.

Difficulties with handwriting fluency appear not only to impair the speed with which children can take notes or copy but also to affect compositional fluency and quality (55). For example, Berninger et al. found that instruction in handwriting increased students’ scores on a writing composition test (56).

With more widespread use of computers in classrooms, word processing tools are increasingly being used to address the writing problems of children with LDs (57). When writing fluency is a problem, word processing may be used as a text entry strategy on its own, or it can be combined with word prediction programs (58). Voice recognition software has improved to the point that it may be a practical text entry strategy for many students with writing disabilities (59). However, research on the efficacy of these tools remains sparse. In one of the few studies to compare the efficacy of different word processing strategies for improving writing fluency, accuracy, and composition in students with LDs, Lewis et al. compared groups of students after a year of writing instruction using either keyboarding, keyboarding with word prediction software, or keyboarding with word prediction and synthesized speech software (60). All groups using word processing tools showed *decreases* in speed of text entry over handwriting, although the keyboarding with text prediction group showed the smallest decrease. There were no improvements in composition skills in any of the treatment groups.

Treatment Monitoring

The unexpected results of the foregoing study by Lewis et al. reinforce the need to monitor response to treatment and to verify that interventions for children with LDs achieve their intended results. *Curriculum-based measurement* (CBM) is a relatively new development in special education and provides a useful tool for continuous monitoring of children’s response to treatment in a number of academic areas (61 ,62 ,63 and 64). CBM involves the collection of brief samples of students’ performance on basic skills on a weekly or monthly basis. For example, CBM procedures in reading involve the administration of short reading probes (e.g., passages of 200 words) to children once or twice per week. The number of correct responses per passage is charted, and slope is then used as an indicator of a child’s response to treatment. Slopes that do not differ from zero are an obvious indicator of the need for a new treatment approach. However, estimates of typical response to treatment for students with LDs are also available and can be used as a basis for deciding whether a given treatment is producing sufficient progress (65). When formative evaluation strategies such as CBM are incorporated into treatment strategies, outcomes for students with disabilities improve markedly (66).

PSYCHOPHARMACOLOGY

Part of "44 - Learning Disorders "

Psychostimulants

Early studies of psychostimulants in children with LDs suggested strong immediate effects in enhancing reading, spelling, and arithmetic as well as in laboratory measures of learning (67 ,68 ,69 ,70 and 71). However, reviews concluded that lasting educational gains resulting from psychotropic drugs have not been demonstrated (72 ,73). Stimulant drug effects have generally been dose related, with linear increases in performance with higher doses (74 ,75 ,76 and 77). Drug-induced changes reflect increased output, accuracy, efficiency, and improved learning acquisition. There is also evidence of increased effort and self-correcting behaviors (78). Some studies suggest a positive effect of stimulants on memory consolidation that is not accounted for by concomitant effects on acquisition (79). Because most studies involve students with comorbid ADHD, measures of specific effects of stimulants on LDs are rare. However, because improvement in learning acquisition occurs in both clinical cases and neurologically normal persons treated with amphetamine (80), it seems likely that stimulant effects on learning are nonspecific with respect to diagnosis.

Stimulants have been widely used in rehabilitation of memory and LDs in brain injuries and encephalopathies

secondary to medical X-irradiation of the brain. Animal models of selective exposure to X-irradiation during infancy show enhanced learning from amphetamine treatment (81).

Nootropics

Piracetam (Nootropil, Nootropyl, 2-oxo-l-pyrrolidone acetamide) was originally developed as a molecular analogue of γ-aminobutyric acid (GABA) for the purpose of altering vestibular function in motion sickness, but it is probably neither a GABA receptor agonist nor antagonist. Numerous analogues of the piracetam molecule are currently under study, including oxiracetam (Neuromet), pramiracetam, etiracetam, nefiracetam, aniracetam, and rolziracetam. This group of nootropics is commonly referred to as the "racetams." Piracetam has virtually no detectable peripheral effects at any dose in animals or humans and does not affect cerebral blood flow, unlike other putative cerebral enhancers. It appears to alter cellular brain metabolism, however, because it increases the concentration ratio of brain adenosine triphosphate. In neurologically normal volunteers, a single dose of piracetam was found to change brain global functional state as measured by multichannel electroencephalographic recordings (82). Investigators have suggested that the defining characteristics of nootropics should include lack of peripheral effect, absence of action on blood flow, and an increase in brain metabolism (83).

Animal research indicates that memory deficits induced by epileptogenic kindling procedures are prevented by pretreatment with piracetam (84). Piracetam (100 mg/kg, IP) and oxiracetam (10 mg/kg, IP) prevented the negative effects of microwaves on memory processes in exposed rats (85). Hypobaric hypoxia of pregnant rats is followed by the reduction of weight gain of the newborn pups, delayed impairment of memory (passive and active tasks), and changes of extrapolative water escape. Piracetam (200 mg/kg/d) administered at early postnatal period (from the eighth to the twentieth day of life) corrected behavioral disturbances and physical development in rats (86). Piracetam (800 mg/kg) administered orally once daily for 5 days before training completely antagonized the scopolamine-provoked amnesia in step-through–trained mice, and piracetam (600 mg/kg) administered orally once daily for 5 days before training abolished the memory-impairing effect of clonidine in shuttle-box–trained rats and the amnestic effect of methergoline in step-down–trained rats.

Early studies by Dimond and Brouwers suggested that piracetam could facilitate transfer of information across the callosal pathways and hence is a "superconnector" drug (87). Numerous studies with neurologically normal and dyslexic adults indicated that the drug could enhance verbal learning. These studies were reviewed by Wilshire (88). An early report on reading involved 16 dyslexic men matched with 14 student volunteers for a 21-day trial of piracetam. It was found, using a double-blind crossover technique, that the dyslexic men significantly increased their verbal learning by approximately double that of control students (89). Early uncontrolled trials with a broader group of LDs were followed by a series of systematic studies of learning, memory, and reading (90).

Studies of 60 dyslexic boys 8 to 14 years old, who were carefully selected for exclusion of intellectual, sensory, psychiatric, and neurologic impairment and educational deprivation, were conducted to determine the efficacy of piracetam, over a 12-week period, in improving reading and other related skills (91). There were no changes at the end of 12 weeks to distinguish the groups in accuracy or comprehension of prose reading. Short-term memory gains, however, were recorded for the treated group on two different tests, digit span, and a test (Neimark) of immediate and delayed recall. The mean digit span scaled score for the entire group was 1 SD below their mean IQ. Considering only the performance of children whose digit span scaled scores were 1 SD or below the mean (7 or less), the treated group made a significant gain at the end of 12 weeks. On the Neimark test, the treated group was significantly superior to the untreated group on first trial learning, and they also lost significantly fewer object names after a delay. Improved retrieval from long-term storage could be demonstrated for the treated group on the rapid automatized naming test. Although there was no significant difference between the groups at screening, the treated group was significantly faster on letter naming at the end of the drug trial. The treated group also improved their single word reading on the Wide Range Achievement Test (WRAT).

After previous research suggested that piracetam improves performance on tasks associated with the left hemisphere, a 12-week, double-blind, placebo-controlled study of developmental dyslexics was conducted. Six study sites treated 257 dyslexic boys between the ages of 8 and 13 years who were significantly below their potential in reading performance. The children were of at least normal intelligence, had normal findings on audiologic, ophthalmologic, neurologic, and physical examination, and were neither educationally deprived nor emotionally disturbed. Piracetam was found to be well tolerated in this study population. Children treated with piracetam showed improvements in reading speed. No other effects on reading were observed. In addition, improvement in auditory sequential short-term memory was observed in those piracetam-treated patients who showed relatively poor memory at baseline (92).

Piracetam was given in a 3,300-mg daily dose to half of a group of 55 dyslexic boys aged 8 to 13 years, in a 12-week, double-blind, placebo-controlled study. The other half of the subjects received placebo. Compared with the placebo control group, the boys treated with piracetam did not show statistically significant improvements above their baseline scores on measures of perception, memory, language, reading accuracy or comprehension, or writing accuracy. However, reading speed and numbers of words written in a timed

period were significantly enhanced in subjects treated with piracetam as compared with placebo. Effective reading and writing ability, taking both rate and accuracy into consideration, were also significantly improved in the piracetam group as compared with the placebo treatment group (93).

Two hundred twenty-five dyslexic children between the ages of 7 years 6 months and 12 years 11 months whose reading skills were significantly below their intellectual capacity were enrolled in a multicenter, 36-week, double-blind, placebo-controlled study. Piracetam-treated children showed significant improvements in reading ability (Gray Oral Reading Test) and reading comprehension (Gilmore Oral Reading Test). Treatment effects were evident after 12 weeks and were sustained for the total period (36 weeks) (94).

The neurophysiologic mechanisms involved in the effects of piracetam were examined in studies using event-related potentials. Eight- to 12-year-old dyslexic boys were randomly assigned to 3.3 g of piracetam or matching placebo per day in two divided doses over a 12-week period. Children performed a vigilance task in which they pressed a key when two alphabetic letters or shapes occurred in sequence. Event-related potentials to letters and shapes, for active and passive responses, were recorded at the vertex and left and right parietal areas of the scalp. Performance measures included letter and form hits, misses, commission errors, and reaction times. Piracetam increased the amplitude of a late positive component (believed to correspond to P300) at the vertex for letter hits. Piracetam also increased the latency of this component in both hemispheres, but only for active responses (letter hits) in the left hemisphere and passive responses (correct rejections and misses) in the right hemisphere. Reaction time to letter hits was significantly correlated with the latency of the P300 component, a finding suggesting that letters created increased effort or attentional demand on the subjects compared with forms. An early event-related potential component (P225) also showed increased amplitude to piracetam in both hemispheres, and effects were limited to form hits. These effects were thought possibly to reflect slow negative potentials arising from stimulus anticipation in the CNV-like paradigm. The results were cautiously interpreted as indicating a facilitation of verbal processing mechanisms responsible for analyzing the verbal significance of visual stimuli (95).

In a subsequent study, 29 dyslexic children (aged 7 to 12 years) were assigned to piracetam or matching placebo for 36 weeks. Event-related potentials were obtained at the end of treatment from a vigilance paradigm that required a response to letter or form matches. The drug group showed a significant advantage in letter hits compared with placebo and a reduced variance in reaction time. The drug increased the amplitude of three factors from a principal components analysis of event-related potentials and was interpreted as increasing a processing negativity when stimuli were letters. Piracetam was interpreted as enhancing feature analysis and increasing attentional resources among dyslexic children when the stimuli are recognized as having linguistic significance (96). These effects were shown to be dose related in a subsequent study (97).

One negative study examined the interaction of piracetam and tutoring (98). Sixty children with dyslexia (41 boys, 19 girls; ages 9 to 13 years) were enrolled in a 10-week summer tutoring program that emphasized word-building skills. They were randomly and blindly assigned to receive either placebo or piracetam. The children were subtyped as “dysphonetic” or “phonetic” on the basis of scores from tests of phonologic sensitivity and phoneme-grapheme correspondence skills. Of the 53 children who completed the program, 37 were classified as dysphonetic and 16 as phonetic. The phonetic group improved significantly more in word-recognition ability than the dysphonetic group. Overall, the children taking medication did not improve more than the nonmedicated children in any aspect of reading. However, within the medication-treated group, the phonetic subgroup gained most in word recognition.

SUMMARY AND CONCLUSIONS

Part of "44 - Learning Disorders "

Significant difficulties continue to bedevil the definition of LD, including problems surrounding various criteria such as an IQ/learning discrepancy or low absolute achievement level. Approaches that define the disorder by resistance to high-quality instruction may be the most valid for purposes of identifying persons with LDs in genetic, neuroimaging, and pharmacologic studies. The high degree of comorbidity with many psychiatric disorders raises further issues for studies requiring a homogeneous symptom pattern, and it seems likely that further advances will require replacing broad clinical patterns with more specific processing deficits based on cognitive neuroscience. Despite these limitations, existing research is encouraging regarding the possibility of precise genetic and neuroanatomic localization of LDs, particularly for RDs. Again, however, subtyping issues at the phenotypic level require elucidation before further progress is likely.

Much of the pharmacologic work has been confounded by the comorbidity of LD with ADHD and other childhood disorders. Evidence generally supports the finding that psychostimulants (e.g., dextroamphetamine and methylphenidate) have positive effects on immediate learning *performance* but less impact on long-term academic gains. Work with nootropic drugs shows intriguing effects on verbal learning, single-word reading, and left-hemisphere processing of alphabetic stimuli. Good controlled trials indicate that piracetam may be a safe and effective enhancer of reading in school-aged children, with gains double the rate expected in seriously impaired readers. LD remains a large public health problem, is significantly undertreated, has devastating lifetime outcomes, and therefore merits greater

research efforts to understand its neurobiology and treatment needs.

REFERENCES

1. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington, DC: American Psychiatric Association, 1994.

2. Seymour PHK. Individual cognitive analysis of competent and impaired reading. *Br J Psychol* 1987;78:483–506.

3. Seymour PHK, MacGregor CJ. Developmental dyslexia: a cognitive experimental analysis of phonological, morphemic, and visual impairments. *Cogn Neuropsychol* 1984;1:43–82.

4. Rourke BP, ed. *Neuropsychology of learning disabilities: essentials of subtype analysis*. New York: Guilford, 1985.

5. Rourke BP. Reading and spelling disabilities: a developmental neuropsychologcial perspective. In: Kirk U, ed. *Neuropsychology of language, reading, and spelling*. New York: Academic, 1983:209–234.

6. Filipek PA. Neurobiologic correlates of developmental dyslexia: how do dyslexics' brains differ from those of normal readers? *J Child Neurol* 1995;10:S62–S69.

7. Shaywitz SE, Shaywitz BA, Fletcher JM, et al. Prevalence of reading disability in boys and girls: results of the Connecticut Longitudinal Study. *JAMA* 1990;264:998–1002.

8. Ward TJ, Ward SB, Glutting JJ, et al. Exceptional LD profile types for the WISC-III and WIAT. *School Psychol Rev* 1999;28:629–643.

9. Fletcher JM, Shaywitz SE, Shankweiler DP, et al. Cognitive profiles of reading disability: comparisons of discrepancy and low achievement definitions. *J Educ Psychol* 1994;86:1–18.

10. Stanovich KE. Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read Res Q* 1986;21:360–406.

11. Fletcher JM, Francis DJ, Shaywitz SE, et al. Intelligent testing and the discrepancy model for children with learning disabilities. *Learn Disabil Res Pract* 1998;13:186–203.

12. Vellutino FR, Scanlon DM, Sipay ER, et al. Cognitive profiles of difficult-to-remediate and readily remediated poor readers: early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *J Educ Psychol* 1996;88:601–638.

13. Geary DC. *Children's mathematical development: research and practical applications*. Washington, DC: American Psychological Association, 1994.

14. Fletcher JM, Shaywitz SE, Shaywitz BA. Comorbidity of learning and attention disorders: separate but equal. *Pediatr Clin North Am* 1999;46:885–897.

15. Smith SB, Simmons DC, Kameenui EJ. Phonological awareness: research bases. In: Simmons DC, Kameenui EJ, eds. *What reading research tells us about children with diverse learning needs*. Mahwah, NJ: Erlbaum, 1998:61–127.

16. Torgesen JK, Wagner RK. Alternative diagnostic approaches for specific developmental reading disabilities. *Learn Disabil Res Pract* 1998;13:220–232.

17. Wagner RK, Torgesen JK, Rashotte CA. *Comprehensive test of phonological processing: examiner's manual*. Austin, TX: PRO-ED, 1999.

18. Wagner RK, McBride-Chang C. The development of reading-related phonological processes. *Ann Child Dev* 1996;12:177–206.

19. Wagner RK, Torgesen JK. The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol Bull* 1987;101:192–212.

20. Shaywitz BA, Pugh KR., Fletcher JM, et al. What cognitive and neurobiological studies have taught us about dyslexia. In: Greenhill LL, ed. *Learning disabilities: implications for psychiatric treatment*, vol 19. Washington, DC: American Psychiatric Press, 2000:59–95.

21. Posner MI, Raichle ME. *Images of mind*. New York: WH Freeman, 1994.

22. Geary DC. Mathematical disabilities: cognitive, neuropsychological, and genetic components. *Psychol Bull* 1993;114:345–362.

23. Wolf M, Bowers PG. The double-deficit hypothesis for the developmental dyslexias. *J Educ Psychol* 1999;91:415–438.

24. Berninger V, Mizokawa DT, Bragg R. Theory-based diagnosis and remediation of writing disabilities. *J School Psychol* 1991;29:57–79.

25. Wadsworth SJ, DeFries JC, Stevenson J, et al. Gender ratios among reading-disabled children and their siblings as a function of parental impairment. *J Child Psychol Psychiatry* 1992;33:1229–1239.

26. Deb S. Structural neuroimaging in learning disability [see Comments]. *Br J Psychiatry* 1997;171:417–419.

27. Eden GF, Zeffiro TA. Neural systems affected in developmental dyslexia revealed by functional neuroimaging. *Neuron* 1998;21:279–282.

28. Filipek PA. Neuroimaging in the developmental disorders: the state of the science. *J Child Psychol Psychiatry* 1999;40:113–128.

29. Shapleske J, Rossell SL, Woodruff PW, et al. The planum temporale: a systematic, quantitative review of its structural, functional and clinical significance. *Brain Res Brain Res Rev* 1999;29:26–49.

30. Rumsey JM, Donohue BC, Brady DR, et al. A magnetic resonance imaging study of planum temporale asymmetry in men with developmental dyslexia. *Arch Neurol* 1997;54:1481–1489.

31. Price CJ, Howard D, Patterson K, et al. A functional neuroimaging description of two deep dyslexic patients. *J Cogn Neurosci* 1998;10:303–315.

32. Best M, Demb JB. Normal planum temporale asymmetry in dyslexics with a magnocellular pathway deficit. *Neuroreport* 1999;10:607–612.

33. Demb JB, Boynton GM, Heeger DJ. Brain activity in visual cortex predicts individual differences in reading performance. *Proc Natl Acad Sci USA* 1997;94:13363–13366.

34. Kavale DA, Forness SR. *Efficacy of special education*. Washington, DC: American Association on Mental Retardation, 1999.

35. Daly EJ, Martens BK, Hamler KR, et al. A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. *J Appl Behav Anal* 1999;32:83–94.

36. Ball EW, Blachman BA. Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Read Res Q* 1991;24:49–66.

37. Brady S, Fowler A, Stone B, et al. Training phonological awareness: a study with inner-city kindergarten children. *Ann Dyslexia* 1994;44:26–51.

38. Felton RH. Effects of instruction on the decoding skills of children with phonological processing problems. *J Learn Disabil* 1993;26:583–589.

39. Foorman BR, Francis DJ, Fletcher JM, et al. The role of instruction in learning to read: preventing reading failure in at-risk children. *J Educ Psychol* 1998;90:37–55.

40. Bus AG, van IJzendoorn MH. Phonological awareness and early reading: a meta-analysis of experimental training studies. *J Educ Psychol* 1999;91:403–414.

41. Torgesen JK. Individual differences in response to early interventions in reading: the lingering problem of treatment resisters. *Learn Disabil Res Pract* 2000;15:55–64.

42. Manis FR, Custodio R, Szeszulski PA. Development of phonological and orthographic skill: a 2-year longitudinal study of dyslexic children. *J Exp Child Psychol* 1993;56:64–86.

43. Meyer MS, Felton RH. Repeated reading to enhance fluency: old approaches and new directions. *Ann Dyslexia* 1999;49:283–306.

44. Mastropieri MA, Leinart A, Scruggs TE. Strategies to increase reading fluency. *Intervent School Clin* 1999;34:278–283, 292.

45. Gough PB, Tumner WE. Decoding, reading, and reading disability. *Remedial Spec Educ* 1986;7:6–10.

46. Torgesen JK, Wagner RK, Rashotte CA, et al. Preventing reading failure in young children with phonological processing disabilities: group and individual responses to instruction. *J Educ Psychol* 1999:579–593.

47. Talbott E, Lloyd JW, Tankersley M. Effects of reading comprehension interventions for students with learning disabilities. *Learn Disabil Q* 1994;17:223–232.

48. Mastropieri MA, Scruggs TE, Bakken JP, et al. Reading comprehension: a synthesis of research in learning disabilities. In: Scruggs TE, Mastropieri MA, eds. *Advances in learning and behavioral disabilities*, vol 10. Greenwich, CT: JAI Press, 1996:277–303.

49. Mastropieri MA, Scruggs TE, Shiah S. Mathematics instruction for learning disabled students: a review of the research. *Learn Disabil Res Pract* 1991;6:89–98.

50. Cawley JF, Parmar RS, Yan W, et al. Arithmetic computation performance of students with learning disabilities: implications for curriculum. *Learn Disabil Res Pract* 1998;13:68–74.

51. Xin YP, Jitendra AK. The effects of instruction in solving mathematical word problems for students with learning disabilities: a meta-analysis. *J Spec Educ* 1999;32:207–225.

52. Englert CS, Raphael TE, Anderson LM, et al. Making strategies and self-talk visible: writing instruction in regular and special education classrooms. *Am Educ Res J* 1991;28:337–372.

53. Graham S, Harris KR Troia G. Writing and self-regulation: cases from the self-regulated strategy development model. In: Schunk D, Zimmerman B, eds. *Developing self-regulated learners: From teaching to self-reflective practice*. New York: Guilford, 1998:20–41.

54. La Paz S. Strategy instruction in planning: teaching students with learning and writing disabilities to compose persuasive and expository essays. *Learn Disabil Q* 1997;20:227–248.

55. Weintraub N, Graham S. Writing legibly and quickly: a study of children's ability to adjust their handwriting to meet common classroom demands. *Learn Disabil Res Pract* 1998;13:146–152.

56. Berninger V, Vaughn K, Abbott R, et al. Treatment of handwriting problems in beginning writers: from handwriting to composition. *J Educ Psychol* 1998;89:652–666.

57. La Paz S. Composing via dictation and speech recognition systems: compensatory technology for students with learning disabilities. *Learn Disabil Q* 1999;22:173–182.

58. Bos CS, Vaughn S. *Teaching students with learning and behavior problems*. Boston: Allyn & Bacon, 1998.

59. Gore A. Speak easy. *Macworld* 2000;17:23.

60. Lewis RB, Graves AW, Ashton TM, et al. Word processing tools for students with learning disabilities: a comparison of strategies to increase text entry speed. *Learn Disabil Res Pract* 1998;13:95–108.

61. Deno SL. Curriculum-based measurement: the emerging alternative. *Except Child* 1985;52:219–232.

62. Fuchs LS, Fuchs D. Combining performance assessment and curriculum-based measurement to strengthen instructional planning. *Learn Disabil Res Pract* 1996;11:183–192.

63. Fuchs LS, Hamlett CL, Fuchs D. *Monitoring basic skills progress [Computer programs]*. Austin, TX: PRO-ED, 1990.

64. Shinn MR, ed. *Curriculum-based measurement: assessing special children*. New York: Guilford, 1989.

65. Fuchs LS, Fuchs D, Hamlett CL, et al. Formative evaluation of academic progress: how much growth can we expect? *School Psychol Rev* 1993;22:27–48.

66. Fuchs LS, Fuchs D. Effects of formative evaluation on student achievement: a meta-analysis. *Except Child* 1986;53:199–208.

67. Conners CK, Eisenberg L. The effects of methylphenidate on symptomatology and learning in disturbed children. *Am J Psychiatry* 1963;120:458–464.

68. Conners CK, Eisenberg L, Barcai A. Effect of dextroamphetamine on children: studies on subjects with learning disabilities and school behavior problems. *Arch Gen Psychiatry* 1967;17:478–485.

69. Conners CK, Eisenberg L, Sharpe L. The effects of methylphenidate on paired-associate learning and porteous maze performance in emotionally disturbed children. *J Consult Psychol* 1964;28:14–22.

70. Conners CK, Rothschild G, Eisenberg L, et al. Dextroamphetamine sulfate in children with learning disorders: effects on perception, learning, and achievement. *Arch Gen Psychiatry* 1969;21:182–190.

71. Conners CK, Rothschild GH. Drugs and learning in children. In: Hellmuth J, ed. *Learning disorders*, vol 3. Seattle: Special Child Publications, 1968:193–223.

72. Aman MG. Psychotropic drugs and learning problem: a selective review. *J Learn Disabil* 1980;13:87–97.

73. Conners CK. Review of stimulant drugs in learning and behavior disorders. *Psychopharmacol Bull* 1971;7:39–40.

74. Rapport MD, Stoner G, DuPaul GJ, et al. Methylphenidate in hyperactive children: differential effects of dose on academic, learning, and social behavior. *J Abnorm Child Psychol* 1985;13:227–243.

75. Rapport MD, Kelly KL. Psychostimulant effects on learning and cognitive function: findings and implications for children with attention deficit hyperactivity disorder. *Clin Psychol Rev* 1991;11:61–92.

76. Rapport MD, Kelly KL. Psychostimulant effects on learning and cognitive function: findings and implications for children with attention deficit hyperactivity disorder. *Clin Psychol Rev* 1995;15:1–32.

77. Rapport MD, Quinn SO, DuPaul GJ, et al. Attention deficit disorder with hyperactivity and methylphenidate: the effects of dose and mastery level on children's learning performance. *J Abnorm Child Psychol* 1989;17:669–689.

78. Douglas VI, Barr RG, O'Neill ME, et al. Short term effects of methylphenidate on the cognitive, learning and academic performance of children with attention deficit disorder in the laboratory and the classroom. *J Child Psychol Psychiatry* 1986;27:191–211.

79. Evans RW, Gualtieri CT, Amara I. Methylphenidate and memory: dissociated effects in hyperactive children. Fourteenth Annual Meeting of the International Neuropsychological Society (1986: Denver, Colorado). *Psychopharmacology (Berl)* 1986;90:211–216.

80. Rapoport JL, Buchsbaum MS, Weingartner H, et al. Dextroamphetamine: its cognitive and behavioral effects in normal and hyperactive boys and normal men. *Arch Gen Psychiatry* 1980;37:933–943.

81. Highfield DA, Hu D, Amsel A. Alleviation of X-irradiation–based deficit in memory-based learning by D-amphetamine: suggestions for attention deficit-hyperactivity disorder. *Proc Natl Acad Sci USA* 1998;95:5785–5788.

82. Wackermann J, Lehmann D, Dvorak I, et al. Global dimensional complexity of multi-channel EEG indicates change of human brain functional state after a single dose of a nootropic drug. *Electroencephalogr Clin Neurophysiol* 1993;86:193–198.

83. Deleted in press.

84. Becker A, Grecksch G. Nootropic drugs have different effects on kindling-induced learning deficits in rats. *Pharmacol Res* 1995;32:115–122.

85. Krylov IN, Iasnetsov VV, Dukhanin AS, et al. Pharmacologic correction of learning and memory disorders induced by exposure to high-frequency electromagnetic radiation. *Biull Eksp Biol Med* 1993;115:260–262.

86. Trofimov SS, Ostrovskaia RU, Kravchenko EV, et al. Behavior disorders in rats exposed to intrauterine hypoxia, and their correction by postnatal treatment with piracetam. *Biull Eksp Biol Med* 1993;115:43–45.

87. Dimond S, Brouwers EYM. Improvement of human memory through the use of drugs. *Psychopharmacology (Berl)* 1976;49:307–309.

88. Deleted in press.

89. Wilsher C, Atkins G, Manfield P. Piracetam as an aid to learning in dyslexia: preliminary report. *Psychopharmacology (Berl)* 1979;65:107–109.

90. Simeon J, Waters B, Resnick M. Effects of piracetam in children with learning disorders. *Psychopharmacol Bull* 1980;16:65–66.

91. Helfgott E, Rudel RG, Kairam R. The effect of piracetam on short- and long-term verbal retrieval in dyslexic boys. *Int J Psychophysiol* 1986;4:53–61.

92. Di Ianni M, Wilsher CR, Blank MS, et al. The effects of piracetam in children with dyslexia. *J Clin Psychopharmacol* 1985;5:272–278.

93. Tallal P, Chase C, Russell G, et al. Evaluation of the efficacy of piracetam in treating information processing, reading and writing disorders in dyslexic children. *Int J Psychophysiol* 1986;4:41–52.

94. Wilsher CR, Bennett D, Chase CH, et al. Piracetam and dyslexia: effects on reading tests. *J Clin Psychopharmacol* 1987;7:230–237.

95. Conners CK, Blouin AG, Winglee M, et al. Piracetam and event-related potentials in dyslexic males. *Int J Psychophysiol* 1986;4:19–27.

96. Conners CK, Reader M, Reiss A, et al. The effects of piracetam upon visual event-related potentials in dyslexic children. *Psychophysiology* 1987;24:513–521.

97. Conners CK, Conrey J, Reader M, et al. A double-blind, placebo-controlled, randomized block design study of the effects of four doses of piracetam on visual event-related potentials in children with developmental reading disorder (Report to UCB, Brussels, Belgium, and to G.H. Besselaar Associates). Washington, DC: George Washington University Medical School Department of Psychiatry, 1987.

98. Ackerman PT, Dykman RA, Holloway C, et al. A trial of piracetam in two subgroups of students with dyslexia enrolled in summer tutoring. *J Learn Disabil* 1991;24:542–549.

99. Comings DE, Comings BG. A controlled study of Tourette syndrome. I. Attention-deficit disorder, learning disorders, and school problems. *Am J Hum Gen* 1987;41:701–741.

100. Comings DE, Comings BG. A controlled family history study of Tourette's syndrome. I. Attention-deficit hyperactivity disorder and learning disorders. *J Clin Psychiatry* 1990;51:275–280.

101. Comings DE, Gade-Andavolu R, Gonzalez N, et al. Additive effect of three noradrenergic genes (ADRA2a, ADRA2C, DBH) on attention-deficit hyperactivity disorder and learning disabilities in Tourette syndrome subjects. *Clin Genet* 1999;55:160–172.

102. DeFries JC, Fulker DW, LaBuda MC. Evidence for a genetic aetiology in reading disability of twins. *Nature* 1987;329:537–539.

103. Fagerheim T, Raeymaekers P, Tonnessen FE, et al. A new gene (DYX3) for dyslexia is located on chromosome 2. *J Med Genet* 1999;36:664–669.

104. Field LL, Kaplan BJ. Absence of linkage of phonological coding dyslexia to chromosome 6p23–p21.3 in a large family data set [published erratum appears in *Am J Hum Genet* 1999;64:334]. *Am J Hum Genet* 1998;63:1448–1456.

105. Fisher SE, Marlow AJ, Lamb J, et al. A quantitative-trait locus on chromosome 6p influences different aspects of developmental dyslexia. *Am J Hum Genet* 1999;64:146–156.

106. Gayan J, Olson RK. Reading disability: evidence for a genetic etiology. *Eur Child Adolesc Psychiatry* 1999;8:52–55.

107. Gayan J, Smith SD, Cherny SS, et al. Quantitative-trait locus for specific language and reading deficits on chromosome 6p. *Am J Hum Genet* 1999;64:157–164.

108. Gillis JJ, DeFries JC, Fulker DW. Confirmatory factor analysis of reading and mathematics performance: a twin study. *Acta Genet Med Gemellol* 1992;41:287–300.

109. Knopik VS, Alarcon M, DeFries JC. Comorbidity of mathematics and reading deficits: evidence for a genetic etiology. *Behav Genet* 1997;27:447–453.

110. Knopik VS, DeFries JC. Etiology of covariation between reading and mathematics performance: a twin study. *Twin Res* 1999;2:226–234.

111. Petryshen TL, Kaplan BJ, Liu MF, et al. Absence of significant linkage between phonological coding dyslexia and chromosome 6p23-21.3 as determined by use of quantitative-trait methods: confirmation of qualitative analyses. *Am J Hum Genet* 2000;66:708–714.

112. Reynolds CA, Hewitt JK, Erickson MT, et al. The genetics of children's oral reading performance. *J Child Psychol Psychiatry* 1996;37:425–434.

113. Klingberg T, Hedehus M, Temple E, et al. Microstructure of temporo-parietal white matter as a basis for reading ability: evidence from diffusion tensor magnetic resonance imaging [see Comments]. *Neuron* 2000;25:493–500.

114. Fersten E, Luczywek E, Zabolotny W, et al. Dynamics of blood flow velocity in middle cerebral arteries in dyslexic persons. *Neurol Neurochir Pol* 1999;33:1099–1108.

115. Duncan CC, Rumsey JM, Wilkniss SM, et al. Developmental dyslexia and attention dysfunction in adults: brain potential indices of information processing. *Psychophysiology* 1994;31:386–401.

116. Georgiewa P, Rzanny R, Hopf JM, et al. fMRI during wordprocessing in dyslexic and normal reading children. *Neuroreport* 1999;10:3459–3465.

117. Rumsey JM, Horwitz B, Donohue BC, et al. A functional lesion in developmental dyslexia: left angular gyral blood flow predicts severity. *Brain Lang* 1999;70:187–204.

118. Helenius P, Tarkiainen A, Cornelissen P, et al. Dissociation of normal feature analysis and deficient processing of letter-strings in dyslexic adults. *Cereb Cortex* 1999;9:476–483.

119. Nicolson RI, Fawcett AJ, Berry EL, et al. Association of abnormal cerebellar activation with motor learning difficulties in dyslexic adults. *Lancet* 1999;353:1662–1667.

120. Pennington BF, Filipek PA, Lefly D, et al. Brain morphometry in reading-disabled twins. *Neurology* 1999;53:723–729.

121. Green RL, Hutsler JJ, Loftus WC, et al. The caudal infrasylvian surface in dyslexia: novel magnetic resonance imaging-based findings. *Neurology* 1999;53:974–981.

122. Richards TL, Dager SR, Corina D, et al. Dyslexic children have abnormal brain lactate response to reading-related language tasks. *AJNR Am J Neuroradiol* 1999;20:1393–1398.

123. Demb JB, Boynton GM, Heeger DJ. Functional magnetic resonance imaging of early visual pathways in dyslexia. *J Neurosci* 1998;18:6939–6951.

124. McPherson WB, Ackerman PT, Holcomb PJ, et al. Eventrelated brain potentials elicited during phonological processing differentiate subgroups of reading disabled adolescents. *Brain Lang* 1998;62:163–185.

125. Shaywitz SE, Shaywitz BA, Pugh KR, et al. Functional disruption in the organization of the brain for reading in dyslexia. *Proc Natl Acad Sci USA* 1998;95:2636–2641.

126. Richardson AJ, Cox IJ, Sargentoni J, et al. Abnormal cerebral phospholipid metabolism in dyslexia indicated by phosphorus-31 magnetic resonance spectroscopy. *NMR Biomed* 1997;10:309–314.

127. Halperin JM, Newcorn JH, Koda VH, et al. Noradrenergic mechanisms in ADHD children with and without reading disabilities: a replication and extension. *J Am Acad Child Adolesc Psychiatry* 1997;36:1688–1697.

128. Rumsey JM, Casanova M, Mannheim GB, et al. Corpus callosum morphology, as measured with MRI, in dyslexic men. *Biol Psychiatry* 1996;39:769–775.

129. Paulesu E, Frith U, Snowling M, et al. Is developmental dyslexia a disconnection syndrome? Evidence from PET scanning. *Brain* 1996;119:143–157.

130. Eden GF, VanMeter JW, Rumsey JM, et al. The visual deficit theory of developmental dyslexia. *Neuroimage* 1996;4:S108–S117.

131. Shapleske J, Rossell SL, Woodruff PW, et al. The planum temporale: a systematic, quantitative review of its structural, functional and clinical significance. *Brain Res Brain Res Rev* 1999;29:26–49.

45

Psychosis in Childhood and Its Management

Paramjit T. Joshi

Kenneth E. Towbin

Paramjit T. Joshi: Division of Behavioral Medicine, Department of Psychiatry and Behavioral Sciences, George Washington University School of Medicine Children's National Medical Center, Washington, DC.

Kenneth E. Towbin: Complex Developmental Disorders Clinic, Department of Psychiatry and Behavioral Sciences, George Washington University School of Medicine Children's National Medical Center, Washington, DC.

The appearance of *psychotic symptoms in childhood*, albeit rare, is an important clinical entity. This importance extends beyond their clinical prevalence and has begun to influence our understanding of the principal psychotic conditions. The term *psychosis* is generally categoric and includes subgroups within it. It is clear that the peak onset of the most common psychotic disorders, schizophrenia and bipolar disorder, is in adolescence (1 ,2). This points directly toward developmental events in biological, social, and psychological domains of late childhood and adolescence that set the stage for activating psychotic disorders. However, in addition, it appears increasingly likely that certain early childhood characteristics and developmental deficits may presage psychosis and are related to the outcome of psychotic disorders.

For the purposes of this chapter, *psychosis* is defined as the presence of disruptions in thinking, accompanied by delusions or hallucinations, along with an alteration in the thought processes, termed a *thought disorder*. Delusions and hallucinations are considered to be positive psychotic symptoms. *Delusions* are fixed, false, idiosyncratic beliefs that the child cannot be deterred from, with logical reasoning, whereas *hallucinations* are percepts that arise in the absence of external sensory stimulus. Psychotic symptoms always encompass a broad range of conditions, but it is particularly so when they appear in children and adolescents. Psychotic symptoms in children present distinctive diagnostic and clinical challenges because of the powerful influences of immaturity and the moving target produced by development.

Although there may at one time have been confusion about whether children are capable of having psychotic symptoms, it is now certain that children, like adults, can and do experience psychoses (i.e., disruptions in the form of mental life). Children and adolescents experience the same range and types of psychotic symptoms as do adults. They can lose the connections between their thoughts (formal thought disorder) and have perceptions without external stimuli (hallucinations). The term psychosis as described by McHugh and Slavney (3) is intended simply to indicate that mental life has been disrupted in its capacities or forms, as a result of a process that generates new forms of psychological experience.

Psychotic symptoms can be considered as general or nonspecific phenomena emerging with different disorders and etiologic possibilities. Modern psychiatry eschews the misleading dichotomy of functional versus organic causes and recognizes that some of these disorders stem from known brain or metabolic disorders, whereas for other conditions the pathophysiologic sources have yet to be discovered. The interplay between environmental and biological forces is at work across the spectrum of these conditions. The psychiatrist who must determine whether a young patient suffers from a psychotic disorder faces a challenging array of possibilities, more extensive than when the patient is an adult. The influences of development, environment, and cognition are greater for young or developmentally immature patients than for adults. Nonbiological events are clearly more influential because, in most respects, children are more vulnerable to their surroundings. Immaturity makes children more susceptible to environmental stressors and cognitive distortions. Children routinely have intrusions of fantasy into ordinary mental life; determining when this becomes pathologic can be a matter of degree. Children learn and experiment with imitation, and they can acquire habits and strategies used by those around them. They have not developed the cognitive abilities that permit them to observe and compare their experiences in an objective manner. The range of normal functioning is greater in childhood, so the child's behavior may simply be a result of immaturity, rather than a deviation from a normal pathway.

With the advent of categoric classification of psychiatric disorders, the criteria for psychotic disorders have become more stringent, and the concepts have been defined more narrowly. When one examines a 5-year old child who claims that he is “superman and can fly,” the challenge is to determine whether the child has a delusion. Similarly, in a child who complains about hearing a voice telling her to “do bad things,” one must determine whether she is talking about her conscience or is experiencing auditory hallucinations. This must be distinguished from *make-believe* (e.g., having an imaginary friend). Children can describe this makebelieve phenomenon, and clinicians need to discern the differences as they work with children with symptoms of psychosis. Such characteristics are sought by the clinician in the child’s answers to particular questions. The task and challenge as child and adolescent psychiatrists are to ask the right questions, to differentiate delusions and hallucinations from other forms of thought, such as a vivid imagination in a young child.

HISTORY

Part of "45 - Psychosis in Childhood and Its Management "

Interest in childhood psychosis can be traced to the nineteenth century, when Maudsley first wrote a description of the “insanity of early life” in 1874 in his textbook, *Physiology and Pathology of Mind* (4). He took a developmental approach by noting that the mental faculty of children was not organized, and hence the insanity in children must be of the simplest kind, influenced more by “reason of bad descent or of baneful influences during uterine life.” However, De Sanctis may be credited first with setting out childhood schizophrenia as different from mental deficiency and from certain neurologic disorders, such as epilepsy or postinfectious encephalopathy (5). It was not until 1919, that Kraeplin introduced the concept of *dementia praecox* and noted its onset in late childhood and adolescence (6). Given the insidious onset of the disorder, Kraeplin cautiously suggested that 3.5% of patients with schizophrenia had the onset of their illness before the age of 10 years. This led to an increased interest in understanding the developmental aspects of psychosis. Historically, despite this early description of the syndrome by Kraeplin that is now recognized as schizophrenia, other diagnostic terms were put forward as well. These included *dementia praecossima* (5) and *dementia infantilis* (7). Potter offered clearer descriptions of schizophrenia, with consideration of the child’s developmental age, and offered specific diagnostic criteria for children (8).

Despite efforts to recognize childhood schizophrenia as a distinct clinical entity, during the decades between 1920 and 1970, the term *childhood psychosis* comprised all forms of severe mental disorders in children, including schizophrenia and autism. Kanner’s description of early infantile autism catalyzed an alternative view of the conceptualization of these disorders.

Beginning with the works of Loretta Bender (9), Leo Kanner (10), and others (11), all considered childhood schizophrenia to fall under the broader category of childhood psychoses. Nevertheless, there came an acknowledgment and new awareness of major developmental differences in the perception of reality (12) and that developmentally or culturally appropriate beliefs (e.g., imaginary playmates and fantasy figures) did not, of themselves, suggest psychosis. This cluster of syndromes, including infantile autism, was defined by developmental lags in the maturation of language, perception, and motility (11). Although psychotic speech and thoughts were initially considered inherent components of childhood schizophrenia, hallucinations and delusions were not required criteria (6 ,13 ,14 and 15). DSM-II adopted this nosology and grouped all childhood psychoses under childhood schizophrenia. As a result of this broad grouping, the literature regarding childhood schizophrenia from this period overlaps with that of autism and does not differentiate autism from other psychotic disorders. With further development of psychiatric taxonomy and elucidation of the phenomenology (course, onset, family history, and associated features), the distinctiveness of the various childhood psychoses and the similarity between child and adult schizophrenia were demonstrated (16 ,17). This change had a pronounced influence on the nosology of these disorders and led eventually to changes with the DSM-III (18). Schizophrenia arising in childhood and infantile autism came to be recognized as distinct clinical syndromes, each with its unique and distinct psychopathologic phenomenology, theories about causes, and longitudinal course. Research since the advent of DSM-III generally validated this decision (19 ,20). This distinction has had an impact on how children with these disorders are currently evaluated, managed, and treated.

COGNITIVE ASPECTS

Part of "45 - Psychosis in Childhood and Its Management "

Although children do not describe disorders, they nevertheless may complain of changes in their mental and cognitive states. To these changes, clinicians add signs, based on observations derived from the mental state examination of the children and data obtained from laboratory or cognitive tests. Subsequently, a distinctive pattern may emerge over the course of the child’s illness. A collection of symptoms and signs occurring in a certain temporal pattern is then used to categorize the child’s problem. Psychotic symptoms can be attributed to distinct mental illnesses (functional psychoses), which are contrasted with the psychotic symptoms that usually result from a demonstrable underlying pathologic mechanism and organic origin (organic psychoses), such as delirium. Cognitive impairments, particularly impaired concentration and ability to focus, usually accompany psychosis in children. However, when the psychosis is secondary to an organic origin, there is often accompanying

impairment in the sensorium presenting as confusion and disorientation, as is typical of delirium.

From a cognitive and developmental standpoint, certain clinical features in children create diagnostic challenges. One problem is distinguishing true psychotic phenomena in children from nonpsychotic idiosyncratic thinking, perceptions caused by developmental delays, exposure to disturbing and traumatic events, and overactive and vivid imaginations. Furthermore, because the onset of childhood schizophrenia is insidious, with a lifelong history of developmental and personality abnormalities, differentiating between the premorbid state and the active psychotic state can be difficult. It has also been suggested that the development of psychotic conditions during childhood may have major adverse effects on development, a feature further complicating diagnostic assessment (21).

Investigators have noted that social withdrawal, “shyness,” and disturbances in adaptive social behavior seem to be the first signs of dysfunctional premorbid development. Eggers et al. suggested that these should be considered vulnerability factors, indicative of a risk of psychotic illness (22). Recent work has also pointed to early language deficits and motor impairments as being significant for very early-onset schizophrenia, in children younger than 12 years (23). However, a socially odd child is not usually schizophrenic. In fact, most children who have hallucinations are not schizophrenic (24 ,25 and 26), because they lack the requisite persistence and associated symptoms. Intellectual delays have long been considered as general risk factors for psychopathology and psychosis in children (27). In fact, the estimated rates may be low, because most studies examining psychosis in children exclude patients with mental retardation (28 ,29).

CLINICAL AND DEVELOPMENTAL CONSIDERATIONS

Part of "45 - Psychosis in Childhood and Its Management "

Developmental factors influence the detection, form, and context of psychotic symptoms in children. One problem of assessing psychotic disorders in very young children compared with older children is that these symptoms in young children tend to be more fluid and less complex. Isolated hallucinations can occur in acutely anxious but otherwise developmentally intact preschool children. In older children, hallucinations may occur in the absence of other signs of psychosis, but they are usually associated with other psychopathologic conditions, such as depression, severe anxiety, and posttraumatic stress disorder.

Often, there is an underestimation of the subtle differentiation among age-related cognitive preoccupations, pseudohallucinations, and imaginative experiences. Further, it is often too difficult to tease out the physiognomic-animistic interpretations of the inner and outer world on one hand and the first prominent psychotic phenomena such as delusions and hallucinations on the other.

It is critical to avoid rushing to a premature conclusion about unusual behaviors and beliefs in children. Such atypical mental experiences in children can be recognized as prodromal or prepsychotic signs only after the manifestation of frank psychotic symptoms. Odd beliefs and unusual behaviors deserve close observation, but they cannot be ascribed to psychosis without the concomitant presence of a thought disorder.

For example, a young schizophrenic girl lived by the railroad tracks all her life. At age 11 years, about the time when her disorder had its onset, she noted that the sound of the train whistle changed, and she began to wonder why. She came to believe that it had a specific purpose and meaning—that it beckoned her. Until that time, such events were inconsequential and unimportant, but at about age 11 years, she started to attach a different meaning to them. She was uncomfortable with these thoughts and realized that it was not her usual pattern of thinking. Things around her started to have special meaning, her thoughts were “strange,” and she was puzzled and bewildered.

This may be considered a *predelusional phenomenon*, idiosyncratic but not yet fixed. Over the next several years, she developed ideas of reference, thought broadcasting, and thought insertion. She believed that the train whistle was sending special messages from God to her. She no longer questioned these perceptions and believed them to be real. By age 14 years, she was diagnosed with childhood-onset schizophrenia.

A formal thought disorder in a child is more ominous and requires careful psychiatric and neurologic evaluation. Distinguishing between the formal thought disorder of schizophrenia and that of developmental disorders, personality disorders, and speech and language disorders also presents diagnostic problems (30). Symptoms such as thought disorder have been noted to arise in persons with pervasive developmental disorders, particularly those with good language skills, such as (often referred to as “high functioning”) autistic persons and those with Asperger syndrome (31 ,32).

Although loose associations and incoherence are valid diagnostic signs of early-onset schizophrenia, these symptoms are also sometimes seen in schizotypal children (33). The inclusion criteria of disorganized speech according to DSM-IV (34), rather than a formal thought disorder, presents a particular challenge when assessing children, because disorganized speech is an inherent component of many of the developmental disorders. Clearly, the assessment and ascertainment of delusions, hallucinations, and thought disorder in linguistically impaired children are difficult and complicated.

Therefore, developmental disorders must be considered in the differential diagnosis of a child presenting with psychotic symptoms. The use of comparable criteria across the age span facilitates analyses of progressive symptoms from

childhood to adulthood. However, one of the difficulties in assessing psychotic disorders in very young children is to determine whether nonspecific behavioral disturbances represent an incipient psychosis or are signs of autism or pervasive developmental disorders (35 ,36).

Further, the conceptualization of psychoses in childhood as a neurodevelopmental disorder has drawn increasing attention, especially as it relates to childhood-onset schizophrenia (37). Therefore, another alternative in the conceptualization of psychotic episodes is a grouping of symptoms that are not part of the formal DSM or International Classification of Diseases (ICD) scheme. For decades, clinicians recognized that a pattern of brief psychotic episodes, affective dysregulation, and poor social abilities occurs in children. Early references on schizophrenia (17) and later writings (38 ,39) noted the diagnostic problem of children with poor social development and psychosis. Now, the absence of a formal single diagnostic address for this syndrome has produced a wide variety of terms applied to the same phenomena. The older literature suggested that such children may be considered to have "borderline syndrome of childhood" (40), and then later "schizotypal disorder of childhood" was considered (38). In 1986, Cohen et al. suggested the term *multiplex developmental disorder* and proposed that this condition was best understood as a developmental deviation within the group of pervasive developmental disorders (41). Towbin and co-workers offered operational criteria and preliminary validating evidence for the concept and criteria and used the term *multiple complex developmental disorder* (42). Following Cohen et al., the view of Towbin and co-workers was that multiple complex developmental disorder was a higher-functioning type of pervasive developmental spectrum disorder. Rather than pointing to one particular outcome, Towbin and coworkers suggested that multiple complex developmental disorder was a nonspecific risk factor for a poor adaptation in adult life but with a myriad of adult diagnostic outcomes such as schizophrenia, bipolar illness, or any of the more severe unstable personality disorders. Further elaboration of these criteria have shown support for the concept and validated that children with pervasive developmental disorder not otherwise specified and autism can be meaningfully separated from those with multiple complex developmental disorder (43). Further exploration of the concept of multiple complex developmental disorder has received support from neurophysiologic studies as well (44).

The National Institute of Mental Health (NIMH) project on early schizophrenia culled children for a study of clozapine. The most common referrals were children whose symptoms closely resembled those of multiple complex developmental disorder. The NIMH group suggested the term *multidimensionally impaired* (45 ,46) and offered criteria that were analogous to those described by Towbin and co-workers. However, despite findings that many of these children met partial criteria for pervasive developmental disorder not otherwise specified, the NIMH group preferred to consider the constellation a *forme fruste* of schizophrenia (46). Yet longitudinal studies suggested that the constellation remains stable and does not progress to schizophrenia (46). Other work reporting on similar children has used terms like such as "pervasive developmental disorder plus bipolar" (47) or "obsessive difficult temperament" (48). As further exploration now points to "high rates of speech and language, motor, and social impairments in patients with childhood-onset schizophrenia," the association with pervasive developmental spectrum disorders is drawn even closer for this very early-onset subgroup (23).

In addition to the developmental factors and disorders described earlier, the other differential diagnoses of childhood psychoses can be classified as described in the following sections.

Functional Psychoses

Childhood-Onset Schizophrenia

Schizophrenic psychoses with onset before age 11 years are rare. The prevalence in this age group is about 0.01 to 0.05 per 1,000. In addition, developmental status can affect the expression of the disorder. The earliest descriptions by DeSanctis (5), Bleuler (49), and Kraeplin (6) reported the onset and occurrence during childhood and considered schizophrenic psychoses to be an early onset of the same disease, which appeared to be on a continuum phenomenologically, that they observed in adolescents and adults. Furthermore, it has been shown that schizophrenic psychoses can be diagnosed reliably in children using the same criteria as for adults (20 ,36 ,50 ,51). Very few studies to date have dealt with the long-term outcome in childhood-onset schizophrenia. Most studies have followed children for between 1 and 5 years (52). Because of methodologic difficulties, there is a striking absence of data before the age of 11 years on the long-term course of psychosis. Asarnow emphasized the "crucial importance" of long-term follow-up data for establishing the validity of psychotic symptoms manifested in early childhood (53). This is especially important because children often describe "hearing voices," especially in clinical populations. An astute clinician will delve into this symptom in greater depth, to obtain a qualitative appreciation of these "voices." Often, the child will describe this "voice" inside his head as if hearing his own voice or that of an adult in his life. He most likely will not hear this voice through his ears and seems affectively not to be too troubled by it. Conversely, a child experiencing true auditory hallucinations is frightened, puzzled, and unable to be reassured. This differentiation is especially important because management of these youngsters often includes the use of psychotropic medications, which, in and of themselves, require serious consideration because of their long-term adverse effects. If the phenomenology of these so-called psychotic

symptoms is not clarified, many youngsters with pseudohallucinations will be prescribed psychotropic medication needlessly. In addition, they will wrongly be labeled with a psychotic disorder.

Premorbid developmental peculiarities have been reported in children with childhood-onset schizophrenia who have been followed into their thirties. These peculiarities are primarily internalizing such as shyness, isolatory behaviors, lack of interest, awkwardness, being fickle with peculiar facial expression, aggression, paranoia, anxious thoughts, and being mistrustful of others, along with symptoms of depression. These signs have been reported to be much more common than externalizing, acting-out behaviors such as temper tantrums, aggression, opposition, and hostility (22). From a developmental standpoint, the age of first manifestation of nonpsychotic symptoms is younger than the age of onset of schizophrenic symptoms (53 ,54 ,55 and 56). However, the predictive relevance in prepsychotic symptoms in children seems to be extremely uncertain because of the high variability of developmental peculiarities.

The nature of the diagnostic subtypes varies markedly across the course of the illness. In patients with continuous predominantly catatonic symptoms, the outcome is poor. Eggers et al. suggested that detailed case description helps to illuminate the heterogeneous psychopathology of childhood-onset schizophrenia (22). These investigators found that various temporary premorbid behavioral peculiarities were precursors of childhood-onset schizophrenia. Children with early-onset schizophrenic psychosis develop a phenomenology of positive and negative psychotic symptoms that are similar to those seen in adult patients with schizophrenia, and the course variability is perhaps even greater than in adult patients. Their findings contradicted the assumption that childhood-onset schizophrenia is characterized only by negative symptoms, because a differentiation between premorbid and prodromal signs proved to be arbitrary.

Since Kraeplin's first description of *dementia praecox* in 1889, the onset and course of schizophrenia relied heavily on first admission data and on the subsequent course of the disease. However, Hafner et al. argued that items taken from the preadmission phase of the disease were often incorrectly used as premorbid characteristics (57). In an attempt systematically to account for the age and gender distribution of the true onset and the symptoms and pattern of the early and later course, Hafner et al. developed an Interview for the Retrospective Assessment of the Onset of Schizophrenia (IRAOS) (57). This instrument allows an objective, reliable, and valid assessment of the symptoms, psychological impairments, demographic and social characteristics, and the referring points in time of the early course of psychosis. Their findings suggested that the IRAOS provides information on the earliest course of the disease and enables them to separate premorbid characteristics, possibly the most powerful predictors of the later course and outcome, from contamination with symptoms and deficits belonging to the early phase of the disease. The influence of certain life events on the early course is also made accessible to empiric research. Other instruments that have been used for assessing psychotic symptoms in youngsters have been the Interview Schedule for Children (58), the Diagnostic Interview Schedule for Children (59), the Schedule for Affective Disorders and Schizophrenia for School-Aged Children (60 ,61).

Several nonspecific and nondiagnostic neurobiological abnormalities have been reported in patients with schizophrenia. These include deficits in smooth pursuit eye movements and autonomic responsivity (62 ,63). Neuroimaging findings include a progressive increase in ventricular size and a fourfold greater decrease in cortical gray matter volume during adolescence, with the greatest differences occurring in the frontal and temporal regions (64 ,65 ,66 and 67). Others findings reported in the literature are a smaller total cerebral volume, correlated with negative symptoms (37), and frontal lobe dysfunction (68).

Schizophrenia with childhood onset is usually a severe and chronic disorder with a more guarded prognosis and poorer therapeutic response to neuroleptic agents than schizophrenia with adolescent or adult onset. New research and data will help to clarify the origin and pathogenesis of schizophrenia in children. Subsequently, development of more effective treatments and preventive measures may reduce its severity.

Mood Disorders

Mood disorders such as major depression and acute mania can often be accompanied by psychotic symptoms. Over the past several decades, the prevalence of mood disorders appears to have been increasing (69). Although information on the epidemiology of psychotic depression in children is limited, Chambers et al. described the occurrence of psychotic depression in children (61). The psychotic symptoms usually are mood congruent, but at times they can be quite like those seen in childhood schizophrenia (20 ,70 ,71 and 72). This overlap in symptoms increases the likelihood of incorrect diagnosis, especially at the time of onset. Sometimes, the negative symptoms of schizophrenia in children can be mistaken for those of depression. However, it has been shown that children with schizophrenia have poorer premorbid adjustments, lower IQs, and more chronic dysfunction, when compared with children who suffer from a depressive disorder (50). It is therefore prudent to make only a tentative diagnosis at the outset that must be confirmed longitudinally. Careful follow-up of psychotic patients is needed to detect diagnostic errors. This issue can be compounded, however, if the symptoms resolve with antipsychotic medications. It becomes unclear whether the child improves because of treatment or spontaneous remission. Approximately one-half of adolescents with bipolar disorder may

be originally diagnosed as having schizophrenia (20 ,70). Therefore, it is extremely important that longitudinal reassessment is needed to ensure accuracy of the diagnosis. Despite an increased family history of depression in schizophrenic youth (20), family psychiatric history can be an extremely helpful differentiating factors. However, the opposite is not true, that is, an increased family history of schizophrenia in depressed or bipolar youngsters. Often, the rule of thumb is first to rule out mood disorder in a child or adolescent before the diagnosis of schizophrenia is more strongly considered.

Even though there is an overlap of the quality of psychotic symptoms in children with mood disorders and childhood-onset schizophrenia, often with careful examination, some of the mood-congruent symptoms can be ascertained. As clinicians, it is important that we ascertain chronologically what came first, that is, a change in mood and then the onset of delusions or hallucinations, or a disturbance in thought followed by a change in mood. For example, the child who first starts to have “strange thoughts” and to hear voices over time becomes puzzled, fearful, distraught, and depressed. This is quite different from the child who first starts to lose interest in activities, to feel irritable or depressed, to not want to play with friends, and who demonstrates neurovegetative symptoms, such as a decrease in appetite, sleep disturbance, and lethargy. Subsequently, the child starts to think he is a bad and evil person and then hears a voice that tells him he is a bad boy and that he should kill himself. The phenomenology in this instance is quite different. However, it is not always this clear, and there is a high rate of misdiagnosis in both directions (72 ,73).

Psychotic symptoms during a manic episode have been recognized for many years, although misdiagnoses of schizophrenia were, and remain, relatively common (74). Bipolar disorder eventually develops in a minority of children initially hospitalized for major depression (1). This is particularly so if the child has a positive family history of bipolar disorder, psychomotor retardation, rapid onset of symptoms, mood-congruent psychotic symptoms, or pharmacologically induced mania or hypomania. The characteristics of the delusions and hallucinations are often mood congruent (expansiveness, grandiosity, and euphoria). Therefore, a child experiencing mania may have delusions of being “superman” with special powers, of being able to fly and leap from high places. Conversely, the child may believe that he or she has special skills playing baseball, even though the child perhaps may have problems with gross motor skills and is clumsy and uncoordinated. Similarly, the child may hear voices, the content of which are mood congruent, with the altered state in mood (i.e., grandeur), and may believe that the voice is saying that he or she is superior and can do anything.

Brief Reactive Psychosis

Occasionally, children and adolescents suddenly develop psychotic symptoms that can last from a few hours or days. The child experiences these symptoms when under tremendous stress, such as after a death in the family, witnessed acts of violence or destruction, or physical or sexual abuse. The *acute psychotic symptoms* often resolve quickly, with total recovery in a few days. These youngsters may suddenly become disorganized, confused, agitated, or withdrawn. At times, their speech becomes nonsensical and incomprehensible. They may also experience delusions and hallucinations. These, too, are usually short-lived.

Anxiety Disorders

Children who experience *acute anxiety* or who have a history of maltreatment, abuse or neglect report significantly higher rates of psychotic symptoms when compared with controls (75). Several studies have documented psychotic-like symptoms in children with *posttraumatic stress disorder*. In such instances, the psychotic symptoms actually represent intrusive thoughts or worries regarding the traumatic event (73 ,76 ,77). Mental status examination usually reveals the lack of a formal thought disorder, and the psychotic-like symptoms are more akin to derealization or depersonalization, as is often observed in traumatized children. Furthermore, there is often a qualitative difference in the way children with anxiety disorders and those with childhood-onset schizophrenia relate. The former have better-developed relationship and prosocial skills compared with the socially isolated, awkward, and odd behaviors of a child with schizophrenia. An identifiable traumatic event, abuse, or neglect in the child’s history, in and of itself, does not necessarily rule out a psychotic disorder, because children with both schizophrenia and mood disorders may have had such experiences (73).

Organic Psychoses

Neurologic Conditions

Seizure Disorder

Children with *seizure disorders* can experience hallucinations as part of the seizure activity. Complex partial seizures, especially those with a temporal focus, may be associated with interictal psychotic symptoms of delusions, hallucinations, and unusual preoccupations. Caplan and co-workers described a formal thought disorder in children with partial complex seizures (78 ,79), although their way of defining thought disorder makes it intertwine closely with language organization deficits. However, they did emphasize that these epileptic children usually do not display negative symptoms such as those seen in schizophrenia. Hallucinations in children with epilepsy typically are brief. Therefore, these children experience mainly positive symptoms, which

are often short-lived. Caplan and co-workers also described a higher incidence of formal thought disorder in those children who have lower IQs, earlier onset of the seizure disorder, and poor seizure control. They postulated that these symptoms may either reflect the underlying neuropathology that produces the seizures or result from the “kindling phenomenon” as a secondary effect of the seizure activity.

Deteriorative Neurologic Disorders

Psychotic symptoms have been described in children who have a *deteriorative and degenerative neurologic disorders* such as subacute sclerosing panencephalitis (80). Other disorders include Wilson disease, lipid storage disorders, and Huntington chorea. These are usually differentiated from childhood-onset schizophrenia by the presence of neurologic findings on physical examination of the child, further corroborated by abnormal findings on laboratory testing. Children suffering from such neurologic deterioration often have a gradual, persistent, but global decline in their neurologic condition.

Central Nervous System Lesions

These conditions include brain tumors, congenital malformations, and head trauma.

Metabolic and Hormonal Disturbances

Various *metabolic and hormonal conditions* can be responsible for psychotic symptoms in children. Endocrinopathies may include disorders of the adrenal, thyroid, or parathyroid glands. Exogenous metabolic disturbances leading to psychotic symptoms can include exposure to heavy metals.

Toxic Psychoses

Toxic psychosis or delirium usually occurs secondary to bacterial or viral infections, high fevers, and exogenous toxins including medications, illicit drugs, alcohol, and poisonings. Unlike childhood schizophrenia or other psychotic disorders, in which impaired thinking and communication are the most salient symptoms, toxic psychosis is more likely to cause vivid, disturbing visual or tactile hallucinations and other perceptual problems. Auditory hallucinations can also occur, but their content is qualitatively different from those experienced in childhood schizophrenia or mood disorders. These sensory experiences may be extremely frightening and may be accompanied by agitation or by uncontrolled or even aggressive behaviors. Children and adolescents often describe the experience as “losing their mind”—a frightening concept, and they can become disoriented, unable to orient to person or place, or comprehend why they are behaving in an unusual manner. They may also experience fluctuating levels of alertness.

In children, infections (bacterial or viral) can cause encephalitis, meningitis, and human immunodeficiency virus–related syndromes, which can result in delirious states. High fevers, regardless of origin, have been known to cause delirious states with perceptual disturbances. In addition, chronic liver and kidney disease may cause delirious states associated with psychotic symptoms in children, manifested by states of confusion, distortions in perceptions, and frank hallucinations.

The best example of medication-induced psychosis is that resulting from high doses of stimulants (the most commonly prescribed group of medications in this age group). In young children, normal doses of common medications, such as over-the-counter antihistamines and decongestants, can induce similar symptoms. Some of the other medications that can have a similar result are steroids, which can cause not only a disturbance in mood (depression and manic symptoms), but also delirium. Children prescribed anticholinergic drugs are also vulnerable to developing delirium, presenting with psychotic symptoms.

Other causes, especially in older children and adolescents, are alcohol intoxication, amphetamine-like drugs (“speed” and cocaine), hallucinogenic drugs (LSD and psilocybin), solvents, and cannabis. Most children who develop drug-induced psychosis recover once the drugs are out of their system.

The psychotic symptoms sometimes experienced by patients after anesthesia should be included in the category of toxic psychoses. Although usually short-lived, this phenomenon is reported by patients to be a very frightening experience. Support, reassurance, and ensuring safety at the time are usually sufficient in the management of patients after anesthesia.

MANAGEMENT AND TREATMENT

Part of "45 - Psychosis in Childhood and Its Management "

Assessment

Effective treatment requires knowledge of the psychotic disorders, diagnostic criteria, symptoms, and longitudinal course, in addition to an understanding of the youngster’s developmental, social, educational, and psychological needs. Treatment strategies therefore need to focus on the clinical symptoms and morbidity of the underlying disorder, while also addressing any comorbid disorders or biopsychosocial stressors. The physician must prioritize symptoms and diagnoses, so a reasonable treatment plan addresses multiple problems. A clinician examining a child for psychoses must first ascertain whether the child comprehends the clinician’s question about delusions and hallucinations and whether the child endorses the psychotic symptoms only to please the interviewer or to get attention. In addition, it is important to determine whether the child acts on the basis of the delusional or hallucinatory perceptions—associated with an affective response of fear, dread, avoidance, or elation.

The assessment of the child with psychotic symptoms should include a careful, comprehensive, and thoughtful

evaluation. The history is often obtained from multiple informants, and several sessions may be required to gain accurate assessment of the child's mental status. The assessment should include a detailed evaluation of the symptom presentation, course of illness, and phenomenology. A developmental history of the child and a detailed family psychiatric history are invaluable components of the evaluation and assessment. A positive family history, especially for an affective disorder or schizophrenia because these disorders tend to run in families, often helps the clinician with the differential diagnosis in the child.

Once it is determined that the child is experiencing psychotic symptoms, it is important foremost to ascertain the cause of such symptoms. This will, in large part, determine the management and treatment of the child presenting with psychosis. A thorough physical examination is essential, and pertinent tests and procedures may be necessary, as clinically indicated. These may include imaging studies, an electroencephalogram, toxicology screens, and renal and liver function tests. Some children may require consultation with other pediatric specialists.

Psychological and projective testing are not indicated as a method of diagnosing specific disorders causing the psychosis. However, they can be helpful for intellectual assessment and to determine developmental delays, because these deficits may influence the presentation or interpretation of symptoms. Routine use of adaptive function measures is important for understanding actual function in social, daily living, and communication domains. These can be quite helpful in planning and maintaining developmentally relevant treatment goals. Similarly, speech and language evaluations are often helpful, especially with a child who appears to have linguistic impairments on examination.

Treatment

If it is deemed that the cause is organic, then the first step is to diagnose and treat the underling cause of the psychotic symptoms. This may include treating a partial complex seizure disorder, managing a metabolic imbalance, or treating an underlying infection or reducing a fever. Conversely, if it is determined that there is no medical cause for the psychotic symptoms, then the next step is to ascertain whether the psychosis is functional. If so, is it secondary to severe depression or acute mania with psychotic symptoms or secondary to a schizophrenic illness?

Adequate treatment requires a combination of pharmacotherapy and various psychosocial interventions that target the child's specific difficulties. Some of this depends on the phase of the underlying illness (81):

Stage 1 (prodromal phase): The child may experience some period of deteriorating function, which may include social isolation, idiosyncratic preoccupations and behaviors, and academic difficulties.

Stage 2 (acute phase): This is usually the time when the child comes to the attention of a mental health professional, when the clinical picture is dominated by frank delusions and hallucinations and other positive symptoms such as a formal thought disorder or strange and idiosyncratic behaviors.

Stage 3 (recovery phase): The symptoms usually begin to remit and dissipate. However, often there may still be the presence of some psychotic symptoms, although they are less disturbing to the child. In this phase, the child may continue to experience some levels of confusion, disorganization, or lability in mood.

Stage 4 (residual phase): The positive symptoms continue to subside, but the child continues to experience apathy, lack of motivation, withdrawal, and restricted or flat affect.

Unfortunately, some children remain symptomatic and chronically impaired, despite what would be considered adequate treatment. Usually, such impairment is characterized by persistent symptoms, which occur especially if the psychosis is secondary to a schizophrenic illness, rather than the result of depression or mania.

Psychosocial interventions should include working with both the parents and the child. Interventions targeted at improving family functioning, problem solving, communication skills, and relapse prevention have been shown to decrease relapse rates in adults (82). Children may benefit from social skills training and may require specialized educational programs, academic adjustments, and support at school. Ongoing illness teaching and medication education, are important to promote compliance with treatment and to help in coping with the daily and sometimes long-term implications of the child's illness. Every effort should be made for the child to be maintained in the least restrictive setting, such as home. However, in some cases, the severity and chronicity of the underlying illness may warrant long-term placement in a hospital or residential facility.

Pharmacotherapy is instituted in an attempt to treat the underlying cause of the psychosis, or for symptom control, in those children who have psychotic symptoms secondary to a known origin. Informed consent from the parents or guardian should be obtained before treatment with psychopharmacologic agents is instituted.

It is not in the purview of this chapter to discuss each medication in detail. For the treatment of major depression, the following antidepressants have been used in children:

Tricycle antidepressants (nortriptyline, imipramine, desipramine)

Selective serotonergic reuptake inhibitors (fluoxetine, paroxetine, sertraline, fluvoxamine, citalopram)

Nonselective serotonergic reuptake inhibitors (nefazodone, mirtazapine)

Monoamine oxidase inhibitors (phenelzine, tranylcypromine) (seldom used currently)

Others: bupropion, venlafaxine

Mood stabilizers that have been used for the treatment of manic-depressive illness in children include the following:

Anticonvulsants (divalproex sodium, carbamazepine, gabapentin)
Lithium

Often, the use of antipsychotic medications in addition to the use of antidepressants or mood stabilizers is indicated in functional psychosis.

If the suspicion is that of early-onset schizophrenia, then antipsychotics are first-line medications. Although children may metabolize neuroleptics more rapidly than adolescents and adults, optimum doses for children are typically less than those required in adolescents and adults.

First-line agents include traditional neuroleptic medications that block dopamine receptors or the atypical antipsychotic medications that have a variety of effects including antagonism of serotonergic receptors. The atypical antipsychotic medications are reported to be at least as effective for positive symptoms and may even be more helpful for negative symptoms. Further, there is some suggestion that they have fewer adverse effects. Except for clozapine, the novel agents also appear to produce tardive dyskinesia. Experience with novel antipsychotic agents is too scant to determine whether the risk of tardive dyskinesia is equal to or less than with the older antipsychotics. Newer antipsychotic medications that have been used in children are risperidone and olanzapine. They may be less sedating than the traditional neuroleptic agents such as haloperidol, fluphenazine, thioridazine, and chlorpromazine. There have been some case reports in the literature of the use of clozapine for children and adolescents with schizophrenia in whom normally adequate treatment with other traditional antipsychotic medications has failed.

For a child suffering from acute reactive psychosis, support and safety are the two primary considerations. If the child is extremely stressed and acutely ill, hospitalization may be necessary to provide a safe and structured environment. Brief treatment with antipsychotic mediations has often been effective for the alleviation of psychotic symptoms in some children. However, medications will not eliminate the problem that originally caused the brief psychosis. Thus, psychotherapy is often helpful in helping the child learn to cope with the emotional trauma that may have precipitated the episode.

Toxic psychosis requires immediate medical intervention to identify the cause and to provide appropriate treatment. Identifying the cause may include laboratory tests such as serum electrolytes, liver function tests, toxicology screens, blood alcohol level, serum levels of prescribed medications including theophylline, tricyclic antidepressants (nortriptyline, amitriptyline, imipramine, desipramine), or mood stabilizers (valproic acid, lithium, carbamazepine). Neuroleptics, because of their common usage, comprise another important group of medications that needs to be considered in the psychiatric population. The rare but possible development of *neuroleptic malignant syndrome*, manifesting as a disturbance of sensorium, fever, rigidity, and high blood pressure, should be considered. A history of treatment with neuroleptics and an elevated creatinine phosphokinase usually enable one to determine this cause (83). Most children who develop drug-induced psychosis recover once the drugs are discontinued and out of their system. The gravest danger occurs during the psychotic episode, when a child may cause serious harm to himself or herself or to others, because judgment is impaired. Some children may need brief hospitalization until the cause is determined and the psychotic symptoms dissipate. Except for the presence of neuroleptic malignant syndrome as the cause, brief treatment with antipsychotic medications may be necessary to decrease the child’s agitation and to control the distorted perceptions.

In addition to the foregoing treatment strategies, other interventions and services may be needed to address either comorbid conditions or associated sequelae of the underlying disorder causing the psychosis, such as substance abuse, depression, and suicidal tendencies.

CONCLUSIONS

Part of "45 - Psychosis in Childhood and Its Management "

From the clinical perspective, the rapid change and development of childhood have immediate implications for diagnosis and intervention. When one is treating children, it is important to maintain diagnostic fluidity and to tolerate the pressure of uncertainty.

In the realm of childhood-onset psychopathology, we have a great deal to learn from the psychoses. The stability of a diagnostic category over time is usually considered to be one measure of the construct validity of that category. One possibility is that lack of stability of a diagnosis during childhood implies that it lacks validity. This is only important if one is trying to establish a unique direct link with later-onset disorders and to apply the same terminology. However, another possibility is that some childhood diagnoses are only risk factors for development of more enduring adult conditions, such as the relationship between conduct disorder and antisocial personality disorder. Variability of normal and psychopathologic development and the heavy influence of environmental features and familial functioning during childhood make it difficult to be certain about diagnoses in children. Although categoric classification has its advantages and at times is necessary (34), dimensional perspectives can be important for understanding these phenomena as well.

It is useful to recognize the close, reciprocal relationship between diagnostic classification and biological or genetic advances. Advances in genetic and imaging studies should open the way to a different classification system that links

symptoms, neural circuitry, and biological (genetic) markers more closely than any current system. We should not be too surprised to discover that conditions once considered to be quite distinct are now closely linked, as has been found for Tourette syndrome, obsessive-compulsive disorder, and attention-deficit/hyperactivity disorder (84). “Diagnostic stability” may refer to the stability of these biological features, rather than the stability of clinical signs and symptoms. One could expect that developments in molecular biology would shed light on the natural history, protective factors, and risk factors for a specific biological risk. However, these developments will depend on increasingly reliable, reproducible diagnoses. The success of these strategies demands that probands or cohorts have been reliably diagnosed according to the most valid criteria at the time. Thus, the diagnoses inform the biological work, and the biological work, in turn, influences our classification and criteria. This process fosters a more accurate understanding of the natural history, pathophysiology, and etiology of disorders and the relationships among disorders. It will inform us about prenatal psychoneurohormonal factors that influence development, sexual differentiation, and maturation of the central nervous system. In time, we stand to gain a much clearer understanding of the complex, diverse contributions to the development and maintenance of psychotic disorders. We look forward to developing treatments that will offer more comfort and better functioning to the patients and families afflicted with these chronic conditions. There is, in addition, a real prospect of finding protective factors and preventive interventions that can avert the worst manifestations of these disorders.

REFERENCES

1. Strober M, Carlson G. Bipolar illness in adolescents with major depression. *Arch Gen Psychiatry* 1982;39:549–555.

2. Chambers WJ, Puig-Antich J, Tabrizi MA, et al. Psychotic symptoms in prepubertal major depressive disorder. *Arch Gen Psychiatry* 1982;39:921–927.

3. McHugh PR, Slavney PR. *The perspectives of psychiatry*. Baltimore: Johns Hopkins University Press, 1998.

4. Maudsley H. *The physiology and pathology of the mind*. London: Macmillan, 1867.

5. DeSanctis S. Sopra aclune varieta della demenzi precocce. *Rev Sper Feniatr Med Leg* 1906;32:141.

6. Kraeplin E. *Dementia praecox and paraphrenia*. Edinburgh: Livingstone, 1919.

7. Heller T: Uber Dementia infantilis. *Z Kinderforsch* 1930;37:661. [Reprinted in Howells JG, ed. *Modern perspective in international child psychiatry* Edinburgh: Oliver & Boyd, 1969.]

8. Potter HW. Schizophrenia in children. *Am J Psychiatry* 1933;89:1253.

9. Bender L. Childhood schizophrenia. *Nerv Child* 1943;1:138–140.

10. Kanner L. Autistic disturbance off affective contact. *Nerv Child* 1943;2:217–250.

11. Fish B, Ritvo E. Psychoses of childhood. In: Noshpitz JD, Berllin I, eds. *Basic handbook of child psychiatry*. New York: Basic, 1979:249–304.

12. Piaget J. *The child's construction of reality*. London: Routledge and Kegan, 1955.

13. Bender L. Childhood schizophrenia. *Am J Orthopsychiatry* 1947;17:40–56.

14. Despert JL. Delusional and hallucinatory experiences in children. *Am J Psychiatry* 1948;104:528–537.

15. Despert JL. Schizophrenia in children. *Psychiatr Q* 1938;12:366–371.

16. Kolvin I. Studies in the childhood psychoses. *Br J Psychiatry* 1971;6:209–234.

17. Rutter M. Childhood schizophrenia reconsidered. *J Autism Child Schizophr* 1972;2:315–337.

18. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, third ed. Washington, DC: American Psychiatric Association, 1980.

19. Beitchman JH. Childhood schizophrenia: a review and comparison with adult onset schizophrenia. *Psychiatr Clin North Am* 1985;8:793–814.

20. Werry JS, McClellan J, Chard L. Early onset schizophrenia, bipolar and schizoaffective disorders: a clinical follow-up study. *J Am Acad Child Adolesc Psychiatry* 1991;30:457–465.

21. Volkmar F. Childhood and adolescent psychosis: a review of the past 10 years. *J Am Acad Child Adolesc Psychiatry* 1996;35:843—851.

22. Eggers C, Bunk D, Krause D. Schizophrenia with onset before the age of eleven: clinical characteristics of onset and course. *J Autism Dev Disord* 2000;30:29–40.

23. Nicolson R, Lenane M, Singaracharlu S, et al. Premorbid speech and language impairments in childhood-onset schizophrenia: association with risk factors. *Am J Psychiatry* 2000;157:794–800.

24. Del Beccaro MA, Burke P, McCauley E. Hallucinations in children: a follow-up study. *J Am Acad Child Adolesc Psychiatry* 1988;27:462–465.

25. Garralda ME. Hallucinations in children with conduct and emotional disorders. I. The follow-up study. *Psychol Med* 1984;14:597–604.

26. Garralda ME. Hallucinations in children with conduct and emotional disorders: the clinical phenomena. *Psychol Med* 1984;14:589–594.

27. McClellan J, McCurry C. Neuro-cognitive pathways in the development of schizophrenia. *Semin Clin Neuropsychiatry* 1998;3:320–332.

28. Bettes B, Walker E. Positive and negative symptoms in psychotic and other psychiatrically disturbed children. *J Child Psychol Psychiatry* 1987;28:555–567.

29. Russell AT, Bott L, Sammons C. The phenomenology of schizophrenia occurring in childhood. *J Am Acad Child Adolesc Psychiatry* 1989;28:399–407.

30. Caplan R, Tanguay P. Development of psychotic thinking in children. In: Lewis M, ed. *Child and adolescent psychiatry: comprehensive text book*. Baltimore: Williams & Wilkins, 1991:310–317.

31. Dykens E, Volkmar F, Glick M. Thought disorder in high-functioning autistic adults. *J Autism Dev Disord* 1991;21:291–301.

32. Ghaziuddin M, Leininger L, Tsai L. Brief report: thought disorder in Asperger’s syndrome. Comparison with high functioning autism. *J Autism Dev Disord* 1995;25:311–318.

33. Caplan R, Guthrie D, Gish B, et al. The Kiddie Formal Thought Disorder Scale: clinical assessment, reliability, and validity. *J Am Acad Child Adolesc Psychiatry* 1989;28:408–416.

34. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington, DC: American Psychiatric Association, 1994.

35. Cantor S, Evans J, Pearce J. Childhood schizophrenia, present but not accounted for. *Am J Psychiatry* 1982;139:758–762.

36. Watkins JM, Asarnow RF, Tanguay PE. Phenomenology and classification of the childhood psychoses. *Psychol Med* 1988;18:191–201.

37. Alaghband-Rad J, McKenna K, Gordon CT, et al. Childhood onset schizophrenia: biological markers in relation to clinical characteristics. *Am J Psychiatry* 1997;154:64–68.

38. Russell AT. Schizophrenia. In: Hooper S, Hynd R, Mattison R, eds. *Assessment and diagnosis of child and adolescent psychiatric disorders: current issues and procedures*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1992:23–63.

39. Cantor S, Pearce J, Pezzot-Pearce T, et al. The group of hypotonic schizophrenia: present but not accounted for. *Schizophr Bull* 1981;7:1–11.

40. Lewis M. Borderline disorders in childhood. *Child Adolesc Psychiatr Clin North Am* 1994;3:31–43.

41. Cohen DJ, Paul R, Volkmar FR. Issues in the classification of pervasive and other developmental disorders: toward DSM-IV. *J Am Acad Child Psychiatry* 1986;25:213–220.

42. Towbin KE, Dykens ED, Pearson GS, et al. Conceptualizing “borderline syndrome of childhood” and “childhood schizophrenia” as a developmental disorder. *J Am Acad Child Adolesc Psychiatry* 1993;32:775–782.

43. Buitelaar J, Van der Gaag R. Diagnostic rules for children with PDDNOS and multiple complex developmental disorder. *J Child Psychol Psychiatry* 1998;39:911–919.

44. Lincoln AJ, Bloom D, Katz M, et al. Neuropsychological and neurophysiological indicies of auditory processing impairment in children with multiple complex developmental disorder. *J Am Acad Child Adolesc Psychiatry* 1998;37:100–112.

45. McKenna K, Gordon CT, Rapoport JL et al. Childhood-onset schizophrenia: timely neurobiological research. *J Am Acad Child Adolesc Psychiatry* 1994;33:771–781.

46. Kumra S, Jacobsen LK, Lenane M, et al. “Multidimensionally impaired disorder”: is it a variant of very early-onset schizophrenia? *J Am Acad Child Adolesc Psychiatry* 1998;37:91–99.

47. Wozniak J, Biederman J, Faraone S, et al. Mania in children with pervasive developmental disorder revisited. *J Am Acad Child Adolesc Psychiatry* 1997;36:1552–1559.

48. Garland EJ, Weiss M. Obsessive-difficult temperament and its response to serotonergic medication. *J Am Acad Child Adolesc Psychiatry* 1996;35:916–920.

49. Bleuler E. *Dementia praecox or the group of schizophrenias*. Zinkin J, trans. New York: International Universities Press, 1911:195.

50. Asarnow JR, Ben-Meir S. Children with schizophrenia spectrum and depressive disorders: a comparative study of pre-morbid adjustment, onset pattern and severity of impairment. *J Child Psychol Psychiatry* 1988;29:477–488.

51. McClellan JM, Werry J. Practice parameters for the assessment and treatment of children and adolescents with schizophrenia. *J Am Acad Child Adolesc Psychiatry* 1994;33:616–635.

52. Mattanah JJF, Becker DF, Levy KN, et al. Diagnostic stability in adolescents followed upto 2 years after hospitalization. *Am J Psychiatry* 1995;12:889–894.

53. Asarnow JR, Thompson MC, Goldstein MJ. Childhood onset schizophrenia: a follow-up study. *Schizophr Bull* 1994;20:647–670.

54. Green WH, Padron-Gayol M, Hardesty AS, et al. Schizophrenia with childhood onset: phenomenological study of 38 cases. *J Am Acad Child Adolesc Psychiatry* 1992;31:968–976.

55. Russell AT. The clinical presentation of childhood-onset schizophrenia. *Schizophr Bull* 1994;20:631–646.

56. Maziade M, Gingras N, Rodrigue C, et al. Long-term stability of diagnosis and symptom dimensions in a systematic sample of patients with onset of schizophrenia in childhood and early adolescence. I. Nosology, sex and age of onset. *Br J Psychiatry* 1996;169:361–370.

57. Hafner H, Riecher-Rossler, Hambrecht M, et al. IRAOS: an instrument for the assessment of onset and early course of schizophrenia. *Schizophr Res* 1992;6:209–223.

58. Kovacs M. *The interview schedule for children (ISC)*, 10th rev. Pittsburgh: University of Pittsburgh School of Medicine, 1978.

59. Shaffer D, Fisher P, Dulcan M, et al. The second version of the Diagnostic Interview Schedule for Children (DISC-2). *J Am Acad Acad Child Adolesc Psychiatry* 1996;35:865–877.

60. Puig-Antich J, Chambers W. *The Schedule for Affective Disorders and Schizophrenia for School-Aged Children (Kiddie-SADS)*. New York: New York State Psychiatric Association, 1978.

61. Chambers WJ, Puig-Antich J, Hirsch M, et al. The assessment of affective disorders in children and adolescents by semistructured interview: test-retest reliability of the Schedule of Affective Disorders and Schizophrenia for School Age Children, Present Episode Version. *Arch Gen Psychiatry* 1985;34:583–591.

62. Zahn TP, Jacobsen LK, Gordon CT, et al. Autonomic nervous system markers of psychopathology in childhood onset schizophrenia. *Arch Gen Psychiatry* 1997;54:904–912.

63. Jacobsen LK, Giedd JN, Berquin PC, et al. Quantitative morphology of the cerebellum and fourth ventricle in childhood-onset schizophrenia. *Am J Psychiatry* 1997;154:1663–1669.

64. Rapoport JL, Giedd J, Kumra S, et al. Childhood onset schizophrenia: progressive ventricular change during adolescence. *Arch Gen Psychiatry* 1997;54:897–903.

65. Rapoport JL, Giedd J, Blumenthal J, et al. Progressive cortical change during adolescence in childhood onset schizophrenia: a longitudinal magnetic resonance imaging study. *Arch Gen Psychiatry* 1999;56:649–654.

66. Gordon CT, Frazier JA, McKenna K, et al. Childhood onset schizophrenia: an NIMH study in progress. *Scizophr Bull* 1994;20:697–712.

67. Frazier JA, Alaghband-Rad J, Jacobsen L et al. Pubertal development and onset of psychosis in childhood onset schizophrenia. *Psychiatry Res* 1997;70:1–7.

68. Thomas MA, Ke Y, Levitt J. Preliminary study of frontal lobe 1H MR spectroscopy in childhood onset schizophrenia. *J Magn Reson Imaging* 1998;8:841–846.

69. Klerman GL. The current age of youthful melancholia: evidence for increase in depression among adolescents and young adults. *Br J Psychiatry* 1988;152:4–14.

70. Carlson GA. Child and adolescent mania: diagnostic considerations. *J Child Psychol Psychiatry* 1990;31:31–342.

71. Joyce PR. Age of onset in bipolar affective disorder and misdiagnosis of Schizophrenia. *Psychol Med* 1984;14:14–149.

72. McClellan JM, Werry JS, Ham M. A follow-up study of early onset psychosis: Comparison between outcome diagnoses of schizophrenia, mood disorders and personality disorders. *J Autism Dev Disord* 1993;23:243–262.

73. McClellan J, McCurry C. Early onset psychotic disorders: diagnostic stability and clinical characteristics. *Eur Child Adolesc Psychiatry* 1999;8[Suppl 2]:1S–7S.

74. Weller EB, Weller RA. Assessing depression in prepubertal children. *Hillside J Clin Psychiatry* 1986;8:193–201.

75. Famularo R, Kinscherff R, Fenton T. Psychiatric diagnoses of maltreated children: preliminary findings. *J Am Acad Child Adolesc Psychiatry* 1992;31:863–867.

76. Hornstein JL, Putnam FW. Clinical phenomenology of child and adolescent dissociative disorders. *J Am Acad Child Adolesc Psychiatry* 1992;31:1077–1085.

77. Altman H, Collins M, Mundy P. Subclinical hallucinations and delusions in nonpsychotic adolescents. *J Child Psychol Psychiatry* 1997;38:413–420.

78. Caplan R, Shields WD, Mori L, et al. Middle childhood onset of interictal psychosis: case study. *J Am Acad Child Adolesc Psychiatry* 1991;30:893-896.

79. Caplan R, Guthrie D, Shields WD, et al. Formal thought disorder in pediatric complex partial seizure disorder. *J Child Psychol Psychiatry* 1992;33:1399–1412.

80. Caplan R, Tanguay PE, Szekely AG. Subacute sclerosing panencephalitis presenting as childhood psychosis: a case study. *J Am Acad Child Psychiatry* 1987;26:440–443.

81. American Academy of Child and Adolescent Psychiatry. *Practice parameters for the assessment and treatment of children and adolescents with schizophrenia*, revised ed. Washington, DC: American Academy of Child and Adolescent Psychiatry, 2000.

82. American Psychiatric Association. Practice guideline for the treatment of patients with schizophrenia. *Am J Psychiatry* 1997;154:1–63.

83. Joshi PT, Capozzoli JT, Coyle JT. Neuroleptic malignant syndrome: a life-threatening complication of neuroleptic treatment in adolescents with affective disorder. *Pediatrics* 1991;87:2.

84. Towbin KE, Peterson BS, Cohen DJ, et al. Differential diagnosis. In: Leckman JF, Cohen DJ, eds. *Tourette's syndrome: tics, obsessions, compulsions: developmental psychopathology and clinical care*. New York: John Wiley, 1999:118–139.

46

Behavioral Phenotypes of Neurodevelopmental Disorders: Portals into the Developing Brain

James C. Harris

James C. Harris: Departments of Psychiatry and Behavioral Sciences, Pediatrics, and Mental Hygiene, Johns Hopkins University School Medicine, Baltimore, Maryland.

HISTORICAL BACKGROUND

Part of "46 - Behavioral Phenotypes of Neurodevelopmental Disorders: Portals into the Developing Brain "

Increasing evidence indicates that specific neurodevelopmental disorders may be associated with particular patterns of behavior. A description of behavior was included by Langdon Down in the first published description of a specific mental retardation syndrome, Down syndrome (1). In his description, Down observed: “They have considerable powers of imitation, even bordering on being mimics. Their humorousness and a lively sense of the ridiculous often colors their mimicry.” Later, he added: “Several patients who have been under my care have been wont to convert their pillow cases into surplices (vestments) and to imitate, in tone and gesture, the clergymen or chaplain which they have recently heard.” He also commented on personality traits, saying: “Another feature is their great obstinacy—they can only be guided by consummate tact.” Although these stereotypes were not confirmed in subsequent studies (2 ,3), the prospect of linking behavior and genetics was introduced in this first description of a neurogenetic disorder. Subsequent early clinical descriptions, such as that of tuberous sclerosis complex by Critchley and Earl (4), identified peculiar, and severe, behavioral problems in children and adults with that condition. Yet despite the early recognition of syndrome-specific behavioral and psychiatric features, neurogenetic disorders were not empirically investigated for behavioral deficits until 1990s, when new conceptual and methodologic procedures were introduced (5).

Two main reasons may explain this lack of interest after the early reports by Down and others. First, there was a general negative reaction against eugenics and claims for genetic bases of personality (6). This negative reaction established a climate in which it was not considered appropriate for academic investigators to emphasize the genetics of behavior. Second, there has been a major emphasis on learning theory and its applications to the field of mental retardation, in which most genetic disorders are found. Tremendous strides have been made in the education of even the most severely mentally retarded persons. Advances in academic and social adaptive education, in conjunction with motor treatment, have placed greatest emphasis on how severely and multiply handicapped people could attain greater degrees of independence and social integration. With the emphasis on normalization, research into severe disorders in learning tended to be deemphasized. Moreover, the occurrence of associated psychiatric and behavioral problems was interpreted more in terms of learning theory rather than in being unlearned behaviors associated with behavioral phenotypes. The focus has been on addressing the potential of the individual person and the developmental possibilities. Yet this focus could not continue to ignore reports from families and clinical observations of characteristic patterns of behavior and stereotypes.

With the establishment of active and refined learning-based approaches and a better understanding of the interpretation of genetic findings, reappraisal and revision of attitudes toward research with behavioral phenotypes have begun. O'Brien suggested three reasons for this shift (7). First, research findings have been reliably reported with various syndromes. Second, there are continued reports from family members as large family organizations have developed in the United States and other countries that describe characteristic behavioral patterns and interpersonal responses. In meetings, parent groups frequently report similar behavior problems and difficulties in management across syndromes. The interest in parent groups in improving the life of their children has led to additional hypotheses and more refined observations on behavioral characteristics. Third, new techniques in genetics provide new insights into

the extent and mechanisms of the human genome as the basis of behavior. Advances in other aspects of neuroscience, including neurophysiology and neuroanatomy, provide additional means of designating brain mechanisms that may be involved. With the establishment of these new methods of evaluation and the identification of rating scales to measure behavioral phenotypes, there is now an increased focus on behavioral phenotypes in developmental neuropsychiatry. Finally, comprehensive study of children with different developmental disabilities may increase our appreciation of the relative contribution of genetic variables in the pathogenesis of specific affective and behavioral disorders.

Nyhan introduced the term *behavioral phenotype* to describe outwardly observable behavior so characteristic of children with genetic disorders that its presence suggests the underlying genetic condition (8). In speaking of compulsive self-injury in Lesch–Nyhan disease (LND), a disorder that he initially described, Nyhan noted: “We feel that these children have a pattern of unusual behavior that is unique to them. Stereotypical patterns of behavior occurring in syndromic fashion in sizable numbers of individuals provide the possibility that there is a concrete explanation that is discoverable. In these children, there are so many anatomical abnormalities, from changes in hair and bones to dermatoglyphics, that it is a reasonable hypothesis that their behaviors are determined by an abnormal neuroanatomy that would be discoverable, possibly neurophysiologically, ultimately anatomically… these children all seem self-programmed. These stereotypical patterns of unusual behavior could reflect the presence of structural deficits in the central nervous system” (8).

Such observations have led to greater emphasis on assessment of behavior, and the recognition of behavioral phenotypes in some disorders has led to closer scrutiny of known neurodevelopmental conditions. Initially, the focus was on documenting the patterns of behavior because the study of brain and behavior requires the identification of welldefined syndromes for investigation. Now that developments in the neurosciences provide a means to understand the biological bases of such behavioral patterns, the focus has shifted to understanding the neurobiological mechanisms underlying characteristic behavioral patterns, including cognitive processes and social interactions. Such patterns are reported in numerous syndromes arising from genetic or chromosomal abnormalities. Thus, molecular analysis of the underlying genetic disorder has been initiated in several syndromes with the hope of revealing the biological basis of the behavioral phenotype. However, because of the rarity of many of these syndromes and the complexity of their genetic basis, establishing the validity of the association between syndrome and behavioral phenotype is difficult. Nevertheless, Flint pointed out that evidence from animal studies with relevance to human behavioral phenotypes shows that the pathway from genotype to phenotype may be accessible after careful delineation of each of the features of the behavioral phenotypes (9 ,10). However, in regard to the study of cognition, he suggested that we require a greater integration of different levels of understanding of cognition to exploit the genetic discoveries, “a rapprochement between molecular and systems neuroscience” (10).

Much of the research in behavioral genetics uses a “top-down” approach to the qualitative analysis of complex traits such as novelty seeking, memory, personality traits, and intelligence (11). Linkage or association strategies are used to examine naturally occurring alleles of candidate genes in a “wild-type population.” These alleles are usually functional polymorphisms rather than mutations and, if they are quantitative trait loci, may be associated with individual differences in the trait in question. However, Tully suggested that such genes may have minor effects on the phenotype of the individual because alleles with a more striking effect could reduced fitness and would be selected against in evolution (12). Specifically, chromosomal deletions that may have cognitive and behavioral consequences may be associated with monosomy (13). The loss of one copy of genes that are dose sensitive may be significant in brain development. Such genes may play a fundamental role in development of the functional organization of the brain, but they may not be as important for individual differences in the general population. Moreover, partial variants of disorders such as LND that result in a range of enzyme levels may allow study of dose response to enzyme deficits.

This chapter uses a developmental perspective to provide a definition and characterization of behavioral phenotypes in neurodevelopmental disorders, and it discusses etiologic factors, methods to understand underlying mechanisms, and natural history. It addresses the question: What do behavioral phenotypes that occur in specific neurogenetic disorders teach us, and how may they provide a portal to understand the developing brain? This question is considered by reviewing studies of neurogenetic disorders with behavioral phenotypes: (a) LND, an X-linked disorder, that results from the absence of an enzyme, hypoxanthine-guanine phosphoribosyltransferase (HPRT), that is involved in purine metabolism; (b) Prader–Willi syndrome (PWS) and Angelman syndrome (AS), in which the parental origin of the genes involved (uniparental disomy or UPD) is an important factor in the cause; (c) fragile X syndrome, a disorder caused by unstable trinucleotide repeat expansion that results in the absence of a gene that encodes an RNA-binding protein thought to play a role in translational regulation of selective messenger RNA transcripts; and (d) Williams syndrome (WMS), a contiguous gene disorder with an unusual cognitive phenotype in which language is preserved but the patient has severe visual spatial disabilities. Each of these neurogenetic disorders provides a portal to understand neurodevelopment. Other nongenetic disorders that are environmentally induced, such as fetal alcohol syndrome, are not discussed but also offer keys to understanding the developing brain (23).

DEFINITION AND CHARACTERIZATION

The study of behavioral phenotypes emphasizes the discovery, among individuals with known chromosomal, genetic, or neurodevelopmental disorders, of those mental and behavioral features causally related to the underlying condition. Examples are the characteristic self-mutilation of fingers and lips in LND, the hyperphagia and compulsive behaviors in PWS, gaze aversion in fragile X syndrome, and the superficial sociability, hyperlalia, and language disorder in WMS. When present, the behavior suggests the syndrome. As Nyhan suggested, these are “syndromes of behavior” (14). Still, despite their behavioral presentations, not all individuals with the disorder show the classic behavioral features, but the probability is greater that they will. The essential issue is that the behavior suggests the diagnosis.

Efforts to define what is meant by a behavioral phenotype are continuing. Harris proposed that behavioral phenotypes are stereotypic patterns of behavior that are reliably identified in groups of individuals with known neurodevelopmental disorders and are “not learned” (15 ,16). They may be the consequence of neurodevelopmental abnormalities that are potentially discoverable. This approach to definition is a phenonic approach that takes as its starting point observations of the behavior itself rather than beginning with a discrete and genetically identifiable condition, such as Down syndrome. Using the phenomic approach, the behavioral phenotype of Rett syndrome (17), with its characteristic hand and hand-to-mouth stereotypies, identified it as disorder with a behavior phenotype many years before the genetic origin was recognized. Moreover, the phenomic approach does not discount acquired disorders, such as fetal alcohol syndrome, as having behavioral phenotypes. The impact of alcohol on cellular signaling is now well known, with its consequences of cell death, abnormal midline brain development, behavioral problems, and learning disabilities (16 ,23).

Such considerations led Flint and Yule to propose the following definition that includes the characteristic types of behaviors: “The behavioral phenotype is a characteristic pattern of motor, cognitive, linguistic, and social abnormalities that is consistently associated with a biological disorder” (18). This does not mean that the behavior is present in all instances but that the probability of its occurrence is increased. In the future, more may be learned about brain mechanisms by comparing persons with behavioral involvement with others who have the same syndrome but without the behavioral features.

Although some investigators have sought to limit the study of behavioral phenotypes to known genetic disorders (11), knowledge of the genetic disorder is only the first step. Links from gene to behavior are complicated in that one gene may lead to the encoding of many, perhaps ten or more, different proteins; the number of genes and type of mutation determine complexity. For example, in LND, the disorder of purine metabolism clearly leads to the overproduction of uric acid and renal stones, but the pathway to the movement disorder and self-injury is not direct and may be mediated through effects on the arborization of dopamine neurons (19). Moreover, there are variants of LND with different degrees of enzyme deficit, ranging up to 20%, that have clinical effects.

Thus, several caveats are necessary as we consider pathways from genes to behavior (11): (a) the behavioral descriptions, like other physical features of neurodevelopmental disorders, have increased probability of occurring and do not occur in all cases; they are not be fully expressed in all affected individuals; (b) the genetic background of the individual may affect the phenotypic expression; (c) the possibility exists that environmental factors may modify expression; (d) the behavioral presentation may be modified by the extent of mental retardation associated with the disorder; and (e) variability occurs in mouse models in which there may be species-specific factors so mutant mouse models do not replicate the behavioral features. One must consider the genetic background, strain differences, and the differences in the rodent physiology. In LND, the HPRT-deficient mouse has a uricase enzyme that breaks down uric acid. Therefore, it is not a model for the hyperuricemic metabolic disorder, but it still may be a useful model to study dopamine deficiency in the brain. Thus, aspects of the disorder may be modeled in transgenic mice or in other species.

In some animal models, links to specific pathophysiology have been established. The canine model of narcolepsy (20) is an interesting example of an approach to a human clinical disorder. Mutations for canine, autosomal recessive, narcolepsy were identified by linkage analysis in canine backcrosses, and homology was demonstrated between human chromosome 6 and canine chromosome 12. Canine narcolepsy is caused by a disruption of a G-protein–coupled receptor, the hypocretin (orexin) receptor 2 gene (Hcrt2) in three canine breeds. However, human narcolepsy is not associated with frequent hypocretin gene mutation (20). Nonetheless, most humans with narcolepsy have undetectable hypocretin 1 levels in cerebrospinal fluid.

PSYCHOPATHOLOGY AND BEHAVIORAL PHENOTYPES

Part of "46 - Behavioral Phenotypes of Neurodevelopmental Disorders: Portals into the Developing Brain "

Numerous neurogenetic disorders are associated with nonspecific behaviors that may be found in several syndromes. These include attention problems, hyperactivity, impulsivity, self-injury, aggression, autistic-like behavior, and preservative behaviors. Such presentations indicate vulnerability of the developing brain and perturbation of brain systems resulting in these clinical conditions. However, because these behaviors occur across many syndromes, they lack specificity and do not qualify as specific behavioral phenotypes.

Still, these behavioral features should be included in the description of the disorders. For example, the relationship between aggression and antisocial behavior has been suggested in monoamine oxidase A (MAOA) deficiency. Brunner et al. described an association between abnormal behavior and MAOA deficiency in several males from a single large Dutch kindred (21). The affected males differed from unaffected males in that they tested in the borderline range of mental retardation and demonstrated increased impulsive behavior, that is, aggressive behavior, abnormal sexual behavior, and arson. Yet a specific psychiatric diagnosis was not made in four affected males who were examined by psychiatrists. Because MAOA deficiency leads to increased 5-hydroxytryptamine (5-HT) levels, the aggressive behavior in these persons may be an exception to studies linking low 5-HT with impulsive aggression. Brunner et al. suggested that even if a possible association between MAOA deficiency and abnormal behavior is confirmed in other kindreds, “the data do not support the hypothesis that MAOA constitutes an ‘aggression gene’.” These investigators noted that genes are essentially simple and code for proteins, whereas behavior is complex; thus, a direct causal relationship between a single gene and a specific behavior is highly unlikely. In MAOA deficiency, complexity is shown by the variability in the behavioral phenotype and by the highly complex consequences of MAOA deficiency on neurotransmitter function. Thus, the full pathway from gene to complex behavior must be considered; the concept of a gene that directly encodes behavior is simplistic (21). Still, a great deal may be learned by considering such pathways in neurogenetic syndromes.

PREVALENCE

Part of "46 - Behavioral Phenotypes of Neurodevelopmental Disorders: Portals into the Developing Brain "

With increasing attention to neurogenetic disorders, the number of identifiable behavioral phenotypes is increasing. Careful observations of behavior are necessary when considering intervention for neurogenetic disorders. Although standardized rating scales and personality profiles have been developed to measure behavioral phenotypes (22 ,23), profiles pertinent to the specific disorder are needed. Besides behavioral phenotypes, isolated special abilities that occur in genetically based syndromes require assessment. These include special abilities in calculation and in music (24). These special abilities may potentially be related to the proposed modular organization of the central nervous system.

BEHAVIORAL PHENOTYPES OF SPECIFIC NEURODEVELOPMENTAL DISORDERS

Part of "46 - Behavioral Phenotypes of Neurodevelopmental Disorders: Portals into the Developing Brain "

The sections that follow discuss four syndromes in which behavioral phenotypes have been identified: LND, PWS/AS, fragile X syndrome, and WMS. Characteristic behaviors are highlighted, findings on origin are discussed, and potential neurochemical and neuroanatomic abnormalities are reviewed. Behavioral and pharmacologic therapies have had limited success in many of these conditions, so better characterization of the individual condition is essential to establish treatment. Neuroanatomic studies, brain imaging studies, and continuing investigations of neurotransmitter systems, endocrine rhythms, and sleep studies may provide information that will be helpful in the future in treatment.

Lesch–Nyhan Disease

LND is a rare (1:380,000) sex-linked recessive disease caused by an inborn error of purine nucleotide metabolism. It is caused by an almost complete deficiency of the enzyme HPRT, which is involved in the purine salvage (purine base recycling) pathway (25). Self-injury is the major behavioral manifestation; this behavior was sufficiently characteristic that Nyhan introduced the term “behavioral phenotype” as a descriptor (8). LND is of psychosocial and psychiatric importance because of the lifelong suffering experienced by the involved child and his family, the uniqueness of the behavioral phenotype, and the resources needed for lifelong patient supervision. Moreover, an understanding of the neurobiological basis of this disease may contribute to a better understanding of brain mechanisms involved in self-injurious and compulsive behaviors.

Genetic and Metabolic Aspects

The HPRT-encoding gene is located on the X chromosome in the q26-q27 region and is made up of nine exons and eight introns totaling 57 kilobases (kb). The *HPRT* gene is transcribed to produce a mRNA of 1.6 kb that contains a protein-encoding region of 654 nucleotides. More than 270 mutations throughout the coding regions have been identified (79). Techniques that provide information on the three-dimensional structure of the HPRT protein make it possible to correlate structure and function of the enzyme (26). Eads et al. reported the effects of single amino acid substitutions on the stability and activity of HPRT (26).

The gene involved in LND is on the X chromosome, so the disorder occurs almost entirely in males; occurrence in females is extremely rare. The metabolic abnormality is the result of an abnormal gene product—a deficiency in the enzyme HPRT. This enzyme is normally present in each cell in the body and is highest in the brain, especially in the basal ganglia. Its absence prevents the normal metabolism of hypoxanthine and results in excessive uric acid production and manifestations of gout without specific drug treatment (i.e., allopurinol). The full disease requires the virtual absence of the enzyme. Other syndromes with partial HPRT deficiency are associated with gout without the neurologic and behavioral symptoms. Page and Nyhan reported that HPRT levels are related to the extent of motor symptoms,

the presence or absence of self-injury, and possibly the level of cognitive function (27). Hypoxanthine accumulates in the cerebral spinal fluid, but uric acid does not because it is not produced in the brain and does not cross the blood–brain barrier.

Behavioral Phenotype

Self-injurious behavior usually is expressed as self-biting; however, other patterns of self-injurious behavior may emerge with time. It is not uncommon for self-injury to progress to deliberate self-harm (19 ,28). Characteristically, the fingers, mouth, and buccal mucosa are mutilated. The biting pattern is often asymmetric, so the patient may mutilate the left or right side of the body and may become anxious if he perceives that this side of the body is threatened. Other associated maladaptive behaviors include head or limb banging, eye poking, pulling of fingernails, and psychogenic vomiting (28).

Self-mutilation in LND is conceptualized as a compulsive behavior that the child tries to control but generally is unable to resist. With increasing age, the affected child becomes more adept at finding ways to control his self-injury. He may enlist the help of others to protect him against these impulses or may learn self-restraint.

A language pattern that consists of repeated ambivalent statements with anxiety and coprolalia (vulgar speech) is characteristic. Moreover, the patient may be compulsively aggressive and may inflict injury on others through pinching, grabbing, or using verbal forms of aggression. Frequently, he will apologize for this behavior immediately afterward and will say that the behavior was out of his control.

Etiologic Factors

The cause of the neurologic and behavioral symptoms is not clearly established; however, abnormalities in dopamine function have been demonstrated in three autopsied cases (29). The behavior is not caused by either hyperuricemia or by excess hypoxanthine because LND partial variants whose HPRT levels are greater than 2 do have hyperuricemia but they do not self-injure. Moreover, infants treated for hyperuricemia from birth whose uric acid level is normalized still develop self-injury despite having normal levels of uric acid.

Wong and Harris et al. used positron emission tomography to investigate how dopamine dysfunction contributes to the self-injurious behavior (30). These authors documented reductions in dopamine transporter density of 68% in putamen and 42% in caudate in six patients with classic LNS and self-injurious behavior. To clarify the relationship between presynaptic dopamine transporter binding in the striatum and self-injurious behavior further, Harris, Jinnah and Wong (30a) studied seven patients with Lesch–Nyhan variants (HPRT levels 1.8% to 20.0%) and two patients with HPRT levels less than 1.5%, all nine without self-injurious behavior (age range, 12 to 37 years). The extent of motor findings was documented on quantitated neurologic examination. Two patients with HPRT levels less than 1.5% and two patients with HPRT levels of l.8% and 2.5% with severe movement disorder were not different in WIN 35,428 dopamine transporter binding in positron emission tomography imaging than the previously described classic patients with LND who did injure themselves. The study of variant cases with motor symptoms but with no self-injurious behavior suggests that reductions in dopamine receptor density are not a sufficient explanation of the self-injury. However, these authors found that HPRT level and the extent of motor deficit were correlated with dopamine transporter binding in caudate and putamen in the nine cases. Dopamine transporter binding was significantly correlated with HPRT levels in whole cells. Moreover, when the movement disorder was rated on the Fahn-Marsden dystonia rating scale, putamen dopamine transporter density was significantly correlated with symptom severity. These findings suggest that dopamine reduction is linked to the extent of the movement disorder, but it may not be a sufficient explanation for self-injurious behavior, and other neurotransmitters need to be examined. Moreover, these variant subjects with levels from 2% to 20% showed cognitive deficit profiles similar to those of classic LND (31).

Future investigation will need to take into account the existence of a variety of mutations in the *HPRT* gene structure. Why partial HPRT deficiency does not lead to neurologic and behavioral symptoms remains unclear; perhaps neurotrophic factors are active with minute amounts of the enzyme. It is advisable to study combined drug and behavioral treatment. An emphasis on parental training is of particular importance for drug compliance and generalization of treatment effects. As in other inborn errors, continuous family support is essential. Harris provides a description of a comprehensive treatment program for LND (19).

Prader–Willi Syndrome

PWS is a neurodevelopmental disorder characterized by obesity, short stature, cryptorchidism, mental retardation, hyperphagia, learning disability, short stature, hypogonadism, hypotonia, small hands and feet, and dysmorphic facies. Patients have an increased prevalence of daytime sleepiness, scoliosis, and other orthopedic abnormalities. Because of the obesity, heart failure and diabetes may occur as complications. Although it is a rare disorder (1 in 10,000 to 15,000), its behavioral phenotype has assumed prominence in genetics because of its relationship with AS, which has a different behavioral phenotype, although both disorders involve genomic imprinting of the same region of chromosome 15.

Genetics

PWS may result from both chromosomal deletion and maternal UPD. In UPD, two copies of the maternal chromosome

are inherited with no paternal contribution (32). Without the presence of the chromosome donated by the father, the normal imprinting of the two maternally donated chromosomes leads to absence of gene expression in this interval. This results in a functional abnormality that is essentially equivalent to the structural abnormality found in a deletion in the 15q11-q13 region that is associated with the disorder. Moreover, in about 5% of cases, abnormalities in the mechanism of imprinting may occur when the imprinting control center itself has a mutation.

Several genes are included in the most commonly deleted region in PWS. Some are paternally imprinted, and others are maternally imprinted (33). Among these, *ZNF 127, NDN, SNURF-SMRPN, IPW* are paternally imprinted. Another gene, *UBE3A* (E6-AP ubiquitin lipase), is maternally imprinted. Others genes in this region that are expressed from both maternal and paternal chromosomes include three γ-aminobutyric acid (GABA) receptor subunits *(GABRB3, GABRA5, GABRG3)* (33). Because similar phenotypes result from deletions and from imprinting in PWS, it is less likely that nonimprinted genes play a role in PWS or AS. Among these genes, a specific gene for PWS has not been established, so several of these genes may contribute to the phenotype. For example, the *SMRPN* gene is involved in protein slicing and is expressed throughout the brain; however, it is not thought that the PWS is the direct outcome of this deficit. The *NCD* (necdin) gene does lead to failure to thrive in certain mouse strains, so it may be a factor; however, those mice that survive do develop into apparently normal adults. Thus, the disorder is most likely linked to the loss of more than one gene in this region. Conversely, the mutation of a single gene, *UBE3A*, has been found in cases of AS (34).

Behavioral Phenotype

The extent of cognitive impairment is variable in PWS. Some patients test in the normal range of intelligence, but most test in the mild to moderate range of mental retardation. Others may test in the severe range of mental retardation. The behavioral phenotype includes unusual food-related behavior (compulsive food seeking, hoarding, gorging), skin picking, irritability, anger, low frustration tolerance, and stubbornness. Standardized methods of assessment have substantiated increased rates of depression, anxiety, and compulsive behavior. Up to 50% of children and adults with PWS demonstrate behavioral disorders.

Compulsive eating is the most disabling of these behavioral manifestations and leads to obesity and the complications of severe obesity, such as respiratory impairment and diabetes. The hyperphagia, which has been consistently found, has received the most systematic behavioral evaluation. When not carefully supervised, patients may steal food and, in some instances, eat unpalatable food, although this can be avoided with appropriate supervision. Holm and Pipes evaluated food-related behavior in the PWS (36). They found that behavioral problems were most commonly related to food and included food stealing, foraging for food, gorging, and indiscriminate eating with little food selectivity. No special circumstances that resulted in food stealing or gorging were identified.

Besides the food-related compulsions, emotional lability with temper tantrums, stubbornness, negativism, skin picking and scratching, and non–food-related obsessions have been examined. A questionnaire survey involving 369 cases identified compulsive and impulsive aggressive behavior (37). These authors used the Overt Aggression Scale, the Yale-Brown Obsessive-Compulsive Disorder Scale, a clinical global rating, and DSM-III-R criteria to diagnose self-stimulation and self-injury, compulsive behavior, and obsessive behaviors. These investigators found that skin picking was the most common form of self-injury, observed in 19.6% of this sample. Other types of self-injury with lower frequency were nose picking, nail biting, lip biting, and hair pulling. The second behavioral problem area was compulsive behavior; food hoarding was the most severe manifestation and occurred in 17.7%. Other compulsive behaviors included counting, symmetric arrangements of objects, checking, and hand washing, but they were less common. Obsessive thinking was far less characteristic, with only 1.4% rated in the severe range on an item dealing with concerns about contamination. State et al. reviewed the evidence in regard to compulsive behaviors in PWS and the relationship with obsessive compulsive disorder (38). Behavioral problems identified in the preschool years persist throughout the school years and continue into adolescence and adulthood.

Etiologic Factors

Investigators have proposed that the genetic abnormality in PWS leads to hypothalamic dysfunction that results in aspects of the clinical phenotype, such as dysregulation of feeding, delay in sexual development, sleep disorder, and abnormality of thermoregulation. In support of hypothalamic dysfunction, Swaab et al., in a postmortem study, found reduction in oxytocin cells in certain regions of the hypothalamus (35). However, other brain regions and neuropeptides may be involved in PWS. Because the loci of GABA subunits is in the area around the 15q11-13 region, GABA has been measured in PWS, and abnormalities have been reported in plasma levels in some patients.

To clarify the mechanism leading to the behavioral phenotype further, differences between deletion and maternal UPD causes have been assessed (39). Similar studies have been completed in AS (40). Differences in intellectual functioning in PWS with a paternal 15q11-q13 deletion versus maternal UPD of chromosome 15 were evaluated using measures of intelligence and academic achievement in 38 patients with PWS (24 with deletion and 14 with UPD).

The patients with UPD had significantly higher verbal IQ scores than those with deletion ($p < .01$). The magnitude of the difference in verbal IQ was 9.1 points (69.9 versus 60.8 for UPD and deletion PWS patients, respectively). Only 17% of subjects with the 15q11-q13 deletion had a verbal IQ greater than or equal to 70, whereas 50% of those with UPD had a verbal IQ greater than or equal to 70. Performance IQ scores did not differ between the two PWS genetic subtype groups. This report documents the difference between verbal and performance IQ score patterns among patients with PWS of the deletion versus the UPD subtype. Comprehensive treatment of behavioral problems in PWS is described by Holm et al. (41).

Angelman Syndrome

In contrast to PWS, investigators have shown that one gene in the deleted region can lead to AS (34). AS is a neurologic disorder with a heterogeneous genetic origin. It most frequently results from a *de novo* interstitial deletion in the 15q11-q13 region, but it is also caused by paternal UPD or an imprinting mutation. The remaining 20% to 30% of patients with AS exhibit biparental inheritance and a normal pattern of allelic methylation in the 15q11-q13 region. In this biparental inheritance group, mutations in the *UBE3A* gene have been shown to be a cause of AS. Moncla et al. described the phenotypic expression in 14 patients with AS involving eight *UBE3A* mutations (34). These were made up of 11 familial cases from five families and three sporadic cases. Some subtle differences from the typical phenotype of AS were noted. Consistent features were psychomotor delay, a happy disposition, a hyperexcitable personality, EEG abnormalities, and mental retardation with severe speech impairment. The other main features of AS—ataxia, epilepsy, and microcephaly—were either milder or absent in various combinations among these cases. Moreover, myoclonus of cortical origin was commonly observed with severe myoclonic seizures. Most of these patients were overweight. This study showed that ataxia, myoclonus, EEG abnormalities, speech impairment, characteristic behavioral phenotype, and abnormal head circumference are attributable to a deficiency in the maternally inherited *UBE3A* allele. Finally, analysis of mutation transmission showed an unexpectedly high rate of somatic mosaicism in normal carriers. These clinical findings have important consequences for genetic counseling in AS.

Fragile X Syndrome

The fragile X syndrome is characterized by mental retardation, behavioral characteristics, and the physical findings of a long face with large, protruding ears and macroorchidism (42). Fragile X syndrome is the most common known cause of inherited mental retardation, and it may also result in learning disabilities and social deficits in those who do not test in the mentally retarded range. After the identification of the fragile X mental retardation *(FMR1)* gene, the cytogenetic marker (a fragile site at Xq27.3) was replaced by molecular diagnosis. Recognition of this gene has broadened our understanding of the spectrum of the fragile X syndrome.

Genetics

Fragile X syndrome is caused by massive expansion of CGG triplet repeats located in the 5′-untranslated region of the *FMR1*. The cloning of the *FMR1* gene led to the characterization of its protein product FMRP. The full mutation is associated with a process of methylation; the addition of methyl groups along the “backbone of the DNA helix” (42). In patients with fragile X syndrome, the expanded CGG triplet repeats are hypermethylated, and the expression of the *FMR1* gene is repressed, which leads to the absence of FMR1 protein (FMRP) and subsequent mental retardation. The encoded protein is a ribosome-associated, RNA-binding protein thought to play a role in translational regulation of selective messenger RNA transcripts. FMRP is an RNA-binding protein that shuttles between the nucleus and cytoplasm. This protein has been implicated in protein translation because it is found associated with polyribosomes and the rough endoplasmic reticulum (43). A similar mechanism is proposed for *FMR2*, which encodes a large protein of 1,311 amino acids and is a member of a gene family encoding proline-serine–rich proteins that have properties of nuclear transcription factors (44).

The fragile X syndrome was one of the first examples of a “novel” class of disorders caused by a trinucleotide repeat expansion in the X chromosome. In the genetically normal population, the CGG repeat varies from six to 54 units. Affected subjects have expanded CGG repeats (more than 200) in the first exon of the *FMR1* gene (the full mutation). Phenotypically normal carriers of the fragile X syndrome have a repeat in the 43 to 200 range (the premutation). The process of methylation silences transcription so a fully methylated full mutation results in no FMR1 protein’s being produced. The absence of FMR1 protein results in fragile X syndrome. Two additional disorders result in a fragile site at Xq27.3; there are *FRAXE*, which is usually associated with a milder form of mental retardation, and *FRAXF*, which is not consistently associated with mental retardation. These two mutations also have CGG repeat expansions and are distal to the *FMR1* site. The transcriptional silencing of the *FMR2* gene also has been implicated in *FRAXE* mental retardation. *FRAXE* individuals have been shown to exhibit learning deficits, including speech delay and reading and writing problems.

The frequency of the premutation and mutation may be variable in different populations because of founder effects (42). Thus, the prevalence in an English study was 1 in 2,200, and in an Australian study it was 1 in 4,000, but it

was higher in Finland, where it is proposed that the initial settlers included one or more fragile X carriers.

Behavioral Phenotype

There is a substantial degree of genetic and phenotypic heterogeneity in the physical, cognitive, and behavioral phenotype. The behavioral phenotype has been the subject of considerable study and includes mental retardation and learning disabilities, language impairment, hand flapping, gaze aversion, perseveration, and neuropsychiatric disturbance, principally attention-deficit/hyperactivity disorder and pervasive developmental disorder–like symptoms. These patients are more interested in social interactions than those with autistic disorder; the avoidance of social contact may be secondary to hyperarousal or increased sensitivity to stimuli associated with social situations. The behavioral phenotype may be more helpful than the physical phenotype in diagnosis because most prepubertal patients do not have macroorchidism or the characteristic long face.

Attentional difficulty and concentration problems are commonly associated, and hyperactivity may be a presenting symptom in nonretarded boys with fragile X syndrome. Self-injury, most commonly hand biting and scratching, may be elicited by excitement and by frustration. Female patients with fragile X syndrome may be unaffected, although abnormalities in social interaction, thought process, and affect regulation have been reported in carriers. Both schizotypal features and depression have also been found in carriers.

Most girls with the full mutation show shyness and social anxiety. In women with the full mutation, the social anxiety is associated with social awkwardness and schizotypal features. Anxiety disorders, avoidance disorder, and mood disorder symptoms are common (42).

Gaze Aversion

Gaze aversion is a striking feature of affected males with fragile X syndrome. There is consistency in gaze aversion over repeated trials in the same individual; nearly all male patients with fragile X syndrome who are more than 8 or 9 years old avert their gaze on greeting another person. Their unusual greeting is characterized by both head and gaze aversion along with an appropriate recognition of the social partner (45). This greeting response is qualitatively different from gaze aversion that is described in autistic patients. Those with Down syndrome and nonspecific mental retardation do not show this behavioral pattern on greeting. The idiosyncratic gaze behavior in fragile X syndrome may disrupt social interactions. Despite their apparent social anxiety and aversion to eye contact, male patients with fragile X syndrome are otherwise socially responsive and can be affectionate.

Speech and Language

Speech and language in fragile X syndrome is generally delayed, even thought the IQ may be in the normal range. Deficits in both receptive and expressive language include dysfluency, production of incomplete sentences, echolalia, palilalia (reiteration of the speaker's own words and phrases in a perseverative manner), verbal perseveration, and poor fluency in conversation. Compulsive utterances and shifts in speech pitch are common, and auditory processing and memory deficits are present.

Etiology

The FMR1 protein is expressed most abundantly in neurons and testes with the localization primarily in the cytoplasm. High concentrations of FRM1 mRNA have been found at the synapse in rat brains, especially in areas involved in synaptogenesis in the hippocampus, cerebral cortex, and cerebellum (46). Hinton et al. found thin and immature dendritic branches with small synapses in neuroanatomic studies of the neocortex in three male patients with fragile X syndrome (47). The expression of the FMR2 protein also has been characterized. To characterize the expression of the FMR2 protein, polyclonal antibodies were raised against two regions of the human FMR2 protein and were used in immunofluorescence experiments on cryosections of mouse brain. The FMR2 protein is localized in neurons of the neocortex, Purkinje cells of the cerebellum, and the granule cell layer of the hippocampus. FMR2 staining is shown to co-localize with the nuclear stain 4,6-diamidino-2-phenylindole (DAPI) and confirms that FMR2 is a nuclear protein. The localization of FMR1 and FMR2 protein to the mammalian hippocampus and other brain structures involved with cognitive function is consistent with the learning deficits seen in patients with fragile X syndrome. Comprehensive treatment of fragile X syndrome is described by Hagerman and Cronister (48).

Williams (Williams–Beuren) Syndrome

WMS is a rare (1 in 25,000), genetically based neurodevelopmental disorder associated with a characteristic physical, linguistic, cognitive, and behavioral phenotype. This syndrome provides a unique opportunity to study personality development, linguistic functioning, and visuospatial development. The syndrome is characterized by congenital facial and cardiovascular anomalies (supravalvular aortic stenosis and peripheral pulmonary stenosis), failure to thrive, and mental retardation that may be accompanied by transient idiopathic infantile hypercalcemia (49). Adolescents with WMS have expressive language abilities that are better than expected for their mental age. Because of their hyperverbal speech, the investigation of WMS allows the study of the dissociability of components of language and other cognitive brain systems. In mentally retarded patients with WMS, linguistic

abilities may be selectively spared, unlike language learning disability occurring in normally intelligent children (50).

In WMS, a deletion of 1.5 Mb on one copy of chromosome 7 results in the specific physical, cognitive, and behavioral features. Molecular dissection of the WMS phenotype may lead to identification of genes important in human cognition and behavior.

Genetics

WMS is caused by a chromosomal deletion at 7q11.23. A contiguous gene deletion disorder, it results from hemizygous deletion of about 20 genes (51). This chromosomal region is highly repetitive, and the deletion arises from recombination between misaligned repeat sequences flanking the WMS region. The deletion breakpoints cluster within the repeats, so most patients with WMS have similar, although not identical, deletions of 1.5 Mb.

The first deleted gene identified in the critical region was that for elastin *(ELN)*. Studies of patients having deletions or point mutations confined to this gene showed that hemizygosity for *ELN* causes supravalvular aortic stenosis but not the other typical features of WMS.

Several other genes have now been identified that are deleted in most patients with WMS. These include the following: *LIMK1*, which codes for a protein tyrosine kinase expressed in the developing brain; that for syntaxin 1A *(STX1A)*, which encodes a component of the synaptic apparatus; *RFC2*, which codes for a subunit of the replication factor C complex involved in DNA replication; and *FZD3*, homologous to the *Drosophila* tissue-polarity gene, “frizzled” (51).

Cognitive Phenotype

Bellugi et al. proposed that “the cognitive hallmark of WMS is dissociation between language and face processing (relative strengths) and spatial cognition (profound impairment)” (50). The WMS phenotype demonstrates specific dissociations in the higher cognitive functions. These investigators proposed that general cognitive deficits are present but linguistic abilities are spared. They found extreme spatial cognitive deficits with intact face processing. Of special interest is the social phenotype in WMS: an overly friendly, engaging personality and excessive sociability with strangers (52). WMS subjects show an unusual positive response in their social judgments of unfamiliar persons.

Howlin et al. investigated cognitive, linguistic, and academic assessments in a representative sample of 62 adults with WMS (average age of the group was 26 years; mean full-scale IQ was 61) (53). Less difference was found in verbal and performance IQ and between receptive and expressive language skills in the adults than that found in children. Still, subtest scores documented an almost identical cognitive profile to that found in children. Reading, spelling, arithmetic, and social adaptation remained at a low level, with functioning around a 6- to 8-year age equivalent. The consistency in intellectual abilities in both child and adult studies of patients with WMS supports the notion of a syndrome specific pattern of cognitive, linguistic, and adaptive functioning.

The use of adult neuropsychological models to explain developmental disorders of genetic origin such as WMS has been challenged (54 ,55). It is assumed that uneven cognitive profiles found in childhood or adulthood in WMS characterize infant starting states and that modules underlying these abilities start out either intact or impaired. However, findings from two experiments with infants with WMS (selected for study based on claims of innate modularity) suggest a within-syndrome double dissociation: for numerosity judgments, WMS subjects do well in infancy but poorly in adulthood, whereas for language, WMS subjects show poor performance in infancy but do well in adulthood. The theoretic and clinical implications of these findings in WMS emphasize the importance of an developmental approach to neurogenetic disorders. Karmiloff-Smith et al. previously proposed that in WMS, language follows a different path to normal acquisition and may turn out to be more like second language learning (56).

Finally, Tager-Flusberg et al. tested the hypothesis that the WMS phenotype involves sparing abilities involved in the domain of understanding other minds (mentalizing or theory of mind) (57). They compared a group of mentally retarded adults with WMS to an age-, IQ-, and language-matched group of adults with PWS, and a group of age-matched normal adults, on a task that tests mentalizing ability. The task involved identifying the correct labels to match photographs of complex mental state expression focused on the eye region of the face. The adults with WMS performed significantly better than the adults with PWS on this task, and about half the group performed in the same range as the normal adults. Such findings provide support for the proposal that mentalizing is a distinct cognitive domain. The authors proposed that this sparing of cognitive capacity could be “linked to the relative sparing of limbic-cerebellar neural substrate in WMS, which is also connected to cortico-frontal regions that are known to be involved in understanding complex mental states.”

Linking Genes and Cognition

An approach to studying cognition is to carry out genetic and psychometric testing of patients who have small deletions within the WMS critical region. *LIMK1* and *STX1A* are good candidate genes to investigate cognitive or behavioral aspects of WMS. The gene for *LIMK1* was implicated as a cause of the visuospatial characteristics of WMS (58); however, other investigators were unable to substantiate this association in three further cases (59). The genes for *STX1A*

and *FZD9* were proposed as involved, based on brain-specific gene expression in the developing *(FZD9)* or adult *(STX1A)* central nervous system. However, when these genes were underexpressed by 50%, as is expected in WMS, Korenberg et al. reported that deletion of these genes was not associated with significant effects on overall cognition (51). However, these authors did propose that genes responsible for mental retardation and other features of the disorder are “located in the region telometric to RFC2 through GTF21 at the telometric border of the deletion.” Moreover, mild cognitive deficits reported in a subject deleted for elastin and *LIMK1* genes (59) were consistent with findings in those with deletion of genes in the *WMSTF* through *LIMK1* region having mild cognitive deficits. Thus, studies of patients with rare and atypical deletions may be informative in identifying candidate genes to understand the cognitive deficit.

Linking Anatomic and Behavioral Changes

WMS is associated with specific neuromorphologic and neurophysiologic findings. There is proportional sparing of frontal, limbic, and neocerebellar structures on magnetic resonance imaging (60). Abnormal functional organization of the neural systems that underlie language processing is revealed through studies using event-related potentials (61). Event-related potential studies suggest abnormal cerebral specialization for spared cognitive functions in WMS. The lack of uniformity in the cognitive, neuromorphologic, and neurophysiologic domains of WMS makes it a compelling model for elucidating the relationships among cognition, the brain, and, ultimately, the genes.

Another approach is to investigate anatomic changes in brain regions in WMS that may the result of gene deletions. In WMS, Galaburda and Bulligi found that the overall shape of the brain is not consistently abnormal (62), although in some cases abnormal brain shape is apparent. The most consistent anatomic finding is abnormal length of the central sulcus producing an unusual configuration of the dorsal central region. This includes the distal portion of the superior-parietal lobule and dorsal frontal gyrus. These regions may be linked to abnormal behavior in patients with WMS. Cytoarchitecture of WMS forebrain appears mostly normal, although subtle dysplastic changes are noted. Abnormal neuronal size of cortical neurons was suggested in one region and may be linked to increased subcortical connectivity. Elastin does not stain in the cerebellum, whereas Lim kinase does stain in cortical neurons.

Thus, in WMS, the link of neuroanatomy and behavior seems to fit a dorsal ventral dichotomy and not a frontal-caudal, left-right, or cortical subcortical dichotomy. Gallaburda and Bellugi proposed that the dorsal portions of the hemispheres, the frontal and parietal-occipital regions, may be involved (62). They noted that some language functions are preserved that are linked to ventral systems. Face recognition, also a ventral function, is preserved despite severe visuospatial dysfunction, a dorsal function. Anatomic findings also suggest possible involvement of the visually linked lateral nucleus of the amygdala. Galaburda and Bellugi speculated that this could be related to the lack of appropriate fear in WMS of new and unfamiliar faces, perhaps also threatening ones. Moreover, because this region may receive auditory projections, WMS subjects may not be sensitive to threatening voice and speech. Further work is needed at architechtonic and histologic levels to confirm sparing of ventral regions. To understand the linking of genes with neuroanatomy, it is necessary to find more genes with brain developmental effects. Of particular interest in this regard is the proposal that the region deleted in WMS may be a hotspot in mammalian brain evolution (51). Hagerman outlined a comprehensive approach to treatment of WMS (63).

ANIMAL MODELS: SIMULATIVE OR SUBSTITUTIVE

Part of "46 - Behavioral Phenotypes of Neurodevelopmental Disorders: Portals into the Developing Brain "

Animal models may be used to elucidate critical brain mechanisms involved in disorders with behavioral phenotypes. Early animal models focused on the impact of traumatic events during the developmental period, as exemplified by the social isolation and chronic stress (learned helpless) models of depression (64). These animal models generally simulated rather than substituted for the disorder. Animal models have used pharmacologic challenges to study neurochemical mechanisms linked to aberrant behavior or have introduced transgenic mice as substitutive models of conditions with behavioral phenotypes. Examples of these models include pemoline models of stereotyped self-biting behavior in the rat (65), SNAP mutant mouse model of attention-deficit/hyperactivity disorder (66), and transgenic and knockout mouse models for fragile X syndrome (67 ,68), MAOA deficiency (69), and LND (70 ,71 and 72). These models may contribute to the understanding of psychopathology. However, despite genetic replication of a disorder in the mouse, the behavior may not be replicated, so even these animal models often simulate aspects of the condition and are not fully substitutive.

Molecular genetic techniques combined with techniques to manipulate the developing mouse embryo make it feasible to produce such genetic animal models. Embryonic stem cells are isolated from a pregnant mouse with identifiable coat color that acts as a donor. The embryonic stem cells are grown in cell culture and then are genetically modified with the insertion of genetic material or through mutation of endogenous genes. Modified embryonic stem cells are microinjected into a blastula that is isolated from another mouse that ordinarily has a different coat color. The blastula is then reimplanted into a female host mouse and develops *in utero*. The inserted stem cells are incorporated into the

developing fetuses, and progeny that contain genetically altered cells are chimeras that can be identified by their mosaic coat colors. As adults, these chimeras, in which genetically modified cells have been involved in the establishment of the germ cell line, may then transmit the altered gene to their own offspring. It takes several generations to produce an affected animal, by using these embryonic stem cell techniques that depend on whether the needed phenotype can be produced in the heterozygous, homozygous, or hemizygous condition.

To illustrate these animal models, we may contrast animal models of LND based on the use of the neurotoxin 6-hydroxydopamine (6-OHDA), transgenic mouse models of LND, calcium channel blocker models of self-injury, and the mutant mouse model of fragile X syndrome.

Lesch–Nyhan Disease 6-OHDA (Rat Model)

After the demonstration of dopamine deficiency in striatum in autopsies of brain in three human patients with LND (35), Breese et al. administered 6-OHDA to neonatal and adult rats to test the effects of dopamine depletion in an animal model (73 ,74). These authors demonstrated that the age at which neural function is disrupted is an important factor in the type of motor and behavioral symptoms observed after a neural insult to basal ganglia structures. They documented a relationship between dopaminergic supersensitivity and self-injurious behavior. Rats treated with 6-OHDA in the neonatal period demonstrated self-biting with mutilation when they were challenged as adults with L-DOPA or a D1 dopamine agonist, but no such self-injurious behavior was found in the adult rats treated with 6-OHDA. Because of the self-biting, the neonatal 6-OHDA–treated rat was proposed as a model for LND, and dopamine deficiency was linked to self-injury. In these studies, rats that were not HPRT deficient were given injections of 6-OHDA at 5 days of age to denervate basal ganglia regions. These brain regions developed supersensitive dopamine receptors. Self-biting was documented in the lesioned animals when they were challenged as adults with a dopamine agonist; however, untreated adult rats did not show this behavior.

HPRT-Deficient Mouse

For LND, molecular techniques were used to produce two HPRT-deficient strains of mice. One strain was produced by retroviral interruption of the human *HPRT* gene in the embryonic stem cells (72). Another model was produced through the selection of embryonic stem cells for spontaneous mutations in the *HPRT* gene (70 ,71). In both instances, the mouse strains produced had nondetectable levels of HPRT. However, neither strain showed the spontaneous behavioral abnormalities or neurologic presentation seen in patients with LND. Tests of both cognitive functions and motor functions were intact in these animals. Similar findings were documented in a double knockout that is HPRT/APRT deficient (75).

This HPRT-deficient transgenic mouse model of LND (70) may still contribute to our understanding of LND. Reductions in dopamine of 40% or more (76) have been documented (77) in these animals. Because of questions about strain differences, Jinnah et al. studied the caudate nucleus in five HPRT-deficient strains of mice and made comparisons to littermate controls (77). Reductions of dopamine and also of the dopamine transporter of 35% to 40% were found in these animals. These results indicate an abnormality in the dopamine system despite apparently normal spontaneous behavior.

The absence of behavioral changes in the HPRT-deficient mice was unexpected. Originally, it was thought that uricase, which is present in rodents but is not present in primates, may act in a protective manner to lessen behavioral manifestations because uric acid, which normally builds up in the blood in LND, and hypoxanthine, which accumulates in cerebrospinal fluid, would not do so in mice because of the presence of uricase. This explanation is consistent with the inability of treatment with allopurinol, a xanthine oxidase inhibitor that prevents the accumulation of uric acid, to improve the behavior disorder in patients with LND. However, the mice were still found to have reduced dopamine in brain (76). Thus, it is the consequence of the HPRT deficits on dopamine, and possibly other neurotransmitter systems, that leads to the behavior in humans. This mouse differs from the Breese rat model in that dopamine depletion is complete in the Breese rat model but only 40% to 50% reduced in the HPRT-deficient mouse. Dopamine depletion is substantially greater in human subjects (76) than in the mice; therefore, this difference may account for the differences in behavior.

Bay K 8644 Model of Self-Injurious Behavior

Another approach to study self-injury is the calcium channel blockers model proposed by Jinnah et al. (78). The L-type calcium channel agonist (+/-) Bay K 8644 causes motor abnormalities in adult mice. These authors showed that administration of this drug could also cause the self-injurious biting, particularly when it was given to young mice. Self-biting was provoked by injecting small quantities of (+/-) Bay K 8644 directly into the lateral ventricle of the brain, a finding indicating a central effect of the drug. Similar behaviors can be elicited by administration of another L-type calcium channel agonist, FPL 64176. The self-biting elicited by (+/-) Bay K 8644 can be inhibited by pretreating the mice with dihydropyridine L-type calcium channel antagonists such as nifedipine, nimodipine, or nitrendipine. Moreover, self-biting is not inhibited by nondihydropyridine

antagonists including diltiazem, flunarizine, or verapamil. The known actions of (+/-)Bay K 8644 as an L-type calcium channel agonist, the reproduction of similar behavior with another L-type calcium channel agonist, and the protection from this behavior that results from certain L-type calcium channel antagonists implicate calcium channels, and their possible association with neurotransmitter deficits, in the mediation of the self-biting behavior (78).

Fragile X Mouse Model

The mutant mouse model of fragile X syndrome demonstrates another use of an animal model for a neurogenetic disorder. Transgenic fragile X knockout mice were developed to provide an animal model to study the physiologic function of the fragile X gene *(FMR1)* and to understand better the clinical phenotype caused by the absence of the fragile X protein.

The fragile X mouse model demonstrates macroorchidism and cognitive, affective, and behavioral features similar to the human condition (79). In the Morris water maze test, the Fmr1 knockout mice learned to find the hidden platform nearly as well as the control animals, but they showed impaired performance after the position of the platform was modified. The fragile X knockout mouse exhibited subtle deficits in spatial learning but normal early-phase long-term potentiation.

Jin and Warren expanded these studies by examination of late-phase hippocampal long-term potentiation, the protein synthesis–dependent form of long-term potentiation, in the Fmrl knockout mice (43). Initially, they found that late-phase long-term potentiation was normal and proposed that either absence of fragile X mental retardation protein has no influence on long-term potentiation or that any such influence is too subtle to be demonstrated by this technique. Moreover, when they examined spatial learning in this knockout mouse using the hippocampus-dependent Morris water maze, near-normal performance was observed. However, because the knockout mouse strain they used differed from that used in the earlier investigations that did show learning deficits, their studies were repeated using the same mouse knockout line that showed the deficit. Now significant, but subtle, increased swim latencies on the Morris maze test in reversal trials were found to be in agreement with the earlier studies. Thus, strain differences among mouse strains influence the behavior in the Fmrl knockout phenotype. Because the finding were subtle, these authors chose to investigate a paradigm less dependent on hippocampal function, one using the conditional fear paradigm. In this paradigm, the knockout animals showed significantly less freezing behavior than their wild-type littermate with two types of stimuli, contextual and conditional fear stimuli. These researchers concluded that that amygdala dysfunction may also be involved in fragile X syndrome.

These examples from LND and fragile X syndrome illustrate how animal models may contribute to our understanding of behavioral phenotypes: self-biting in LND and learning and fear responses in fragile X syndrome.

CONCLUSION

Part of "46 - Behavioral Phenotypes of Neurodevelopmental Disorders: Portals into the Developing Brain "

The study of behavioral phenotypes in neurodevelopmental disorders demonstrates the complexity in mapping pathways from genes to cognition and complex behavioral phenotypes. Behavioral phenotypes occur in disorders with mendelian inheritance (LND) and nonmendelian inheritance (PWS/AS, FRX). An investigation of these syndromes demonstrates that recognition of the involved gene is only the first step. Identification of the involved protein and of its expression in brain is critical. To clarify the mechanism, the use of animal models, neuroanatomic study, brain imaging techniques, systems neuroscience, and detailed descriptions of behavior are needed. The study of partial variants of the disorder (LND, WMS), comparison of deletion versus UPD (PWS, AS), and the study of atypical subjects who exhibit some but not all features of the disorder (WMS) are essential in understanding developmental pathways. Moreover, a neurodevelopmental model is essential because brain modularity of function cannot be assumed. Animal models must be carefully chosen because genetic background may influence the expression of the disorder. Such models may be important in simulating aspects of the disorder, but they may not substitute for the human condition. Flint proposed that success in the study of behavioral phenotypes requires a screen for regions of monosomy, the use of a sophisticated battery of neuropsychological and behavioral tests to describe the phenotype, a transcript map to identify quickly the genes that are likely to affected by the deletion, and a way to of deciding which genes are dosage sensitive (5). These are challenges that lie ahead as we continue to investigate behavioral phenotypes as portals to understanding the developing brain.

REFERENCES

1. Down JL. *Mental affectations of childhood and youth*. London: JA Churchill, 1887.

2. Gunn, P, Berry P, Andrews RJ. The temperament of Down’s syndrome in infants: a research note. *J Child Psychol Psychiatry* 1981;22:189–194.

3. Gath A, Gumley D. Behavior problems in retarded children with special reference to Down’s syndrome. *Br J Psychiatry* 1986;149:156–161.

4. Critchley M. Earl CJC. Tuberous sclerosis and allied conditions. *Brain* 1932;55:311–346.

5. Flint J. Behavioral phenotypes: conceptual and methodological issues. *Am J Med Genet* 1998;81:235–240.

6. Bax M. *Behaviours seen in children with mucopolysaccharidosis in behavioural phenotypes*. Welshpool, UK: Abstracts and Syndrome Information Publications, Society for the Study of Behavioural Phenotype, 1990.

7. O'Brien G. Behavioural phenotypy in developmental psychiatry. *Eur J Child Adolesc Psychiatry* 1992;1[Suppl]:1–61.

8. Nyhan W. Behavioral phenotypes in organic genetic disease: presidential address to the Society for Pediatric Research, May 1, 1971. *Pediatr Res* 1972;6:1–9.

9. Flint J. Annotation: behavioural phenotypes. A window onto the biology of behaviour. *J Child Psychol Psychiatry* 1996;37:355–367.

10. Flint J. The genetic basis of cognition. *Brain* 1999;122:2015–2031.

11. Skuse DH. Behavioural phenotypes: what do they teach us? *Arch Dis Child* 2000;82:222–225.

12. Tully T. Discovery of genes involved with learning and memory: an experimental synthesis of Hirshian and Benzerian perspectives. *Proc Natl Acad Sci USA* 1996;93:13460–13467.

13. Fischer E, Scambler P. Human haploinsufficiency: one for sorrow, two for joy. *Nat Genet* 1994;2:142–155.

14. Nyhan W. In: O'Brien G, Yule W, eds. *Behavioural phenotypes*. Cambridge: Cambridge University Press, 1995:ix.

15. Harris JC. Behavioral phenotypes in mental retardation syndromes. In: Barrett R, Matson J, eds. *Advances in developmental disorders*, vol 1. New York: Jai Press, 1987:77–106.

16. Harris J. Introduction to behavioral phenotypes. In: Harris J, ed. *Assessment, diagnosis and treatment of the developmental disorders*. New York: Oxford University Press, 1998:245–249.

17. Rett Syndrome Diagnostic Criteria Work Group. Diagnostic criteria for Rett syndrome. *Ann Neurol* 1988;23:425–428.

18. Flint J, Yule W. Behavioural phenotypes. In: Rutter M, Taylor E, Hersov L, eds. *Child and adolescent psychiatry*, third ed. Oxford: Blackwell Scientific, 1994:666–687.

19. Harris J. Lesch–Nyhan disease. In: Harris J, ed. *Assessment, diagnosis and treatment of the developmental disorders*. New York: Oxford University Press, 1998:306–318.

20. Nishino S, Ripley B, Overeem S, et al. Hypocretin (orexin) deficiency in human narcolepsy. *Lancet* 2000;355:39–40.

21. Brunner HG, Nelen MR, van Zandvoort P, et al. X-linked borderline mental retardation with prominent behavioral disturbance: phenotype, genetic localization, and evidence for disturbed monoamine metabolism.*Am J Hum Genet* 1993;52:1032–1039.

22. O'Brien G. *Behavioural measurement in mental handicap: a guide to existing schedules*. Oxford: Society for the Study of Behavioural Phenotypes, 1991.

23. Ikonomidou C, Bittigau P, Ishimaru MJ, et al. Ethanol-induced apoptotic neurodegeneration and fetal alcohol syndrome. *Science* 2000;287:1056–1058.

24. Hill AL. Savants: mentally retarded individuals with special skills. *Int Rev Res Ment Retard* 1978;9:277–298.

25. Seegmiller JE, Rosenbloom FM, Kelley WN. Enzyme deficit associated with a sex-linked human neurological disorder and excessive purine synthesis. *Science* 1967;155:1682.

26. Eads JC, Scapin G, Xu Y, et al. The crystal structure of human hypoanthine-guanine phosphoribosyltransferase with bound GMP. *Cell* 1994;78:325–334.

27. Page T, Nyhan WL. The spectrum of HPRT deficiency: an update. *Adv Exp Med Biol* 1989;253A:129–132.

28. Anderson L, Ernst M. Self-injury in Lesch–Nyhan disease. *J Autism Dev Disord* 1994;24:67–81.

29. Lloyd KG, Hornykiewicz O, Davidson, L et al. Biochemical evidence of dysfunction of brain neurotransmitters in the Lesch–Nyhan syndrome. *N Engl J Med* 1981;305:1106–1111.

30. Wong DF, Harris JC, Naidu, S et al. Dopamine transporters are markedly reduced in Lesch–Nyhan disease *in vivo. Proc Natl Acad Sci USA* 1996;93:5539–5543.

30a. Harris JC, Wong DF, Jinnah HA, et al. Dopamine transporter binding of WIN 35,428 correlates with HPRT level and extent of movement disorder but not with self-injurious behavior. *Abstr Soc Neurosci* 1999;25.

31. Shretelen D, Harris J, Park K. Cognitive function in Lesch–Nyhan disease (in press).

32. Nicholls RD, Saitoh S, Horsthemke B. Imprinting in Prader–Willi and Angelman syndromes. *Trends Genet* 1998;14:194–200.

33. State MW, Dykens EM. Genetics of childhood disorders: XV. Prader–Willi syndrome: genes, brain, and behavior. *J Am Acad Child Adolesc Psychiatry* 2000;39:797–800.

34. Moncla A, Malzac P, Livet MO, et al. Angelman syndrome resulting from UBE3A mutations in 14 patients from eight families: clinical manifestations and genetic counselling. *J Med Genet* 1999;36:554–560.

35. Swaab DF, Purba JS, Hofman MA. Alterations in the hypothalamic paraventricular nucleus and its oxytocin neurons (putative satiety cells) in Prader–Willi syndrome: a study of five cases. *J Clin Endocrinol Metab* 1995;80:573–579.

36. Holm VA, Pipes PL. Food and children with the Prader–Willi syndrome. *Am J Dis Child* 1976;130:1063–1067.

37. Stein DJ, Keating J, Zar H. Compulsive and impulsive symptoms in Prader–Willi syndrome. In: *Abstracts in new research (NR33)*. San Francisco: Annual Meeting of the American Psychiatric Association, 1993.

38. State MW, Dykens EM, Rosner B, et al. Obsessive-compulsive symptoms in Prader–Willi and "Prader–Willi-Like" patients. *J Am Acad Child Adolesc Psychiatry* 1999;38:329–334.

39. Roof E, Stone W, MacLean W, et al. Intellectual characteristics of Prader–Willi syndrome: comparison of genetic subtypes. *J Intellect Disabil Res* 2000;44:25–30.

40. Moncla A, Malzac P, Voelckel MA, et al. Phenotype-genotype correlation in 20 deletion and 20 non-deletion Angelman syndrome patients. *Eur J Hum Genet* 1999;7:131–139.

41. Holm VA, Sulzbacher S, Pipes PL. *The Prader–Willi syndrome*. Baltimore: University Park Press, 1981.

42. Hagerman R. Clinical and molecular aspects: fragile X syndrome. In: Tager-Flusberg H, ed. *Neurodevelopmental disorders*. Cambridge, MA: MIT Press, 1999:27–42.

43. Jin P, Warren ST. Understanding the molecular basis of fragile X syndrome. *Hum Mol Genet* 2000;9:901–908.

44. Miller WJ, Skinner JA, Foss GS, et al. Localization of the fragile X mental retardation 2 (FMR2) protein in mammalian brain. *Eur J Neurosci* 2000;12:381–384.

45. Wolff PH, Gardner J, Paccia J, et al. The greeting behavior of fragile X males. *Am J Ment Retard* 1989;93:406–411.

46. Witt RM, Kaspar BK, Brazelton AD, et al. Developmental localization of fragile X mRNA in rat brain [Abstract 293.6]. *Soc Neurosci Abstr* 1995;21:1.

47. Hinton VJ, Brown WT, et al. Analysis of neocortex in three males with fragile X syndrome. *Am J Med Genet* 1991;17:123–131.

48. Hagerman RJ, Cronister AC, eds. *The fragile X syndrome: diagnosis, treatment, and research*, second ed. Baltimore: Johns Hopkins University Press, 1996.

49. Williams JCP, Barratt-Boyes BG, Lowe JB. Supravalvular aortic stenosis. *Circulation* 1961;24:1311–1318.

50. Bellugi U, Lichtenberger L, Jones W, et al. The neurocognitive profile of Williams syndrome: a complex pattern of strengths and weaknesses. *J Cogn Neurosci* 2000;12[Suppl]:7–29.

51. Korenberg JR, Xiao-Ning C, Hirota H, et al. Genome structure and cognitive map of Williams syndrome. *J Cogn Neurosci* 2000;12[Suppl]:89–107.

52. Jones W, Bellugi U, Lai Z. Hypersociability in Williams syndrome. *J Cogn Neurosci* 2000;12[Suppl]:30–46.

53. Howlin P, Davies M, Udwin O. Cognitive functioning in adults with Williams syndrome. *J Child Psychol Psychiatry* 1998;39:183–189.

54. Paterson SJ, Brown JH, Gsodl MK, et al. Cognitive modularity and genetic disorders. *Science* 1999;286:2355–2358.

55. Bishop DV. Perspectives: cognition. An innate basis for language? *Science* 1999;286:2283–2284.

56. Karmiloff-Smith A, Grant J, Berthoud I, et al. Language and Williams syndrome: how intact is “intact”? *Child Dev* 1997;68:246–262.

57. Tager-Flusberg H, Boshart J, Baron-Cohen S. Reading the windows to the soul: evidence of domain-specific sparing in Williams syndrome. *J Cogn Neurosci* 1998;10:631–639.

58. Frangiskakis JM, Ewart AK, Morris CA, et al. LIM-kinase l hemizygosity implicated in impaired visuospatial constructive cognition. *Cell* 1996;86:5969.

59. Tassabehji M, Metcalfe K, Karmiloff-Smith A, et al. Williams syndrome: use of chromosomal microdeletions as a tool to dissect cognitive and physical phenotypes *Am J Hum Genet* 1999;64:118–125.

60. Reiss AL, Eliez S, Schmitt E, et al. Neuroanatomy of Williams syndrome: a high resolution MRI study. *J Cogn Neurosci* 2000;12[Suppl]:65–73.

61. Mills DL, Alvarez TD, St. George M, et al. Electrophysiological studies of face processing in Williams syndrome. *J Cogn Neurosci* 2000;12[Suppl]:47–64.

62. Galaburda AM, Bellugi U. Multi-level analysis of cortical neuroanatomy in Williams syndrome. *J Cogn Neurosci* 2000;12[Suppl]:74–88.

63. Hagerman RJ. *Neurodevelopmental disorders: diagnosis and treatment*. In: Harris J, series ed. *Developmental perspectives in psychiatry*. New York: Oxford University Press, 1999.

64. Harris J. Experimental animal modeling of depression and anxiety. *Psychiatr Clin North Am* 1989;12:815–836.

65. King BH, Cromwell HC, Lee HT, et al. Dopaminergic and glutamatergic interactions in the expression of self-injurious behavior. *Dev Neurosci* 1998;20:180–187.

66. Hess EJ, Collins KA, Wilson MC. Mouse model of hyperkinesis implicates SNAP-25 in behavioral regulation. *J Neurosci* 1996;16:3104–3111.

67. D’Hooge R, Nagels G, Franck F, et al. Mildly impaired water maze performance in male Fmr1 knockout mice. *Neuroscience* 1997;76:367–376.

68. Fisch GS, Hao HK, Bakker C, et al. Learning and memory in the FMR1 knockout mouse. *Am J Med Genet* 1999;84:277–282.

69. Cases O, Lebrand C, Giros B, et al. Plasma membrane transporters of serotonin, dopamine, and norepinephrine mediate serotonin accumulation in atypical locations in the developing brain of monoamine oxidase A knock-outs. *J Neurosci* 1998;18:6914–6927.

70. Hooper M., Hardy K, Handyside A, et al. HPRT-deficient (Lesch–Nyhan) mouse embryos derived from germ line colonization by cultured cells. *Nature* 1987;326:292–295.

71. Williamson DJ, Hooper ML, Melton DW. Mouse models of hypoxanthine phosphoribosyltransferase deficiency. *J Inherit Metab Dis* 1992;15:665–673.

72. Kuehn MR, Bradley A, Robertson EJ, et al. A potential animal model for Lesch–Nyhan syndrome through introduction of HPRT mutations into mice. *Nature* 1987;326:295–298.

73. Breese GR, Baumeister AA, McCown TJ. Behavioral differences between neonatal and adult-6-hydroxydopamine treated rats to dopamine gonists: relevance to neurological symptoms in clinical synromes with reduced brain dopamine. *J Pharmacol Exp Ther* 1984;231:343–354.

74. Breese GR, Criswell E, Duncan GE. A dopamine deficiency model of Lesch–Nyhan disease: the neonatal-6-OHDA–lesioned rat. *Brain Res Bull* 1990;25:477–484.

75. Engle SJ, Womer DE, Davies PM, et al. HPRT-APRT-deficient mice are not a model for Lesch–Nyhan syndrome. *Hum Mol Genet* 1996;5:1607–1610.

76. Jinnah HA, Gage FH, Friedmann T. Animal models of Lesch–Nyhan syndrome. *Brain Res Bull* 1990;25:467–475.

77. Jinnah HA, Jones MD, Wojcik BE, et al. Influence of age and strain on striatal dopamine loss in a genetic mouse model of Lesch–Nyhan disease. *J Neurochem* 1999;72:225–229.

78. Jinnah HA, Yitta S, Drew T, et al. Calcium channel activation and self-biting in mice. *Proc Natl Acad Sci* 1999;96:15228–15232.

79. Kooy RF, D’Hooge R, Reyniers E. Transgenic mouse model for the fragile X syndrome. *Am J Med Genet* 1996;64:241–245.

SUGGESTED READINGS

Harris J. Behavioral phenotypes. In: Harris J, ed. *Assessment, diagnosis and treatment of the developmental disorders*. New York: Oxford University Press, 1998:251–376.

O'Brien G, Yule W, eds. *Behavioural phenotypes*. Cambridge: Cambridge University Press, 1995.

Section VI
Schizophrenia and related disorders

Daniel R. Weinberger

Schizophrenia and related disorders - Introduction

The schizophrenia section of the *Fifth Generation* reflects both a subtle shift in emphasis in schizophrenia research and a refinement of earlier trends. Molecular and systems neuroscience and genetics are the basic biological sciences of psychiatry, and their increasingly important role in research is reflected in the schizophrenia section. Schizophrenia research is inexorably on a path from clinical phenomenology to cellular and molecular pathogenesis. Although this complex disease will never be reducible to one gene or one signal-processing pathway, the molecular origins of risk and of manifest psychopathology are being elucidated. The main topics covered in the *Fourth Generation* volume, including treatment of acute psychosis, mechanisms of antipsychotic drug action, basic models of pathophysiology, application of brain imaging technologies, and efforts to characterize symptoms in terms of failures of specific neuronal systems, are represented here, but the presentations are more scientifically mature. Consistent with the completion of the reference sequence for the human genome and the increasing sophistication of approaches to complex genetic disorders such as schizophrenia, clinical and molecular genetics are now reviewed as a new chapter in this section. Because of the introduction of more effective and much better tolerated antipsychotic drugs, the economics of treating schizophrenia has changed, and this topic also is now covered as a new chapter. Finally, a new chapter is devoted to the issue of developing informative animal models of schizophrenia. The traditional view that a disorder of human perception, cognition, and comportment could not be modeled in the experimental animal has yielded in the face of heuristic models of neurochemical, physiological, and molecular aspects of schizophrenia. These models promise to open new pathways for drug discovery and to test specific hypotheses about etiology and pathophysiology.

Neuroimaging has been a popular tool for asking specific questions about brain anatomy and function that may help to unravel the pathophysiology of the illness. There is no doubt, after more than 2 decades of neuroimaging research in schizophrenia, that the schizophrenic brain is abnormal in a variety of experimental domains. The details of the abnormalities are still to be fully clarified, but the level of analysis has been meaningfully elevated. New technologies in brain imaging have allowed testing of much more sophisticated hypotheses. Magnetic resonance imaging has replaced positron emission tomography as the primary functional mapping tool, and new approaches to *in vivo* chemistry with nuclear medicine resonance spectroscopy and to anatomic tract tracing with diffusion tensor imaging have revealed abnormalities not approachable with earlier methods. Positron emission tomography has become primarily a neurochemical assay technique. Positron emission tomography imaging of dopamine systems has refined our understanding of the role of dopamine in the pathophysiology of a symptomatic episode and of the mechanism of an antipsychotic response.

Progress in schizophrenia research has followed discoveries in the basic neurosciences. This is illustrated in several chapters that address the development and plasticity of the prefrontal cortex and the importance of intrinsic prefrontal circuitry and prefrontocentric distributed networks to the pathophysiology of schizophrenia. The study of postmortem

tissues has been revived by the availability of cellular and molecular assay tools that permit surveys of gene and protein expression at an industrial scale. Numerous provocative molecular findings have emerged from these preliminary forays into the cell biology of the illness. Although the mystery of this devastating illness is still unsolved, the clues that have emerged in the past 5 years and the opportunities for discovery that now exist make it very likely that much more of schizophrenia will be understood at the cellular level in the next 5 years.

47

Schizophrenia: Course Over the Lifetime

Philip D. Harvey

Michael Davidson

Philip D. Harvey: Mount Sinai School of Medicine, New York, New York.

Michael Davidson: Sheba Medical Center, University of Tel Aviv, Tel Aviv, Israel.

Among the lifelong remitting and relapsing illnesses, the course of schizophrenia is among the most widely debated. At the core of the debate are the following questions:

1. What is the best way to investigate the course of schizophrenia?
2. What are the manifestations preceding and shortly after the first psychotic episode?
3. How do the manifestations of the illness, both clinical and biological, change over the life span, especially during later life and senescence?
4. Should schizophrenia be conceptualized as the response of a stable encephalopathy to different stages of the life cycle (1), as a progressive, degenerating disease (2), or as a hybrid of the two concepts (3)?

This chapter attempts to provide a critical assessment of these questions in light of the latest empiric data and current conceptualization of this disease.

The results of the major studies on the course of the illness over 20 to 40 years of follow-up are consistent in reporting a chronic, generally persistent course of illness for 50% to 70% of the patients who receive an initial diagnosis of schizophrenia (4 ,5 ,6 ,7 ,8 and 9). However, a more careful examination of the reports reveals marked heterogeneity in course both between and within cohorts (10 ,11 and 12). The reason for this heterogeneity may be that different studies have examined widely diverse samples of subjects and may also be related to the different definitions of what constitutes a good outcome. These definitions range from disease-free for the majority of life to simply not floridly psychotic at the time of last assessment (13). Very few of these studies included elderly patients in their samples or accounted for attrition, and even fewer examined longitudinal biological changes. This is unfortunate because accurate information on the course of schizophrenia is essential to plan the delivery of care, to evaluate treatment effectiveness, to provide information to newly diagnosed patients and their families, and to advance schizophrenia research. The paucity of data on the course of schizophrenia is mostly the result of limitations inherent in studying a relatively low-incidence illness of unknown origin and pathophysiology, with an insidious onset and a course affected by a multitude of personal and social factors.

WHAT IS THE BEST WAY TO INVESTIGATE THE COURSE OF SCHIZOPHRENIA?

Part of "47 - Schizophrenia: Course Over the Lifetime "

Study Design

The ideal way to determine the course of schizophrenia is to follow a randomly sampled birth cohort throughout the entire age of risk for schizophrenia and then continue to follow the incident cases, and appropriately selected controls, through the entire life span. A related but less informative strategy is to follow-up apparently healthy persons hypothesized to be at high-risk of schizophrenia such as first-degree relatives of affected persons (14 ,15 ,16 and 17). An alternative strategy is the prospective follow-up of patients from the first time they seek help for psychosis (18 ,19 ,20 ,21 ,22 ,23 ,24 ,25 ,26 ,27 ,28 and 29). Unfortunately, the birth-cohort strategy is impractical because schizophrenia is a very low-incidence disease (.87%) (30), the age of risk spans over more than 4 decades of life, and the age of risk appears different for males and females. Thus, following a birth cohort of 10,000 individuals for 40 years, starting at age 5 years, would detect approximately 90 cases of schizophrenia (not accounting for attrition), a number that is insufficient to make any statement regarding the course of a heterogeneous syndrome such as schizophrenia. Similarly, the high-risk strategy is limited in scope because it excludes most future patients with schizophrenia who do not have affected first-degree relatives, in addition to the problems of investing in a very large research effort for a

relatively small number of persons who will convert into cases.

Therefore, the most often employed strategy to map out the course of schizophrenia has been to start the follow-up only after a diagnosis of psychosis is established (first psychotic episode cohorts). However, these are selected cohorts in that they include only persons who seek help (often in an academic center) and consent to participate in research (31). Furthermore, this strategy does not provide prospectively collected information on events preceding the first psychotic episode. Moreover, it is conceivable that some of the patients recruited for the first-episode studies have been chronically ill for several years but undiagnosed (32 ,33). This, in turn, could explain the discrepancies between studies finding differences between first-episode patients and patients hospitalized on a long-term basis and other studies that do not find such differences (34).

Regardless of the study design, all prospectively followed schizophrenic cohorts will be characterized by high attrition. There are many reasons for attrition: lack of insight, disappointment with the care received, recovery accompanied by the wish to forget and conceal the experience of illness, and being too sick to maintain contact. Whether following-up 100% of the cohort would find the outcome to be better, worse, or the same (and in which aspect) remains unclear. Even if the debate on the best way to follow patients to describe the course of the illness could be settled, and the attrition rate could be reduced (both of which are unlikely), it still is not clear what defines a case of schizophrenia and therefore who should be followed to elucidate the long-term course of illness.

What Is a Case of Schizophrenia?

The debate on the nosologic boundaries of schizophrenia is as old as the term itself, and the last hundred years of research have done little to settle this debate (35 ,36 ,37 ,38 and 39). On the contrary, the pendulum has swung back and forth between a discrete nuclear definition of schizophrenia based mostly on psychotic features and severely impaired functioning and a continuum that includes questionable psychotic manifestations, schizophrenia spectrum personality disorders, and moderate to severe deviations from normality based on psychometric indicators (32 ,40).

Despite the contribution of the modern diagnostic classifications to the definition of schizophrenia, the abandonment of the continuum-based concept and adoption of the dichotomous model have not occurred, largely because the continuum model has received considerable pathophysiologic support. On the contrary, the schizophrenia-nonschizophrenia distinction is a matter of operational convenience brought about by treatment and economic developments emerging since the 1960s. The emergence in the 1950s of antipsychotic drug treatments that ameliorated psychosis while producing severe adverse effects called on the medical community to distinguish between patients who should and who should not be treated with these medications. Throughout the 1980s, as the accounting between providers of health care and health insurance organizations was becoming more thoughtful, the latter began to demand definitions of which patients were entitled to reimbursement and which were not. Finally, clinical investigators into the biology of schizophrenia also supported a model clearly distinguishing between schizophrenia and nonschizophrenia as more amenable to research. Needless to say, the course of illness of a cohort of schizophrenic patients depends on the definition of the cases enrolled in the cohort (41). Hence, until objective biological markers can be combined with phenomenologic criteria to define a case, the question of the course of “exactly what illness?” will continue to be raised.

It would also be reasonable to assume that regardless of the degree of the cohort’s heterogeneity, part of the variability in the course of illness is determined by the interaction between the affected individual and a wide array of societal, familial, and personal interactive influences (42 ,43 ,44 ,45 ,46 ,47 and 48). A few of these influences can be captured by careful collection of demographic and treatment information. However, the effects of changes that occur over many decades, such as changes in health care delivery, changes in the public perception of severe mental illness, and the interactions between these changes and aging, may not be amenable to survey and quantification. For example, how will deinstitutionalization, reduction of stigma, intensive community care, managed care, novel neuroleptics, open international borders and resultant migration, and the influence of advocacy groups affect the definition and course of schizophrenia?

In summary, the inherent limitations of studying birth and high-risk cohorts, coupled with the observation that many of the dynamic changes occur over a time span of 3 to 5 years before and immediately after the first diagnosed episode of psychosis, have been the impetus for the proliferation of first-episode studies in the 1990s. These studies can provide some useful information about schizophrenia, particularly because most patients experiencing schizophrenic symptoms in Western societies are likely to be diagnosed and treated at least once by mental health professionals.

WHAT ARE THE MANIFESTATIONS PRECEDING AND SHORTLY FOLLOWING THE ONSET OF THE FIRST PSYCHOTIC EPISODE?

Part of "47 - Schizophrenia: Course Over the Lifetime "

Premorbid Phase

The observation that that some schizophrenic patients have premorbid abnormalities dates back to Bleuler (49). History taken on the first contact with a mental health professional

often reveals subtle or flagrant motor, cognitive, emotional, and behavioral deviations during childhood, social withdrawal and mood and personality changes during adolescence, and attenuated psychotic symptoms several months to several years before the first treatment contact and the diagnosis of psychosis (51 ,52 ,53 ,54 ,55 ,56 ,57 ,58 ,59 ,60 ,61 and 62). The period immediately preceding the onset of psychosis, during which behavior and functioning deteriorate from a stable, premorbid level of functioning, as well as the behavioral changes that identify it, is referred to as the *prodrome*. However, the factors that precipitate the transition from prodrome to the first incident of help seeking and the resultant diagnosis are not necessarily distinctly related to the illness itself.

Factors such as the educational level of patients and their families, socioeconomic status, and availability of health care may all determine when the first contact occurs (63 ,64 ,65 ,66 ,67 and 68). Moreover, events such as the sudden unavailability of a caregiver able to maintain a highly symptomatic patient in the community or any change in the threshold of abnormal behavior tolerated by the community can precipitate treatment contact, hospitalization, and diagnosis. Hence, the presence of the premorbid manifestation, the onset of the prodrome, the emergence of the symptoms that define an episode of the illness, and ascertainment of the full syndrome of illness including formal diagnosis do not necessarily coincide and are not always clearly distinct points in time (31). Methods employed to investigate the phenomena preceding the first contact for help and the diagnosis of schizophrenia are the high-risk method, the birth-cohort method, and the historical prospective (or follow-back) method.

The high-risk studies that followed-up children and siblings of patients affected by schizophrenia into adulthood demonstrated that these relatives were more likely than the general population to be affected by emotional and behavioral abnormalities and abnormal psychophysiologic reactions (69 ,70 ,71 ,72 ,73 ,74 ,75 ,76 ,77 ,78 ,79 ,80 and 81). For instance, one study compared cognitive and behavioral assessments of twin pairs healthy at the time of testing and discordant for psychoses later on with twin pairs who both remained healthy. The healthy twin from the ill pair performed better than the ill co-twin but worse than the average of the twins from the healthy pair (82) (Fig. 47.1). Thus, abnormalities were found to be associated both with schizophrenia and with being a nonpsychotic identical twin of a schizophrenic patient. Even though the increased risk can be demonstrated in targeted populations, this strategy has not been completely successful in defining the premorbid aspects of schizophrenia. This is because most persons who belong to the high-risk groups represent a small, atypical subgroup of patients with schizophrenia and because of the relatively small number, approximately 10% to 15% at most (30), of high-risk persons who eventually develop schizophrenia.

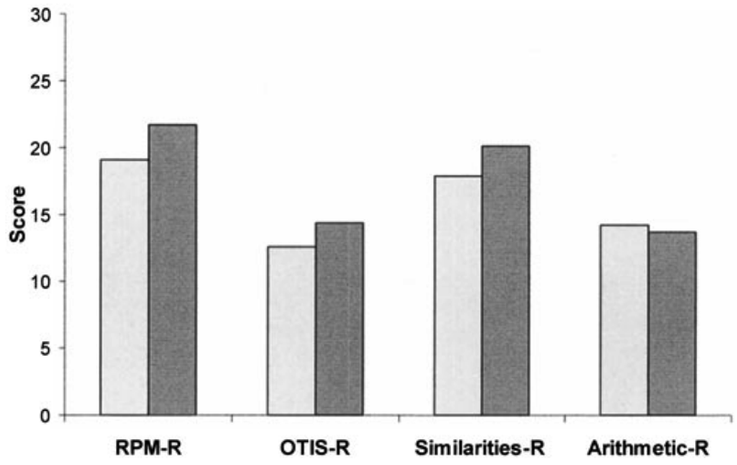


FIGURE 47.1. Intellectual functioning in members of twin pairs concordant and discordant for schizophrenia.

National health authorities have conducted follow-up studies of persons born in a geographically defined area over a specified period (birth cohort) to study protective and risk factors for healthy development and disease. Among the most publicized and complete studies are two British studies: the Medical Research Council National Survey of Development, covering all births during the week of 3 to 9 March 1946, and the National Child Development Study, covering all births during 3 to 9 March 1958 (83 ,84). Persons born during these 2 weeks were interviewed and assessed, together with their parents, several times during early childhood and adolescence. Developmental and scholastic achievement data collected on these cohorts were later linked to the data in a registry containing diagnoses of patients discharged from psychiatric hospitals. An overview of these studies indicates that, as a group, persons with future schizophrenia cases had delayed developmental milestones, speech and behavioral difficulties, and lower IQ scores compared with noncases (individuals who did not appear in the psychiatric registry). Although future cases were overrepresented in the lowest third of the IQ scores, these future cases had scores that were distributed over the entire range. The decline in IQ was not limited to a particular test, and the magnitude of decline ranged between 0.25 and 0.75 standard deviations (SD). Thus, the level of performance seen was not necessarily even outside the average range of IQ scores (defined as IQs between 90 and 110, which is 0.67 SD above or below the average score of 100).

Follow-back or historical prospective studies examine the archival premorbid histories of individuals who are already diagnosed as suffering from schizophrenia. They can be based on the linkage of databases containing routine psychometric tests administered by educational or military authorities to large numbers of healthy adolescents with national psychiatric registries. This strategy takes advantage of large-scale, readily available data enabling the testing of hypotheses with high statistical power. The disadvantage of the strategy is that, like birth-cohort studies, the data contained in the archival assessments are not aimed at the detection

of schizophrenia or its premorbid manifestations, which may be responsible for the low predictive specificity found in many of these studies. Several follow-back studies have produced results very similar to the birth-cohort studies, findings confirming, both quantitatively and qualitatively, the cognitive and behavioral abnormalities of future schizophrenic patients (85).

For instance, one study based on a national population of adolescents called by the nonselective Israeli Draft Board revealed that apparently healthy persons who several years later developed schizophrenia had lower mean group scores than their healthy classmates by about 1 SD on items reflecting social adjustment and IQ (53). The differences derived from a “shift to the left” of the future patients, one that was clearly more pronounced on social adjustment than on IQ.

Despite the consistency between the studies’ results, their interpretation remains uncertain. The premorbid signs of the illness are widely variable, and a single “typical prodrome” cannot be identified. For example, for some persons, the premorbid manifestations consist of shyness detectable in elementary school, many years before the manifestation of psychosis. For others, the premorbid manifestations consist of IQ scores 0.67 SD lower than expected, detected in adolescence, or in nonpsychotic paranoid thoughts manifested several years before the first psychotic episode in persons with unimpaired IQ. Yet for others, the premorbid manifestations consist of withdrawn behavior and depressed mood preceding psychosis only by a few months. Furthermore, for some patients, the prodrome is manifested as a crescendo of progressive, continuous deterioration during childhood and adolescence and for others as the barely detectable presence of a few minor cognitive abnormalities. Finally, it is possible that some of the variability in the quality and time of manifestations of premorbid manifestations reflects limitations of the study designs, which are often cross-sectional assessments (34). It is conceivable that a true prospective follow-up study, specifically designed to detect signs of premorbid schizophrenia and conducted from birth through age of risk, would reveal that the same person who manifests mild delay in developmental milestones as a toddler (56), shyness and learning difficulties in elementary school (50 ,52), restricted peer interaction as a teenager (86), and depressed mood and unusual thoughts in adolescence (87) would have psychosis in early adulthood (30). Alternatively, a particular premorbid manifestation could lead to a particular subtype of schizophrenia (86). It is uncertain whether these various premorbid or prodromal manifestations, which differ in quality, severity, and time of onset, bear the same relation to the first psychotic exacerbation or to the course of the schizophrenic illness.

Despite these uncertainties and even though with current psychometric tools, premorbid abnormalities are detected only in a few persons with future schizophrenia, their presence has opened up both conceptual and practical lines of investigation. Conceptually, it would be interesting to explain the pathophysiologic relationship between the premorbid symptoms and the manifestation of the illness. Practically, it would be helpful if the prodrome could be developed into a reliable predictor of future illness, based on which a secondary prevention strategy could be implemented.

Because the clinical manifestation of schizophrenia could represent an accumulation of genetic and environmental risk factors (or lack of environmental protective factors), the premorbid abnormalities, particularly the early-life ones, could be conceptualized as markers of vulnerability. This is consistent with a “multiple-hit” hypothesis by which, in addition to the genetic and environmental factors that have led to the premorbid manifestations, an environmental insult or a gene expressed later in life may be necessary to develop the full syndrome of schizophrenia. A corollary hypothesis would suggest that, depending on the additional, later insults, the same early-life manifestations (e.g., avoidant personality traits) could remain stable through life with no pathologic implication, could evolve into milder mental disorders such as a schizophrenia spectrum personality disorder, or could lead to schizophrenia. If indeed the phenotype of schizophrenia reflects the consequences of an accumulation of genetic and environmental risk factors, studying the course of the disease from birth through the end of the age of risk may be required to identify specific etiologic patterns. Alternatively, it is possible that a subgroup of these persons who manifest certain premorbid abnormalities may be inevitably destined to manifest schizophrenia in the future, and for these, and only these persons, the prodromal manifestations are obligatory precursors of the illness.

Is Secondary Prevention a Realistic Goal?

From a practical point of view, it would be tempting to use the occurrence of the premorbid and prodromal manifestations of the illness to identify persons at imminent risk of developing schizophrenia and to intervene before the onset of the first psychotic episode, in an attempt to delay or ameliorate it (47 ,57 ,88 ,89 ,90 ,91 ,92 and 93). It would be reasonable to hypothesize that any intervention that would delay or attenuate the first psychotic episode would have a major impact on the long-term outcome of the illness. This idea draws support from studies indicating that patients with shorter duration of untreated psychosis have more rapid symptomatic remission and may incur less deterioration in the long run (94 ,95 and 96).

However, the relatively low specificity of the premorbid symptoms such as subtle cognitive deficits, poor social adjustment, changes in personality, and depressed mood has given rise to concerns that an excessive number of persons could be exposed unnecessarily to the stigma of a provisional diagnosis of severe mental illness. Although it is possible to improve the specificity of prediction, for example, by

targeting only persons at very high risk (e.g., first-degree relatives of schizophrenic patients who also manifest putatively prodromal symptoms), this strategy would exclude the 90% of future patients who do not have an affected relative. Furthermore, even if the prediction could be improved, it is not certain that effective prevention exists. Antipsychotic drugs proven to reduce symptoms and to prevent exacerbation in patients who already experienced psychosis may or may not be effective in delaying the onset of psychosis.

Moreover, the notion that psychosis exerts a toxic effect on the brain and that longer duration of untreated psychosis should result in worse outcome has been challenged (20 ,97). It has been argued that (a) duration of untreated psychosis cannot be accurately assessed, (b) the delay in requesting and obtaining treatment is not the cause of a worse outcome but the result of an insidious-onset illness that is a more severe form, and (c) long duration of persistent untreated psychosis and persistence of psychosis despite treatment both reflect the same psychosis-severity phenomena without proving a causal relationship between the two. Finally, the proof that the duration of untreated psychosis correlates with the more relevant indices of outcome such as quality of life or overall illness outcome is still equivocal.

For all these reasons, the question of treating persons who are not yet floridly psychotic has stirred public debate beyond the professional community. Yet because of the potential benefits of secondary prevention on one hand and the risks and ethical implications associated with it on the other, it is essential to search for rational strategies to assess the risk-to-benefit ratio. Examining such ratios in an area where preventive measurements are already an accepted reality would be such a strategy. For example, even though after remission from the first psychotic episode, only 60% of drug-free patients have an exacerbation of their illness within the first year, 100% of patients are routinely treated with neuroleptics. Hence 40% are exposed to the adverse effects of neuroleptics, although they are not likely to experience a worsening of their symptoms. Similarly, seven families of schizophrenic patients must go through the effort, expense, and potential adverse effects of intensive family therapy for 1 year, to prevent relapse on the part of one of seven recently discharged patients with schizophrenia (98).

The dilemma of preventive treatment is not limited to psychiatry. For instance, approximately 70 elderly patients with moderate hypertension must be treated with antihypertensive drugs for 5 years to save one life, and 100 men with no evidence of coronary heart disease must be treated with aspirin for 5 years to prevent one heart attack (99). In a study using the number needed to treat method, which is the number of persons who need to receive treatment to prevent one bad outcome, it was calculated that one must administer antipsychotics to 35 adolescents with paranoid or schizotypal personality disorder for 1 to 3 years to delay hospitalization for schizophrenia by 6 months to 1 year in a single patient. This calculation assumes that approximately 5% of these adolescents will convert to schizophrenia, and it also assumes a 60% treatment success rate in delaying conversion, which is the same rate by which neuroleptics can induce extended remission in first-episode patients (100).

The early detection and treatment strategy is supported by preliminary results from a community clinic where youths with prodromal symptoms were treated with open-label neuroleptics plus supportive measures or supportive measures alone (101 ,102). The results indicated that more members of the neuroleptic-treated group were symptom-free for a longer period than similar youths given only supportive therapy or those who refused to enroll in the trial. In a different study, nonpsychotic, first-degree relatives of patients complaining mostly of cognitive deficit also were found to benefit from neuroleptic treatment (103). In summary, although there is much interest in the events leading to the first psychotic episode and a strong appeal for secondary prevention, the information currently available is still tentative. In contrast, much information and a few solid practical implications regarding the first episode of psychosis are known.

First Episode of Psychosis

Most studies of patients followed for 2 to 5 years after the first episode of illness have provided highly informative data regarding the early course of positive (3 ,34 ,105) and negative (19 ,104 ,106 ,107 and 108) symptoms, cognitive functioning (109 ,110 ,111 ,112 ,113 ,114 ,115 and 116), functional status (117 ,118 and 119), and response to treatment (95). Furthermore, some of the studies have described radiologic changes in the brain after the first episode (120 ,121 and 122).

Often, the appearance or worsening of psychotic symptoms constitutes the trigger for the first contact with a mental health professional and subsequent diagnosis and hospitalization. Hence, it is no wonder that what is described as the first episode of schizophrenia is dominated by the presence of positive symptoms, mostly fully formed delusions and hallucinations. Almost 90% of first-episode patients treated with neuroleptics experience a rapid, albeit transient, remission of their psychotic symptoms. Despite the good initial response to treatment, relapse with reoccurrence of psychotic symptoms is common. Predominance of negative symptoms and hebephrenic, catatonic presentations are not part of the characteristic presentation of the first episode. Occasionally, however, negative symptoms of insidious onset are present on the first episode, and the response of these symptoms to treatment is very limited. Cognitive deficits are common and relatively severe at the time of the first episode. Performance on most cognitive tests is approximately 1 SD below age- and education-adjusted expectations, with more that 50% of the first-episode patients performing even worse (123). The impairment affects almost

all aspects of cognition; however, specific areas of impairment are distributed unevenly. For example, deficits in memory, abstraction, and attention are more severe than deficits in verbal or perceptual skills (124). This impairment, measured on remission from the first episode, goes beyond the one-third to two-thirds SD deficit that characterizes the premorbid cognitive performance of the schizophrenic patients, and it raises the question whether it reflects a progressively deteriorating process. In a cross-sectional comparison of Raven Progressive Matrixes scores (a valid measure of IQ), it was found that apparently healthy adolescents closer to their first hospitalization for psychosis performed more poorly than adolescents who were tested several years before their first exacerbation, but better than patients whose disease had already exacerbated (125) (Fig. 47.2). Furthermore, cognitive performance appears to be slightly worse in patients with chronic disease (114) in comparison with first-episode patients, a finding providing indirect evidence of further cognitive deterioration beyond the first episode (110 ,111). In contrast to psychotic symptoms, cognitive functions are less responsive to the neuroleptic treatment administered for schizophrenia (126).

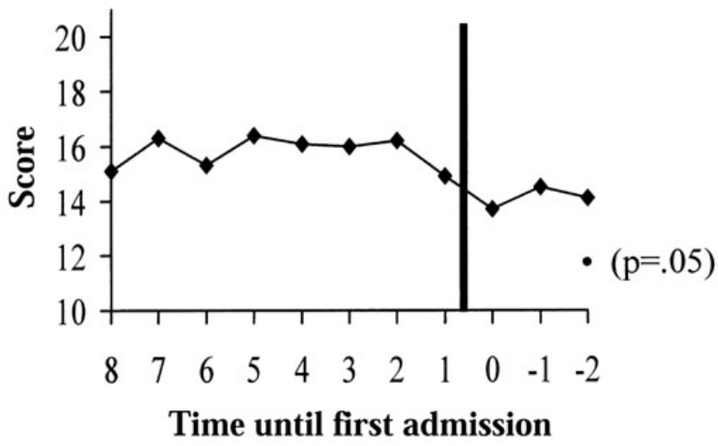


FIGURE 47.2. Scores on the Ravens Progressive Matrices as a function of time until first admission for schizophrenia.

In contrast to the evidence from studies of conventional antipsychotic treatments that suggest little improvement in cognition with treatment, two separate studies demonstrated modest longitudinal improvements in certain areas of cognitive functioning (111 ,127). These findings suggest diversity in the course of cognitive deficit even early in the illness, although they also indicate that there is no consistent pattern of specific dimensions of improvement. Furthermore, even though an improvement in cognition was seen in these studies, no research to date has demonstrated that many first-episode patients show evidence of normalization in their cognitive functioning. Thus, although evidence of worsening in cognitive functioning associated with duration of illness was collected from the study of patients with a longer duration of illness (124), patients with multiple psychotic episodes (116), and elderly patients with continuous psychosis (reviewed later), there is still marked heterogeneity of recovery of cognitive functioning immediately after the first episode.

Despite the good remission of psychosis achieved by most first-episode patients (95), and even though negative symptoms and cognitive impairments are not very severe at this stage of the illness, most patients are already affected by persistent social and vocational decline in the first psychotic episode. For instance, in a study reported by Ho and colleagues (128), more than half of a sample of first-episode patients with schizophrenia were found to be supported by public funds within 12 months of their first episode of illness, and fewer than 25% of them had a job or went to school. Despite evidence of improvement in cognition on the part of some patients at the time of the first episode, continuing cognitive and functional deficit is the rule.

Taken together, the premorbid and first-episode studies indicate that many of the manifestations of schizophrenia, including psychosis, are present many months to few years before the formal diagnosis, and most, but not all, patients respond well to treatment in terms of their positive symptoms and have a better course of illness in several different domains for the first year or 2 of illness than later. Occupational and cognitive deficits are clearly disproportionate compared with the severity of psychotic symptoms in most cases, despite evidence of improvement on the part of some patients. However, these results may be biased, because most first-episode studies enroll patients who (a) were sufficiently sick to need hospitalization, but (b) became sufficiently well to be able and willing to consent to be followed-up after discharge, yet (c) are not sufficiently recovered to be completely out of the treatment network. More important, most first-episode studies last less than 5 years because of attrition, funding, or other factors.

HOW DO THE ILLNESS' MANIFESTATIONS CHANGE DURING LATER LIFE AND SENESCENCE?

Part of "47 - Schizophrenia: Course Over the Lifetime "

Middle Course of Schizophrenia

Until the early 1990s, the characteristics of schizophrenia in patients older than 55 years were largely the subject of speculation. As of 1993, it was estimated that less than 5% of all of the research ever performed on patients with schizophrenia had included any patients older than 55 years (129). It was "common knowledge" that by age 55 to 60 years the illness has run its course, psychotic symptoms had burned out, and most patients did not need or did not benefit from medications. Since the early 1990s, however, a substantial amount of research on this topic has been completed, with this area one of the fastest developing aspects of research on schizophrenia. This research has considered all the topic areas covered by studies on younger patients

and has yielded a considerable amount of information that has helped to refine thinking about schizophrenia in general.

One of the sources of the common knowledge that the course of schizophrenia was established into old age was the consistent findings of symptomatic, cognitive, and functional stability on the part of patients after their first few episodes. Although many patients experience multiple psychotic episodes through middle age and many patients experience continuous psychotic symptoms, there was little evidence of change in cognitive or functional status on the part of these patients. Most research on the course of functional status suggests that the impairments noted at the time of the first episode are rarely reduced. Estimates of the proportion of patients with schizophrenia who are employed are in the range of about 40%, with most patients employed in noncompetitive, sheltered settings (130). Likewise, independent living is the exception for patients with schizophrenia. There is also no significant evidence that functional status in patients with schizophrenia changes markedly over time or is altered by treatment with older antipsychotic medications (131). This large body of data raises issues of importance when older patients are studied, including whether changes seen in later life are part of the natural course of the illness or whether they are the result of additional comorbidities.

Cognitive and Functional Deficits in Older Patients

It has been consistently reported, however, that many patients older than 65 years who have a lifelong course of schizophrenia, especially those with a history of long-term institutional care, have marked deficits in cognitive and functional status (132 ,133 and 134). Similar findings have been reported at different research sites in the United States and in the United Kingdom (135). Because of the lack of data regarding the lifetime course of functional and cognitive deficits in schizophrenia, it is not clear whether the presence of severe deficits in functioning seen in these elderly institutionalized patients with schizophrenia is the result of deterioration in their cognitive and functional status or is a lifelong feature of their adjustment. There are multiple potential methodologic issues associated with the study of older patients, particularly patients with a history of long-term institutional stay. Among these issues are the difficulty in identifying the patients' "true premorbid status," long-term treatment with antipsychotic medications, extremely invasive somatic treatments, and institutionalization and demoralization, potentially leading to poor motivation to cooperate with testing. There is no question, as would be expected from studies of younger patients, that chronically institutionalized patients have low levels of premorbid functioning, in domains of educational, social, occupational, and independent living skills (132 ,133 and 134). However, the functional history of these institutionalized patients is inconsistent with the idea that their current, grossly impaired status could possibly be their lifelong level of functioning.

Many of these questions are being addressed by a longitudinal cohort study carried out by the Mt. Sinai School of Medicine group since the late 1980s, as well as other investigators who have become increasingly interested in this population. A study of the baseline characteristics of the Mt. Sinai sample demonstrated that these elderly patients manifested moderate to severe negative and positive schizophrenic symptoms not dissimilar to the symptoms present in younger institutionalized patients (133). Many of these patients had cognitive and social performance compatible with dementia (136) that could not be accounted for by somatic treatment, lengthy institutionalization, poor motivation and education, or comorbidity. For example, in the original publication on this population (133), it was demonstrated that psychosurgery, insulin coma, electroconvulsive therapy, and the severity of negative symptoms were not the factors accounting for cognitive deficits. Relevant to the issue of motivational deficits, in a subsample of the patients from that study (137), the average level of education was found to be more than 11 years, and their reading performance was higher than the tenth grade level. In contrast to these indicators of educational achievement, the current average Mini-Mental Status Examination (MMSE) score was 20 (consistent with moderate dementia). Thus, some elderly institutionalized patients with schizophrenia appear to manifest decline in their functioning relative to premorbid functioning.

Studies of the cognitive performance of elderly schizophrenic patients have identified "double dissociation" performance profiles that discriminate them from patients with clearly identified dementia (138 and 139), and a profile of differential deficits has been identified. Differential deficits cannot be caused by a single constant factor, such as failing to provide adequate effort when assessed. These data suggest that studies of very poor-outcome long-stay patients, although clearly reflecting the most seriously ill subset of the population, are not hugely biased by the obvious factors associated with long institutional stay.

Longitudinal Course of Cognition and Functional Status in Late-Life Schizophrenia: Patients with Chronic Illness

As noted earlier, some elderly institutionalized patients with schizophrenia appear to manifest decline in their functioning well past premorbid levels at some time in the course of their illness. The time course, prevalence, and correlates of this decline are as yet undiscovered. There is surprisingly little longitudinal research on cognitive functioning and functional skill deficits in schizophrenia. One metaanalysis suggested that indicators of cognitive performance were

largely stable over time in 15 follow-up studies of patients with schizophrenia (140). The total sample size in all these previous studies was only 639, and 225 of those patients were chronically institutionalized patients studied by the Mt. Sinai group in a short-term (1- to 2-year) follow-up study (141). In contrast, in two separate published longitudinal studies of the course of cognitive and functional status in elderly poor-outcome patients with schizophrenia (142 ,143), the Mt. Sinai group found that about 15% of these patients per year showed evidence of cognitive and functional worsening. The second study also demonstrated statistically significant cognitive and functional decline over an average of 2.5 years in 57 geriatric schizophrenic patients who entered the study as chronically hospitalized but were reassessed after discharge to nursing home care (143). These data suggest that some proportion of elderly patients with schizophrenia with a history of long-term institutional care experience a notable decline in their functioning over a relatively brief follow-up period. These data suggest the possibility of some adverse effect of aging after a lifetime of poor functional outcome and extensive cognitive deficits.

The Mt. Sinai group completed a larger follow-up study based on 1,102 patients. Some of these patients were unavailable for later study at each of the subsequent reassessments, because they had died or were discharged to nursing home care, where follow-up could not be performed. The primary analyses examined the development of new-onset severe cognitive and functional impairment. Patients were divided on the basis of their baseline Clinical Dementia Rating (CDR) (148) score, such that patients with baseline scores of 1.0 or less were considered less impaired. Worsening in cognitive and functional status was defined as having a CDR score at a subsequent follow-up of 2.0 or greater. At baseline, there were 456 patients with CDR scores of 1.0 or less, whose average MMSE score was 20.8.

The actuarial life-table method, with discrete interval procedures, was used to assess the cumulative risk of cognitive and functional decline over the three intervals between assessments, while considering subject attrition. The cumulative “survival” (i.e., no worsening in cognitive and functional impairment) was then calculated, and survival curves were constructed. At the first follow-up time beginning at 15 months, 17% of patients met criteria for worsening (corrected), and at the second follow-up time beginning at 30 months, 20.4% of the remaining patients met criteria for worsening. At the third follow-up time beginning at 48 months, 25.3% of the remaining patients manifested worsening of their cognitive and functional deficits. Thus, over the entire follow-up period, a corrected rate of cognitive decline of 51% was noted.

The influences of potential risk factors on rates of cognitive and functional decline over the entire follow-up period for the initially higher-functioning patients were examined. The Wilcoxon statistic was used to measure the difference in survival curves as a function of risk factor status. Gender was unassociated with risk for cognitive and functional decline, as were neuroleptic treatment status, age, and age at first psychiatric admission. In contrast, three risk factors were associated with increases in risk for cognitive and functional decline. Patients with more severe positive (Wilcoxon statistic [1df] = 4.28; $p < .05$) and negative (Wilcoxon statistic [1df] = 17.03; $p < .0001$) symptoms were found to be at higher risk for decline. In addition, patients with more education were less likely to experience a cognitive and functional decline than were patients with lower levels of formal education (Wilcoxon statistic [1df] = 8.65; $p < .01$).

In the final analysis, presented in Fig. 47.3 , the influences the risk factors previously demonstrated as significant predictors of risk for decline were examined for their influence on the rate of cognitive and functional decline over the entire follow-up period. First, patients were divided at the median level for both baseline severity of symptoms and education levels and were assigned to one of four groups

on the basis of their status for symptom severity and education level. The Wilcoxon statistic was then used to measure the difference in survival curves as a function of combined risk factor status. Patients below the median level for education and above the median for severity of symptoms were at the highest risk of cognitive and functional decline. This group's level of risk was significantly greater than those with similarly high levels of positive symptoms but higher levels of education (Wilcoxon statistic [1df] = 8.49; $p < .001$) and those with lower levels of positive symptoms and higher levels of education (Wilcoxon statistic [1df] = 15.31; $p < .001$). Similarly, a significant interaction was seen between the influences of negative symptom severity and education level on risk rates for cognitive and functional decline.

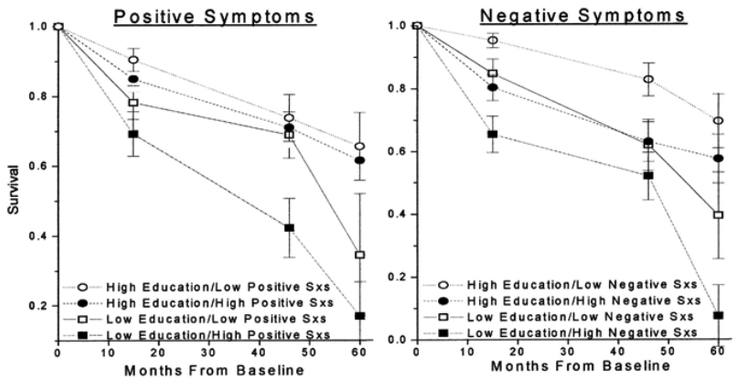


FIGURE 47.3. Effects of combined symptom (positive and negative) severity and educational level factors on survival from cognitive and functional decline over a 60-month follow-up period.

Certain potential limitations that must be addressed in the interpretation of these data. No control for institutionalization as a direct risk factor was used in this study. Despite the difficulty in identification of such as a group, there is no other direct way to index institutionalization effects. Similarly, these data do not control specifically for the development of subtle new-onset medical conditions. This is a less difficult question to address in later research. Finally, as noted earlier, multiple additional factors, including subtle environmental changes, may interact with the easily measured risk factors examined in this study.

Longitudinal Course of Cognition and Functional Status in Late-Life Schizophrenia: Better-Outcome Patients

Although the studies just reviewed indicate that some proportion of poor-outcome patients experience cognitive and functional decline, there is no evidence to date of cognitive decline in patients with a history of better lifetime functional outcome. Cross-sectional comparisons of older and younger better-outcome patients conducted by the University of California at San Diego (UCSD) group found little evidence of relatively poorer cognitive performance on the part of older patients (138 ,144 and 145). It is impossible to determine, of course, from cross-sectional data that these older better-outcome patients, with minimal evidence of previous decline in their cognitive and functional status, would never experience a decline at a later date. Furthermore, the proportion of patients in the UCSD samples older than 65 years was only about 15%, a finding suggesting that if the risk of cognitive and functional decline increases with age, these patients may only be entering the period of increased risk. Finally, few of these patients had a history of symptom severity consistent with extended periods of treatment-refractory psychosis, and very few would have met the criteria for kraeplinian status previously demonstrated to be associated with very poor lifetime functional outcome (146 and 147). These data suggest the need to determine whether long-term institutionalization or the patient characteristics that cause institutionalization are the operant factors in the cognitive decline seen in worse-outcome patients. One of the best possible strategies could be to perform a longitudinal comparison of outpatients with and without a prior history of long-term institutional stay, to separate patient characteristics from current environmental factors.

Persistent Symptoms Revisited: Duration of Continuous Psychosis

The data from follow-up studies of poor-outcome patients suggest that persistent schizophrenic symptoms, combined with evidence of premorbid educational underachievement, are associated with marked increases in risk for functional decline over relatively short follow-up periods. These data again raise the issue of persistent symptoms as a risk factor for the later course of illness and also suggest that evidence of lifelong intellectual compromise may increase this risk. These data may help to address some of the differences in findings between previous studies of ambulatory patients and these extremely impaired, continuously refractory patients. First, these institutionalized patients have persistent symptoms that have kept them hospitalized for decades and clearly distinguish them from ambulatory samples. As previously demonstrated, functional deficits and negative symptoms do not interfere with discharge to nursing home care in this population, whereas persistent positive symptoms do (149 ,150). Second, these patients are all older than 65 years. In previous longitudinal studies, even institutionalized patients younger than 65 years old had essentially no risk of cognitive and functional decline over a 6-year follow-up period (151). It would not be a surprising finding that ambulatory patients in this age range who have never been institutionalized would not have elevations in their risk for decline either.

These data may provide a heuristic for understanding the variance in outcome, measured by cognitive performance and ratings of functional status, in older patients with schizophrenia. In institutionalized patients with similar periods of institutional stay, MMSE scores range from 0 to 30, and functional limitations range from moderate deficits in social skills to incontinence and complete dependence on others for feeding and bathing. In addition, better-outcome patients clearly have indications of higher levels of premorbid and current cognitive functioning. These data suggest that the interaction of reduced levels of educational attainment, often referred to as a marker of cognitive reserve (152), and particularly persistent symptoms of illness, may predict functional decline. The previous suggestion that education attainment is an indicator of a cognitive risk-protective factor for dementia (153) appears relevant to schizophrenia. Thus, patients with schizophrenia in late life who have severe and persistent psychotic symptoms, as well as reduced levels of educational attainment, appear to have a much greater risk of worsening in functional status than

patients whose positive symptoms are less treatment refractory and whose cognitive reserve may be greater.

The length of time that some of these patients have experienced continuous psychotic symptoms, despite conventional antipsychotic treatment, is staggering. Some of these patients have been treated since the 1950s with conventional medications, with little relief of their symptoms. The duration of untreated psychosis seen in typical samples of first-episode patients with schizophrenia pales in comparison with these histories of continuous psychosis. This duration of continuous psychosis is much more similar to that typically seen at the time of the initial introduction of antipsychotic medication in the 1950s. At that time, long duration of untreated psychosis was found to be associated with risk for greater functional deficit after initiation of antipsychotic treatment than for patients whose symptoms were treated sooner after the development of illness (96). Much later research will need to address the issues of the impact of continuous psychotic symptoms, in terms influence on the course of illness and whether continuous psychosis despite treatment has the same impact on development as lengthy periods of untreated psychosis at the outset of the illness.

SCHIZOPHRENIA: STABLE ENCEPHALOPATHY OR PROGRESSIVE DISEASE WITH CORRESPONDING LIFELONG BIOLOGICAL CHANGES?

Part of "47 - Schizophrenia: Course Over the Lifetime "

A most controversial aspect of schizophrenia is whether the few biological and many phenomenologic abnormalities reported are consistent with a degenerative, progressively deteriorating course of the illness (154 ,155 ,156 ,157 and 158) or a static course for accounted by an early (developmental) insult (1 ,159 ,160 and 161). The neurodevelomental models suggest that a perinatal neuronal insult disrupts normal neural maturation and results in disruption of neuronal circuits and thus abnormal neuronal function. It is further postulated that the clinical manifestation of symptoms is triggered by interaction between the initial defect with neuronal maturation processes such as neuronal migration, glial proliferation, and synaptic pruning. This maturation process, in turn, accounts for the gap between the hypothesized early-life insult and later clinical manifestation.

The neurodevelopment concept has prevailed mostly because schizophrenia lacks specific biochemical and histologic changes (gliosis, cellular debris, or amyloid deposits) closely paralleling behavioral abnormalities that define progressive degenerative disorders. Furthermore, because Alzheimer disease has been seen as the prototype of a progressive neurodegenerative disorder, the absence of fast and relentless worsening of illness has been taken as evidence against a degenerative hypothesis in schizophrenia. However, an overview of the data regarding the course and the biology of schizophrenia reveals no sufficient evidence to settle this debate, and the same behavioral evidence can be interpreted to support either of the two hypotheses. For example, subtle cognitive, behavioral, and motor deviation from norms are present in childhood, are amplified in adolescence, and exacerbate shortly before and after the first psychotic episode. This can be interpreted a classic interaction between an early defect and brain maturation or as the behavioral consequence of a slowly progressive degenerative brain process. In addition, lack of consistent worsening of psychosis across episodes argues for the static hypothesis, whereas progressively poorer antipsychotic response after each additional episode could be interpreted as evidence of a slowly progressive degenerative process.

Similarly, biological findings, mostly structural neuroimaging studies, have produced results compatible with both hypotheses (120 ,121 ,162 ,163). Some investigators reported no evidence of progressive brain disease, in either the domains of overall cerebral size (i.e., cortical atrophy) or the size of the cerebral ventricles (i.e., ventricular enlargement) (164 ,165). However, some cross-sectional and longitudinal studies have produced different results. There are several limitations, of course, in using neuroimaging to make direct inferences about changes in the brain, particularly in reference to whether these changes are degenerative.

Brain Structure Immediately after the First Episode

One of the interesting recent topics in the area of the course of schizophrenia is that of changes in brain structure after the first episode. Research by DeLisi and colleagues suggested that some patients recovering from the first episode of the illness have evidence of progression in the size of their cerebral ventricles (166 ,167). This progressive ventricular enlargement is consistent with that seen in patients with more chronic illness, both during adolescence for childhood-onset patients (168) and during middle age for poor-outcome patients with a more typical age of onset (169). These changes are modest, but they are also detected in relatively short follow-up periods. Because the patients in the studies by DeLisi et al. experienced an increase in the ventricular size of about 3.5% in 5 years, the amount of change that would be expected over a lifetime, if this change were continuous, could be substantial.

Changes in Brain Function in Patients with Established Illness

Changes in cerebral structure have also been noted in patients with an established illness. In a 5- year prospective study (169) comparing middle-aged patients with schizophrenia who varied in their lifetime functional outcome from chronic “kraeplinian” patients with community dwellers, the kraeplinian patients demonstrated progressive ventricular enlargement. These data, consistent with those of

the first-episode and childhood-onset studies, suggest that the cerebral change is dynamic over the life span in patients with schizophrenia. In addition, two separate sets of cross-sectional studies, examining p300 prolongation, suggested that older patients have longer latencies (170 ,171). The finding of prolonged p300 latency can be associated with the presence of neurodegenerative diseases (172). Finally, cross-sectional studies have also found that older patients show relatively greater atrophic changes in the size of the olfactory bulb (173), and this has been found to correspond to a concurrent deterioration in olfactory sensitivity (174). Because olfactory deficits are also detected with consistency in patients with degenerative conditions, these data are consistent with the p300 data just reviewed. The data to date on the processes of dynamic cortical change in schizophrenia are hardly conclusive. There are, however, multiple, albeit indirect, suggestions that the idea that brain structure and function in schizophrenia are immutably stable over the life span in all patients is open to question. These are important issues that will shape future research in this area.

Is Integration Possible?

In an attempt to account for the phenomenology, course, and epidemiology of schizophrenia, McGlashan and Hoffman postulated that reduced synaptic connectedness resulting from early, developmental disturbance of synaptogenesis or faulty synaptic pruning is at the root of schizophrenia (3). More specifically, an innately sparse synaptic substrate combined with normal pruning in childhood and adolescence or a normal substrate combined with abnormally accelerated pruning could theoretically reach a critical threshold at which deficient synaptic connectivity manifests as abnormal perceptions and ideas. Because the model incorporates both static and dynamic components, it can account for some of the apparently competing postulations of schizophrenia pathophysiology. Caution is required, however, because integrative models of the development of schizophrenia have been foiled in the past by the remarkable heterogeneity of the illness.

CONCLUSIONS

Part of "47 - Schizophrenia: Course Over the Lifetime "

From the time schizophrenia was defined, it has been viewed as a chronic, lifelong disease. Yet its specific manifestations along the life cycle have been poorly described, mostly because of inherent methodologic difficulties. Improved medical record keeping and the realization that understanding the course of illness is essential to the understanding the pathophysiology of the illness have been behind the modern long-term follow-up studies. Determining, for example, the earliest manifestations of the illness or the trigger and the length of the window of deterioration has implications for preventive and palliative treatment. Even if otherwise possible, the identification of clinical correlates of progressive cortical atrophy by computed tomography, magnetic resonance imaging, or the biological meaning of spectroscopic abnormalities in synaptic connectivity requires detailed description of the illness course. As more research is focused on these issues, important information about the nature of schizophrenia itself will result.

DISCLOSURE

Part of "47 - Schizophrenia: Course Over the Lifetime "

Dr. Harvey has received research support from Janssen and Lilly and has served as a consultant or on a speakers bureau for the following companies: Pfizer, Janssen, Lilly, HMR, BMS, and Astra-Zeneca.

REFERENCES

1. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 1987;44:660–669.

2. Lieberman JA. Is schizophrenia a neurodegenerative disorder? A clinical and neurobiological perspective. *Biol Psychiatry* 1999;46:729–739.

3. McGlashan TH, Hoffman RE. Schizophrenia as a disorder of developmentally reduced synaptic connectivity. *Arch Gen Psychiatry* 2000;57:637–648.

4. Fenton WS, McGlashan TH. Natural history of schizophrenia subtypes. I. Longitudinal study of paranoid, hebephrenic, and undifferentiated schizophrenia. *Arch Gen Psychiatry* 1991;48:969–977.

5. DeSisto MJ, Harding CM, McCormick RV, et al. The Maine and Vermont three-decade studies of serious mental illness. I. Matched comparison of cross-sectional outcome. *Br J Psychiatry* 1995;167:331–338.

6. Huber G, Gross G, Schuttler R. A long-term follow-up study of schizophrenia: psychiatric course of illness and prognosis. *Acta Psychiatr Scand* 1975;52:49–57.

7. Stephens JH, Richard, P, McHugh PR. Long-term follow-up of patients hospitalized for schizophrenia, 1913 to 1940. *J Nerv Ment Dis* 1997;185:715–721.

8. Marneros A, Steinmeyer EM, Deister A, et al. Long-term outcome of schizoaffective and schizophrenic disorders: a comparative study. III. Social consequences. *Eur Arch Psychiatry Neurol Sci* 1989;238:135–139.

9. Steinhausen HC, Meier M, Angst J. The Zurich long-term outcome study of child and adolescent psychiatric disorders in males. *Psychol Med* 1998;28:375–383.

10. Davidson L, McGlashan TH. The varied outcomes of schizophrenia. *Can J Psychiatry* 1997;42:34–43.

11. Hubschmid T, Perrin N, Ciompi L. The dynamics of outcome predictors in the early phase of schizophrenic psychoses: an empirical study. *Psychiatr Prax* 1991;18:196–201.

12. Kolakowska T, Williams AO, Ardern M, et al. Schizophrenia with good and poor outcome. I. Early clinical features, response to neuroleptics and signs of organic dysfunction. *Br J Psychiatry* 1985;146:229–239.

13. McGlashan TH, Fenton WS. Subtype progression and pathophysiologic deterioration in early schizophrenia. *Schizophr Bull* 1993;19:71–84.

14. Cornblatt B, Obuchowski M, Roberts S, et al. Cognitive and behavioral precursors of schizophrenia. *Dev Psychopathol* 1999;11:487–508.

15. Erlenmeyer-Kimlin, L. Neurobehavioral deficits in offspring of schizophrenic parents: liability indicators and predictors of illness. *Am J Med Genet* 2000;97:65–71.

16. Marcus J, Hans SL, Auerbach JG, et al. Children at risk for schizophrenia: the Jerusalem Infant Development Study. II. Neurobehavioral deficits at school age. *Arch Gen Psychiatry* 1993;50:797–809.

17. Squires-Wheeler E, Skodol AE, Bassett A, et al. DSM-III-R schizotypal personality traits in offspring of schizophrenic disorder, affective disorder, and normal control parents. *J Psychiatr Res* 1989;23:229–239.

18. Bromet EJ, Jandorf L, Fennig S, et al. The Suffolk County Mental Health Project: demographic, pre-morbid and clinical correlates of 6-month outcome. *Psychol Med* 1996;26:953–962.

19. Hafner H, Maurer K, Loffler W, et al. The ABC Schizophrenia Study: a preliminary overview of the results. *Soc Psychiatry Psychatr Epidemiol* 1998;33:380–386.

20. Ho BC, Andreasen NC, Flaum M et al. Untreated initial psychosis: its relation to quality of life and symptom remission in first-episode schizophrenia. *Am J Psychiatry* 2000;157:808–815.

21. Hoff AL, Riordan H, O'Donnell DW, et al. Neuropsychological functioning of first-episode schizophreniform patients. *Am J Psychiatry* 1992;149:898–903.

22. Jarskog LF, Mattioli MA, Perkins DO, et al. First-episode psychosis in a managed care setting: clinical management and research. *Am J Psychiatry* 2000;157:878–884.

23. Johnstone EC, Crow TJ, Johnson AL, et al. The Northwick Park Study of first episodes of schizophrenia. I. Presentation of the illness and problems relating to admission. *Br J Psychiatry* 1986;148:115–120.

24. Johnstone EC, Frith CD, Crow TJ, et al. The Northwick Park "Functional" Psychosis Study: diagnosis and outcome. *Psychol Med* 1992;22:331–346.

25. Keshavan MS, Schooler NR. First-episode studies in schizophrenia: criteria and characterization. *Schizophr Bull* 1992;18:491–513.

26. Larsen TK, Johannessen JO, Opjordsmoen S. First-episode schizophrenia with long duration of untreated psychosis. pathways to care. *Br J Psychiatry Suppl* 1998;172:45–52.

27. Larsen TK, McGlashan TH, Moe LC. First-episode schizophrenia. I. Early course parameters. *Schizophr Bull* 1996;22:241–256.

28. Lieberman JA, Alvir JM, Woerner M, et al. Prospective study of psychobiology in first-episode schizophrenia at Hillside Hospital. *Schizophr Bull* 1992;18:351–371.

29. Thara R, Eaton WW. Outcome of schizophrenia: the Madras longitudinal study. *Aust NZ J Psychiatry* 1996;30:516–522.

30. Gotteman II, Shields JM. *Schizophrenia: the epigenetic puzzle*. New York, Oxford University Press, 1982.

31. Maurer K, Hafner H. Methodological aspects of onset assessment in schizophrenia. *Schizophr Res* 1995;15:265–276.

32. Chapman LJ, Chapman JP, Kwapil TR, et al. Putatively psychosis-prone subjects 10 years later. *J Abnorm Psychol* 1994;103:171–183.

33. Ganguli R, Brar JS. Generalizability of first-episode studies in schizophrenia. *Schizophr Bull* 1992;18:463–469.

34. McGlashan TH. Duration of untreated psychosis in first-episode schizophrenia: marker or determinant of course? *Biol Psychiatry* 1999;46:899–907.

35. Berenbaum H, Fujita F. Schizophrenia and personality: exploring the boundaries and connections between vulnerability and outcome. *J Abnorm Psychol* 1994;103:148–158.

36. Levav I, Kohn R, Dohrenwend BP, et al. An epidemiological study of mental disorders in a 10-year cohort of young adults in Israel. *Psychol Med* 1993;23:691–707.

37. Lieberman JA. Schizophrenia: comments on genes, development, risk factors, phenotype, and course. *Biol Psychiatry* 1999;46:869–870.

38. Phelan J. Six-month stability of psychiatric diagnoses in first-admission patients with psychosis. *Am J Psychiatry* 1994;151:1200–1208.

39. Pulver AE, Carpenter WT, Adler L, et al. Accuracy of the diagnoses of affective disorders and schizophrenia in public hospitals. *Am J Psychiatry* 1988;145:218–220.

40. Tsuang MT, Faraone SV. The concept of target features in schizophrenia research. *Acta Psychiatr Scand Suppl* 1999;395:2–11.

41. Vazquez-Barquero JL, Cuesta Nunez MJ, Herrera CS, et al. Sociodemographic and clinical variables as predictors of the diagnostic characteristics of first episodes of schizophrenia. *Acta Psychiatr Scand* 1996;94:149–155.

42. Agid O, Shapira B, Zislin J, et al. Environment and vulnerability to major psychiatric illness: a case control study of early parental loss in major depression, bipolar disorder and schizophrenia. *Mol Psychiatry* 1999;4:163–172.

43. Birchwood M, Cochrane R, Macmillan F, et al. The influence of ethnicity and family structure on relapse in first-episode schizophrenia: a comparison of Asian, Afro-Caribbean, and white patients. *Br J Psychiatry* 1992;161:783–790.

44. Brekke JS, Ansel M Long J et al. Intensity and continuity of services and functional outcomes in the rehabilitation of persons with schizophrenia. *Psychiatr Serv* 1999;50:248–256.

45. Browne S, Clarke M, Gervin M, et al. Determinants of quality of life at first presentation with schizophrenia. *Br J Psychiatry* 2000;176:173–176.

46. van Os J, Driessen G, Gunther N, et al. Neighbourhood variation in incidence of schizophrenia: evidence for person-environment interaction. *Br J Psychiatry* 2000;176:243–248.

47. Garbharran HN, Zibin T. The influence of ethnicity and family structure on relapse in first-episode schizophrenia. *Br J Psychiatry* 1993;163:127.

48. Phelan JC, Bromet EJ, Link BG. Psychiatric illness and family stigma. *Schizophr Bull* 1998;24:115–126.

49. Bleuler E. *Dementia praecox, or the group of schizophrenias*. International Universities Press, 1911.

50. Cannon M, Jones P, Huttunen MO, et al. School performance in Finnish children and later development of schizophrenia: a population-based longitudinal study. *Arch Gen Psychiatry* 1999;56:457–463.

51. Gupta S, Rajaprabhakaran R, Arndt S, et al. Premorbid adjustment as a predictor of phenomenological and neurobiological indices in schizophrenia. *Schizophr Res* 1995;16:189–197.

52. Jones PB, Bebbington P, Foerster A, et al. Premorbid social underachievement in schizophrenia. Results from the Camberwell Collaborative Psychosis Study. *Br J Psychiatry* 1993;162:65–71.

53. Davidson M, Reichenberg A, Rabinowitz J, et al. Behavioral and intellectual markers for schizophrenia in apparently healthy male adolescents. *Am J Psychiatry* 1999;156:1328–1335.

54. Done DJ, Leinon E, Crow TJ, et al. Linguistic performance in children who develop schizophrenia in adult life: evidence for normal syntactic ability. *Br J Psychiatry* 1998;172:130–135.

55. Kwapil TR, Miller MB, Zinser MC, et al. Magical ideation and social anhedonia as predictors of psychosis proneness: a partial replication. *J Abnorm Psychol* 1997;106:491–495.

56. Walker E, Lewine RJ. Prediction of adult-onset schizophrenia from childhood home movies of the patients. *Am J Psychiatry* 1990;147:1052–1056.

57. Yung AR, Phillips LJ, McGorry PD, et al. Prediction of psychosis: a step towards indicated prevention of schizophrenia. *Br J Psychiatry Suppl* 1998;172:14–20.

58. Gilvarry C, Takei N, Russell A, et al. Premorbid IQ in patients with functional psychosis and their first-degree relatives. *Schizophr Res* 2000;41:417–429.

59. Iacono WG, Ficken JW, Beiser M. Electrodermal activation in first-episode psychotic patients and their first-degree relatives. *Psychiatry Res* 1999;88:25–39.

60. Nicolson R, Lenane M, Singaracharlu S, et al. L. Premorbid speech and language impairments in childhood-onset schizophrenia: association with risk factors. *Am J Psychiatry* 2000;157:794–800.

61. Isohanni I, Jarvelin MR, Nieminen P, et al. School performance as a predictor of psychiatric hospitalization in adult life: a 28-year follow-up in the Northern Finland 1966 Birth Cohort. *Psychol Med* 1998;28:967–974.

62. Olin SS, John RS, Mednick SA. Assessing the predictive value of teacher reports in a high risk sample for schizophrenia: a ROC analysis. *Schizophr Res* 1995;16:53–66.

63. Basu D, Malhotra A, Bhagat A, et al. Cannabis psychosis and acute schizophrenia: a case-control study from India. *Eur Addict Res* 1999;5:71–73.

64. Bhugra D, Hilwig M, Mallett R, et al. Factors in the onset of schizophrenia: a comparison between London and Trinidad samples. *Acta Psychiatr Scand* 2000;101:135–141.

65. Hambrecht M, Hafner H, Loffler W. Beginning schizophrenia observed by significant others. *Soc Psychiatry Psychiatr Epidemiol* 1994;29:53–60.

66. Hinrichsen GA, Lieberman JA. Family attributions and coping in the prediction of emotional adjustment in family members of patients with first-episode schizophrenia. *Acta Psychiatr Scand* 1999;100:359–366.

67. McClellan J, Werry J. Practice parameters for the assessment and treatment of children and adolescents with schizophrenia: American Academy of Child and Adolescent Psychiatry. *J Am Acad Child Adolesc Psychiatry* 1997;36:177S–193S.

68. Varma VK, Wig NN, Phookun HR, et al. First-onset schizophrenia in the community: relationship of urbanization with onset, early manifestations and typology. *Acta Psychiatr Scand* 1997;96:431–438.

69. Amin F, Silverman JM, Siever LJ, et al. Genetic antecedents of dopamine dysfunction in schizophrenia. *Biol Psychiatry* 1999;45:1143–1150.

70. Keefe RS, Silverman JM, Mohs RC, et al. Eye tracking, attention, and schizotypal symptoms in nonpsychotic relatives of patients with schizophrenia. *Arch Gen Psychiatry* 1997;54:169–176.

71. Lauer CJ, Bronisch T, Kainz M, et al. Pre-morbid psychometric profile of subjects at high familial risk for affective disorder. *Psychol Med* 1997;27:355–362.

72. Lyons MJ, Toomey R, Faraone SV, et al. Correlates of psychosis proneness in relatives of schizophrenic patients. *J Abnorm Psychol* 1995;104:390–394.

73. Lyons MJ, Toomey R, Faraone SV, Tsuang MT. Comparison of schizotypal relatives of schizophrenic versus affective probands. *Am J Med Genet* 1994;54:279–285.

74. Maier W, Lichtermann D, Minges J, et al. Personality disorders among the relatives of schizophrenia patients. *Schizophr Bull* 1994;20:481–493.

75. Maier W, Minges J, Lichtermann D, et al. Personality variations in healthy relatives of schizophrenics. *Schizophr Res* 1994;12:81–88.

76. Parnas J, Cannon TD, Jacobsen B, et al. Lifetime DSM-III-R diagnostic outcomes in the offspring of schizophrenic mothers: results from the Copenhagen High-Risk Study. *Arch Gen Psychiatry* 1993;50:707–714.

77. Rosenberg DR, Sweeney JA, Squires-Wheeler E, et al. Eye-tracking dysfunction in offspring from the New York High-Risk Project: diagnostic specificity and the role of attention. *Psychiatry Res* 1997;66:121–130.

78. Varma SL, Sharma, I. Psychiatric morbidity in the first-degree relatives of schizophrenic patients. *Br J Psychiatry* 1993;162:672–678.

79. Varma SL, Zain AM, Singh S. Psychiatric morbidity in the first-degree relatives of schizophrenic patients. *Am J Med Genet* 1997;74:7–11.

80. Webb CT, Levinson DF. Schizotypal and paranoid personality disorder in the relatives of patients with schizophrenia and affective disorders: a review. *Schizophr Res* 1993;11:81–92.

81. Zivanovic O, Borisev L. Genetic factors in the onset of schizophrenia. *Med Pregl* 1999;52:25–28.

82. Reichenberg A, Rabinowitz J, Weiser M, et al. Premorbid functioning in a national population of male twins discordant for psychoses. *Am J Psychiatry* 2000;157:1514–1516.

83. Jones P, Rodgers B, Murray R, et al. Child development risk factors for adult schizophrenia in the British 1946 birth cohort. *Lancet* 1994;344:1398–1402.

84. Done DJ, Crow TJ, Johnstone EC, et al. Childhood antecedents of schizophrenia and affective illness: social adjustment at ages 7 and 11. *BMJ* 1994;309:699–703.

85. David AS, Malmberg A, Brandt L, et al. IQ and risk for schizophrenia: a population-based cohort study. *Psychol Med* 1997;27:1311–1323.

86. Cuesta MJ, Peralta V, Caro F. Premorbid personality in psychoses. *Schizophr Bull* 1999;25:801–811.

87. Moller P. First-episode schizophrenia: do grandiosity, disorganization, and acute initial development reduce duration of untreated psychosis? An exploratory naturalistic case study *Compr Psychiatry* 2000;41:184–190.

88. Falloon IR, Coverdale JH, Laidlaw TM, et al. Early intervention for schizophrenic disorders: implementing optimal treatment strategies in routine clinical services. OTP Collaborative Group. *Br J Psychiatry Suppl* 1998;172:33–38.

89. Dalery J. Schizophrenia: from prediction to prevention: a challenge for the 21st century [Editorial]. *Encephale* 1999;25:193–194.

90. Johannessen JO. Early intervention and prevention in schizophrenia: experiences from a study in Stavanger, Norway. *Seishin Shinkeigaku Zasshi* 1998;100:511–522.

91. Klosterkotter J. From fighting to preventing disease: is such a paradigm possible for schizophrenic disorders? *Fortschr Neurol Psychiatr* 1998;66:366–377.

92. Munoz RF, Mrazek PJ, Haggerty RJ. Institute of Medicine report on prevention of mental disorders: summary and commentary. *Am Psychol* 1996;51:1116–1122.

93. Oguru T. Prevention of schizophrenia: from a projection to practical planning. *Seishin Shinkeigaku Zasshi* 1998;100:501–506.

94. Loebel AD, Lieberman JA, Alvir JM, et al. Duration of psychosis and outcome in first-episode schizophrenia. *Am J Psychiatry* 1992;149:1183–1188.

95. Robinson DG, Werner MG, Alvir JM, et al. Predictors of treatment response from a first episode of schizophrenia or schizoaffective disorder. *Am J Psychiatry* 1999;156:544–549.

96. Wyatt RJ. Neuroleptics and the natural course of schizophrenia. *Schizophr Bull* 1991;17:325–351.

97. Craig TJ, Bromet EJ, Fennig S, et al. Is there an association between duration of untreated psychosis and 24-month clinical outcome in a first-admission series? *Am J Psychiatry* 2000;157:60–66.

98. Pharoah FM, Mari JJ, Streiner D. Family intervention for schizophrenia. *Cochrane Database Syst Rev* 2000;(2):CD000088.

99. Sackett DL, Richardson S, Rosenberg W. *Evidence based medicine: how to practice and teach EBM*. London: Churchhill-Livingstone, 1997.

100. Emsley RA. Risperidone in the treatment of first-episode psychotic patients: a double-blind multicenter study. Risperidone Working Group. *Schizophr Bull* 1999;25:721–729.

101. Birchwood M, McGorry P, Jackson H. Early intervention in schizophrenia. *Br J Psychiatry* 1997;170:2–5.

102. Mihalopoulos C, McGorry PD, Carter RC. Is phase-specific, community-oriented treatment of early psychosis an economically viable method of improving outcome? *Acta Psychiatr Scand* 1999;100:47–55.

103. Tsuang MT, Stone WS, Seidman LJ, et al. Treatment of nonpsychotic relatives of patients with schizophrenia: four case studies. *Biol Psychiatry* 1999;45:1412–1418.

104. Gupta S, Andreasen NC, Arndt S, et al. The Iowa Longitudinal Study of Recent Onset Psychosis: one-year follow-up of first episode patients. *Schizophr Res* 1997;23:1–13.

105. McGorry PD. The influence of illness duration on syndrome clarity and stability in functional psychosis: does the diagnosis emerge and stabilise with time? *Aust NZ J Psychiatry* 1994;28:607–619.

106. House A, Bostock J, Cooper J. Depressive syndromes in the year following onset of a first schizophrenic illness. *Br J Psychiatry* 1987;151:773–779.

107. Husted JA, Beiser M, Iacono WG. Negative symptoms in the course of first-episode affective psychosis. *Psychiatry Res* 1995;56:145–154.

108. Mayerhoff DI, Loebel AD, Alvir JM, et al. The deficit state in first-episode schizophrenia. *Am J Psychiatry* 1994;151:1417–1422.

109. Albus M, Hubmann W, Mohr F, et al. Are there gender differences in neuropsychological performance in patients with first-episode schizophrenia? *Schizophr Res* 1997;28:39–50.

110. Bilder RM, Goldman RS, Robinson D, et al. Neuropsychology of first-episode schizophrenia: initial characterization and clinical correlates. *Am J Psychiatry* 2000;157:549–559.

111. Hoff AL, Sakuma M, Wieneke M, et al. Longitudinal neuropsychological follow-up study of patients with first-episode schizophrenia. *Am J Psychiatry* 1999;156:1336–1341.

112. Hutton SB, Puri BK, Duncan LJ, et al. Executive function in first-episode schizophrenia. *Psychol Med* 1998;28:463–473.

113. Mohamed S, Paulsen JS, O'Leary D, et al. Generalized cognitive deficits in schizophrenia: a study of first-episode patients. *Arch Gen Psychiatry* 1999;56:749–754.

114. Saykin AJ, Shtasel DL, Gur RE, et al. Neuropsychological deficits in neuroleptic naive patients with first-episode schizophrenia. *Arch Gen Psychiatry* 1994;51:124–131.

115. Sobizack N, Albus M, Hubmann W, et al. Neuropsychological deficits in the initial acute episode of schizophrenia: a comparison with chronic schizophrenic patients. *Nervenarzt* 1999;70:408–415.

116. Sweeney JA, Haas GL, Keilp JG, et al. Evaluation of the stability of neuropsychological functioning after acute episodes of schizophrenia: one-year followup study. *Psychiatry Res* 1991;38:63–76.

117. Abu-Akel A. A study of cohesive patterns and dynamic choices utilized by two schizophrenic patients in dialog, pre- and post-medication. *Lang Speech* 1997;40:331–351.

118. Erickson DH, Beiser M, Iacono WG, et al. The role of social relationships in the course of first-episode schizophrenia and affective psychosis. *Am J Psychiatry* 1989;146:1456–1461.

119. Vazquez-Barquero JL, Cuesta MJ, Herrera Castanedo S et al. Cantabria first-episode schizophrenia study: three-year follow-up. *Br J Psychiatry* 1999;174:141–149.

120. Hirayasu Y, Shenton ME, Salisbury DF, et al. Hippocampal and superior temporal gyrus volume in first-episode schizophrenia. *Arch Gen Psychiatry* 2000;57:618–619.

121. James AC, Crow TJ, Renowden S, et al. Is the course of brain development in schizophrenia delayed? Evidence from onsets in adolescence. *Schizophr Res* 1999;40:1–10.

122. Velakoulis D, Pantelis C, McGorry PD, et al. Hippocampal volume in first-episode psychoses and chronic schizophrenia: a high-resolution magnetic resonance imaging study. *Arch Gen Psychiatry* 1999;56:133–141.

123. Gold JM, Harvey PD. Cognitive deficits in schizophrenia. *Psychiatr Clin North Am* 1993;16:295–312.

124. Saykin AJ, Gur RC, Gur RE, et al. Neuropsychological function in schizophrenia: selective impairment in memory and learning. *Arch Gen Psychiatry* 1991;48,618–623.

125. Rabinowitz J, Reichenberg A, Weiser M, et al. Cognitive and behavioral functioning in schizophrenia before the first hospitalization and shortly after: a cross-sectional analysis of registry data. *Br J Psychiatry* 2000;177:26–32.

126. Medalia A, Gold J, Merriam A. The effects of neuroleptics on neuropsychological test results of schizophrenics. *Arch Clin Neuropsychol* 1988;3:249–271.

127. Gold S, Arndt S, Nopoulos P, et al. Longitudinal study of cognitive functioning in first episode and recent onset schizophrenia. *Am J Psychiatry* 1999;156:1342–1348.

128. Ho BC, Miller D, Nopoulos P, et al. A comparative effectiveness study of risperidone and olanzapine in the treatment of schizophrenia. *J Clin Psychiatry* 1999; 60:658–663.

129. Belitsky R, McGlashan T. The manifestations of schizophrenia in late life: a dearth of data. *Schizophr Bull* 1993;19:683–685.

130. Harvey PD McGurk SR. Cost of schizophrenia: a focus on occupational impairment. *Econ Neurosci* 2000;1:64–78.

131. Hegarty JD, Baldessarini RJ, Tohen M. One hundred years of schizophrenia: a meta-analysis of the outcome literature. *Am J Psychiatry* 1994;151:1409–1416.

132. Bartels SJ, Mueser KT, Miles KM. Functional impairments in elderly patients with schizophrenia and major affective disorders living in the community: social skills, living skills, and behavior problems. *Behav Ther* 1997;28:43–63.

133. Davidson M, Harvey PD, Powchik P, et al. Severity of symptoms in chronically institutionalized geriatric schizophrenic patients. *Am J Psychiatry* 1995;152:197–207.

134. Arnold, SE, Gur, RE, Shapiro, RM, et al. Prospective clinicopathological studies of schizophrenia: accrual and assessment of patients. *Am J Psychiatry* 1995;152:731–737.

135. Harvey PD, Leff J, Trieman N, et al. Cognitive impairment and adaptive deficit in geriatric chronic schizophrenic patients: a cross national study in New York and London. *Int J Geriatr Psychiatry* 1997;12:1001–1007.

136. Harvey PD, Davidson M, Powchik P, et al. Assessment of dementia in elderly schizophrenics with structured rating scales. *Schizophr Res* 1992;7:85–90.

137. Harvey PD, Moriarty PJ, Friedman J, et al. Differential preservation of cognitive functions in geriatric patients with lifelong chronic schizophrenia: less impairment in reading compared to other skill areas. *Biol Psychiatry* 2000;47:962–968.

138. Heaton R, Paulsen J, McAdams LA. Neuropsychological deficits in schizophrenics: relationship to age, chronicity and dementia. *Arch Gen Psychiatry* 1994;51:469–476.

139. Davidson M, Harvey P, Welsh KA, et al. Cognitive functioning in late-life schizophrenia: a comparison of elderly schizophrenic patients and patients with Alzheimer's disease. *Am J Psychiatry* 1996;153:1274–1279.

140. Rund BR. A review of longitudinal studies of cognitive functions in schizophrenia patients. *Schizophr Bull* 1998;24:425–435.

141. Harvey PD, White L, Parrella M, et al. The longitudinal stability of cognitive impairment in schizophrenia: mini-mental state scores at one- and two-year follow-ups in geriatric in-patients. *Br J Psychiatry* 1995;166:630–633.

142. Harvey PD, Silverman JM, Mohs RC, et al. Cognitive decline in late-life schizophrenia: a longitudinal study of geriatric chronically hospitalized patients. *Biol Psychiatry* 1999;45:32–40.

143. Harvey PD, Parrella M, White L, et al. The convergence of cognitive and adaptive decline in late-life schizophrenia. *Schizophr Res* 1999;35:77–84.

144. Eyler Zorrilla LT, Heaton RK, McAdams LA, et al. Cross-sectional study of older outpatients with schizophrenia and healthy comparison subjects: no differences in age-related cognitive decline. *Am J Psychiatry* 2000;157:1324–1326.

145. Palmer BW, Heaton RK, Paulsen JS. Is it possible to be schizophrenic yet neuropsychologically normal? *Neuropsychology* 1997;11:437–446.

146. Keefe RSE, Frecksa, E, Apter, S, et al. Clinical characteristics of kraeplinian schizophrenia: a replication and extension of previous findings. *Am J Psychiatry* 1996;153:806–811.

147. Keefe RS, Mohs RC, Losonczy MF, et al. Characteristics of very poor outcome schizophrenia. *Am J Psychiatry* 1987;144:889–895.

148. Berg L. Clinical dementia rating (CDR). *Psychopharm Bull* 1988;24:637–639.

149. White L, Parrella M, McCrystal-Simon J, et al. Characteristics of elderly psychiatric patients retained in a state hospital during downsizing: a prospective study with replication. *Int J Geriatr Psychiatry* 1997;12:474–480.

150. Trieman N, Leff J. Difficult to place patients in a psychiatric hospital closure programme: the TAPs project 24. *Psychol Med* 1996;26:765–774.

151. Friedman JI, Harvey PD, Coleman T, et al. A six-year follow-up across the life span in schizophrenia: a comparison with Alzheimer's disease and healthy subjects. *Am J Psychiatry* 2001;158:1441–1448.

152. Satz P. Brain reserve capacity on symptom onset after brain injury: a formulation and review of evidence for threshold theory. *Neuropsychology* 1993;3:273–295.

153. Schmand B, Smit JH, Geerlings MI, et al. The effects of intelligence and education on the development of dementia: a test of the brain reserve hypothesis. *Psychol Med* 1997;27: 1337–1344.

154. DeQuardo JR, Tandon R, Goldman R, et al. Ventricular enlargement, neuropsychological status, and premorbid function in schizophrenia. *Biol Psychiatry* 1994;35:517–524.

155. Woods BT. Is schizophrenia a progressive neurodevelopmental disorder? Toward a unitary pathogenetic mechanism. *Am J Psychiatry* 1998;155:1661–1670.

156. Lieberman JA. Pathophysiologic mechanisms in the pathogenesis and clinical course of schizophrenia. *J Clin Psychiatry* 1999;60[Suppl 12]:9–12.

157. Selemon LD, Goldman-Rakic PS. The reduced neuropil hypothesis: a circuit based model of schizophrenia. *Biol Psychiatry* 1999;45:17–25.

158. Rapoport JL, Giedd JN, Blumenthal J, et al. Progressive cortical change during adolescence in childhood-onset schizophrenia: a longitudinal magnetic resonance imaging study. *Arch Gen Psychiatry* 1999;56:649–654.

159. Goldberg T, Hyde TM, Kleinman JE, et al. Couse of schizophrenia: neuropsychological evidence for static encephalopathy. *Schizophr Bull* 1993;19:787–804.

160. Harrison PJ. The neuropathology of schizophrenia: a critical review of the data and their interpretation. *Brain* 1999;122:593–624.

161. Gur RE, Turetsky BI, Bilker WB, et al. Reduced gray matter volume in schizophrenia. *Arch Gen Psychiatry* 1999;56:905–911.

162. Kim JJ, Mohamed S, Andreasen NC, et al. Regional neural dysfunctions in chronic schizophrenia studied with positron emission tomography. *Am J Psychiatry* 2000;157:542–548.

163. Lawrie SM, Whalley H, Kestelman JN, et al. Magnetic resonance imaging of brain in people at high risk of developing schizophrenia. *Lancet* 1999;353:30–33.

164. Jaskiw G, Juliano, D, Goldberg TE, et al. Cerebral ventricular enlargement in schizophreniform disorder does not progress: a seven-year followup study. *Schizophr Res* 1994;14:23–29.

165. Sponheim S, Iacono WG, Beiser M. Stability of ventricular size after the onset of psychosis in schizophrenia. *Psychiatry Res* 1991;40:21–29.

166. DeLisi L, Sakuma M, Tew W, et al. Schizophrenia as a chronic active brain process: a study of progressive brain structural change subsequent to the onset of psychosis. *Psychiatry Res* 1997;74:129–140.

167. DeLisi LE, Sakuma M, Ge S, et al. Association of brain structural change with heterogenous course of schizophrenia from early childhood through five years subsequent to a first hospitalization. *Psychiatry Res* 1998;84:75–88.

168. Rapoport JL, Giedd J, Kumra S, et al. Childhood onset schizophrenia: progressive ventricular change during adolescence. *Arch Gen Psychiatry* 1997;54:897–903.

169. Davis KL, Buchsbaum MS, Shihabuddin L, et al.Ventricular enlargement in poor outcome schizophrenia. *Biol Psychiatry* 1998;43:783–793.

170. O'Donnell BF, Faux SF, McCarley RW, et al. Increased rate of p300 prolongation with age in schizophrenia: electrophysiological evidence for a neurodegenerative process. *Arch Gen Psychiatry* 1995;52:544–549.

171. Mathalon DH, Ford JM, Rosenbloom M, et al. P300 reduction and prolongation with illness duration in schizophrenia. *Biol Psychiatry* 2000;47:413–427.

172. Pfefferbaum A, Wenegrat B, Ford J, et al. Clinical application of the p3 component of relate potentials: dementia, depression, and schizophrenia. *Electroencephalogr Clin Neurophysiol* 1984;49:104–124.

173. Turetsky BI, Moberg PJ, Yousem DM, et al. Reduced olfactory bulb volume in patients with schizophrenia. *Am J Psychiatry* 2000;157:828–830.

174. Moberg PJ, Doty RL, Turetsky BI, et al. Olfactory identification deficits in schizophrenia: correlation with duration of illness. *Am J Psychiatry* 1997;154:1016–1018.

48

Neurocognitive Functioning in Patients with Schizophrenia: An Overview

Terry E. Goldberg

Michael F. Green

Terry E. Goldberg: Clinical Brain Disorders Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland.

Michael F. Green: Department of Psychiatry and Biobehavioral Sciences, University of California School of Medicine, Los Angeles, California.

Increasingly, neurocognitive paradigms are used to study patients with schizophrenia. With such paradigms, the cognitive abnormalities in schizophrenia are characterized by means of experimental and clinical tests. These techniques have indicated that some types of cognitive impairment are not only reliably present in schizophrenia, but are also central and enduring features of the disease. This chapter, a revision of the one published in 1995, focuses on certain recent advances in characterizing the precise nature of cognitive impairments in schizophrenia, on understanding the implications of these for treatment given the course and relationship to outcome of these variables, and on novel applications of neurocognitive approaches to the genetics of schizophrenia.
Cognitive abnormalities were noted by early investigators of schizophrenia. In the original clinical descriptions of schizophrenia made by Kraepelin (64), he commented, “Mental efficiency is always diminished to a considerable degree. The patients are distracted, inattentive… they cannot keep the thought in mind.” Some years later, Shakow (95) began a series of studies in which he examined abnormalities in patients’ reaction time in response to different types of readiness information and imperative stimuli. Hunt and Cofer (54) noted the intellectual quotient (IQ) of schizophrenic patients to be lower than that of normal controls. However, the increasing influence of psychodynamic theory tended to minimize the significance of the cognitive deficits of schizophrenia. It was thought that the deficits displayed on formal psychologic testing were secondary to impaired motivation or cooperation, gross breakdowns in reality testing, or disordered thought processes.

This view changed rapidly with the advent of *in vivo* techniques of brain imaging. First, it became evident that the lateral cerebral ventricles of patients with schizophrenia are larger than those of controls on computed tomography (96). Second, functional brain imaging suggested that the frontal lobe blood flow or metabolism of schizophrenic patients is decreased. Moreover, it was shown that one type of cognitive impairment, poor performance on the Wisconsin Card Sorting Test (WCST), is directly linked to impaired activation of the prefrontal cortex in regional cerebral blood flow (106). It was within this context that a series of studies in which broad neuropsychological test batteries were used demonstrated that patients with chronic schizophrenia could not be reliably discriminated from heterogeneous brain-damaged populations (69). These findings led to a reinterpretation of the original neuropsychological studies; it was increasingly realized that patients with schizophrenia perform in the range typically found in brain-damaged populations because schizophrenia involves structural and functional abnormalities of the brain that are, in some sense, primary, and compromise to a differential degree frontal lobe and temporal lobe function. From this perspective, schizophrenia is viewed as a disease of cortex in which information processing dysfunction is an obligatory concomitant.

We examine certain crucial, conceptually driven issues that derive from this view: What is the course of global cognitive impairment in schizophrenia? What is the character of neurocognitive impairments in schizophrenia? What is the relevance of traits like neurocognitive impairment to linkage or association studies in which the goal is to discover susceptibility genes relevant to the etiology of schizophrenia? We conclude this chapter by noting that neurocognitive impairments may be of prognostic significance in schizophrenia because of the importance of such functions in providing orientation to and encoding relevant environmental information, remembering new information, propitiously

retrieving old information, and working on line with old and new information to make responses or decisions. In this account, cognitive impairments in schizophrenia can be considered target symptoms that must be corrected.

COURSE

Part of "48 - Neurocognitive Functioning in Patients with Schizophrenia: An Overview "

Several contrasting views of the course of cognitive function in schizophrenia are extant. One view suggests that cognitive deficits become progressively worse throughout the long duration of the illness. After an insidious onset, patients' intellectual functions become weaker and social skills become coarser (73). A second view suggests that cognitive deficits, once they arise, remain relatively stable; this view is thus consistent with the notion of a static encephalopathy.

It is clear that once the clinical manifestations of the illness become overt, a sharp decline in cognitive ability takes place in many patients. In longitudinal studies spanning the premorbid and morbid periods, Schwartzman and Douglas (92) found a significant decrement in the performance of schizophrenic patients tested on an army intelligence examination (standard deviation of nearly 0.5), whereas the score of controls improved. (The patients were similar to controls in the premorbid period.) In the study of Weickert et al. (105), about 50% of a large series of treatment-refractory patients exhibited a large decline in IQ (> 10 points) from estimated premorbid levels, although a minority of patients had marked cognitive limitations from early on (see ref. 89). This is not say that subtle premorbid deficits do not exist in the majority of patients. Recent population-based studies have demonstrated attenuations in intelligence measures in schizophrenic patients-to-be (13 ,16), in addition to delays in the attainment of some early developmental milestones (58).

Studies of patients during their first episode of schizophrenia substantiate the view that marked cognitive abnormalities are present at the very onset of the illness. In several studies (8 ,36), the neuropsychological profile of first-episode schizophrenic patients was remarkably similar to that of patients with chronic schizophrenia and did not show a decline at 1- to 2-year follow-up. Both groups of patients performed poorly on a wide range of tests, including tests assessing memory, executive functioning, and attentional abilities.

A number of cross-sectional studies searched for evidence of decline during the chronic phases of the illness. Davidson et al. (14) reported a decline of two to three points per decade in a global measure of cognitive functioning, the Mini-Mental State Examination, across the range of 25 to 95 years. (To place this in perspective, the decline in patients with Alzheimer disease is 1.5 points per year.) Consistent with these results, Harvey et al. (49) recently found that elderly patients in this cohort display a marked decline on a clinical global rating of functioning. However, when patients were followed longitudinally on a variety of cognitive measures for 1- to 2-year periods, little change was evident. As Harvey noted, the changes were not continuous, nor did they occur in the sample as a whole. It is possible the effects observed were secondary to the interaction between compromised cognitive reserve in schizophrenia and normal aging; it is also possible that high doses of neuroleptics and long-term institutionalization had a significant effect on daily living skills and some cognitive functions.

In contradistinction to these findings, Goldstein and Zubin (41) found no differences in performance on the complex cognitive tasks of the Halstead Reitan Battery between large samples of younger and older patients with chronic schizophrenia. Heaton et al. (51) demonstrated that a large sample of older schizophrenic outpatients did not manifest deterioration in performance above and beyond that of normal aging. Hyde et al. (55) used a cross-sectional approach in which successive cohorts of schizophrenic patients were assessed. The study design allowed comparison over an extremely wide range of duration of illness (patients ranged in age from 18 to 70 years). In addition, each cohort was matched on a measure of premorbid intellectual ability, and patients with confounding neurologic or systemic diseases were excluded. No significant differences between age cohorts were noted on tests known to be sensitive to progressive dementias: the Mini-Mental State Examination, Dementia Rating Scale, verbal list learning, and semantic fluency. Thus, over five decades of illness, no progression was noted. A synthesis of these results suggests that in the modal patient, a sharp decline in cognitive ability, including general intellectual efficiency, occurs around the time of the onset of clinical symptoms (± 3 to 5 years), which is followed by an arrest in deterioration and a long period of impaired but stable cognitive function.

This view of the natural history of schizophrenia is consistent with a neurodevelopmental perspective (107) in that a prenatal lesion remains silent for years before manifesting itself in overt symptomatology and cognitive impairment. Contrary to some interpretations, Kraepelin (64) held to this account, stating, "As a rule, if no essential improvement intervenes in at most two or three years after the appearance of the more striking morbid phenomena, a state of weak mindedness will be developed which usually changes slowly and insignificantly." At the very least, the set of findings suggests that cognitive impairment in schizophrenia is an enduring feature of the disorder.

COGNITIVE IMPAIRMENTS

Part of "48 - Neurocognitive Functioning in Patients with Schizophrenia: An Overview "

It is possible that nearly every cognitive function of a schizophrenic patient is impaired, and to an equivalent degree (1 ,2). However, we examine three functions in detail because (a) evidence has been found of differential impairments,

especially in the cognitive domains related to frontal system executive and attentional systems and medial temporal memory systems; (b) these measures are important in regard to outcome (see below); (c) ongoing and systematic experimental work indicates that these cognitive functions can be mapped onto neural systems in a principled manner in normal and schizophrenic persons (30); and (d) such measures are useful in intermediate phenotyping.

Attention

Early descriptions of the clinical phenomenology of schizophrenia emphasized impairment of volitional attention. This clinical observation has been amply supported by many years of experimental study with the use of a wide variety of tasks. Recent models have sharpened the lines between selective attention, shifting attention, and biasing for and encoding relevant target information. We investigate some of these functions by examining three tasks: the Continuous Performance Test (CPT), the Covert Visual Orienting test, and the Stroop Test.

The classic test of selective attention is the Stroop color–word task, in which a word (e.g., red) can be printed in incongruent colors (e.g., green). Depending on instructions, the task is either to name the actual word or name the ink color in which the word is written. The attentional task requires the subject to focus selectively on one dimension of the stimulus and ignore or inhibit contextually inappropriate response tendencies. Normal subjects are slowed when they have to name a color of ink that is incongruent with the word because they have to inhibit their overlearned tendency of reading the word (see ref. 68 for review). Schizophrenic patients may have differential problems on this task in reaction time or accuracy, a finding that has been taken to suggest that they have disproportionate difficulty in inhibiting overlearned tendencies (of reading the word), and may be susceptible to failure in conditions of cognitive conflict more generally, because they are unable to use the contextual information appropriately (e.g., by focusing) (22 ,83).

Another type of task requires covert shifts of attentional resources in response to task instructions or cues, but this time in anticipation of a target in a particular location. It was pioneered by Posner and Dehaene (84). In this paradigm, participants view a central fixation point flanked by two small squares, within which a target is to appear. Participants are to respond as quickly as possible to the target. The reliability with which a preceding cue predicts the location of the target is manipulated and thus provides a measure of two components of selective attention: *engagement* (the benefit of a valid cue as evidenced by a fast response) and *disengagement* (the cost of focusing on an invalid cue followed by orientation and response to the actual location of the target). Although qualitative problems have been reported in patients in this domain (e.g., hemifield-dependent RT effects or a disproportionately slow response to invalid cues), several other studies have not found differences beyond general slowing (35).

A test of “sustained” attention, the CPT, has been used to demonstrate consistently that patients with schizophrenia “miss” targets (77). This task involves monitoring a random series of numbers or letters that are represented continuously, often at a rate of approximately one per second. Participants are asked to detect a target event by pressing a response button and to avoid responding to foils or distracting stimuli.

In an important study of the CPT, Servan-Schreiber et al. (94) showed that when the delay interval between the cue and the stimulus to which a response is to be made was increased to 5 seconds, patients were disproportionately inaccurate in their responding. It is possible that the use of rather long delays changes the basic nature to one of delayed response. Thus, in a recent study, Elvevaag et al. (26) were unable to replicate these findings of delay-induced impairment. Indeed, of the numerous errors made by the patients with schizophrenia, disproportionately more were omission errors at short delay intervals and low target probabilities, a finding taken to suggest a specific problem in rapidly encoding and acting on the imperative stimulus (i.e., constructing a representation of the stimuli to be attended to) under certain unengaging situations (i.e., when few responses are required) or in biasing perceptual representations for target recognition, presumably by an executive system. Based on work on a very different paradigm involving short-term memory in the auditory system, Javitt et al. (56) also proposed that precision or efficacy of encoding is impaired in schizophrenic patients.

Memory

Memory impairment is often the most striking feature of neurocognitive impairment in schizophrenia. Newer work has sought to determine if patients with schizophrenia have qualitative abnormalities in specific stages of mnemonic processing. Toward this end, Elvevaag and colleagues conducted an encoding study in which subjects had to state whether the letter *a* was present in a word (shallow level) or make a decision as to whether the word represented a living thing or not (deep level). Much previous work has demonstrated that words are recalled better when they are encoded deeply. Preliminary results indicated that although patients’ performance was worse than that of controls, they showed the same benefit of deep encoding (B. Elvevaag and T. E. Goldberg, 2001, *unpublished observations*). Although Kareken et al. (60) noted that failures in strategy-driven semantic encoding on this task contributed to impaired performance, their measures were indirect. Elvevaag and colleagues (24) also examined schizophrenic patients’ susceptibility to “false recognition.” They found that patients not only did patients make fewer false-positive errors by incorrectly recognizing semantic lures when poor general memory

(e.g., impaired recall or recognition) was covaried, but they also again were not differentially impaired (24). Consistent with this finding, another study found that susceptibility to interference effects in patients with schizophrenia in so-called AB-ABr paradigms (in which an initial list of paired associates is presented, followed by representation after the items have been "shuffled") is not a differential problem, but rather one that is confounded by general memory problems (B. Elvevaag and T. E. Goldberg, *unpublished observations*). Together, these findings demonstrate that patients with schizophrenia respond in a systematic and lawful manner to a variety of manipulations that target specific mnemonic encoding and orthogonalization (e.g., resistance to interference) processes. Thus, patients may have subtle impairments in different mnemonic processing stations that additively or interactively produce effects of large magnitude.

Moreover, the memory problem in schizophrenia does not appear to be one of binding (the ability to learn associations between various items and distinguish those items from other items that may be similar). This has implications for those who premise aberrant consciousness based on so-called binding abnormalities (12).

Working Memory

Patients with schizophrenia often seem unable to maintain some form of volitional control over the maintenance and manipulation of even basic information. They appear to have difficulty in formulating plans, initiating them, and flexibly changing a strategy once it is no longer effective; they also have difficulty in using feedback efficiently. Moreover, patients sometimes have problems when interrupted; they appear to forget what they were doing after only short periods of interference. One construct that attempts to capture these types of processing failures is working memory, which can involve not only the storage of information over brief delays, but the simultaneous storage and processing of information in a capacity-limited store or computational workspace. These types of behavior have been investigated in various laboratory-based neurocognitive tasks, including the Brown–Peterson test, digit span, WCST, Intradimensional/Extradimensional Set Shifting Test, and various delayed-response tasks.

Patients with schizophrenia have difficulty on the Brown–Peterson test, in which words have to be remembered over short delays during which covert rehearsal is prevented, presumably because of a compromised executive component. Patients are differentially sensitive to longer delays and larger memory sets (38a). However, patients perform abnormally even on basic short-term verbal working memory tasks, including digit span (100)

Several investigators demonstrated that schizophrenic patients exhibit deficits on the WCST, which demands set shifting, response to feedback, and abstraction (28). Patients seem to have difficulty abstracting concepts, and they also make perseverate responses to incorrect responses. Shallice et al. (95a) stressed the consistency of executive deficits in their detailed analyses of single cases, as most patients in their series displayed difficulties in generating rules for the WCST or solving puzzles of the Tower of Hanoi type.

Strong evidence indicates that the WCST may involve the working memory system. For instance, Sullivan et al. (101) found that WCST perseveration is strongly associated with other tests that are thought to require working memory, including self-ordered pointing (in which a subject monitors his or her own series of responses). Gold et al. (34) found the WCST to be highly correlated with a letter–number span task that involves information maintenance and manipulation over short delays. Statistical differences between normal and schizophrenic subjects on the WCST were eliminated when letter-number span performance was covaried, which suggests that both tasks are performed in a similar multimodal or all-purpose cognitive workspace

Much recent work has focused on a task requiring both intradimensional and extradimensional set shifting, in effect a componential version of the WCST. In intradimensional shifts, subjects are required to change their response set to an alternative design within a category (e.g., a new exemplar of a line design) while an irrelevant dimension (e.g., shape) introduced earlier continues to be ignored. In a later stage, an extradimensional shift is demanded as new exemplars are introduced, but subjects are now required to respond to the previously irrelevant dimension (e.g., shapes rather than lines). Subjects make decisions based on feedback after each trial. Patients with chronic schizophrenia display markedly impaired attentional set shifting on the intradimensional/extradimensional task. They demonstrated a significantly higher rate of attrition at the intradimensional shift stage in comparison with patients with frontal lobe lesions, and they were similarly impaired in comparison with patients with frontal lobe lesions at the extradimensional shift stage (79). Patients with chronic disease also showed impairments in regard to Tower tasks, spatial memory span, and spatial working memory tasks. Thus, patients with schizophrenia showed an overall deficit in executive function, often greater than that observed in patients with frontal lobe lesions (80).

Several studies have indicated that an impairment of working memory is present in schizophrenia, even in patients who are relatively intellectually intact. For instance, Pantelis et al. (79) found that although patients with a high IQ performed better than patients with low IQ on the intradimensional/extradimensional task, their performance was still remarkably abnormal, especially in the extradimensional shifts. Elliot et al. (21) were able to confirm these results, even in patients with preserved intellectual function (i.e., IQs > 100). Weickert et al. (105) used a different methodology to reach similar conclusions. They found that nearly all patients—irrespective of whether they exhibited

developmentally compromised intellectual function, normal premorbid intellectual function that declined significantly (the modal subgroup in this study), or preserved intellectual function (i.e., both current and putative premorbid IQ was normal)—displayed deficits in comparison with a normal control group on the WCST measure of perseveration. These results indicate that working memory may represent a core deficit in schizophrenia.

Another set of studies also argues for a deficit in working memory's "visual scratchpad." They are particularly important because failure on this class of delayed response tasks is often taken to be the signature of abnormalities in dorsolateral prefrontal cortex, an area uniquely positioned and designed to exert control over a wide variety of information processing. Using an ocular motor-delayed response paradigm developed by Goldman-Rakic (40) for use in primates, Park et al. (82) found that patients with schizophrenia have grave difficulties maintaining information for location over a brief delay in which they have to perform an interference task. Fleming et al. (29) replicated the findings of this study by using short-term memory for visual patterns. Because patients in these studies also were impaired on control tasks that did not have delays, encoding problems may have contributed to overall level of performance.

NEUROCOGNITION AS AN INTERMEDIATE PHENOTYPE

Part of "48 - Neurocognitive Functioning in Patients with Schizophrenia: An Overview "

Although schizophrenia is a heritable condition, linkage studies in which diagnosis is used as a phenotype have been disappointing, as few significant or replicable chromosomal loci have been identified (20). Using psychiatric diagnosis as the major phenotype may be a major confound. One possible reason is that people may not inherit schizophrenia *per se*, but rather a variety of information-processing deficits from which schizophrenia emerges. In other words, although impairment in any given cognitive process may exact only a small cost in social and vocational functioning, a constellation of impairments may be disabling and result in the emergence of psychosis. Thus, understanding the genetic architecture of individual processes may well be critical for understanding the genetics of "schizophrenia." This account is consistent with a polygenic model of schizophrenia, which implies that the genetic complexity of schizophrenia qua schizophrenia can be reduced by determining affected status based on neurobiological or neurocognitive dimensions; the genetic architecture of these dimensions is simpler than that of schizophrenia but segregates both illness and family risk for illness. This approach involves identifying abnormalities that (a) are quantitative, stable, and enduring; (b) have a pathophysiology that involves neural systems implicated in the disorder; and (c) have a clear effect on outcome. Certain cognitive functions may meet these criteria.

A spate of work has examined early stimulus processing and cognition in relatives of patients with schizophrenia. This so-called high-risk approach has several strengths; for example, abnormalities cannot be attributed to florid psychopathology, cooperation, and medication. It can also be used to identify cognitive processes that may serve as intermediate phenotypes. Several recent studies, in addition to many older ones (65), have produced strong evidence that relatives of patients have subtle impairments in select cognitive functions. In a study of Cannon et al. (6), siblings showed deficit profiles intermediate between those of patients and controls in verbal memory, abstraction, attention, and language. Faraone et al. (27) examined neurocognitive performance in 35 relatives (sibs and children) and 72 normal controls and found deficits in abstraction, attention, and verbal memory in relatives. Classification analysis was highly significant. Studies with other paradigms, including backward masking, delayed response, and verbal working memory, also revealed differences between sibling and controls (10 ,46 ,82). In an important study, Cassens et al. (7) showed that a variety of tasks demanding frontal lobe processing, including complex verbal working memory, semantic encoding, and source monitoring, are not only heritable but are impaired in a stepwise genetically-at-risk fashion in the monozygotic and dizygotic co-twins of schizophrenic persons.

However, simply examining group differences does not directly address issues of familiality/heritability. A newer approach uses computations of relative risk (RR) (88) in necessarily large samples of controls, sibling, and index cases. One type of RR is based on comparisons of concordance rates for impairment on a given trait within sibships with the rate of impairment in the general population. The statistic indicates whether a given quantitative trait is familial and by inference heritable. It is important for predicting the strength of genetic effects on a given phenotype.

In an earlier study in which Goldberg et al. (39) examined monozygotic twins discordant for schizophrenia, they found subtle attenuations of performance in otherwise-well co-twins when they were compared with normal twins on neurocognitive measures indexing working memory, speed of information processing, and episodic memory. The concordance for these traits was thus higher than the concordance for illness. Based on these results, Egan and colleagues (19a) have used a variety of paradigms to assess specific cognitive functions in a sample of schizophrenic index cases, their well siblings, and healthy controls. Specific tests were selected because they reliably measure impairments in schizophrenic patients, are stable, and, in many cases, are known to be heritable. These criteria are obviously of key importance in determining if a person is impaired because of genetic or environmental factors, or simply because of measurement error.

They first assessed RR of the CPT, given that prior work from other groups had suggested that this type of test might be sensitive to certain cognitive impairments in relatives of

patients (19). In the study of Chen et al. (9), who reported an extremely high RR in a Chinese cohort when degraded and nondegraded versions of a CPT were used, seemingly minor features of the methodology, including a sample in which siblings and parents with attendant age differences and a low educational level were combined and psychotic relatives were included, might have led to artifactual inflation of risk computations. Egan et al. (19a) examined 147 patients with schizophrenia, 193 of their siblings, and 47 controls. They did not include parents, and the educational of the groups was high and equivalent in siblings and controls (above grade 13). The IQ of index cases was 94, for their siblings it was 107, and for controls it was 108. The percentage of siblings carrying the schizophrenic spectrum diagnosis was relatively low—under 5%. In a version of the CPT that had flanking distracters, they found that 50% of patients, 24% of siblings, and 18% of controls performed one standard deviation below the control mean when *d′* was used as a dependent measure. The RR for this phenotype was 2.1. This finding suggested that the cognitive demands that this test imposes are under genetic control, the alleles that control this type of information process may be overrepresented in some families of schizophrenic patients, and that this finding is not redundant with diagnosis. However, it was not clear whether CPT impairment is a disease-modifying variable or a susceptibility trait, given that the sibling group as a whole did not differ from controls. In contrast, examination with a test of continuous working memory (the so-called *n*-back task, which demands rapid encoding of stimuli, temporal coding, interference by a restricted set of stimuli, and maintenance) revealed that at "2-back" the RR was above 7.5 and that the sibling group as a whole was significantly impaired in comparison with normal controls, which suggests that the genetic structure that underlies impaired performance may also confer liability (37).

Impairments in several other domains of cognition have also been examined. To assess the suitability of cognitive function for use as a phenotype in genetic studies, Egan et al. estimated RR (19a) in the aforementioned cohort of siblings. They hypothesized that the RR of cognitive dysfunction would be moderate and that different subgroups of families would demonstrate different patterns of impairment. A set of instruments measuring these constructs included IQ, set shifting and working memory, memory, speed, and fluency. RR was estimated by using cutoff scores of one and two standard deviations below the control mean. Patients performed markedly worse than controls on all tests except a measure of premorbid intelligence. The entire sibling group showed impaired performance on the WCST, letter fluency, and Trails B. Siblings of patients with impaired performance also showed deficits on the CVLT, Wechsler Memory Scale-Revised (WMS-R), and Trails A. When one standard deviation was used as the cutoff, the RR of siblings was elevated on the Trails B (RR, 3.8). Trends (p = .01 to .05) toward an increased RR were also seen with the California Verbal Learning Test (CVLT), WCST, letter fluency, memory for stories, and Wide Range Achievement Test (WRAT) (RR, 1.7 to 2.8). When two standard deviations was used as the cutoff, the RRs were generally higher, ranging from 4.3 to more than 13. Correlations between tests of different cognitive functions were weak, which suggests they measure relatively independent processes; factor analysis confirmed this. Multiple regression analysis also demonstrated that impairment on one test did not predict impairment on another test in the sibling group. Thus, cognitive dysfunction along several dimensions is familial and probably genetic. The use of cognitive phenotypes may reduce clinical and genetic heterogeneity and improve the power of genetic studies of schizophrenia.

NEUROCOGNITIVE DEFICITS AND FUNCTIONAL OUTCOME IN SCHIZOPHRENIA

Part of "48 - Neurocognitive Functioning in Patients with Schizophrenia: An Overview "

By any standard, schizophrenia is a remarkably disabling illness. Among young adults in developed countries, it ranks near the top of causes of disability in both men and women (75). There is now increasing support for the idea that key aspects of disability, such as reductions in social competence and the capacity for independent living and vocational success, are the result of neurocognitive compromise.

Although the neurocognitive deficits of schizophrenia have been long recognized, their functional consequences have only recently been appreciated. Throughout most of the twentieth century, studies of the neurocognition of schizophrenia focused rather narrowly on attempts to define and characterize the deficits. However, initial forays to study the implications of these deficits for daily living suggested that neurocognitive deficits may be critical for functional outcome (50). Starting in the early 1990s, a large number of studies examined the associations between rather specific neurocognitive measures and functional outcome in schizophrenia. This being said, individual studies were underpowered with small sample sizes and were mainly atheoretic. To make inferences even more difficult, there was little overlap in either the neurocognitive or the functional outcome measures. Nonetheless, some conclusions from this literature can be drawn.

The literature generally supports the conclusion that neurocognitive deficits are related to functional outcome in schizophrenia (42 ,45), including skill acquisition in psychosocial rehabilitation programs, laboratory assessments of social problem-solving ability or analogue measures of instrumental skills, and broader aspects of behavior in community outcome and activities of daily living.

Indeed, using intrapair differences in twins concordant for schizophrenia, Goldberg et al. (38) observed that virtually

all the variance on the Global Assessment Scale could be accounted for by differences in the performance of four neuropsychological variables: IQ, memory for stories, fluency, and card sorting. In this design, the experience of illness, institutionalization, medication, psychotic symptomatology, and, of course, genome is shared. Although in one sense the design "stacks the deck" because of its artificiality, it does illustrate the importance of neurocognition in predicting level of functioning. This is not to say that symptoms do not have an impact on social and vocational outcome; they do, at least in the short term. What is important to note is that cognitive impairment may also contribute in a unique manner to outcome. These results suggest that patients' deficits in learning new information, rapidly completing tasks, purposefully recalling old information, and generating novel plans or hypotheses may have an impact on their capacity to perform a job efficiently, take part in social transactions, and make decisions.

It is not clear which neurocognitive measures are the most useful predictors and correlates of functional outcome. Despite this lack of consensus, studies on the relationships between neurocognition and functional outcome have frequently included assessments for one or more of the neurocognitive constructs listed in Table 48.1 .

TABLE 48.1. NEUROCOGNITIVE CONSTRUCTS EXAMINED IN STUDIES OF FUNCTIONAL OUTCOME

Secondary memory	*Secondary* (also called *episodic* or *strategic*) *memory* refers to the ability to acquire and store information over a period of time that lasts for at least several minutes. Typically, this type of memory is assessed with a list of words or passages of text. The amount of information in the words or passages exceeds the immediate memory span.
Working or immediate memory	*Immediate memory* refers to the ability to maintain a limited amount of information for a brief time (usually a few seconds). Immediate memory is considered to be a component of working memory. Most of the studies of neurocognition and functional outcome have used passive tasks instead of more typical working memory tasks that require the information to be both maintained and manipulated.
Attention/vigilance	Sometimes called *sustained attention*, this ability involves maintaining a readiness to respond to a particular target stimulus and inhibiting responses to nontargets over a period of time. It requires one to distinguish signal (targets) from noise (nontargets), an ability known as *sensitivity*. It is typically measured with a CPT in which a series of briefly presented stimuli appear on a computer screen; subjects are asked to respond only to selected targets.
Executive functioning/ card sorting	*Executive functioning* refers to volition, planning, purposive action, and self-monitoring of behavior. Problem solving tests such as the WCST are frequently used to assess executive functioning. These tests assess the subject's ability to attain, maintain, and shift cognitive set.

CPT, Continuous Performance Test; WCST, Wisconsin Card Sorting Test.

This literature includes a substantial number of replicated findings (Fig. 48.1). A tally of replications by themselves is not entirely useful because they do not indicate how many times an association was sought, nor the strength of the relationships. Metaanalysis is more useful for examining the strengths of associations across studies. Table 48.2 shows the results of metaanalyses of four key neurocognitive constructs collapsed across the three outcome domains. In this type of analysis, the combined sample sizes are large and the relationships between neurocognition and functional outcome are highly significant. The metaanalyses demonstrate that these four neurocognitive constructs are significantly related to functional outcome and that the effect sizes for these relationships are generally in the medium range.

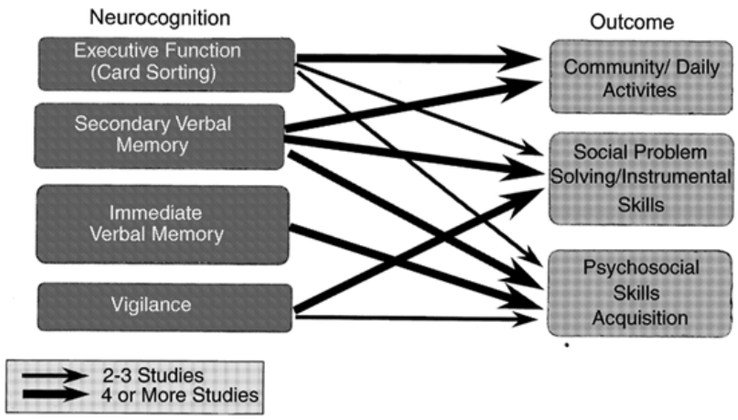


FIGURE 48.1. Neurocognitive constructs and functional outcome. Two levels of replication are represented. A *heavy arrow* indicates that at least four studies found a significant relationship between the neurocognitive construct and the outcome domain; a *thin arrow* indicates that significant relationships were uncovered in two or three studies.

TABLE 48.2. METAANALYSES: NEUROCOGNITIVE PREDICTION OF FUNCTIONAL OUTCOME

Domain	Total Sample Size	Pooled Estimated r^a	Effect Size	*p* Value
Secondary verbal memory	727	.29	Medium	<.0001
Immediate verbal memory	188	.40	Medium–Large	<.0001
Executive functions (card sorting)	1002	.23	Small–Medium	<.0001
Attention/vigilance	682	.20	Small–Medium	<.0001

[a]Estimates weighted by sample size.
From Green MF, Kern RS, Braff DL, et al. Neurocognitive deficits and functional outcome in schizophrenia: are we measuring the "right stuff"? *Schizophr Bull* 2000; 26:119–136, with permission.

Most of the studies in this area have used rather specific measures of neurocognition, and it is the results of these

studies that are reflected in Fig. 48.1 and Table 48.2 . Although effect sizes for the individual constructs are mainly in the medium range, they can become quite large when global or composite measures of neurocognition are used instead of individual measures (48 ,104). Such composite measures indicate that neurocognition can explain between 20% and 60% of the variance in outcome.

How do these relationships compare with those for clinical symptoms? In general, psychotic symptoms (hallucinations and delusions) fare rather poorly as predictors and correlates of functional outcome (43). Negative symptoms are more highly correlated with functional outcome, but across studies, the relationships are neither stronger nor more consistent than those for neurocognitive deficits (17 ,48 ,104). Little is known about disorganized symptoms, which often constitute a separate syndromal dimension that includes formal thought disorder, although recent studies suggest that this type of symptom may be related to functional outcome (76 ,86).

The relative contributions of symptoms and neurocognition to functional outcome have only rarely been tested with appropriate statistical analyses, including multiple regression (48 ,71). These studies do, however, support the idea that the neurocognitive contributions to outcome are stronger than those of symptoms. In one study (104), sophisticated path analyses were used to test the associations among positive symptoms, negative symptoms, cognition, and activities of daily living in two separate samples of schizophrenic patients. A global measure of cognition had strong relationships with activities of daily living (48% and 42% of the variance in the activities of daily activities for the two samples). Various causal models were tested in which certain pathways were omitted. The pathway from cognitive impairment to functional outcome was necessary in the model; the fit was poor when it was omitted. To the extent that symptoms were correlated with functional outcome, the relationships seem to be indirect. In other words, although negative symptoms covary to at least a modest extent with neurocognition (17 ,104) and their relationship to function appears to be mediated through this overlap, *in toto* the results suggest that cognitive impairment, rather than symptoms, most strongly influences functional outcome.

Just as the neurocognitive deficits in schizophrenia are not fully specific to schizophrenia, the correlations with functional outcome are unlikely to be specific to schizophrenia. Based on the role of neurocognitive deficits in other disorders, one would not one expect them to be. The functional consequences of neurocognitive deficits have been observed in a variety of neurologic conditions, including head injury, Alzheimer disease, multiple sclerosis, Parkinson disease, and AIDS encephalopathy (51 ,87 ,103). In fact, neurocognitive deficits have been associated with activities of daily living even in a nonclinical sample of elderly persons (74).

The work so far has been aimed at determining *whether* neurocognition is related to functional outcome. At this time, it can be concluded that it is related, and the effect sizes are generally medium for individual constructs and generally large for composite measures. However, rather little is known about *how* neurocognition is related to functional outcome. It is likely that some cognitive domains have direct, causal relationships, although others may be related to functional outcome through mediators, such as social cognition or the application of knowledge and reasoning to problem solving.

EFFECTS OF MEDICATIONS ON NEUROCOGNITIVE DEFICITS

Part of "48 - Neurocognitive Functioning in Patients with Schizophrenia: An Overview "

One of the most surprising aspects of conventional antipsychotic medications is that although they usually have a profound impact on psychotic symptoms, their effects on neurocognitive

deficits tend to be negligible (7 ,11 ,98). Occasionally, treatment with conventional antipsychotic medications has led to improvement in basic perceptual or attentional processes (3 ,98). However, it can be concluded that changes in neurocognition, if they occur, are small compared with the changes in psychotic symptoms. In terms of disability, this presents a rather unfortunate mismatch in which the domain of illness most affected by conventional medications is not the domain most closely linked to functional outcome.

Conventional antipsychotic medications probably do not directly impair neurocognitive abilities, but they can do so indirectly when they involve the simultaneous administration of anticholinergic medications. Anticholinergic medications given for extrapyramidal side effects compromise certain neurocognitive abilities. Although the range of effects of anticholinergic medications is not well characterized, they may disrupt aspects of secondary verbal memory that rely on rehearsal strategies (18). Other aspects of memory, including immediate or working memory, appear to be less affected (4 ,38), and the effects on other neurocognitive abilities, such as visual processing, are relatively unknown.

The situation with newer atypical antipsychotic agents appears to be more promising. Initial interest in the neurocognitive effects of new antipsychotic medications was stimulated by a series of (mainly open-label) studies of clozapine (4 ,38 ,47 ,53 ,67). The results of these studies were surprising in two respects: First, in most of the studies, clozapine treatment resulted in improvement in verbal fluency (i.e., the ability to generate words that begin with a certain letter or belong to a certain semantic category) and possibly psychomotor speed. Second, the initiation of clozapine treatment in some studies appeared to have at least short-term detrimental effects on visual memory and possibly verbal working memory (38 ,47 ,53).

A large number of studies are emerging for recently approved antipsychotic medications: risperidone, olanzapine, and quetiapine (31 ,85 ,90 ,99). These studies (again, mostly open-label) have generally shown that they have benefits for neurocognition in comparison with conventional antipsychotic medications. Indications of short-term detrimental effects, similar to those seen in some clozapine studies, have so far not been reported for the other newer antipsychotic medications. A rather comprehensive review (72) and a metaanalysis (61) of the existing literature have both provided a basis for optimism about the beneficial neurocognitive effects of newer medications. The metaanalysis of Keefe et al. (61) showed significant effects for the new generation of medications in comparison with conventional agents across a range of neurocognitive areas, including attention, executive functions, and verbal fluency.

The emerging optimism in this area should be tempered by the fact that the lion’s share of the studies have been “open-label,” with the associated risks of experimental bias that can accompany such studies. In many of these studies, a single group was assessed at baseline while on a conventional medication and then assessed again after being switched to an atypical medication. Inferences from these types of studies are necessarily tentative because no control is made for repeated testings and possible practice effects. A small number of parallel group blinded studies are emerging for clozapine (4 ,108), risperidone (44 ,62), and olanzapine (85). These studies offer more convincing support for the proposal that new medications convey neurocognitive benefits in comparison with conventional medications.

Even more than clinical trial studies of symptom reduction, studies of neurocognitive effects raise questions about alternative explanations for treatment effects. For the most part, studies have not been designed or analyzed in a way that allows one to rule out indirect effects. For example, if a newer antipsychotic medication has a better clinical effect than a conventional medication, it may improve neurocognition as an indirect benefit of greater symptom reduction. This explanation, although plausible, seems unlikely. Several studies have noted that changes in neurocognition appear to be independent of any changes in symptoms (38 ,44 ,47 ,67 ,62). An alternative explanation is that the neurocognitive benefits of newer medications are mediated by a reduced need for anticholinergic medications. Although this may turn out be true in some instances, the differential use of anticholinergic medications did not explain the effects of risperidone on immediate and secondary verbal memory in one project (44 ,92). The beneficial effects of newer medications on neurocognition may be mediated by lower rates of extrapyramidal symptoms. It is not known whether such effects will be seen in comparisons with very low doses of conventional medications, when side effects are minimal. Moreover, although a number of mechanisms have been proposed (5-hydroxytryptamine subtype 2A antagonism, indirect glutamate release, dopamine D4 antagonism), all are problematic, and simple reduction of D2 blockade or transient D2 blockade remains a viable explanation of “atypicality” (59). Thus, it is possible that the administration of conventional neuroleptics in inappropriately high doses resulted in a lack of improvement, although dose-reduction studies do not support this explanation (93 ,97). The pharmacologic basis for cognitive enhancement remains obscure. In any event, a single neurotransmitter effect seems unlikely to account for the effects, which probably involve a constellation of actions at serotonergic, adrenergic, cholinergic, and dopaminergic receptors (72). However, it is important to recognize that such actions probably are initiators of changes that ultimately effect gene expression for major excitatory and inhibitory transmitters and their receptors and neuroplasticity.

Although these alternative explanations require serious consideration, it remains possible that the neurocognitive effects of the new generation of antipsychotic medications

are a direct result of the medications themselves (rather than indirect effects of a reduction of clinical symptoms or anticholinergic medications). If so, the effects are serendipitous. These medications were not developed or initially evaluated with neurocognition in mind. The possible role of adjunctive pharmacology specifically for neurocognitive deficits is now receiving serious consideration.

A key challenge for directing studies of adjunctive nosotropic medications in schizophrenia is deciding which neurotransmitter systems are most critical in the pathophysiology of neurocognitive deficits in schizophrenia. One that has been implicated is the glutamate system (78). Based on the cognitive and behavioral effects of antagonists of the *N*-methyl-D-aspartate (NMDA) subtype of glutamate receptor, such as phencyclidine and ketamine, it has been suggested that schizophrenia may involve hypofunction of this receptor (66 ,70). Because the NMDA receptor is modulated by glycine, glycinergic agents can provide a useful means for manipulating glutamate function. Glycine itself has been utilized by Javitt et al. (56) with some effect on negative symptoms. The administration of D-cycloserine, a partial glycine agonist, resulted in a benefit in choice reaction time when it was utilized in conjunction with conventional antipsychotic medications (32), but not when added to clozapine (33). A full glycine agonist, D-serine, demonstrated some success in improving WCST performance (102). If these studies of adjunctive agents result in reliable improvement in neurocognition in schizophrenia, it may become routine for schizophrenic patients to receive a medication for each of the major domains of illness, clinical symptoms and neurocognitive deficits.

CONCLUSIONS

Part of "48 - Neurocognitive Functioning in Patients with Schizophrenia: An Overview "

In the lay imagination, schizophrenic patients experience problems in living because they are divided against themselves, out of touch with reality, and disorganized. The view of scientists, once not altogether different, has changed. Not only have the symptoms been defined and codified, but the neurobiological underpinnings of the disorder have begun to be described. Emerging also is a view in which cognitive impairments may be a relatively central feature of the disorder. Cognitive impairments are involved in the genetic etiology of schizophrenia. They seem enduring in that they are present for much of the clinical history and are associated with outcome. Cognitive impairments also may have a relatively well-delineated profile in which executive, memory, and attentional deficits are prominent. This account carries with it implications for treatment, in that cognitive impairments should be considered target symptoms in the same way as hallucinations, delusions, and anergia.

REFERENCES

1. Blanchard JJ, Neale JM. The neuropsychological signature of schizophrenia: generalized or differential deficit? *Am J Psychiatry* 1994;151:40–48.

2. Braff DL, Heaton R, Kuck J, et al. The generalized pattern of neuropsychological deficits in outpatients with chronic schizophrenia with heterogeneous Wisconsin Card Sorting Test results. *Arch Gen Psychiatry* 1991;48:891–898.

3. Braff DL, Saccuzzo DP. Effect of antipsychotic medication on speed of information processing in schizophrenic patients. *Am J Psychiatry* 1982;139:1127–1130.

4. Buchanan RW, Holstein C, Breier A. The comparative efficacy and long-term effect of clozapine treatment on neuropsychological test performance. *Biol Psychiatry* 1994;36:717–725.

5. Deleted in press.

6. Cannon TD, Zorrilla LE, Shtasel D, et al. Neuropsychological functioning in siblings discordant for schizophrenia and healthy volunteers. *Arch Gen Psychiatry* 1994;51:651–661.

7. Cassens G, Inglis AK, Appelbaum PS, et al. Neuroleptics: effects on neuropsychological function in chronic schizophrenic patients. *Schizophr Bull* 1990;16:477–499.

8. Censits DM, Ragland JD, Gur RC, et al. Neuropsychological evidence supporting a neurodevelopmental model of schizophrenia. *Schizophr Res* 1997;24:289–298.

9. Chen WJ, Liu SK, Chang CJ, et al. Sustained attention deficit and schizotypal personality features in nonpsychotic relatives of schizophrenic patients. *Am J Psychiatry* 1998;155:1214–1220.

10. Conklin HM, Curtis CE, Katsanis J, et al. Verbal working memory impairment in schizophrenia patients and their first-degree relatives: evidence from the digit span task. *Am J Psychiatry* 2000;157:275–277.

11. Cornblatt B, Obuchowski M, Schnur DB, et al. Attention and clinical symptoms in schizophrenia. *Psychiatr Q* 1997;68:343–359.

12. Danion JM, Rizzo L, Bruant A. Functional mechanisms underlying impaired recognition memory and conscious awareness in patients with schizophrenia. *Arch Gen Psychiatry* 1999;56:639–644.

13. David AS, Malmberg A, Brandt L, et al. IQ and risk for schizophrenia: a population-based cohort study. *Psychol Med* 1997;27:1311–1323.

14. Davidson M, Harvey PD, Powchik P, et al. Severity of symptoms in chronically institutionalized geriatric schizophrenic patients. *Am J Psychiatry* 1995;152:197–207.

15. Deleted in press.

16. Davidson M, Reichenberg A, Rabinowitz J, et al. Behavioral and intellectual markers for schizophrenia in apparently healthy male adolescents. *Am J Psychiatry* 1999;156:1328–1335.

17. Dickerson F, Boronow JJ, Ringel N, et al. Neurocognitive deficits and social functioning in outpatients with schizophrenia. *Schizophr Res* 1996;21:75–83.

18. Drachman DA, Leavitt J. Human memory and the cholinergic system. *Arch Neurol* 1974;30:113–121.

19. Egan MF, Goldberg TE, Gscheidle T, et al. Relative risk of attention deficits in siblings of patients with schizophrenia. *Am J Psychiatry* 2000;157:1309–1316.

19a. Egan MF, Goldberg TE, Gscheidle T, et al. Relative risk for cognitive impairments in siblings of patients with schizophrenia *Biol Psych* 2001;50:98–107.

20. Egan MF, Weinberger DR. Neurobiology of schizophrenia. *Curr Opin Neurobiol* 1997;7:701–707.

21. Elliott R, McKenna PJ, Robbins TW, et al. Neuropsychological evidence for frontostriatal dysfunction in schizophrenia. *Psychol Med* 1995;25:619–630.

22. Elvevaag B, Duncan, J, McKenna, PJ. The use of cognitive context in schizophrenia: an investigation. *Psychol Med* 2000;30:885–897.

23. Deleted in press.

24. Elvevaag B, Egan, MF, Goldberg TE. Paired-associate learning and memory interference in schizophrenia. *Neuropsychologia* 2000;1565–1575.

25. Deleted in press.

26. Elvevaag B, Weinberger DR, Suter JC, et al. Continuous performance test and schizophrenia: a test of stimulus-response compatibility, working memory, response readiness, or none of the above? *Am J Psychiatry* 2000;157:772–780.

27. Faraone SV, Seidman LJ, Kremen WS, et al. Neuropsychological functioning among the nonpsychotic relatives of schizophrenic patients: a diagnostic efficiency analysis. *J Abnorm Psychol* 1995;104:286–304.

28. Fey ET. The performance of young schizophrenics and young normals on the Wisconsin Card Sorting Test. *J Consult Psychol* 1951;15:311–319.

29. Fleming K, Goldberg TE, Binks S, et al. Visuospatial working memory in patients with schizophrenia. *Biol Psychiatry* 1997;41:43–49.

30. Frackowiak RSJ, Friston KJ, Frith CD, et al., eds. *Human brain function*. New York: Academic Press, 1997.

31. Gallhofer B, Bauer U, Lis S, Krieger S, et al. Cognitive dysfunction in schizophrenia: comparison of treatment with atypical antipsychotic agents and conventional neuroleptic drugs. *Eur Neuropsychopharmacol* 1996;6[Suppl 2]:13–20.

32. Goff DC, Tsai G, Manoach DS, et al. Dose-finding trial for D-cycloserine added to neuroleptics for negative symptoms in schizophrenia. *Am J Psychiatry* 1995;152:1213–1215.

33. Goff DC, Tsai G, Manoach DS, et al. D-Cycloserine added to clozapine for patients with schizophrenia. *Am J Psychiatry* 1996;153:1628–1630.

34. Gold JM, Carpenter C, Randolph C, et al. Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Arch Gen Psychiatry* 1997;54:159–165.

35. Gold JM, Randolph C, Coppola R, et al. Visual orienting in schizophrenia. *Schizophr Res* 1992;7:203–209.

36. Gold S, Arndt S, Nopoulos P, et al. Longitudinal study of cognitive function in first-episode and recent-onset schizophrenia. *Am J Psychiatry* 1999;156:1342–1348.

37. Goldberg TE, Egan MF, Weinberger DR. Relative risk for working memory and reaction time in families of individuals with schizophrenia *(submitted)*.

38. Goldberg TE, Greenberg RD, Griffin SJ, et al. The effect of clozapine on cognition and psychiatric symptoms in patients with schizophrenia. *Br J Psychiatry* 1993;162:43–48.

38a. Goldberg TE, Patterson K, Taqqu Y, et al. Capacity limitations in short-term memory in schizophrenia. *Psychol Med* 1998;28:665–673.

39. Goldberg TE, Torrey EF, Gold JM, et al. Genetic risk of neuropsychological impairment in schizophrenia: a study of monozygotic twins discordant and concordant for the disorder. *Schizophr Res* 1995;17:77–84.

40. Goldman-Rakic PS. Regional and cellular fractionation of working memory. *Proc Natl Acad Sci U S A* 1996;93:13473–13480.

41. Goldstein G, Zubin J. Neuropsychological differences between young and old schizophrenics with and without associated neurological dysfunction. *Schizophr Res* 1990;3:117–126.

42. Green MF. What are the functional consequences of neurocognitive deficits in schizophrenia? *Am J Psychiatry* 1996;153:321–330.

43. Green MF, Kern RS, Braff DL, et al. Neurocognitive deficits and functional outcome in schizophrenia: are we measuring the “right stuff”? *Schizophr Bull* 2000;26:119–136.

44. Green MF, Marshall BD, Wirshing WC, et al. Does risperidone improve verbal working memory in treatment-resistant schizophrenia? *Am J Psychiatry* 1997;154:799–804.

45. Green MF, Nuechterlein KH. Should schizophrenia be treated as a neurocognitive disorder? *Schizophr Bull* 1999;25:309–319.

46. Green MF, Nuechterlein KH, Breitmeyer B. Backward masking performances in unaffected siblings of schizophrenic patients. Evidence for a vulnerability indicator. *Arch Gen Psychiatry* 1997;54:465–472.

47. Hagger C, Buckley P, Kenny JT, et al. Improvement in cognitive functions and psychiatric symptoms in treatment-refractory schizophrenic patients receiving clozapine. *Biol Psychiatry* 1993;34:702–712.

48. Harvey PD, Howanitz E, Parrella M, et al. Symptoms, cognitive functioning, and adaptive skills in geriatric patients with lifelong schizophrenia: s comparison across treatment sites. *Am J Psychiatry* 1998;155:1080–1086.

49. Harvey PD, Parrella M, White L, et al. Convergence of cognitive and adaptive decline in late-life schizophrenia. *Schizophr Res* 1999;35:77–84.

50. Heaton RK, Pendleton MG. Use of neuropsychological tests to predict adult patients’ everyday functioning. *J Consult Clin Psychol* 1981;49:807–821.

51. Heaton RK, Velin RV, McCutchan A, et al. Neuropsychological impairment in human immunodeficiency virus-infection: implications for employment. *Psychosom Med* 1994;56:8–17.

52. Deleted in press.

53. Hoff AL, Faustman WO, Wieneke M, et al. The effects of clozapine on symptom reduction, neurocognitive function, and clinical management in treatment refractory state hospital schizophrenic inpatients. *Neuropsychopharmacology* 1996;15:361–369.

54. Hunt J McV, Cofer CN. Psychological deficit. In: Hunt J McV, ed. *Personality and the behavior disorders*. New York: Ronald Press, 1944.

55. Hyde TM, Nawroz S, Goldberg TE, et al. Is there cognitive decline in schizophrenia? A cross-sectional study. *Br J Psychiatry* 1994;164:494–500.

56. Javitt DC, Liederman E, Cienfuegos A, et al. Panmodal processing imprecision as a basis for dysfunction of transient memory storage systems in schizophrenia. *Schizophr Bull* 1999;25:763–775.

57. Deleted in press.

58. Jones P, Cannon M. The new epidemiology of schizophrenia. *Psychiatr Clin North Am* 1998;21:1–25.

59. Kapur S, Zipursky R, Jones C, et al. A positron emission tomography study of quetiapine in schizophrenia: a preliminary finding of an antipsychotic effect with only transiently high dopamine D2 receptor occupancy. *Arch Gen Psychiatry* 2000;57:553–559.

60. Kareken DA, Moberg PJ, Gur RC. Proactive inhibition and semantic organization: relationship with verbal memory in patients with schizophrenia. *J Int Neuropsychol Soc* 1996;2:486–493.

61. Keefe RSE, Silva SG, Perkins DO, et al. The effects of atypical antipsychotic drugs on neurocognitive impairment in schizophrenia: a review and meta-analysis. *Schizophr Bull* 1999;25:201–222.

62. Kern RS, Green MF, Marshall BD, et al. Risperidone vs. haloperidol on secondary memory: can newer antipsychotic medications aid learning? *Schizophr Bull* 1999;25:223–232.

63. Deleted in press.

64. Kraepelin E. In: Robertson GM, ed. *Dementia praecox and paraphrenia*. Melbourne, FL: Robert E. Krieger, 1971 (originally published in 1919).

65. Kremen WS, Seidmen LJ, Pepple JR, et al. Neuropsychological risk indicators for schizophrenia: a review of family studies. *Schizophr Bull* 1994;20:103–119.

66. Krystal JH, Karper LP, Seibyl JP, et al. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans: psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry* 1994;51:199–214.

67. Lee MA, Jayathilake K, Meltzer HY. A comparison of the effect of clozapine with typical neuroleptics on cognitive function in the neuroleptic-responsive schizophrenic. *Schizophr Res* 1999;37:1–11.

68. MacLeod CM. Half a century of research on the Stroop effect: an integrative review. *Psychol Bull* 1991;109:163–203.

69. Malec J. Neuropsychological assessment of schizophrenia versus brain damage: a review. *J Nerv Ment Dis* 1978;166:507.

70. Malhotra AK, Pinals DA, Weingartner H, et al. NMDA receptor function and human cognition: the effects of ketamine in healthy volunteers. *Neuropsychopharmacol* 1996;14: 301–307.

71. McGurk SR, Moriarty PJ, Harvey PD, et al. The longitudinal relationship of clinical symptoms, cognitive functioning and adaptive life in geriatric schizophrenia. *Schizophr Res* 2000;42:47–55.

72. Meltzer HY, McGurk SR. The effect of clozapine, risperidone, and olanzapine on cognitive function in schizophrenia. *Schizophr Bull* 1999;25:233–255.

73. Miller R. Schizophrenia as a progressive disorder: relations to EEG, CT, neuropathological and other evidence. *Prog Neurobiol* 1989;33:17–44.

74. Moritz DJ, Kasl SV, Berkman LF. Cognitive functioning and the incidence of limitations in activities of daily living in an elderly community sample. *Am J Epidemiol* 1995;141: 41–49.

75. Murray CJL, Lopez AD, eds. *The global burden of disease*. Boston: Harvard School of Public Health, 1996.

76. Norman RMG, Malla AK, Cortese L, et al. Symptoms and cognition as predictors of community functioning: a prospective analysis. *Am J Psychiatry* 1999;156:400–405.

77. Nuechterlein KH. Vigilance in schizophrenia and related disorders. In: Steinhauer SR, Gruzelier JH, Zubin J, eds. *Handbook of schizophrenia. Neuropsychology, psychophysiology and information processing* (vol 5). Amsterdam: Elsevier Science, 1991:397–433.

78. Olney JW, Farber NB. Glutamate receptor dysfunction and schizophrenia. *Arch Gen Psychiatry* 1995;52:998–1007.

79. Pantelis C, Barber FZ, Barnes TR, et al. Comparison of set-shifting ability in patients with chronic schizophrenia and frontal lobe damage. *Schizophr Res* 1999;37:251–270.

80. Pantelis C, Barnes TR, Nelson HE, et al. Frontal-striatal cognitive deficits in patients with chronic schizophrenia. *Brain* 1997;120:1823–1843.

81. Deleted in press.

82. Park S, Holzman PS, Goldman-Rakic PS. Spatial working memory deficits in the relatives of schizophrenic patients. *Arch Gen Psychiatry* 1995;52:821–828.

83. Perlstein WM, Carter CS, Barch DM, et al. The Stroop task and attention deficits in schizophrenia: a critical evaluation of card and single-trial Stroop methodologies. *Neuropsychology* 1998;12:414–425.

84. Posner MI, Dehaene S. Attentional networks. *Trends Neurosci* 1994;17:75–79.

85. Purdon SE, Jones BDW, Stip E, et al. Neuropsychological change in early phase schizophrenia during 12 months of treatment with olanzapine, risperidone, and haloperidol. *Arch Gen Psychiatry* 2000;57:249–258.

86. Racenstein JM, Penn D, Harrow M, et al. Thought disorder and psychosocial functioning in schizophrenia: the concurrent and predictive relationships. *J Nerv Ment Dis* 1999;187:281–289.

87. Rao SM, Leo GJ, Ellington L, et al. Cognitive dysfunction in multiple sclerosis: II. Impact on employment and social functioning. *Neurology* 1991;41:692–696.

88. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;273:1516–1517.

89. Russell AJ, Munro JC, Jones PB, et al. Schizophrenia and the myth of intellectual decline [see Comments]. *Am J Psychiatry* 1997;154:635–639.

90. Sax KW, Strakowski SM, Keck PE. Attentional improvement following quetiapine fumarate treatment in schizophrenia. *Schizophr Res* 1998;33:151–155.

91. Deleted in press.

92. Schwartzman AE, Douglas VI. Intellectual loss in schizophrenia part 11. *Can J Psychol* 1962;16:161–168.

93. Seidman LJ, Pepple JR, Faraone SV, et al. Neuropsychological performance in chronic schizophrenia in response to neuroleptic dose reduction. *Biol Psychiatry* 1993;33:575–584.

94. Servan-Schreiber D, Cohen JD, Steingard S. Schizophrenic deficits in the processing of context. A test of a theoretical model [see Comments]. *Arch Gen Psychiatry* 1998;55:186–188.

95. Shakow D. *Adaptation in schizophrenia: the theory of segmental set*. New York: John Wiley and Sons, 1979.

95a. Shallice T, Burgess PW, Frith CD. Can the neuropsychological case-study approach be applied to schizophrenia? *Psychol Med* 1991;21:661–673.

96. Shelton RC, Weinberger DR. X-ray computerized tomography studies in schizophrenia: a review and synthesis. In: Nasrallah HA, Weinberger DR, eds. *Handbook of schizophrenia*. Amsterdam: Elsevier Science, 1986:207–250.

97. Spohn HE, Coyne L, Lacoursiere R, et al. Relation of neuroleptic dose and tardive dyskinesia to attention, information-processing, and psychophysiology in medicated schizophrenics. *Arch Gen Psychiatry* 1985;42:849–859.

98. Spohn HE, Strauss ME. Relation of neuroleptic and anticholinergic medication to cognitive function in schizophrenia. *J Abnorm Psychol* 1999;98:367–390.

99. Stip E, Lussier I. The effect of risperidone on cognition in patients with schizophrenia. *Can J Psychol* 1996;41[Suppl 2]:35–40.

100. Stirling JD, Hellewell JS, Hewitt J. Verbal memory impairment in schizophrenia: no sparing of short-term recall [see Comments]. *Schizophr Res* 1997;25:85–95.

101. Sullivan EV, Mathalon DH, Zipursky RB, et al. Factors of the Wisconsin Card Sorting Test as measures of frontal-lobe function in schizophrenia and in chronic alcoholism. *Psychiatry Res* 1992;46:175–199.

102. Tsai G, Yang P, Chung L-C, et al. D-Serine added to antipsychotics for the treatment of schizophrenia. *Biol Psychiatry* 1998;44:1081–1089.

103. Van Gorp WG, Baerwal JP, Ferrando SJ, et al. The relationship between employment and neuropsychological impairment in HIV infection. *J Int Neuropsychol Soc* 1999;6:534–539.

104. Velligan DI, Mahurin RK, Diamond PL, et al. The functional significance of symptomatology and cognitive function in schizophrenia. *Schizophr Res* 1997;25:21–31.

105. Weickert TW, Goldberg TE, Gold JM, et al. Cognitive impairments in patients with schizophrenia displaying preserved and compromised intellect. *Arch Gen Psychiatry* 2000;57:907–913.

106. Weinberger DR, Berman KF, Zec RF. Psychologic dysfunction of dorsolateral prefrontal cortex in schizophrenia: 1. Regional cerebral blood flow evidence. *Arch Gen Psychiatry* 1986;43:114–124.

107. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 1987;44:660–669.

108. Zahn TP, Pickar D, Haier RJ. Effects of clozapine, fluphenazine, and placebo on reaction time measures of attention and sensory dominance in schizophrenia. *Schizophr Res* 1994;13:133–144.

49

Molecular and Population Genetics of Schizophrenia

Ming T. Tsuang

Michael J. Owen

Ming T. Tsuang: Department of Psychiatry, Harvard Medical School; Department of Epidemiology, Harvard School of Public Health; Harvard Institute of Psychiatric Epidemiology and Genetics, Boston, Massachusetts.

Michael J. Owen: Department of Psychological Medicine, University of Wales College of Medicine, Cardiff, Wales.

For much of this century, we have believed that genes play a role in the etiology of schizophrenia, but we have been frustrated in the search for specific mutations by diagnostic dilemmas and technologic shortcomings. Slowly, however, we have made substantive strides in the diagnosis and treatment of schizophrenia, and considerable progress toward an explication of its underlying neurobiology. In this chapter, we focus on the epidemiologic and genetic work that underlies much of the current clinical progress and the promises, perhaps to be fulfilled in the not-too-distant future, to facilitate the treatment, and even prevention, of this devastating disorder. We begin with updates on the epidemiology of schizophrenia and related disorders, followed by a review of its molecular genetic bases. We then consider research strategies that promise to explicate the genetic etiology of schizophrenia further in the next few years.

EPIDEMIOLOGY OF SCHIZOPHRENIA

Part of "49 - Molecular and Population Genetics of Schizophrenia "

Psychiatric epidemiology studies the distribution of disorders in well-defined populations. Its methodology emphasizes the use of representative samples with reliable and valid diagnoses, and a specified time of onset (1 ,2), to identify risk factors that explain why some populations are more vulnerable to psychiatric illness than others. Epidemiologic methods are also critical to an understanding of how frequently a disorder occurs, a concept often expressed in terms of prevalence, incidence, and lifetime risk. We consider the risk for schizophrenia by focusing on these measures.

Prevalence Rate

The prevalence rate of a disorder is the number of people in whom the disorder is diagnosed divided by the total number of persons examined in the population under study. The computed rate depends on several factors: the definition of the disorder, the total number of individuals examined in the population, and the procedure used to choose whom to examine. Ideally, the sample used to compute prevalence should be representative of the population as a whole. The prevalence rate is usually expressed as the number of cases per thousand people surveyed within a year, which is called the *1-year prevalence per thousand*. Studies of schizophrenia from around the world usually report these rates to range from a low of 0.6/1,000 to a high of 17/1,000 (3 ,4 ,5 and 6). Lifetime prevalence rates usually range from 0.9/1,000 to 3.8/1,000 (7 ,8 ,9 and 10), depending on the diagnostic criteria and methods of ascertainment.

The prevalence rates for schizophrenia are relatively constant between countries. Whether we consider East versus West, developed countries versus less developed countries, or other classifications, the 1-year prevalence of schizophrenia is approximately 0.5% and the lifetime prevalence is approximately 1.0%. In other words, schizophrenia is found in approximately one-half of one percent of the population at any point in time.

Incidence Rate

Another way of reporting the rate of schizophrenia in a population is to estimate the number of *new* cases that appear in the population during a specified period of time; this is called the *incidence rate*. Prevalence rates (as discussed above) include both new and old cases because once schizophrenia has emerged, it usually demonstrates a chronic, unremitting course. In other words, once patients are classified as schizophrenic, they usually remain schizophrenic. Like prevalence rates, incidence rates vary according to a number of variables, including the standards of diagnosis. In addition, incidence studies of schizophrenia entail the difficulty

of how to define onset. Unlike the onset of some disorders (e.g., stroke or head injury), that of schizophrenia is often insidious and confused easily with other problems. For example, early symptoms like social withdrawal or unusual thinking may be ignored or mistaken for indications of depression or substance abuse. Because the time of onset is difficult to determine, incidence rates are usually based on a patient's first visit to psychiatric services for schizophrenic symptoms. Thus, misdiagnosed and untreated cases can affect the accuracy of incidence rates significantly.

The incidence rate is usually expressed as the number of new cases in a given period per 100,000 population. For schizophrenia, incidence rates range from a low of 0.10 to a high of 0.70 (11 ,12). As was the case with the prevalence figures, the incidence of schizophrenia is generally stable over time and across geographic areas (13).

Lifetime Risk

Most persons with schizophrenia first become ill between 20 and 39 years of age. We call this the *high-risk period* for schizophrenia. Men tend to be younger at the time of onset than women (14 ,15 and 16), although schizophrenia develops in men and women at approximately equal rates (2). Because of the variability of age at onset, prevalence and incidence rates vary according to the age and sex composition of the population studied. The age distribution is particularly important when the probability or risk that schizophrenia will develop in a person during his or her lifetime is estimated (i.e., the lifetime risk). To estimate lifetime risk, the age distribution of the population surveyed should be taken into account (17).

The lifetime risk for schizophrenia ranges from 0.3% to 3.7%, depending on the definition of schizophrenia and the method of survey used (11 ,12 ,18 ,19). The World Health Organization study shows a narrow range of lifetime risks in 10 countries around the world (0.5% to 1.7%) (13), although it is higher in some genetically isolated populations in Palau, Micronesia, and areas of Finland (20 ,21). Taken together, studies of the lifetime risk for schizophrenia in the general population suggest it is around 1%. In other words, a schizophrenic disorder will develop in approximately one in every hundred people at some time in their life.

Risk Factors

Schizophrenia occurs around the world and in all cultures. International differences in rates of the disorder are usually attributed to diagnostic differences rather than to differences in true rates of illness. The use of broad, ambiguous diagnostic criteria before the late 1970s was an important factor underlying artificial differences in rates of mental disorders recorded in different geographic locales. In the 1960s, for example, the hospital incidence of schizophrenia in the United States significantly exceeded that of Great Britain (28.0/100,000 vs. 17.9/100,000 population), whereas the incidence of mood disorders in British hospitals exceeded that in the United States (36/100,000 vs. 7/100,000). These differences disappeared, however, when identical methods of diagnosis and assessment were used (22). Similarly, in the World Health Organization study, differences in incidence among 10 countries diminished when narrow, standardized diagnostic criteria were utilized (13 ,23).

A variety of risk factors have received attention in schizophrenia. These include, among others, a family history of the disorder, low socioeconomic status, complications during pregnancy and childbirth, sex, and fetal viral infection (18 ,24 ,25 ,26 ,27 and 28). A family history of the disorder and a negative relationship to social class are especially strong and frequently replicated risk factors. The familial basis of schizophrenia is considered further in the section on the genetic epidemiology of schizophrenia. It should be emphasized that although the focus of this chapter is the genetics of schizophrenia, many of the risk factors for the illness cited above are environmental, a fact that underscores the combination of genetic and environmental factors underlying the disorder. This point is stressed further in the discussion of twin studies, below; given an identical twin with schizophrenia, the risk that the disorder will develop in the other twin (who has identical genes) is far less than 100%.

GENETIC EPIDEMIOLOGY OF SCHIZOPHRENIA

Part of "49 - Molecular and Population Genetics of Schizophrenia "

Family Studies

There is little question that schizophrenia (and related disorders) runs in families. In a review of 40 European studies, selected for similarities in diagnostic and ascertainment procedures and performed between 1920 and 1987, the lifetime risks for schizophrenia in relatives of schizophrenic patients were as follows: parents, 6.0%; siblings, 9.0%; offspring of one parent with schizophrenia, 13.0%; offspring of two parents with schizophrenia, 46%; and identical twin of a patient with schizophrenia, 46% (18 ,19). Note that the risk to offspring exceeds the risk to parents. Because the biological relationship is the same (i.e., first-degree relatives), the risk to offspring of patients should be identical to the risk to parents of patients. This would be true under any genetic model. The difference occurs because, by definition, parents have reproduced, and the presence of schizophrenia has an adverse affect on the probability of doing so. The risks to second-degree relatives ranged from 6.0% for half-siblings to 2.0% for uncles and aunts. First cousins, a type of third-degree relative, had an average risk of 2.0%. Consistent with a genetic etiology, these figures show that as the degree of biological/genetic relatedness to a schizophrenic patient increases, so does the risk for schizophrenia.

Although recent studies have used more rigorous research methods and narrower, criterion-based definitions of schizophrenia,

they have essentially confirmed the familial basis of schizophrenia and provided evidence consistent with the genetic hypothesis. In their family study of the Roscommon area in Western Ireland, for example, Kendler et al. (29) reported that the risk to siblings of schizophrenic probands is 9.2%, consistent with the rates reported by Gottesman (18 ,19). The risk to parents of schizophrenic probands was 1.3%, somewhat lower than the 6.0% reported by Gottesman. In general, narrower definitions in modern studies result in somewhat lower risk figures than those reported previously. This point was demonstrated by Tsuang et al. (30), who reported the risk for schizophrenia in first-degree relatives of schizophrenics as 3.2%, compared with 0.6% for relatives of nonpsychiatric controls. Similarly, Guze et al. (31) reported rates of 3.6% and 0.56%, respectively. Despite the lower prevalence figures, both these studies reconfirmed that in comparison with the general population, relatives of people with schizophrenia are at significantly greater risk for development of the disorder. This in itself, however, is not sufficient to demonstrate a genetic basis. Additional strategies, such as twin and adoption paradigms, are necessary to parse out genetic and environmental determinants, and these are considered next.

Twin Studies

The two types of twins are monozygotic and dizygotic. Monozygotic twins (i.e., identical twins) share 100% of their genes, whereas dizygotic twins (i.e., ordinary brothers and sisters) share 50% of their genes. If both members of a twin pair have schizophrenia, they are considered *concordant* for the disorder; if one is schizophrenic and the other is not, they are considered *discordant*. If schizophrenia were caused by genetic factors alone, then concordance rates for monozygotic and dizygotic twins would be 100% and 50%, respectively. On the other hand, if it were caused essentially by environmental variables, then concordance rates for monozygotic and dizygotic twins reared in a common environment would be similar.

The empiric data lie between these extremes but provide clear evidence for a genetic component. For example, concordance rates from twin studies pooled by Kendler (32) were about 53% for monozygotic twin pairs and 15% for dizygotic twin pairs. A similar review by Gottesman (18 ,19) demonstrated median concordance rates of 46% and 14% for monozygotic and dizygotic pairs, respectively. Interestingly, monozygotic twins reared apart have about the same concordance rates as do twins reared together (33). At the same time, monozygotic concordance rates lower than 100% demonstrate an important role for environmental factors. When they obtained quantitative estimates of the relative roles of genetic and environmental factors by translating genetic concordance rates into “heritabilities” (the proportion of the variance between individuals that is attributable to genetic factors), Kendler and Diehl (34) found that about 70% of the variance could be attributed to genetic factors in a series of twin studies. Several recent studies in which DSM-III, DSM-III-R, or DSM-IV diagnostic criteria were used provided even higher estimates of heritability, in the range of 80% to 86% (35 ,36 ,37 and 38).

Studies of twins discordant for schizophrenia have also shed light on the role of genetic and environmental factors. Gottesman and Bertelson (39), for example, followed the Danish schizophrenic twin sample of Fischer (40). They reasoned that if genetic liability is transmitted to the unaffected member of the twin pair, but not expressed because of environmental factors, then their offspring should demonstrate the same genetic liability. This hypothesis was supported when 24 children of unaffected co-twins showed a monozygotic concordance rate of 17%, a rate almost identical to that in the offspring of the twins who had schizophrenia. In contrast, although the risk for schizophrenia in the offspring of dizygotic twins with schizophrenia was similar to the risk in monozygotic twins, the risk in the offspring of unaffected dizygotic twins was only about 2%. What type of environmental factors might contribute to the risk for schizophrenia? A variety of possibilities exist, but adverse events occurring early in development (e.g., gestation, birth) have received the most attention recently, in part because the occurrence of such events during that period may have particularly far-reaching biological consequences. McNeil et al. (41) showed, for example, a relationship of smaller hippocampi and larger ventricles to complications of labor and delivery in the ill member of monozygotic twin pairs discordant for schizophrenia. Tsujita et al. (42) showed evidence of postzygotic genomic discordance in discordant twins, which could influence subsequent transcription in one or more genes. It is thus evident from twin studies with a variety of designs that both genetic *and* environmental factors underlie the expression of schizophrenia. Adoption studies, considered below, further support this conclusion.

Adoption Studies

Like twin studies, adoption studies can disentangle genetic and environmental causes of disease (43). An important series of studies was performed in Denmark, thanks in part to its system of national registers. Among these, Kety et al. (44) studied 5,500 children from the Greater Copenhagen area who were separated from their biological families and adopted between 1923 and 1947. Although schizophrenia or a related disorder developed in only 1.9% of the control adoptees, schizophrenia or a related illness was diagnosed in 8.7% of the adoptees who were separated from a schizophrenic parent. When the biological relatives of schizophrenic and control subjects were compared, 5.8% of the relatives of schizophrenic probands had (definite or probable) schizophrenia, in comparison with 0.9% of the control relatives. The rates of schizophrenia did not differ between the adoptive relatives of the schizophrenic and nonschizophrenic

adoptees. Moreover, children born to nonschizophrenic parents but raised by a schizophrenic parent did not show rates of schizophrenia above those predicted for the general population.

Limitations of adoption studies include the possibility of transmission by the mother during pregnancy or delivery of a liability for schizophrenia via nongenetic biological, or other type of environmental, factors, even if the child is adopted away soon after birth. However, Kety et al. (45) found that 13% percent of paternal half-siblings of schizophrenic adoptees had schizophrenia, in comparison with only 2% of paternal half-siblings of nonschizophrenic adoptees. Because paternal half-siblings have different mothers, these results cannot be attributed to *in utero* (environmental) effects. Many of Kety's findings were replicated recently for schizophrenia (45) and for certain schizophrenia spectrum conditions (46) in a sample drawn from the rest of Denmark.

Data from the Finnish adoption studies (47 ,48) provide additional support for both genetic and environmental influence in the transmission of schizophrenia. Together, family, twin, and adoption studies show consistently that the biological relatives of people with schizophrenia themselves show higher rates of schizophrenia and related disorders when compared with appropriate control groups, regardless of whether they are raised by biological or adoptive parents.

EPIDEMIOLOGY OF SPECTRUM DISORDERS

Part of "49 - Molecular and Population Genetics of Schizophrenia "

The term *related disorders* is used to describe schizophrenic illness of (generally) lesser severity. In fact, genetic studies provide evidence for a spectrum of disorders that are similar to schizophrenia and caused by the same genes. A disorder is considered to be in the schizophrenia spectrum if it occurs more frequently among the biological relatives of schizophrenic patients than it does among the relatives of people who do not have schizophrenia. Many of the behavioral genetic methodologies used to delineate genetic and environmental factors in the etiology of schizophrenia (e.g., family, twin, and adoption studies) have also provided evidence of a genetic etiology in schizophrenia spectrum conditions. Evidence for inclusion in the schizophrenic spectrum is considered next for several candidate disorders.

Psychotic Spectrum Disorders

In about 9% of the first-degree relatives of schizophrenic patients, a psychotic disorder develops that does not meet the criteria for either schizophrenia or a mood disorder (18 ,49). Two prominent examples are schizoaffective disorder and psychosis not otherwise specified (NOS). As the name suggests, the term *schizoaffective disorder* describes patients with features of both schizophrenia and affective disorders (also known as *mood disorders*), although subgroups may exist in which either schizophrenia or affective symptoms predominate. *Psychosis NOS* is a residual diagnostic category for patients with psychotic symptoms who do not fit into a more narrowly defined category. In many cases, the NOS designation serves as a temporary diagnosis for patients with new onset of disease until the course of their symptoms reveals their true diagnosis.

Both schizoaffective disorder and psychosis NOS are more common among the relatives of schizophrenic patients than among the relatives of nonschizophrenic persons. For example, in a survey of family history and family, twin and adoption studies, Prescott and Gottesman (33) found that 13 of 15 studies demonstrated evidence of a familial/genetic component for schizoaffective disorder. Consistent with this finding, monozygotic twins show higher concordance rates for schizoaffective disorder than do dizygotic twins (36 ,50).

Nonpsychotic Spectrum Disorders

Personality Disorders

Milder forms of schizophrenic illness are characterized by nonpsychotic symptoms, such as poor social relationships, anxiety in social situations, and limited emotional responses. Less frequently, mild forms of thought disorder, suspiciousness, magical thinking, illusions, and perceptual aberrations are also present. These symptoms are observed most frequently in three personality disorders, including schizotypal, schizoid, and paranoid personality disorders. Several studies found that (DSM) cluster A personality disorder traits often precede the onset of psychosis in subjects in whom schizophrenia subsequently develops (51 ,52). Moreover, in the New York high-risk project (53), offspring of schizophrenic mothers demonstrated elevated rates of these personality disorders when they were considered together, although not separately.

Most studies of familial prevalence in the biological relatives of schizophrenic patients have related schizotypal personality disorder to the schizophrenia spectrum more strongly than they have either schizoid or paranoid personality disorder (54 ,55). Evidence in favor of including schizotypal personality disorder is consistent across family (56 ,57 and 58), adoption (45), and twin studies (59 ,60). Although not all studies have detected a higher rate of schizotypal personality disorder among relatives of schizophrenic probands (54), most investigations, and particularly those with large samples, show higher rates of schizotypal personality disorder among the relatives of index cases with schizophrenia than among the relatives of control subjects (61). The incidence of the disorder in schizophrenic families has been estimated at between 4.2% and 14.6% (57 ,58 ,62 ,63). In contrast, results for schizoid and paranoid personality disorders have been somewhat more controversial and contradictory, with positive findings sometimes occurring in combined paranoid–schizotypal or schizoid–schizotypal

samples (55). Thus, although some symptoms may overlap between schizotypal, schizoid, and paranoid personality disorders, schizotypal personality disorder is currently the strongest candidate among this group for a nonpsychotic, relatively mild condition that is related genetically to schizophrenia.

Schizotaxia

Paul Meehl (64) first used the term *schizotaxia* in 1962 to describe a genetic vulnerability to schizophrenia. He suggested that a subtle but widespread neurointegrative defect results from this vulnerability that predisposes individuals to the development of either schizotypy or schizophrenia, depending on the protection or liability afforded by environmental circumstances. Later, Meehl reformulated the concept to allow for the possibility that some people with schizotaxia would not progress to either schizophrenia or schizotypal personality disorder, although most would (65). Eventually, the term *schizotypy* entered the psychiatric nomenclature in the form of schizotypal personality disorder. *Schizotaxia* did not, although the term was used in a general sense by researchers to describe the liability for schizophrenia. Now, almost 40 years after the concept was introduced, a broad literature shows that the liability for schizophrenia can be characterized clinically by deficits or abnormalities in psychiatric, neuropsychological, neurobiological, and psychosocial domains in nonpsychotic, first-degree relatives of people with schizophrenia.

Psychiatric features in such relatives frequently include negative symptoms (e.g., asociality and anhedonia) that are qualitatively similar to, but quantitatively milder than, those often seen in schizophrenia (66). Positive symptoms, however, are usually less evident in these relatives than they are in schizophrenia or schizotypal personality disorder. Neuropsychological impairments in biological relatives of people with schizophrenia are also similar to those in patients with schizophrenia, but are generally of lesser severity (67 ,68 ,69 ,70 and 71). The most extensively documented of these impairments involve working memory/attention, verbal memory, and concept formation/abstraction.

We recently suggested a reformulation of Meehl's concept of schizotaxia that focuses on these features of negative symptoms and neuropsychological deficits (67). In addition to specifying the clinical consequences of schizotaxia more specifically, our view of the concept differs from Meehl's in a few other respects. Among these is whether schizotaxia always or even usually progresses to schizotypal personality disorder or schizophrenia. Our empiric analyses suggest that the basic symptoms of schizotaxia occur in 20% to 50% of adult relatives of patients with schizophrenia (68 ,69). This rate is considerably higher than the rates of schizophrenia or schizotypal personality disorder likely to develop in first-degree relatives (≤ 10% for each condition), which suggests that schizotaxia does not lead inevitably to schizophrenia or schizotypal personality disorder. Our view of schizotaxia also differs somewhat from Meehl's formulation in that we include both genetic and nongenetic, adverse biological alterations occurring early in development (e.g., pregnancy and birth complications) in our conception of the syndrome.

Tsuang et al. (72) recently described a set of specific criteria for schizotaxia, but the clinical characterization is still evolving and the syndrome requires independent validation and further research before it can be used clinically (73). These criteria underlie a conception of schizotaxia as a neurodevelopmental condition resulting from genetic and adverse environmental (e.g., pregnancy or delivery complications) factors. Current criteria in adult, nonpsychotic, first-degree relatives of schizophrenic patients include moderate or greater levels of negative symptoms and neuropsychological deficits (as described above). The concept of schizotaxia demonstrates considerable utility because it accounts for clinical deficits in a sizable proportion of relatives who may not otherwise meet the criteria for a schizophrenia-related disorder. Moreover, because schizotaxia may be considered as a risk factor for schizophrenia, as well as a clinically meaningful syndrome in its own right, its recognition may eventually facilitate the development of early intervention and prevention strategies.

MODE OF INHERITANCE

Part of "49 - Molecular and Population Genetics of Schizophrenia "

Given the evidence outlined above for a substantial genetic contribution to the etiology of schizophrenia, methods such as complex segregation analysis (74) can be used to identify the most likely mode of inheritance. Commonly, a mixed model (75) comprising both major gene and polygenic effects is compared with the submodels of a single major locus and polygenic inheritance. However, large sample sizes are required to distinguish between models, especially the polygenic and mixed models, so that the practical usefulness of this approach has been limited to date.

Based on complex segregation analysis and related approaches, the pattern of risks in family and twin studies has been found to be incompatible with a single locus accounting for all the genetic liability to schizophrenia (76 ,77 and 78). However, it has not been possible to distinguish between a polygenic and a mixed model (77). The pattern of risks in family studies, in which the risk decreases rapidly as the degree of genetic relatedness decreases, is also compatible with a model of multiple loci with epistasis (interaction between genes) (79). However, the number of susceptibility loci, the disease risk conferred by each locus, and the degree of interaction between loci all remain unknown. The contribution of individual genes to the familiality of a disorder can be expressed in terms of *(s* (i.e., the relative risk to siblings resulting from possession of the disease allele (79). Risch (79) has calculated that the data for recurrence risks

in the relatives of probands with schizophrenia are incompatible with the existence of a single locus having a value of *(s* greater than 3. Unless extreme epistasis (interaction between loci) exists, models with two or three loci having values of *(s* of 2 or less are more plausible. It should be emphasized that these calculations are based on the assumption that the effects of genes are distributed equally across the whole population. It is quite possible that genes of larger effect are operating in a subset of patients—for example, those from families with a high density of illness.

These quantitative genetic investigations provide strong evidence that genetic factors increase the risk for schizophrenia. However, although it is possible to state that, as a group, siblings of individuals with schizophrenia, for example, have a roughly 10-fold increased risk in comparison with the general population, it is not currently possible to translate this figure to the level of risk for a particular sibling in a particular family. Similarly, heritability estimates refer to variations in liability to schizophrenia in the general population and have no simple meaning for an individual. Another important point is that risk to related individuals does not directly equate with genetic risk because some relatives carry one or more susceptibility alleles for schizophrenia but remain unaffected throughout their lives. In other words, the accumulation of susceptibility alleles, environmental risk factors, and complex interactions between risk factors probably all play a role in determining who becomes ill. Therefore, to quantify genetic risk, it is necessary to identify the susceptibility loci themselves at the molecular level.

MOLECULAR GENETICS: LINKAGE STUDIES

Part of "49 - Molecular and Population Genetics of Schizophrenia "

The first generation of systematic molecular genetic studies of schizophrenia effectively ignored the evidence for genetic complexity and targeted large, multiply affected pedigrees for analysis. This was done in the hope that such families, or at least a proportion of them, were segregating genes of sufficiently large effect that they could be detected unequivocally in this way. This approach has been successful in other complex disorders—Alzheimer disease, for example, in which mutations in three genes, *APP, PS1*, and *PS2*, are now known to cause rare forms of the disorder. In such cases, the disease is of unusually early onset and is transmitted through multiplex pedigrees in an autosomal dominant fashion (80 ,81 and 82).

Unfortunately, it has not proved possible to identify a phenotypic trait analogous to age at onset in Alzheimer disease by which to classify multiplex families segregating schizophrenia. Studies of such large families also initially produced positive findings in schizophrenia (83), but unfortunately these could not be replicated. The reasons for this have become clear as data from systematic genome scans have accumulated; highly penetrant mutations causing schizophrenia are at best extremely rare and quite possibly nonexistent (84 ,85). The false-positives were largely the consequence of a combination of multiple testing and the use of statistical methodology and significance levels derived from work on single-gene disorders.

Despite the failure to identify regions of unambiguous linkage in multiply affected families, modest evidence for several regions has been reported in more than one data set. Areas implicated for which supportive data have also been obtained from international collaborative studies include chromosomes 6p24-22, 8p22-21, and 22q11-12 (86 ,87 ; see refs. 84 and 88 for review). A number of other promising areas of putative linkage are also currently under investigation by international consortia. These include 13q14.1-32 (89 ,90 and 91), 5q21-31 (92 ,93), 18p22-21 (94), 10p15-11 (95 ,96 and 97), 6q (98 ,99), and 1q21-22 (100). However, in each case, both negative and positive findings have been obtained, and in only two cases, those of chromosomes 13q14.1-32 and1q21-22, did any single study achieve genome-wide significance at values of *p* of .05 (90 ,100).

These positive findings contrast with those from a large systematic search for linkage in which a sample of 196 affected sibling pairs, drawn typically from small nuclear families rather than extended pedigrees, was used (101). The results of simulation studies suggest that the power of this study is greater than 0.95 to detect a susceptibility locus of *(s* = 3 with a genome-wide significance of 0.05, but only 0.70 to detect a locus of *(s* = 2 with the conservative assumption that a locus lies midway between two adjacent markers. This study yielded evidence at the level of the definition of Lander and Kruglyak (102) of “suggestive” linkage to chromosomes 4p, 18p, and Xcen. However, none of the findings approached a genome-wide significance of 0.05, corresponding to Lander and Kruglyak’s definition of “significant” linkage.

The findings from linkage studies of schizophrenia to date demonstrates several features that are to be expected in the search for genes for complex traits (103 ,104 ,105 and 106). First, no finding is replicated in all data sets. Second, levels of statistical significance are unconvincing and estimated effect sizes are usually modest. Third, chromosomal regions of interest are typically broad [often > 20 to 30 centimorgan (cM)].

At the present time, therefore, the linkage literature supports the predictions made by Risch (79); it is highly unlikely that a commonly occurring locus of effect size [*(s*] greater than 3 exists, but suggestive evidence implicates a number of regions, consistent with the existence of some susceptibility alleles of moderate effect [*(s* = 1.5 to 3]. Moreover, encouraging results in several chromosomal regions suggest that rarer alleles of larger effect may be segregating in some large, multiply affected families.

Linkage methods in sample sizes that are realistically achievable can detect smaller genetic effects than those in the studies to date. For example, it is possible to detect alleles with values of *(s* of 1.5 to 3 in a sample of 600 to

800 affected sibling pairs (107 ,108). We therefore suggest that priority should now be given to collecting such samples with a robust clinical methodology that is comparable across all interested research groups. However, if liability to schizophrenia is entirely a consequence of the operation of many genes of small effect, then even these large-scale studies will be unsuccessful.

CANDIDATE GENE ASSOCIATION STUDIES

Part of "49 - Molecular and Population Genetics of Schizophrenia "

Once genes of smaller effect than *(*s =1.5 are sought, the number of affected family members required becomes prohibitively large (107 ,108 and 109). For this reason, many researchers have tried to take advantage of the potential of candidate gene association studies to identify such loci (109 ,110). Although a potentially powerful means of identifying genes of small effect, association studies are not without their problems. First, for a complex and poorly understood disorder such as schizophrenia, the choice of candidate genes is limited largely by the imagination and resources of the researcher. This places a stringent burden of statistical proof on positive results because of low prior probability and multiple testing (111). Second, case–control association studies have the potential to generate false-positives because of population stratification. This problem can be addressed by using family-based association methods (112), but because of stigma, adult age at onset, and the disruptive effects of mental illness on family relationships, family-based samples may be unrepresentative in addition to limited in size. Consequently, family-based studies may introduce more spurious results than do case–control studies (113). It would seem unwise, therefore, to discard the case–control study design, which has served epidemiology so well through the years. A third problem common to all molecular genetic studies in complex diseases is that they are prone to type 2 errors simply because they are often underpowered, and therefore to draw satisfactory conclusions from negative studies, larger sample sizes are required than have typically been used to date in psychiatric genetics (111). Fourth, even with larger samples, it is by no means certain that a given replication study will be sufficiently powered to replicate a particular effect. This is because variations may be noted in the contribution of a given susceptibility allele in different patient populations as a result of different allele frequencies at the locus of interest or at interacting loci. Further potential for heterogeneity occurs if the association with the marker is a result of tight linkage with the true susceptibility allele, or if different subtypes of the disease exist. Given that all the above factors may influence power, and that none of the above is known in advance, it is difficult to obtain an accurate measure of the power of a replication study. Because we cannot specify accurately the prior probability of a candidate gene, nor know the true power for replication, it is difficult to draw definitive conclusions from conflicting findings. However, the purpose of experiment is to reject a null hypothesis, and in the face of uncertainty, the burden of proof remains with the proponents of a particular candidate gene.

Most candidate gene studies have been based on neuropharmacologic studies, which suggests that abnormalities in monoamine neurotransmission, in particular dopaminergic and serotoninergic systems, play a role in the etiology of schizophrenia. Overall, the results in this extensive literature are disappointing, but it should be noted that the sample sizes in many of the older studies would now generally be regarded as inadequate, particularly in view of the fact that the polymorphic markers in question did not in themselves represent functional variants and that few genes have been systematically screened even for common functional variants. However, more promising reports of candidate gene associations have recently appeared, three of which are considered here.

Serotonin 5-HT$_{2A}$-Receptor Gene

Many novel antipsychotic drugs affect the serotoninergic system. The first genetic evidence that serotoninergic receptors may play a role in schizophrenia came from a Japanese group reporting an association between a T-to-C polymorphism at nucleotide 102 in the 5-HT$_{2A}$-receptor gene in a small sample (114). A large European consortium comprising seven centers and involving 571 patients and 639 controls then replicated this finding (115), which was further replicated with use of a family-based design (116). Although many other studies followed with mixed results, a recent metaanalysis of all available data from more than 3,000 subjects supports the original finding (p = .0009), and this does not appear to be a consequence of publication bias (117).

Since this metaanalysis was undertaken, a few further negative reports have followed, but none has approached the sample sizes required. If we assume homogeneity and if the association is true, the putative odds ratio (OR) for the C allele can be expected to be around 1.2 in any replication sample. Sample sizes of 1,000 subjects are then be required for 80% power to detect an effect of this size, even at a relaxed criterion of p = .05. Thus, the negative studies are effectively meaningless, but it is also true that the evidence for association, even in the metaanalysis (p = .0009), is not definitive if genome-wide significance levels are required (109). At present, all we can conclude is that the evidence favors association between the T102C 5-HT$_{2A}$ polymorphism and schizophrenia, but the most stringent burden of proof has not yet been met.

If the association is real, it is unlikely that the T102C polymorphism is the susceptibility variant because this nucleotide change does not alter the predicted amino acid sequence of the receptor protein, nor is it in a region of obvious significance for regulating gene expression. T102C is

in complete linkage disequilibrium with a polymorphism in the promoter region of this gene, but no evidence has as yet been found that this has a functional effect either (116). Recent evidence of polymorphic monoallelic expression of the 5-HT_{2A} gene points to the possible existence of sequence variation elsewhere that influences gene expression (118), and this may be the true susceptibility variant.

D3 Dopamine-Receptor Gene

Association has been reported between schizophrenia and homozygosity for a Ser9Gly polymorphism in exon 1 of the D3 dopamine-receptor gene (*DRD3*) (119). As with the 5-HT_{2A} association, the results have now been confirmed in several independent samples, including one family-based study (120), but several negative studies have also been reported. Metaanalysis of data from more than 5,000 individuals has revealed a small (OR = 1.23) but significant (p = .0002) association between homozygosity at Ser9Gly and schizophrenia (120). Again, this cannot easily be ascribed to selective publication (120). Because it is a fairly uncommon genotype, we cannot be certain that homozygosity for the 9Gly allele alone is not associated with an increased risk for schizophrenia, and therefore it is not certain that the findings at D3 are an example of heterosis. However, a plausible biological explanation for D3 heterosis has been put forward in that possession of two different molecular forms of the receptor may allow the dopaminergic neuron to respond more flexibly (119).

At present, then, the status of the D3 dopamine-receptor gene is similar to that of the 5-HT_{2A}-receptor gene—that is, the balance of evidence at present favors association, but the null hypothesis still cannot be confidently rejected. Those wishing to replicate or reject these findings should bear in mind that to obtain power greater than 0.80 to detect an effect of this size at a criterion of p = .05, a sample of 1,500 cases and 1,500 controls is required. So far, no other polymorphisms have been found that might explain the putative D3 association, but several new polymorphisms have been identified in previously unknown exons 5′ to the exon referred to above as *exon 1* (121). These are currently being tested to establish whether variants in this region in linkage disequilibrium with the Ser9Gly polymorphism provide a more functionally plausible explanation of the association with schizophrenia.

ANTICIPATION AND TRINUCLEOTIDE REPEATS

Part of "49 - Molecular and Population Genetics of Schizophrenia "

The term *anticipation*, the phenomenon by which the age at onset of disease becomes earlier from one generation to the next, was first described in connection with severe mental disorder (122). A series of recent studies applying modern diagnostic criteria have now confirmed that the inheritance of schizophrenia is at least consistent with the presence of anticipation, although ascertainment biases offer an alternative explanation (123). Because pathogenic expanded trinucleotide repeats are the only known genetic mechanisms for anticipation, these findings have been taken as suggesting that such mutations may account for at least some of the complexity in the pattern of inheritance in this disorder (124). This hypothesis was supported by two groups who observed that the maximum length of the most common known pathogenic trinucleotide repeat, CAG/CTG, was greater in patients with schizophrenia than in unaffected controls (125 ,126). These findings were later replicated in a European multicenter study (127). Unfortunately, the early repeat expansion detection (RED) studies were followed by a series of unsuccessful attempts to identify the relevant repeat-containing loci by a variety of methods, and by several failures to replicate the RED findings (128 ,129 and 130), thus casting doubt over the CAG/CTG repeat hypothesis.

The trinucleotide repeat hypothesis was rejuvenated, however, with the report of an association between schizophrenia and alleles of a member of the family of calcium-activated potassium channel genes, *KCa3* (*hKCa3*/*KCCN3*) (131). For several reasons, *KCa3* seemed a remarkable candidate gene for schizophrenia. First, the gene contained two CAG repeats, one of which is highly polymorphic. Second, the family of genes to which it belongs is thought to play an important role in regulating neuronal activity, and it was therefore considered a functional candidate gene. Third, the gene was also thought to be a positional candidate, as it was believed to map to chromosome 22q11. Ironically, *KCa3* maps not to 22q11 but to 1q21 (132), which is also a region implicated by linkage studies as possibly containing a susceptibility locus for schizophrenia (100). Two further case–control studies have subsequently supported the findings of Chandy and colleagues (132 ,133).

However, although evidence from three case–control studies lends support to the hypothesis of *KCa3* as a susceptibility gene for schizophrenia, in other respects the case for this gene as a candidate is less certain. First, the RED data cited above lend no support to *KCa3* as a candidate because the polymorphic trinucleotide repeat in this gene is far too short to account for the RED associations (133). Second, a series of case–control and family studies have failed to replicate the findings (134 ,135 ,136 ,137 ,138 ,139 and 140). Thus, we are back to our familiar position in candidate gene analysis; although the data are insufficient to draw firm conclusions, we believe at present that the case for *KCa3* remains firmly with the null hypothesis.

What then is the status of the original associations between large CAG/CTG repeats and schizophrenia? It has been reported that large CAG/CTG RED products (repeat size > 40) are explained by repeat size at two autosomal loci, one at 18q21.1 and the other at 17q21.3 (141 ,142). If the explanation is correct, then it follows that one or both of these loci should be associated with schizophrenia.

Unfortunately, data from Vincent and colleagues (128) and unpublished data from Cardiff unequivocally show that expansions at these loci are not responsible for the RED associations. However, in both samples, only around 50% of large CAG/CTG repeats detected by RED could be explained by polymorphisms at these two loci, which suggests that at least one further locus is responsible for the RED data, a possibility supported by two recent studies on protein extracts from schizophrenic tissues (143).

HIGH RATES OF SCHIZOPHRENIA IN ADULTS WITH VELOCARDIOFACIAL SYNDROME

Part of "49 - Molecular and Population Genetics of Schizophrenia "

Velocardiofacial syndrome (VCFS), also known as *DiGeorge* or *Shprintzen syndrome*, is associated with small interstitial deletions of chromosome 22q11 in 80% to 85% of cases (144). First described by Shprintzen and colleagues (145), VCFS has an estimated prevalence of 1/4,000 births (146). Distinctive dysmorphology, congenital heart disease, and learning disabilities characterize the syndrome, although considerable phenotypic variability occurs. As the first recognized cohort of children with VCFS was followed into adolescence and early adulthood, evidence began to accumulate for a high prevalence of major mental illness. Early reports suggested that psychiatric disorders had developed in more than 10% of the cohort, which mostly resembled chronic schizophrenia with paranoid delusions, although operational criteria were not used (147). In a follow-up study of teenagers (age 17 years) in which DSM-III-R criteria were used, Pulver et al. (148) reported that 11 (79%) of their sample of 14 patients had been given a psychiatric diagnosis: 29% had schizophrenia (22%) or schizoaffective disorder (7%), 29% had simple or social phobia, 21% had depression, and 14% had obsessive-compulsive disorder. More recently, Papolos et al. (146) reported that of their sample of 15 children and 10 adults, four (16%) had psychotic symptoms and 16 (64%) met DSM-III-R criteria for a spectrum of bipolar affective disorders. Although none had schizophrenia, the two oldest members of their patient cohort (ages 29 and 34 years) both had schizoaffective disorder.

To try to gain a more precise determination of the prevalence and nature of psychopathology in adults with VCFS, rather than rely on clinical diagnosis, Murphy and colleagues (149) recently evaluated 50 cases with a structured clinical interview to establish a DSM-IV diagnosis. Fifteen patients with VCFS (30%) had a psychotic disorder, with 24% (n = 12) fulfilling DSM-IV criteria for schizophrenia. In addition, six (12%) had major depression without psychotic features. They were unable to replicate the findings of Papolos and colleagues (146) of a high prevalence of bipolar spectrum disorders in VCFS. However, these workers studied a small sample that included few adults, and in view of the fact that their oldest cases satisfied criteria for schizoaffective disorder, it is possible that the psychotic phenotype in VCFS varies with age. Prospective studies are now required to test this hypothesis.

The current balance of evidence favors the view that the high prevalence of psychosis results from hemizygosity for a gene or genes at chromosome 22q11 rather than ascertainment bias or a nonspecific association with a low intelligence quotient (IQ) (149). In particular, the prevalence of psychosis and schizotypy in VCFS appears to be much greater than that seen in most other congenital abnormalities affecting neural development, and appears not to be correlated with the degree of intellectual impairment (149). Many lines of evidence suggest that schizophrenia is a neurodevelopmental disorder (150). In VCFS, defective development and migration of mesencephalic and cardiac neural crest cells are believed to play a significant role in the pathogenesis of midfacial and cardiac abnormalities (151). Consequently, it has been postulated that a gene or genes causing disruption of neural cell migration may be a common neurodevelopmental mechanism for both VCFS and schizophrenia (152).

What then is the importance of 22q11 and VCFS in the etiology of schizophrenia as a whole? Karayiorgou et al. (153) reported that among 100 randomly ascertained patients with schizophrenia, two were found to have a 22q11 deletion. In contrast, when subjects with schizophrenia were selected for the presence of clinical features consistent with VCFS, 22q11 deletions were identified in 20% to 59% of cases (154 ,155). These findings suggest that a small proportion of cases of schizophrenia may result from deletions of 22q11, and that clinicians should be vigilant, especially when psychosis occurs in the presence of dysmorphology, mild learning disability, or a history of cleft palate or congenital heart disease (156). The question of whether genetic variation in 22q11 confers susceptibility to schizophrenia in cases without a deletion is more difficult to answer. As we have seen, the results of some linkage studies suggest the presence of a schizophrenia susceptibility locus on 22q. However, the linkage findings tend to point telomeric to the VCFS region (86 ,157). Nevertheless, modest evidence for linkage to the VCFS region has also been claimed (90 ,158 ,159), and as we have noted above, linkage mapping in complex diseases is somewhat imprecise. It remains possible that the relationship between VCFS and “typical” schizophrenia is less direct, with little common ground between the genetic and neurodevelopmental mechanisms involved but with convergence on identical or at least similar psychopathologic syndromes.

Another important question concerns the factors that determine whether schizophrenia will develop in a person with VCFS. When adults with VCFS were tested with a quantitative measure of schizotypy, the patients with psychosis had the highest scores (149). However, perhaps of greater interest, those without psychosis had intermediate

scores in comparison with controls (149). If schizotypy is a trait marker for increased liability to psychosis, this suggests that the majority if not all of those with 22q11 deletions are at increased risk for psychosis, but that other genetic or environmental factors are required for this risk to be expressed. The genetic loci involved may reside elsewhere in the genome and include those involved more widely in psychosis, or lie within 22q11. The occurrence of psychosis does not appear to be related to the size of the deletion (153). However, it is possible that susceptibility to psychosis reflects allelic variation of a hemizygous gene or genes within the deletion. The gene encoding catechol-*O*-methyltransferase (COMT), an enzyme involved in the catabolism of catecholamine neurotransmitters, maps to the VCFS region and is therefore an obvious candidate for influencing the expression of psychosis in VCFS probands. This gene exists in two allelic forms encoding high- and low-activity isoforms of the enzyme, and it has been suggested that possession of the allele for low-activity COMT may be associated with the occurrence of schizophrenia in VCFS (160). However, Murphy and colleagues (149) found no evidence for an association between the allele for low-activity COMT and either schizophrenia or schizotypy in patients with VCFS.

FUTURE DIRECTIONS

Part of "49 - Molecular and Population Genetics of Schizophrenia "

Refining the Phenotype for Molecular Genetic Studies

The effectiveness of molecular genetic studies depends on the genetic validity of the phenotypes studied. Perhaps if we were better at defining phenotypes, we would be better at finding genes. It is worth reiterating at this point that the commonly used diagnostic criteria define phenotypes with high heritability. In principle, therefore, it should be possible to identify the genes predisposing to schizophrenia, as defined by current diagnostic criteria, if sufficiently large samples are studied. However, perhaps genetic validity could be improved by focusing on aspects of clinical variation, such as age at onset or symptom profiles, or by identifying biological markers that predict degree of genetic risk or define more homogeneous subgroups. Unfortunately, despite much work, it has not been possible to identify genetically distinct subtypes of schizophrenia. Instead, clinical variation is likely to reflect at least in part a combination of quantitative variation in genetic risk for the disorder and the effect of modifying genes that influence illness expression rather than the risk for illness *per se*. Examples of this phenomenon are probably age at onset and symptom pattern in schizophrenia (161 ,162).

The search for trait markers aims to move genetic studies beyond the clinical syndrome by identifying indices of genetic risk that can be measured in asymptomatic persons or by identifying markers of pathophysiologic processes that are closer to the primary effects of susceptibility genes than are clinical symptoms—so-called intermediate phenotypes or endophenotypes. Work in this area is developing fast; for example, candidate trait markers for schizophrenia include schizotypal personality traits, measures of cognitive processing, brain evoked potentials, and abnormalities in eye movements (163). It is also hoped that advances in brain imaging will lead to the identification of genetically valid trait markers. However, it seems unlikely that these phenotypes will provide a rapid solution to the problem. First, we will need to ensure that the measures used are stable and determine the extent to which they are affected by state. Second, to be of use in gene mapping, such measures will have to be practicably applied to a sufficient number of families or unrelated patients. Third, we will need to ensure that the traits identified are highly heritable, which will itself require a return to classic genetic epidemiology and model fitting. Finally, we cannot assume that the genetic architecture of such intermediate phenotypes will be simple.

Efforts to improve the selection of phenotypes are also concerned with enhancing the traditional categoric approach to defining psychiatric disorders by identifying genetically valid phenotypes that can be measured quantitatively. These can be used in quantitative trait locus approaches to gene mapping. Work in this area has begun but still faces problems, particularly those relating to the confounding influence of state-related effects. Perhaps the best hope of taking account of the complexity and heterogeneity of the schizophrenia phenotype comes from new methods of analysis in which aspects of the phenotype can be entered as covariates in linkage analyses (164).

Genome-wide Association Studies

In recent years, interest has increased in the possibility of systematic, genome-wide association studies (109 ,165). These have the potential of allowing systematic searches for genes of small effect in polygenic disorders. Optimism has been fueled by the fact that the most abundant form of genetic variation, the single-nucleotide polymorphism, is usually bi-allelic and potentially amenable to binary, high-throughput genotyping assays such as microarrays (so-called DNA chips). Moreover, as sequence data accumulate, it has become possible to contemplate the construction and application of very dense maps of hundreds of thousands of single-nucleotide polymorphisms (165).

Essentially two types of genome-wide association study have been proposed: direct and indirect. In the former, association is sought between a disease and a comprehensive catalogue of every variant that can alter the structure, function, or expression of every single gene. In contrast, indirect studies seek associations between markers and disease that are caused by linkage disequilibrium between the markers and susceptibility variants. The hope is that if sufficiently dense marker maps can be applied, it will be possible to screen the whole genome systematically for evidence of linkage

disequilibrium without actually having to screen every functional single-nucleotide polymorphism in the genome. However, a number of uncertainties and difficulties remain. These include, in particular, the difficulty of identifying functional single-nucleotide polymorphisms in regulatory rather than coding regions of the genome, uncertainty about the distances over which linkage disequilibrium is maintained, and the lack at the present time of a rapid, accurate, and cheap method for single-nucleotide genotyping (165).

Given these considerations, it seems clear that the era of genome-wide association studies, direct or indirect, is not yet at hand. Instead, studies in the next few years should probably focus mainly on the direct approach utilizing single-nucleotide polymorphisms from the coding sequence that actually alter protein structure in a wide range of functional and positional candidate genes. Preferably, complete functional systems should be dissected by the application of sensitive methods for mutation detection, followed by association studies in appropriately sized samples. We should also use our knowledge of functional pathways to make predictions about likely epistasis. However, given our ignorance of pathophysiology, the expectation should be that most reported associations will be false and resolved only by replication in large, well-characterized samples. At present, although the indirect approach is not widely applicable at a genome-wide level, smaller-scale studies focusing on specific regions indicated by the results of linkage studies may allow us to map loci. Additionally, such studies will generate the sort of data concerning patterns of linkage disequilibrium in typical “association samples” that will be required to determine whether genome-wide studies are likely to be feasible and what density of map will be required.

If the results are encouraging—that is, if linkage disequilibrium exists across useful distances in the genome—then until genotyping technology has sufficient capacity to permit mass genotyping at low cost, linkage disequilibrium analyses at the genome-wide level are most likely to be based on DNA pooling technologies (166 ,167 ,168 and 169).

Identifying Modifying Genes

It is sometimes assumed that variation in the clinical features of a disease or treatment response simply reflects pleiotropic effects of etiologic risk factors. However, it is becoming increasingly apparent that specific genes probably exist that influence clinical features and treatment response independent of those affecting liability. Thus, another potentially fruitful line of inquiry may be to design studies aimed at seeking modifying rather than causative genes, as these genes may in themselves allow novel drug targets to be identified. Evidence has already been found that variations in age at onset and symptom pattern in schizophrenia are probably caused at least in part by modifying genes (161 ,162). Furthermore, interest is increasing in pharmacogenetics in psychiatry (170), although with some trait markers, we need to exercise caution in the absence of genetic epidemiologic evidence that variations in drug response are under genetic control. We should also not forget that there is no *a priori* reason why responses to behavioral and psychological treatments should be less influenced by genetic factors than by pharmacologic treatments.

Animal Models

Another important challenge will be the development of suitable animal models to allow functional studies of putative disease loci (165). Disorders that predominantly involve higher cognitive function, such as schizophrenia, are likely to prove difficult to model in animals. However, certain features of the human phenotype, such as subtle abnormalities of cell migration, enlarged cerebral ventricles, and abnormalities of information processing, including defects in prepulse inhibition, can be detected in animals (171). In fact, a possible approach to producing a mouse model for at least some of the schizophrenia phenotype is suggested by the finding of an increased prevalence of schizophrenia in VCFS. Currently, attempts are under way to produce transgenic mice in which the syntenic region of mouse chromosome 16 is deleted (172). These animals should be investigated closely for neuroanatomic and behavioral phenotypes of possible relevance to schizophrenia, and such studies are already yielding encouraging findings (173).

Functional Studies

The most important and most obvious implication of identifying genetic risk factors for schizophrenia is that it will inspire a new wave of neurobiological studies from which, it is hoped, new and more effective therapies will emerge. However, although the unequivocal identification of associated genetic variants will represent a great advance, many years of work will be required before this is likely to translate to routine clinical practice. An early problem will be to determine exactly which genetic variant among several in linkage disequilibrium within a given gene is actually responsible for the functional variation. Even when a specific variant within a gene can be identified as the one of functional importance, functional analysis, in terms of effect at the level of the organism, is likely to be particularly difficult for behavioral phenotypes in the absence of animal models. An extra level of complexity is that we will need to be able to produce model systems, both *in vivo* and *in vitro*, that allow gene–gene and gene–environment interaction to be studied.

Genetic Nosology

Although the development of new therapies will take time, it is likely that the identification of susceptibility genes will

have an earlier effect on psychiatric nosology. If genetic risk factors are correlated with clinical symptoms and syndromes, it should be possible to study heterogeneity and comorbidity to improve the diagnosis and classification of psychosis. The prospects will also be enhanced for identifying clinically useful biological markers as an aid to diagnosis, so that we can move beyond the current situation of making diagnoses based entirely on clinical signs and symptoms. Improvements in diagnostic validity will clearly facilitate all avenues of research into these disorders. However, in the present context, we should point out that improvements in diagnosis and classification should enhance our ability to detect further genetic and environmental risk factors, so that a positive feedback between nosology, epidemiology, and molecular genetics can be envisaged.

Molecular Epidemiology

The identification of genetic risk factors can be expected to provide a new impetus to epidemiologic studies of schizophrenia by allowing researchers to investigate the ways in which genes and environment interact. Studies of this kind will require large, epidemiologically based samples together with the collection of relevant environmental data. This work could start now with DNA being banked for future use, although in schizophrenia, the identification of plausible environmental measures might require clues from the nature of the genetic risk factors yet to be identified. A major theme in relation to this work will be the bringing together of methodologies from genetics and epidemiology, which have traditionally adopted somewhat differing analytic approaches (174). Treating susceptibility alleles as risk factors in an epidemiologic context will allow estimates of effect sizes within a population to be made. Accounting for specific genetic effects will also facilitate the search for independent environmental factors and the investigation of potential gene–environment interactions. Scientific validity is likely to be enhanced by ensuring as far as possible that control samples are drawn from the same base population as patients. In addition, the use of incident cases should guard against the risk of identifying loci related to confounds, such as chronicity of illness, rather than susceptibility. Phenotypic assessment is likely to benefit from a prospective element to studies, which counteracts the tendency of patients to forget historical details and the difficulty of making observed ratings retrospectively from case records. However, the price of improved scientific rigor is likely to be considerably more expensive studies because of the longer period and larger number of investigators that will be required to ascertain detailed data from thousands of subjects.

Genetic Testing

A further implication concerns genetic testing. This is a complex area that raises a number of ethical issues, which have been discussed elsewhere (175). However, the potential for predictive testing has probably been overstated, given that susceptibility to schizophrenia almost certainly depends on the combined effects of predisposing and protective alleles at a number of loci and their interaction with the environment. Consequently, until we have a comprehensive molecular understanding of the etiology of schizophrenia, the predictive value of genetic testing is likely to be low. This applies, for example, to apolipoprotein E testing for late-onset Alzheimer disease, and has led to a recommendation that such testing not be performed in asymptomatic persons (175). Indeed, even when all the susceptibility genes for schizophrenia have been identified, it will still not be possible to predict the development of disease with certainty until the relevant genetic and environmental risk and modifying factors have also been identified and the nature of the various interactions understood. Such interactions may be complex and unpredictable (165). However, other possible roles of genetic testing are likely to be of greater value to patients and clinicians. For example, it might be possible to optimize treatment choices by testing genes found to influence treatment responses in psychiatric disorders and so provide more individualized treatment.

CONCLUSIONS

Part of "49 - Molecular and Population Genetics of Schizophrenia "

Attempts to identify the genes that predispose to schizophrenia face formidable challenges arising from both genetic and phenotypic complexity. Research to date has largely excluded the possibility that genes of major effect exist even in a subset of families. Evidence has been obtained of the location of some genes of moderate effect, but none of these findings can be regarded as conclusive, and proof in each case will probably have to await identification of the susceptibility locus itself. The clearest molecular genetic risk factor for schizophrenia that has been identified to date is deletion of a gene or several genes on chromosome 22, which can markedly increase the risk for schizophrenia. However, fairly strong data suggest that allelic variation in genes encoding the 5-HT_{2A} and D3 dopamine receptors confer a small degree of susceptibility.

As in other common diseases, it is hoped that advances will come through the use of a new generation of genetic markers and new methods of genotyping and statistical analysis (165). However, successful application of these methods requires access to large, well-characterized patient samples, and the collection of such data is a priority at the present time. We need to focus research on the development and refinement of phenotypic measures and biological markers. Success will also depend on the traditional medical disciplines of clinical description and epidemiology, and on our ability to integrate these with genetic approaches.

ACKNOWLEDGMENTS

Part of "49 - Molecular and Population Genetics of Schizophrenia "

Preparation of this chapter was supported in part by National Institute of Mental Health grants 1 R01MH41879-01, 5 UO1 MH46318-02, and 1 R37MH43518-01 to Dr. Ming T. Tsuang; the Veterans Administration Medical Research, Health Services Research, and Development and Cooperative Studies Programs; and a NARSAD Distinguished Investigator Award to Dr. Tsuang. Dr. Owen is supported by the United Kingdom Medical Research Council. Dr. Tsuang wishes to acknowledge the assistance of Drs. Stephen Faraone and William Stone, and Dr. Owen wishes to acknowledge the assistance of Drs. Michael O'Donovan and Alastair Cardno in the preparation of this manuscript.

REFERENCES

1. Bromet EJ, Dew MA, Eaton W. Epidemiology of psychosis with special reference to schizophrenia. In: Tsuang MT, Tohen M, Zahner GEP, eds. *Textbook in psychiatric epidemiology*. New York: Wiley-Liss, 1995:283–300.

2. Bromet EJ, Fennig S. Epidemiology and natural history of schizophrenia. *Biol Psychiatry* 1999;46:871–881.

3. Babigian HM. Schizophrenia: epidemiology. In: Freedman AM, Kaplan HI, Sadock BJ, eds. *Comprehensive textbook of psychiatry*. Baltimore: Williams & Wilkins, 1975.

4. Beiser M, Iacono WG. An update on the epidemiology of schizophrenia. *Can J Psychiatry* 1990;35:657–668.

5. Eaton WW. Epidemiology of schizophrenia. *Epidemiol Rev* 1985;7:105–126.

6. Jablensky A. Epidemiology of schizophrenia: a European perspective. *Schizophr Bull* 1986;12:52–73.

7. Eaton W, Day R, Kramer M. The use of epidemiology for risk factor research in schizophrenia: an overview and methodologic critique. In: Tsuang MT, Simpson JC, eds. *Handbook of schizophrenia 3: nosology, epidemiology and genetics*. Amsterdam: Elsevier Science, 1988:151–168.

8. Keith SJ, Regier DA, Rae DS. Schizophrenic disorders. In: Robbins LN, Regier DA, eds. *Psychiatric disorders in America: the epidemiologic catchment area study*. New York: The Free Press, 1991:33–52.

9. Kendler KS, Gallagher TJ, Abelson JM, et al. Lifetime prevalence, demographic risk factors, and diagnostic validity of nonaffective psychosis as assessed in a U.S. community sample. *Arch Gen Psychiatry* 1996;53:1022–1031.

10. Torrey F. Prevalence studies in schizophrenia. *Br J Psychiatry* 1987;150:598–608.

11. Tsuang MT, Faraone SV. Schizophrenia. In: Winokur G, Clayton P, eds. *Medical basis of psychiatry*, second ed. Philadelphia: WB Saunders, 1994:87–114.

12. Tsuang MT, Faraone SV, Day M. The schizophrenic disorders. In: Nicholi AM, ed. *The Harvard guide to modern psychiatry*. Cambridge, MA: Belknap Press, 1988:259–295.

13. Jablensky A, Sartorius N, Ernberg G, et al. Schizophrenia: manifestations, incidence and course in different cultures: a World Health Organization ten-country study. *Psychol Med Monogr Suppl* 1992;20:1–97.

14. Faraone SV, Chen WJ, Goldstein JM, et al. Gender differences in the age at onset of schizophrenia. *Br J Psychiatry* 1994;164:625–629.

15. Goldstein JM, Tsuang MT, Faraone SV. Gender and schizophrenia: implications for understanding the heterogeneity of the illness. *Psychiatry Res* 1989;28:243–253.

16. Seidman LJ, Goldstein JM, Goodman JM, et al. Sex differences in olfactory identification and Wisconsin Card Sorting performance in schizophrenia: relationship to attention and verbal ability. *Biol Psychiatry* 1997;42:104–115.

17. Faraone SV, Tsuang MT. Methods in psychiatric genetics. In: Tsuang MT, Tohen M, Zahner GEP, eds. *Textbook in psychiatric epidemiology*. New York: Wiley-Liss, 1995:81–134.

18. Gottesman II. *Schizophrenia genesis: the origin of madness*. New York: Freeman, 1991.

19. Gottesman II. Origins of schizophrenia: past as prologue. In: Plomin R, McClearn GE, eds. *Nature nurture and psychology*. Washington, DC: American Psychological Association, 1993:231–244.

20. Hovatta L, Terwilliger J, Lichtermann D, et al. Schizophrenia in the genetic isolate of Finland. *Am J Med Genet (Neuropsychiatr Genet)* 1997;74:353–360.

21. Myles-Worsley M, Coon H, Iobech J, et al. Genetic epidemiological study of schizophrenia in Palau, Micronesia: prevalence and familiality. *Am J Med Genet (Neuropsychiatr Genet)* 1999;88:4–10.

22. Cooper JE, Kendell RE, Gurland BJ, et al. *Psychiatric diagnosis in New York and London: a comparative study of mental hospital admissions*. London: Oxford University Press, 1972 (Institute of Psychiatry, Maudsley Monographs, No. 20).

23. Crow TJ. Etiopathogenesis and treatment of psychosis. *Annu Rev Med* 1994;45:219–234.

24. Buka SL, Tsuang MT, Lipsitt LP. Pregnancy/delivery complications and psychiatric diagnosis. A prospective study. *Arch Gen Psychiatry* 1993;50:151–156.

25. Goldstein JM. Sex and brain abnormalities in schizophrenia: fact or fiction? *Harvard Rev Psychiatry* 1996;4:110–115.

26. Goldstein JM, Seidman LJ, Goodman J, et al. *Sex differences in structural brain abnormalities in schizophrenia: frontal versus posterior brain regions*. Colorado Springs, CO: International Congress on Schizophrenia Research. 1997.

27. Tsuang MT, Faraone SV. The case for heterogeneity in the etiology of schizophrenia. *Schizophr Res* 1995;17:161–175.

28. Zornberg GL, Buka SL, Tsuang MT. Hypoxic ischemia-related fetal/neonatal complications and risk of schizophrenia and other nonaffective psychoses: a 19-year longitudinal study. *Am J Psychiatry* 2000;157:196–202.

29. Kendler KS, McGuire M, Gruenberg AM, et al. The Roscommon family study. I. Methods, diagnosis of probands, and risk of schizophrenia in relatives. *Arch of Gen Psychiatry* 1993;50:527–540.

30. Tsuang MT, Winokur G, Crowe RR. Morbidity risks of schizophrenia and affective disorders among first-degree relatives of patients with schizophrenia, mania, depression and surgical conditions. *Br J Psychiatry* 1980;137:497–504.

31. Guze SB, Cloninger R, Martin RL, et al. A follow-up and family study of schizophrenia. *Arch Gen Psychiatry* 1983;40:1273–1276.

32. Kendler KS. Overview: current perspective on twin studies of schizophrenia. *Am J Psychiatry* 1983;140:1413–1425.

33. Prescott CA, Gottesman II. Genetically mediated vulnerability to schizophrenia. *Psychiatr Clin North Am* 1993;16:245–267.

34. Kendler KS, Diehl SR. The genetics of schizophrenia: a current, genetic-epidemiologic perspective. *Schizophr Bull* 1993;19:261–285.

35. Cannon T, Kaprio J, Lonnqvist J, et al. The genetic epidemiology of schizophrenia in a Finnish twin cohort. *Arch Gen Psychiatry* 1998;55:67–74.

36. Cardno AG, Marshall EJ, Coid B, et al. Heritability estimates for psychotic disorders. *Arch Gen Psychiatry* 1999;56:162–168.

37. Farmer AE, McGuffin P, Gottesman II. Twin concordance for DSM-III schizophrenia: scrutinizing the validity of the definition. *Arch Gen Psychiatry* 1987;44:634–641.

38. Onstad S, Skre I, Torgersen S, et al. Twin concordance for DSM-III-R schizophrenia. *Acta Psychiatr Scand* 1991;83:395–401.

39. Gottesman II, Bertelsen A. Confirming unexpressed genotypes for schizophrenia. Risks in the offspring of Fischer's Danish identical and fraternal discordant twins. *Arch Gen Psychiatry* 1989;46:867–872.

40. Fischer M. Psychosis in the offspring of schizophrenic monozygotic twins and their normal co-twins. *Br J Psychiatry* 1971;118:43–52.

41. McNeil TF, Cantor-Graae E, Weinberger DR. Relationship of obstetric complications and differences in size of brain structures in monozygotic twins pairs discordant for schizophrenia. *Am J Psychiatry* 2000;157:203–212.

42. Tsujita T, Niikawa N, Yamashita H, et al. Genomic discordance between monozygotic twins discordant for schizophrenia. *Am J Psychiatry* 1998;155:422–424.

43. Heston LL. Psychiatric disorders in foster home-reared children of schizophrenic mothers. *Br J Psychiatry* 1966;112:819–825.

44. Kety SS, Rosenthal D, Wender PH, et al. The types and prevalence of mental illness in the biological and adoptive families of adopted schizophrenics. *J Psychiatr Res* 1968;1:345–362.

45. Kety SS, Wender PH, Jacobsen B, et al. Mental illness in the biological and adoptive relatives of schizophrenic adoptees. Replication of the Copenhagen study in the rest of Denmark. *Arch Gen Psychiatry* 1994;51:442–455.

46. Kendler KS, Gruenberg AM, Kinney DK. Independent diagnoses of adoptees and relatives as defined by DSM-III in the provincial and national samples of the Danish adoption study of schizophrenia. *Arch Gen Psychiatry* 1994;51:456–468.

47. Tienari P, Lyman CW, Moring J, et al. The Finnish adoptive family study of schizophrenia: implications for family research. *Br J Psychiatry* 1994;164[Suppl. 23]:20–26.

48. Wahlberg K-E, Wynne LC, Oja H, et al. Gene-environment interaction in vulnerability to schizophrenia: findings from the Finnish adoptive family study in schizophrenia. *Am J Psychiatry* 1997;154:355–362.

49. Gottesman II, Shields J. *Schizophrenia: the epigenetic puzzle*. Cambridge: Cambridge University Press, 1982.

50. Bertelsen A, Gottesman II. Schizoaffective psychoses: genetic clues to classification. *Am J Med Genet* 1995;60:7–11.

51. Cannon TD, Mednick SA, Parnas J. Antecedents of predominantly negative- and predominantly positive-symptom schizophrenia in a high-risk population. *Arch Gen Psychiatry* 1990;47:622–632.

52. Cuesta MJ, Peralta V, Caro F. Premorbid personality in psychoses. *Schizophr Bull* 1999;25:801–811.

53. Erlenmeyer-Kimling L, Squires-Wheeler E, Adamo UH, et al. The New York high-risk project. Psychoses and cluster A personality disorders in offspring of schizophrenic parents at 23 years of follow-up. *Arch Gen Psychiatry* 1995;52:857–865.

54. Battaglia M, Torgersen S. Review article. Schizotypal disorder: at the crossroads of genetics and nosology. *Acta Psychiatr Scand* 1996;94:303–310.

55. Tsuang MT, Stone WS, Faraone SV. Schizophrenia: a review of genetic studies. *Harvard Rev Psychiatry* 1999;7:185–207.

56. Baron M, Gruen R, Asnis L, et al. Familial transmission of schizotypal and borderline personality disorders. *Am J Psychiatry* 1985;142:927–934.

57. Gunderson JG, Siever LJ, Spaulding E. The search for a schizotype: crossing the border again. *Arch Gen Psychiatry* 1983;40:15–22.

58. Kendler KS, Masterson CC, Ungaro R, et al. A family history study of schizophrenia-related personality disorders. *Am J Psychiatry* 1984;141:424–427.

59. Coryell WH, Zimmerman M. Personality disorder in the families of depressed, schizophrenic, and never-ill probands. *Am J Psychiatry* 1989;146:496–502.

60. Torgersen S. Relationship of schizotypal personality disorder to schizophrenia: genetics. *Schizophr Bull* 1985;11:554–563.

61. Clementz BA, Grove WM, Katsanis J, et al. Psychometric detection of schizotypy: perceptual aberration and physical anhedonia in relatives of schizophrenics. *J Abnorm Psychol* 1991;100:607–612.

62. Kendler KS, McGuire M, Gruenberg AM, et al. The Roscommon family study. III. Schizophrenia-related personality disorders in relatives. *Arch Gen Psychiatry* 1993;50:781–788.

63. Torgersen S. Genetic and nosological aspects of schizotypal and borderline personality disorders: a twin study. *Arch Gen Psychiatry* 1984;41:546–554.

64. Meehl PE. Schizotaxia, schizotypy, schizophrenia. *Am Psychol* 1962;17:827–838.

65. Meehl PE. Schizotaxia revisited. *Arch Gen Psychiatry* 1989;46:935–944.

66. Tsuang MT, Gilbertson MW, Faraone SV. Genetic transmission of negative and positive symptoms in the biological relatives of schizophrenics. In: Marneros A, Tsuang MT, Andreasen N, eds. *Positive vs. negative schizophrenia*. New York: Springer-Verlag, 1991;265–291.

67. Faraone SV, Green AI, Seidman LJ, et al. "Schizotaxia": clinical implications and new directions for research. *Schizophr Bull* 2001;27:1–18.

68. Faraone SV, Kremen WS, Lyons MJ, et al. Diagnostic accuracy and linkage analysis: how useful are schizophrenia spectrum phenotypes? *Am J Psychiatry* 1995;152:1286–1290.

69. Faraone SV, Seidman LJ, Kremen WS, et al. Neuropsychological functioning among the nonpsychotic relatives of schizophrenic patients: diagnostic efficiency analysis. *J Abnorm Psychol* 1995;104:286–304.

70. Faraone SV, Seidman LJ, Kremen WS, et al. Neuropsychological functioning among the nonpsychotic relatives of schizophrenic patients: the effect of genetic loading. *Biol Psychiatry* 2000;48:120–126.

71. Park S, Holzman PS, Goldman-Rakic PS. Spatial working memory deficits in the relatives of schizophrenic patients. *Arch Gen Psychiatry* 1995;52:821–828.

72. Tsuang MT, Stone WS, Seidman LJ, et al. Treatment of nonpsychotic relatives of patients with schizophrenia: four case studies. *Biol Psychiatry* 1999;41:1412–1418.

73. Tsuang MT, Stone WS, Faraone SV. Towards reformulating the diagnosis of schizophrenia. *Am J Psychiatry* 2000;157:1041–1050.

74. Lalouel JM, Morton NE. Complex segregation analysis with pointers. *Hum Hered* 1981;31:312–321.

75. Morton NE, MacLean CJ. Analysis of family resemblance. III: Complex segregation analysis of quantitative traits. *Am J Hum Genet* 1974;26:489–503.

76. O'Rourke DH, Gottesman II, Suarez BK, et al. Refutation of the single locus model for the etiology of schizophrenia. *Am J Hum Genet* 1982;34:630–649.

77. Risch N, Baron M. Segregation analysis of schizophrenia and related disorders. *Am J Hum Genet* 1984;36:1039–1059.

78. McGue M, Gottesman II. Genetic linkage and schizophrenia: perspectives from genetic epidemiology. *Schizophr Bull* 1989;15:453–464.

79. Risch N. Linkage strategies for genetically complex traits: I. Multilocus models. *Am J Hum Genet* 1990;46:222–228

80. Goate A, Chartier-Harlin MC, Mullan ME. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 1991;349:704–706.

81. Sherrington R, Rogaev EI, Liang Y, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 1995;375:754–760.

82. Levy-Lahad E, Wasco W, Poorkaj P, et al. Candidate gene for the chromosome-1 familial Alzheimer's-disease locus. *Science* 1995;269:973–977.

83. Sherrington et al. 1988.

84. McGuffin P, Owen MJ. Molecular genetic studies of schizophrenia. *Cold Spring Harb Symp Quant Biol* 1996;61:815–822.

85. Craddock N, Khodel V, Van Eerdewegh P. Mathematical limits of multilocus models: the genetic transmission of bipolar disorder. *Am J Hum Genet* 1995;57:690–702.

86. Gill M, Vallada H, Collier D, et al. A combined analysis of D22s278 marker alleles in affected sib-pairs—support for a susceptibility locus for schizophrenia at chromosome 22q12. *Am J Med Genet* 1996;67:40–45

87. Levinson DF, Wildenauer DB, Schwab SG, et al. Additional support for schizophrenia linkage on chromosome-6 and chromosome-8—a multicenter study. *Am J Med Genet* 1996;67:580–594.

88. Moldin SO. The maddening hunt for madness genes. *Nat Genet* 1997;17:127–129.

89. Lin MW, Curtis D, Williams N, et al. Suggestive evidence for linkage of schizophrenia to markers on chromosome 13q14.1-q22. *Psychiatr Genet* 1995;5:117–126.

90. Blouin JL, Dombroski BA, Nath SK, et al. Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21. *Nat Genet* 1998;20:70–73.

91. Brzustowicz LM, Honer WG, Chow EWC, et al. Linkage of familial schizophrenia to chromosome 13q32. *Am J Hum Genet* 1999;65:1096–1103.

92. Schwab SG, Eckstein SG, Hallmayer J, et al. Evidence suggestive of a locus on chromosome 5q31 contributing to susceptibility for schizophrenia in German and Israeli families by multipoint affected sib-pair linkage analysis. *Mol Psychiatry* 1997;2:156–160.

93. Straub RE, Maclean CJ, O'Neill FA, et al. Support for a possible schizophrenia vulnerability locus in region 5q22-31 in Irish families. *Mol Psychiatry* 1997;2:148–155.

94. Schwab SG, Hallmayer J, Lerer B, et al. Support for a chromosome 18p locus conferring susceptibility to functional psychoses in families with schizophrenia, by association and linkage analysis. *Am J Hum Genet* 1998;63:1139–1152.

95. Faraone SV, Matise T, Svrakic D, et al. Genome scan of European-American schizophrenia pedigrees. Results of the NIMH Genetics Initiative and Millennium Consortium. *Am J Med Genet* 1998;81:290–295.

96. Schwab SG, Hallmayer J, Albus M, et al. A potential susceptibility locus on chromosome 10p14-p11 in 72 families with schizophrenia. *Am J Med Genet* 1998;81:528–529.

97. Straub RE, Maclean CJ, Martin RB, et al. A schizophrenia locus may be located in region 10p15-p11. *Am J Med Genet* 1998;81:296–301.

98. Cao Q, Martinez M, Zhang J, et al. Suggestive evidence for a schizophrenia susceptibility locus on chromosome 6q and a confirmation in an independent series of pedigrees. *Genomics* 1997;43:1–8.

99. Martinez M, Goldin LR, Cao Q, et al. Follow-up study on a susceptibility locus for schizophrenia on chromosome 6q. *Am J Med Genet* 1999;88:337–343.

100. Brzustowicz L, Hodgkinson K, Chow E, et al. Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22. *Science* 2000;288:678–682.

101. Williams NM, Rees MI, Holmans P, et al. A two-stage genome scan for schizophrenia susceptibility genes in 196 affected sibling pairs. *Hum Mol Genet* 1999;8:1729–1739.

102. Lander E, Kruglyak L. Genetic dissection of complex traits—guidelines for interpreting and reporting linkage results. *Nat Genet* 1995;11:241–247.

103. Suarez BK, Hampe CL, Van Eerdewegh P. Problems of replicating linkage claims in psychiatry. In: Gershon ES, Cloninger CR, eds. *Genetic approaches to mental disorders*. Washington, DC: American Psychiatric Press, 1994:23–46.

104. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994;265:2037–2048.

105. Lander ES. The new genomics: global view of biology. *Science* 1996;274:536–539.

106. Roberts SB, MacLean CJ, Neale MC, et al. Replication of linkage studies of complex traits: an examination of variation in location estimates. Am J Hum Genet 1999;65:876–884.

107. Hauser ER, Boehnke M, Guo SW, et al. Affected-sib-pair interval mapping and exclusion for complex genetic traits—sampling considerations. *Genet Epidemiol* 1996;13:117–137.

108. Scott W, Pericak Vance M, et al. Genetic analysis of complex diseases. *Science* 1997;275:1327, discussion 1329–1330.

109. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;273:1516–1517.

110. Owen MJ, McGuffin P. Association and linkage—complementary strategies for complex disorders. *J Med Genet* 1993;30:638–639.

111. Owen MJ, Holmans P, McGuffin P. Association studies in psychiatric genetics. *Mol Psychiatry* 1997;2:270–273.

112. Schaid DJ, Sommer SS. Comparison of statistics for candidate gene association studies using cases and parents. *Am J Hum Genet* 1994;55:402–409

113. Risch N, Teng J. The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases—I. DNA pooling. *Genome Res* 1998;8:1273–1288

114. Inayama Y, Yoneda H, Sakai T, et al. Positive association between a DNA sequence variant in the serotonin 2A receptor gene and schizophrenia. *Am J Med Genet* 1996;67:103–105

115. Williams J, Spurlock G, McGuffin P, et al. Association between schizophrenia and T102C polymorphism of the 5-hydroxytryptamine type 2A-receptor gene. *Lancet* 1996;347:1294–1296.

116. Spurlock G, Heils A, Holmans P, et al. A family-based association study of T102C polymorphism in 5-HT_{2A} and schizophrenia plus identification of new polymorphisms in the promoter. *Mol Psychiatry* 1998;3:42–49.

117. Williams J, McGuffin P, Nothen M, et al. Meta-analysis of association between the 5-HT2a receptor T102C polymorphism and schizophrenia. EMASS Collaborative Group. European Multicentre Association Study of Schizophrenia. *Lancet* 1997;349:1221.

118. Bunzel R, Blumcke I, Cichon S, et al. Polymorphic imprinting of the serotonin-2A (5-HT_{2A}) receptor gene in human adult brain. *Mol Brain Res* 1998;59:90–92

119. Crocq MA, Mant R, Asherson P, et al. Association between schizophrenia and homozygosity at the dopamine D3 receptor gene. *J Med Genet* 1992;29:858–860.

120. Williams J, Spurlock G, Holmans P, et al. A meta-analysis and transmission disequilibrium study of association between the dopamine D3 receptor gene and schizophrenia. *Mol Psychiatry* 1998;3:141–149.

121. Anney et al. *(submitted)*.

122. Morel BA. Traité des dégénérescences. Paris: JB Bailliére, 1857.

123. O'Donovan MC, Owen MJ. The molecular genetics of schizophrenia. *Ann Med* 1996;28:541–546.

124. Petronis A, Kennedy JL. Unstable genes—unstable mind? *Am J Hum Genet* 1996;152:164–172.

125. O'Donovan C, Guy C, Craddock N, et al. Schizophrenia and bipolar disorder are associated with expanded CAG/CTG repeats. *Nat Genet* 1995;10:380–381.

126. Morris AG, Gaitonde E, McKenna PJ, et al. CAG repeat expansions and schizophrenia—association with disease in females and with early age at onset. *Hum Mol Genet* 1995;4:1957–1961.

127. O'Donovan M, Guy C, Craddock N, et al. Confirmation of association between expanded CAG/CTG repeats and both schizophrenia and bipolar disorder. *Psychol Med* 1996;26:1145–1153.

128. Vincent JB, Petronis A, Strong E, et al. Analysis of genome-wide CAG/CTG repeats, and at *SEF2-1B* and *ERDA1* in schizophrenia and bipolar affective disorder. *Mol Psychiatry* 1999;4:229–234.

129. Li et al. 1998.

130. Laurent C, Zander C, Thibaut F, et al. Anticipation in schizophrenia: no evidence of expanded CAG/CTG repeat sequences in French families and sporadic cases. *Am J Med Genet* 1998;81:342–346

131. Chandy KG, Fantino E, Wittekindt O, et al. Isolation of a novel potassium channel gene *hSKCa3* containing a polymorphic CAG repeat: a candidate for schizophrenia and bipolar disorder? *Mol Psychiatry* 1998;3:32–37.

132. Dror V, Shamir E, Ghanshani S, et al. *hKCa3/KCNN3* potassium channel gene: association of longer CAG repeats with schizophrenia in Israeli Ashkenazi Jews, expression in human tissues and localization to chromosome 1q21. *Mol Psychiatry* 1999;4:254–260.

133. Bowen T, Guy CA, Craddock N, et al. Further support for an association between a polymorphic CAG repeat in the *hKCa3* gene and schizophrenia. *Mol Psychiatry* 1998;3:266–269.

134. Wittekindt O, Schwab SG, Burgert E, et al. Association between *hSKCa3* and schizophrenia not confirmed by transmission disequilibrium test in 193 offspring parents trios. *Mol Psychiatry* 1999;4:267–270.

135. Stöber G, Jatzke S, Meyer J, et al. Short CAG repeats within the *hSKCa3* gene associated with schizophrenia: results of a family-based study. *Neuroreport* 1998;9:3595–3599.

136. Li T, Hu X, Chandy KG, et al. Transmission disequilibrium analysis of a triplet repeat within the *hKCa3* gene using family trios with schizophrenia. *Biochem Biophys Res Commun* 1998;251:662–665.

137. Austin CP, Holder DJ, Ma L, et al. Mapping of *hKCa3* to chromosome 1q21 and investigation of the linkage of CAG repeat polymorphism to schizophrenia. *Mol Psychiatry* 1999;4:261–266

138. Tsai MT, Shaw CK, Hsiao KJ, et al. Genetic association study of a polymorphic CAG repeat array of calcium activated potassium channel *(KCNN3)* gene and schizophrenia among the chinese population from Taiwan. *Mol Psychiatry* 1998;4:271–273

139. Joober R, Benkelfat C, Brisebois K, et al. Lack of association between the *hSKCa3* channel gene CAG polymorphism and schizophrenia. *Am J Med Genet* 1999;88:154–157.

140. Bonnet Brilhault F, Laurent C, Campion D, et al. No evidence for involvement of *KCNN3 (hSKCa3)* potassium channel gene in familial and isolated cases of schizophrenia. *Eur J Hum Genet* 1999;7:247–250

141. Sidransky E, Burgess C, Ikeuchi T, et al. A triplet repeat on 17q accounts for most expansions detected by the repeat-expansion-detection technique. *Am J Hum Genet* 1998;62:1548–1551.

142. Lindblad K, Nylander PO, Zander C, et al. Two commonly expanded CAG/CTG repeat loci: involvement in affective disorders? *Mol Psychiatry* 1998;3:405–410.

143. Ross CA. Schizophrenia genetics: expansion of knowledge? *Mol Psychiatry* 1999;4:4–5

144. Driscoll DA, Spinner NB, Budarf ML, et al. Deletions and microdeletions of 22q11.2 in velo-cardio-facial syndrome. *Am J Med Genet* 1992;44:261–268.

145. Shprintzen R, Goldberg RB, Lewin ML, et al. A new syndrome involving cleft palate, cardiac anomalies, typical facies and learning disabilities: velo-cardio-facial syndrome. *Cleft Palate J* 1978;15:56–62.

146. Papolos DF, Faedda GI, Veit S, et al. Bipolar spectrum disorders in patients diagnosed with velo-cardio-facial syndrome: does a hemizygous deletion of chromosome 22q11 result in bipolar affective disorder? *Am J Psychiatry* 1996;153:1541–1547

147. Shprintzen RJ, Goldberg RB, Golding-Kushner KJ. Late-onset psychosis in the velo-cardio-facial syndrome. *Am J Med Genet* 1992;42:141–142.

148. Pulver et al. 1994.

149. Murphy KC, Jones LA, Owen MJ. High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch Gen Psychiatry* 1999;56:940–945.

150. Weinberger DR. From neuropathology to neurodevelopment. *Lancet* 1995;346:552–557.

151. Scambler P, Kelly D, Lindsay E, et al. Velo-cardio-facial syndrome associated with chromosome 22 deletions encompassing the DiGeorge locus. *Lancet* 1992;339:1138–1139.

152. Chow EWC, Bassett AS, Weksberg R. Velo-cardio-facial syndrome and psychotic disorders: implications for psychiatric genetics. *Am J Med Genet* 1994;54:107–112.

153. Karayiorgou M, Morris MA, Morrow B, et al. Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. *Proc Natl Acad Sci U S A* 1995;92:7612–7616

154. Gothelf D, Frisch A, Munitz H, et al. Velocardiofacial manifestations and microdeletions in schizophrenic patients. *Am J Med Genet* 1997;72:455–461

155. Bassett A, Hodgkinson K, Chow EWC, et al. 22q11 deletion syndrome in adults with schizophrenia. *Am J Med Genet* 1998;81:328–337

156. Murphy KC, Owen MJ. The behavioral phenotype in velo-cardio-facial syndrome. *Am J Hum Genet* 1997;61[4.SS]:15.

157. Pulver AE, Karayiorgou M, Wolyniec PS, et al. Sequential strategy to identify a susceptibility gene for schizophrenia—report of potential linkage on chromosome 22q12-q13.1. Part 1. *Am J Med Genet* 1994;54:36–43.

158. Lasseter VK, Pulver AE, Wolyniec PS, et al. Follow-up report of potential linkage for schizophrenia on chromosome 22q.3. *Am J Med Genet* 1995;60:172–173.

159. Shaw SH, Kelly M, Smith AB, et al. A genome wide search for schizophrenia susceptibility genes. *Am J Med Genet* 1998;81:364–376.

160. Dunham I, Collins J, Wadey R, et al. Possible role for COMT in psychosis associated with velo-cardio-facial syndrome. *Lancet* 1992;340:1361–1362.

161. Kendler KS, Tsuang MT, Hays P. Age at onset in schizophrenia: a familial perspective. *Arch Gen Psychiatry* 1987;44:881–890.

162. Kendler KS, Karkowski-Shuman L, O'Neill FA, et al. Resemblance of psychotic symptoms and syndromes in affected sibling pairs from the Irish study of high-density schizophrenia families: evidence for possible etiologic heterogeneity. *Am J Psychiatry* 1997;154:191–198.

163. DeLisi LE. A critical overview of recent investigations into the genetics of schizophrenia. *Curr Opin Psychiatry* 1999;12:29–39.

164. Almasy L, Blangero J. Linkage strategies for mapping genes for complex traits in man. In: Crusio WE, Gerlai RT, eds. *Handbook of molecular-genetic techniques for brain and behaviour research*. Amsterdam: Elsevier Science, 1999:100–112.

165. Owen MJ, Cardno AG, O'Donovan MC. Psychiatric genetics: back to the future. *Mol Psychiatry* 2000;5:22–31.

166. Barcellos LF, Klitz W, Field L, et al. Association mapping of disease loci by use of a pooled DNA genomic screen. *Am J Hum Genet* 1997;61:734–747.

167. Daniels J, Holmans P, Williams N, et al. A simple method for analyzing microsatellite allele image patterns generated from DNA pools and its application to allelic association studies. *Am J Hum Genet* 1998;62:1189–1197.

168. Fisher PJ, Turic D, Williams NM, et al. DNA pooling identifies QTLs for general cognitive ability in children on chromosome 4. *Hum Mol Genet* 1999;8:915–922.

169. Kirov G, Williams N, Sham P, et al. Pooled genotyping of microsatellite markers in parent-offspring trios. *Genome Res* 2000;10:105–115.

170. Kerwin R, Owen MJ. The genetics of novel therapeutic targets in schizophrenia. *Br J Psychiatry* 1999;174[Suppl 38]:1–4.

171. Wood GK, Tomasiewicz H, Rutishauser U, et al. NCAM-180 knockout mice display increased lateral ventricle size and reduced prepulse inhibition of startle. *Neuroreport* 1998;9:461–466.

172. Skoultchi AI, Puech A, Saint-Jore B, et al. Comparative mapping of the human and mouse VCFS/DGS syntenic region discloses the presence of a large internal rearrangement. *Am J Hum Genet* 1997;61:A296.

173. Kimber WL, Hsieh P, Hirotsune S, et al. Deletion of 150 kb in the minimal DiGeorge/velocardiofacial syndrome critical region in mouse. *Hum Mol Genet* 1999;8:2229–2237.

174. Sham PC. Statistical methods in psychiatric genetics. *Stat Methods Med Res* 1998;7:279–300.

175. Farmer A, Owen MJ, McGuffin P. Bioethics and genetics research in psychiatry. *Br J Psychiatry* 2000;176:105–108.

50

Animal Models Relevant to Schizophrenia Disorders

Mark A. Geyer

Bita Moghaddam

Mark A. Geyer: Department of Psychiatry, School of Medicine, University of California at San Diego, La Jolla, California.

Bita Moghaddam: Departments of Psychiatry and Neurobiology, Yale University School of Medicine, New Haven, Connecticut.

Animal models used to study schizophrenia include both models of the full syndrome and models of specific signs or symptoms. As reviewed elsewhere (1), models are commonly explored initially because of indications of so-called face validity, but they are evaluated scientifically in terms of their construct and etiologic and predictive validity with respect to both clinical phenomena and responsiveness to antipsychotic drugs. Here, models are organized by the manipulations used to mimic the clinical phenomena. Thus, in some of these models, only specific dependent measures are utilized, whereas others are evaluated by using a range of dependent measures.

A model is defined as any experimental preparation developed to study a particular condition or phenomenon in the same or different species. Typically, models are animal preparations that attempt to mimic a human condition, in our case the human psychopathology associated with the group of schizophrenia disorders. In developing and assessing an animal model, it is important to specify the purpose intended for the model because the intended purpose determines the criteria that the model must satisfy to establish its validity. At one extreme, one can attempt to develop an animal model that mimics the schizophrenia syndrome in its entirety. In the early years of psychopharmacology, the term *animal model* often denoted such an attempt to reproduce a psychiatric disorder in a laboratory animal. Unfortunately, the group of schizophrenia disorders is characterized by considerable heterogeneity and a complex clinical course that reflects many factors that cannot be reproduced readily in animals. Thus, the frequent attempts to model the syndromes of schizophrenia in animals usually met with failure and so prompted skepticism regarding this entire approach.

At the other extreme, a more limited use of an animal model related to schizophrenia is to study systematically the effects of antipsychotic treatments. Here, the behavior of the model is intended to reflect only the efficacy of known therapeutic agents and so lead to the discovery of related pharmacotherapies. Because the explicit purpose of the model is to predict treatment efficacy, the principle guiding this approach has been termed *pharmacologic isomorphism* (2). The fact that such models are developed and validated by reference to the effects of known therapeutic drugs frequently limits their ability to identify new drugs with novel mechanisms of action. Similarly, an important limitation inherent in this approach is that it is not designed to identify new antipsychotic agents that might better treat the symptoms of schizophrenia refractory to current treatments.

Because of the complexity of schizophrenia, another approach to the development of relevant animal models relies on focusing on specific signs or symptoms associated with schizophrenia, rather than mimicking the entire syndrome. In such cases, specific observables that have been identified in schizophrenic patients provide a focus for study in experimental animals. The particular behavior being studied may or may not be pathognomonic for or even symptomatic of schizophrenia, but it must be defined objectively and observed reliably. It is important to emphasize that the reliance of such a model on specific observables minimizes a fundamental problem plaguing animal models of the syndrome of schizophrenia. Specifically, the difficulties inherent in conducting experimental studies of schizophrenic patients have limited the number of definitive clinical findings with which one can validate an animal model of schizophrenia. The validation of any animal model can only be as sound as the information available in the relevant clinical literature (3). By focusing on specific signs or symptoms rather than syndromes, one can increase the confidence in the cross-species validity of the model. The narrow focus of this approach generally leads to pragmatic advantages in the conduct of mechanistic studies addressing the neurobiological substrates of the behavior in question. By contrast, in models intended to reproduce the entire syndrome of schizophrenia, the need for multiple simultaneous endpoints

makes it relatively difficult to apply the invasive experimental manipulations required to establish underlying mechanisms.

Another approach to the development of animal models is based more theoretically on psychological constructs believed to be affected in schizophrenia. Such identification of underlying psychological processes or behavioral dimensions (2 ,3) involves the definition of a hypothetical construct and the subsequent establishment of operational definitions suitable for experimental testing of the validity of the construct. Constructs such as selective attention, perseveration, sensorimotor gating, and working memory have been used in this manner in schizophrenia research. This approach is most fruitful when conceptually or procedurally related experiments are undertaken in both the relevant patient population and the putative animal model. In other words, studies of appropriate patients are needed to establish the operational definitions of the hypothetical construct and the relevance of the construct to schizophrenia. In concert, parallel studies of the potentially homologous construct, process, or dimension are required to determine the similarity of the animal model to the human phenomena. An important and advantageous aspect of this approach is that the validation of the hypothetical construct and its cross-species homology can be established by studies of normal humans and animals in addition to studies of schizophrenic patients and experimentally manipulated animals. Thus, this approach benefits from the existing literature relevant to the hypothetical construct on which the model is based. In a sense, this approach explicitly recognizes that the experimental study of schizophrenia in humans involves as much of a modeling process as does the study of the disorder in animals.

PHENOTYPIC CHARACTERIZATION OF ANIMAL MODELS

Part of "50 - Animal Models Relevant to Schizophrenia Disorders "

Behavioral measures have been used extensively for establishing the validity of animal models of schizophrenia. Some of these measures, such as horizontal locomotion, do not correspond to schizophrenic symptomatology and have been primarily useful for providing a functional measure of the antidopaminergic activity of neuroleptics. Other behavioral measures, such as disruption of prepulse inhibition or impaired attentional set shifting, resemble characteristics of schizophrenia. These measures are useful for establishing the construct and predictive validity of putative animal models.

In addition to behavioral assessments, cellular and molecular markers that are based on described changes in human postmortem and imaging studies are potentially useful measures for establishing the validity of animal models. Furthermore, because of the inherent limitations of modeling in laboratory animals some of the most prominent behavioral abnormalities of schizophrenia, such as delusions and hallucinations, cellular and molecular characterizations can complement behavioral investigations in evaluating developmental and genetic models of schizophrenia.

Behavioral Phenotypes

Locomotor and Stereotypy

Changes in locomotor activity in rodents have often been used to assess both models of schizophrenia and the effects of antipsychotic treatments. The original impetus for the use of locomotor activity measures was derived from the psychostimulant models based on the dopamine hypothesis of schizophrenia, as reviewed elsewhere (3 ,4 and 5). These models arose because of the apparent similarity (i.e., “face validity”) between the symptoms of schizophrenia and the effects of high doses of amphetamine in presumably normal humans (6). Cross-species studies in animals treated with psychostimulants revealed both locomotor hyperactivity and, at higher doses, striking stereotyped or perseverative behaviors, which were seen as having face validity for the stereotyped behavior induced by amphetamine in humans (4 ,6 ,7). Measures of locomotor hyperactivity have been used extensively to characterize the effects of both dopaminergic psychostimulants and *N*-methyl-D-aspartate (NMDA) antagonists, such as phencyclidine (PCP), although PCP-induced hyperactivity differs markedly in qualitative features from that produced by dopaminergic psychostimulants. Although patients with schizophrenia are not typically hyperactive, they often exhibit perseverative or stereotyped behaviors. Hence, many studies in rodents have focused on the forms of stereotypy produced by psychostimulants.

Gating Measures

Clinical observations in schizophrenic patients have identified deficiencies in the processing of information, including an inability automatically to filter or “gate” irrelevant thoughts and sensory stimuli to prevent them from intruding on conscious awareness. Hence, theories of schizophrenic disorders often conceptualize the common aspect of these disorders as involving one or more deficits in the multiple mechanisms that enable normal persons to filter or gate most of the sensory stimuli they receive (8 ,9 and 10). In the most classic measure of filtering deficits, numerous studies have observed deficits in the habituation of startle responses in schizophrenic patients (e.g., 12,31,102), which may reflect failures of sensory filtering leading to disorders of cognition. A more specific class of such mechanisms is referred to as *sensory* or *sensorimotor gating*. Theoretically, impairments in either filtering or gating lead to sensory overload and cognitive fragmentation. It is also possible that the mechanisms that subserve experimental examples of filtering or gating are also responsible for the gating of cognitive information. The hypothetical construct of sensorimotor

gating has been operationalized and explored in both human and animal studies. The validity of this gating construct has been assessed most thoroughly by means of an operational measure based on cross-species homologies in the startle reflex—namely, the prepulse inhibition (PPI) of startle paradigm. In a conceptually related approach, analogous measures of event-related potentials are used across species to study sensory gating in the P50 event-related potential condition–test paradigm.

Habituation

Habituation refers to the decrement in responding when the same unimportant stimuli or cognitions occur repeatedly in the absence of any contingencies. Habituation is considered to be the simplest form of learning and is essential for the development of selective attention. Although habituation can be assessed with a variety of behavioral measures, the most common approach has been to study the gradual decrease in the startle response elicited by a series of tactile or acoustic stimuli in humans, rats, or mice. In patients with schizophrenia or schizotypy, deficits in startle habituation have been reported with the use of either modality of startling stimuli (9 ,11 ,12 and 13). A striking advantage of the startle habituation measure is the fact that extremely similar behavioral tests can be conducted in both humans and experimental animals.

Prepulse Inhibition

The PPI paradigm is based on the fact that a weak prestimulus presented 30 to 500 milliseconds before a startling stimulus reduces, or gates, the amplitude of the startle response. The generality and reliability of this robust phenomenon is clear; PPI is observed in many species, PPI is evident both within and between multiple sensory modalities when a variety of stimulus parameters are used, and PPI does not require learning or comprehension of instructions. Virtually all the evidence available supports the belief that PPI is homologous from rodents to humans, unlike most other cross-species comparisons based on often dubious arguments of similarity or, at best, analogy. As reviewed elsewhere (14 ,15), several laboratories have reported significant deficits in PPI in schizophrenic, schizotypal, and presumably psychosis-prone subjects with the use of a variety of testing procedures and stimulus parameters. Nevertheless, PPI deficits are not unique to patients in whom schizophrenia has been diagnosed; they are also observed in other psychiatric disorders involving abnormalities of gating in the sensory, motor, or cognitive domains.

P50 Gating

In the P50 sensory gating paradigm, two acoustic clicks are presented in rapid succession, usually 500 milliseconds apart. In normal persons, the P50 event-related potential to the second click is reduced or gated relative to the event-related potential to the first click. Schizophrenic patients and their first-degree relatives exhibit less sensory gating (16). An analogous form of sensory gating is studied in rodents based on the N40 event-related potential generated from the hippocampus (17).

Latent Inhibition

Latent inhibition is a relatively complex paradigm that is conceptually related to the gating theories of schizophrenic disorders. *Latent inhibition* refers to the observation that repeated exposures to a sensory stimulus (i.e., habituation) retard the rate at which a subject subsequently acquires a stimulus–response association based on this stimulus (18). Deficits in latent inhibition have been reported in schizophrenic patients (19), although it appears that such deficits may be limited to acute episodes of schizophrenia (19 ,20). This limitation has diminished interest in latent inhibition as a model for schizophrenia.

Social Behavior

Social withdrawal is included among the negative symptoms of schizophrenia and is often one of the earliest symptoms to occur. Models of social isolation have been studied in both monkeys (21) and rats (22). Naturally, given the importance of language in human social interactions, the species-specific differences in social behavior limit the direct comparisons that can be made across species.

Cognitive Measures

Cognitive deficiencies played a prominent role in the original description of schizophrenia by Kraepelin and distinguish the diagnosis of schizophrenia from manic-depressive and other forms of psychosis. Cognitive deficits are reported across all subtypes of schizophrenia and include impairments of attention, working memory, verbal memory, set shifting, and abstraction. Severe cognitive deficits appear to be a major factor contributing to impaired social and vocational functioning and treatment outcome (23). Current modes of therapy for schizophrenia (i.e., dopamine D2 or dopamine D2/serotonin 5-HT_2 antagonists, which effectively reduce the positive symptoms of schizophrenia in the majority of patients) have minimal beneficial effects on cognitive functioning. Typical antipsychotic drugs (e.g., haloperidol, chlorpromazine) may in fact lead to a deterioration in cognitive functions in schizophrenia patients (24) by producing so-called secondary deficit symptoms. Although some reports suggest that the atypical antipsychotic drug clozapine and the new generation of antipsychotics (e.g., olanzapine and risperidone, which target several subtypes of dopamine and serotonin receptors) may improve cognitive function, this effect is relatively small and has not been reproducible across laboratories (25 ,26 and 27). Thus, an important future direction of preclinical research relating to schizophrenia

is the design of animal models and novel treatments that target cognitive dysfunctions associated with this disorder. However, establishing the validity of animal models of cognitive deficits of schizophrenia and designing new pharmacologic approaches to the treatment of symptoms depends on appropriate behavioral paradigms for laboratory animals. These must provide as good an analogy as possible to the empiric measures on which schizophrenic patients are impaired. Unfortunately, the ratings of symptoms commonly assessed in clinical studies are of little value in this context.

The limited cognitive capacity of laboratory animals hinders the design of cognitive tasks that can be considered entirely "analogous" to relevant experimental paradigms, such as the Wisconsin Card Sorting Test and Continuous Performance Test. Therefore, most of the research involving cognitive tasks that are relevant to schizophrenia has been conducted in monkeys. Nevertheless, it is possible to design behavioral paradigms in rodents that can evaluate cognitive constructs "comparable" with those measured in many human experimental paradigms (40 ,41 and 42).

Among the most common of animal cognitive tasks are those with a working memory component (i.e., the ability to guide behavior by forming internal representations of stimuli that are no longer present in the environment). Human psychological tests of working memory, such as the "*n*-back" task (28), on which patients with schizophrenia exhibit an impairment (29), provide a measure of delay-dependent retention of mental representation. Several rodent tasks of working memory, such as delayed matching or nonmatching to sample (30) and discrete trial delayed alternation (31), contain important elements of human experimental paradigms and are used routinely to understand the cellular basis of working memory.

Schizophrenic patients, like patients with overt frontal lobe damage, display profound deficits in tasks that require behavioral flexibility, strategy shifting, and response to environmental feedback. These tasks include the Wisconsin Card Sorting Test (32), the Category Test (33), and the Tower of Hanoi Task (34), all of which involve an ability to utilize knowledge or feedback to change or shape behavior. Several interesting tasks have been characterized for evaluating behavioral flexibility and strategy-shifting ability in the rat. Of note is a maze-based strategy-shift task described by Ragozzino et al. (35). This task requires rats first to learn either a response strategy (unidirectional turn) or a visually cued place-discrimination strategy (black vs. white maze arm) to obtain a food reward. Following acquisition of the initial strategy, the rats are required to shift to the alternative strategy (e.g., response strategy to visual cue strategy) or to reverse the reward contingency within a strategy (e.g., right turn response to left turn response) to receive the food reward. This task is particularly useful in that it allows the experimenter to assess task acquisition, behavioral flexibility within and between strategies, and perseverative behavior, all of which appear to be relevant to the cognitive deficits associated with schizophrenia.

Attentional deficits are among the hallmarks of the clinical phenomenology of schizophrenia (36 ,37). One of the best characterized rodent attentional tasks is the Five-Choice Serial Reaction Time task, which was designed by Robbins and co-workers (38) based on the human Continuous Performance Test of Attention (39). The basic form of this task requires the rat to detect brief flashes of light occurring in one of five holes and provides a steady-state procedure in which the effect of manipulations can be determined against a baseline of stable performance. One advantage of this task is that it allows for a number of manipulations that test several variables on which patients with schizophrenia show deficits during the Continuous Performance Test, such as perseverative responding and omission errors.

The behavioral tests outlined here, albeit not without limitations, provide important tools for establishing the construct validity of animal models of schizophrenia. Although relatively few studies (e.g., 19,75,87) have utilized these measures to characterize animal models, their use may be important for defining novel pharmacologic approaches for the treatment of cognitive deficits associated with schizophrenia.

Cellular and Molecular Phenotypes

Cellular and molecular markers, identified by morphologic findings from postmortem studies, are being used increasingly to establish the validity of animal models of schizophrenia. Although a review of the schizophrenia postmortem literature is beyond the scope of this chapter, some of the findings relevant to animal models include abnormalities in the neuronal organization of the prefrontal cortex, such as a reduction in cortical volume (e.g., 7,43), altered laminar distribution of neurons that contain the enzyme nicotinamide adenine dinucleotide phosphate diaphorase (44), and reduced neuropil and elevated neuronal density (45). In addition to cytoarchitectural findings, postmortem evidence of abnormalities in cortical neurotransmitter function has also been demonstrated. These abnormalities include alterations in γ-aminobutyric acid (GABA) neurons or other indices related to GABA-mediated neurotransmission (46 ,47 and 48), decreased density of tyrosine hydroxylase immunoreactive axons (49), and changes in gene expression of the NR_{2B} and NR_1 subunits of the NMDA receptor (50 ,51). These markers have been used primarily in validating developmental animal models. Examples include alterations in cortical cell migration and neurogenesis in the gestational malnutrition model (52) and reduced thickness of neocortex and hippocampus after prenatal exposure to the influenza virus (53) or the antimitotic agent methyazoxymethanol (54).

Functional findings from imaging studies performed on patients with schizophrenia also have the potential of helping

to establish the predictive and construct validity of animal models. Of note are studies demonstrating an exaggerated dopamine release in response to amphetamine in patients with schizophrenia (55 ,56). These findings have been suggested to support the validity of the amphetamine sensitization model of schizophrenia (57).

PHARMACOLOGIC MODELS

Part of "50 - Animal Models Relevant to Schizophrenia Disorders "

The most common approach for developing animal models has been to exploit pharmacologic treatments or “drug-induced states” that produce schizophrenia-like symptoms in nonschizophrenic humans. These models generally have some predictive or construct validity and have been instrumental in establishing three of the most prominent theories of schizophrenia: the dopamine hypothesis, serotonin (or serotonin–dopamine) hypothesis, and glutamate hypothesis.

Dopamine-Agonist Models

The most widely studied class of drug-induced models of schizophrenia is based on the behavioral effects of psychostimulant drugs such as amphetamine. Although the models that evolved from this approach have demonstrated considerable predictive validity in terms of pharmacologic isomorphism, current thinking now indicates that the original appearance of face validity was actually somewhat misleading. In recent years, the dopamine hypothesis of schizophrenia has evolved into the narrower hypothesis that the mesolimbic dopamine system, distinct from the nigrostriatal dopamine system, is most relevant to schizophrenia. The nigrostriatal dopamine system is now seen as most relevant to the dyskinetic side effects of antipsychotic treatments. The mesolimbic system appears to mediate the locomotor-activating effects of lower doses of amphetamine, whereas the nigrostriatal system mediates the stereotypies that predominate at higher doses (4). Thus, the stereotypies originally proposed to have the most face validity for the human condition now appear to be more closely linked neurobiologically to phenomena that are considered side effects of the clinical treatments. Because schizophrenic patients are not generally considered to be motorically hyperactive, the amphetamine-induced hyperactivity that is mediated by what is believed to be the most relevant neurobiological substrate has seldom been considered to mimic the human disorder. Note that the failure of the model to have face validity has in no way weakened its utility in neurobiological research, which is based on the etiologic and predictive validity of the model. In fact, virtually any of the behavioral effects of amphetamine in rodents, including either locomotor hyperactivity or stereotypy, have a high degree of pharmacologic isomorphism as models for the efficacy of dopamine-antagonist treatments for schizophrenia (4 ,5); thus, the predictive validity of these models appears limited to dopaminergic treatments.

Both dopamine agonists and antagonists have been studied in rat paradigms used to assess latent inhibition. Most of these paradigms utilize three distinct stages: preexposure, conditioning, and expression. Drugs are typically given during the preexposure or conditioning stages, or both. Hence, different drugs can produce different profiles depending on the stage at which they exert their effect. Such complexity does not arise in the testing of latent inhibition in patients with schizophrenia, in whom the condition being tested is necessarily present at all stages of the test paradigm. In both rats and healthy humans, amphetamine disrupts latent inhibition in a haloperidol-sensitive manner (18). In contrast, direct dopamine agonists, such as apomorphine, do not alter latent inhibition. One of the most interesting aspects of the latent inhibition paradigm is that both typical and atypical antipsychotics can actually improve latent inhibition in rats when testing parameters are adjusted to yield low levels of latent inhibition in control animals. Hence, the latent inhibition paradigm differs from most other models in rats in that it is possible to assess the effects of a putative antipsychotic drug in a latent inhibition paradigm without first having to disrupt performance by the administration of a drug or some other manipulation.

As reviewed in detail elsewhere (14 ,15), the sensorimotor gating deficits assessed by measures of PPI of startle in schizophrenia-spectrum patients are mimicked in rats by the activation of dopamine systems. Thus, drugs such as the direct dopamine agonist apomorphine and the indirect dopamine agonists D-amphetamine and cocaine impair PPI in rodents. As in patients with schizophrenia (58), the apomorphine-induced disruption of PPI in rats is not modality-specific, being seen when acoustic prepulses are used to inhibit either acoustic or tactile startle (59). The D2-receptor family appears to mediate the apomorphine disruption of PPI in rats because it is blocked by D2 antagonists, largely insensitive to D1 antagonists, and reproduced by the D2 agonist quinpirole, but not by the D1 agonist SKF 38393. Within the D2 family of receptors, the D2 subtype appears to be the most relevant to these effects. In comparisons of knockout mice lacking either D2, D3, or D4 dopamine receptors, the effects of amphetamine on PPI were absent only in the D2-subtype knockouts (60). In addition to being blocked by typical antipsychotics, the apomorphine-induced disruption of PPI is reversed by the atypical antipsychotics clozapine, olanzapine, and quetiapine (15 ,61), which lack neuroleptic properties in some behavioral assays. For example, clozapine fails to reverse amphetamine- and apomorphine-induced stereotypy in rats, or apomorphine-induced emesis in dogs (4). The ability of antipsychotics, including atypical antipsychotics, to restore PPI in apomorphine-treated rats strongly correlates with their clinical potency (r = .99) (61). In addition to its sensitivity, the specificity of the PPI model for compounds with antipsychotic efficacy is supported by

the fact that it predicts no such efficacy for a wide variety of other psychiatric drugs. Thus, the PPI paradigm appears to be sensitive to both typical and atypical antipsychotics, but—when used with the dopamine agonist apomorphine—this paradigm clearly fails to make the important distinction between these two classes of antipsychotic agents. Thus, converging evidence indicates the important involvement of dopaminergic systems, acting via D2-family receptors, in the control of PPI. These findings in rats parallel the deficits in PPI observed in schizophrenic patients (58), which are also reported to be corrected by both typical and atypical antipsychotics (62 ,63).

It has been suggested that the increased stereotypy and submissive behavior produced by amphetamine in monkeys may mimic the stereotypy and paranoid ideation in schizophrenic patients. Accordingly, both these behavioral effects can be prevented by pretreatment with the antipsychotics haloperidol and chlorpromazine (21). In contrast, the social isolation induced by amphetamine in monkeys is generally regarded as related to the social withdrawal seen in patients with schizophrenia. As in schizophrenic patients, the animals actively avoid other animals, and these effects cannot be reversed by typical antipsychotic drugs such as haloperidol and chlorpromazine (21). Although few novel antipsychotics have been tested in this model, both clozapine and quetiapine have been shown to reverse amphetamineinduced social isolation in monkeys.

One aspect of the psychostimulant model that has generated considerable interest involves the dosage regimens required for amphetamine or related drugs to produce psychotic-like behavior in psychiatrically healthy humans. Because of the widespread belief that amphetamine-induced psychosis is produced only by repeated exposure to the drug (6 ,64), many preclinical researchers have directed their attention to the behavioral effects of amphetamine that are augmented or sensitized by repeated administration of the drug. A review of the available clinical literature, however, reveals that chronic exposure is not required and that psychotic episodes can be produced by acute administration of amphetamine or related drugs (5). The complex and limited nature of the clinical data seems to have led to mistaken interpretations that have inordinately influenced a large proportion of the basic research in this area. Although it is clear that tolerance to the psychosis-inducing effects of amphetamine does not occur in humans, it is not clear that sensitization is required for these effects. Hence, although an animal model based on the effects of chronic amphetamine could be invalidated if tolerance were observed, the development of sensitization does not provide evidence supporting the relevance of the model to schizophrenia. Indeed, it appears that the animal models having the greatest amount of predictive validity are those based on the effects of the psychostimulant that are evident after acute administration (4 ,5).

Serotonin-Agonist Models

Soon after the initial reports of the behavioral effects of lysergic acid diethylamide (LSD), researchers began to explore the idea that the class of drugs represented by LSD might appropriately be called *psychotomimetics*, or even *psychotogens*. This hypothesis was engendered by the similarities between the effects of LSD on perception and affective lability and the symptoms of the early stages of psychoses such as schizophrenia (65). Recent studies systematically comparing hallucinogen-induced psychotic states with the early stages of psychotic disorders have confirmed substantial overlap in the two syndromes (66). Despite many clear similarities, two major differences prompted the dubious albeit widely accepted conclusion that this class of drugs does not provide a useful model of schizophrenia (67 ,68). First, tolerance was found to develop rapidly to the subjective effects of LSD-like drugs, whereas the symptoms of schizophrenia persist for a lifetime. Second, the hallucinations produced by LSD and related drugs are typically visual rather than auditory, as is characteristic of schizophrenia. These two observations weaken the predictive validity of the hallucinogen model of the syndrome of schizophrenia.

Initial interest in hallucinogens was spurred by the possibility that abnormalities of biochemistry might lead to the endogenous production of such compounds and hence be responsible for some psychotic symptomatology. For example, the transmethylation hypothesis posited that serotonin could provide a substrate for the endogenous production of hallucinogens similar to *N,N*-dimethyltryptamine (DMT) (68). Initially, this etiologically based model was dismissed because of the rapid tolerance associated with traditional hallucinogens such as LSD and mescaline. Nevertheless, recent studies indicate that no tolerance occurs to the subjective effects of DMT in humans (69), which suggests that DMT may differ from other hallucinogens and that this model may still be viable. Indeed, different mechanisms may be involved in the various actions of the different hallucinogenic drugs, as suggested by the lack of cross-tolerance to DMT in human subjects made tolerant to LSD (70). Hence, further studies are warranted to provide the objective evidence needed to evaluate adequately the model of psychosis based on the hypothesis of an endogenous psychotogen.

Furthermore, it remains possible that these drugs may be psychotomimetic and therefore have relevance as models of some aspects of psychotic episodes in humans. Recent suggestions of serotoninergic abnormalities in schizophrenia (71) and of 5-HT_{2A}-receptor contributions to the clinical efficacy of atypical antipsychotics (72) have revitalized interest in this possibility. Hallucinogens are now believed to produce their characteristic subjective effects by acting as 5-HT_{2A} agonists (73). Many of the newer atypical antipsychotic drugs are clearly potent 5-HT_{2A} antagonists (72). With regard to specific abnormalities exhibited by patients

with schizophrenia, evidence indicates that the study of hallucinogen action may provide useful animal models. For example, both schizophrenic and schizotypal patients exhibit deficits in startle habituation (9 ,11 ,12 and 13). Hallucinogenic 5-HT_{2A} agonists such as LSD and mescaline produce similar deficits in startle habituation in rats (59 ,74). Conversely, opposite behavioral effects are produced by 5-HT_{2A} antagonists (74), including some antipsychotics (72). Similarly, the PPI-disruptive effects of hallucinogenic 5-HT_2-receptor agonists are blocked by the selective 5-HT_{2A} antagonist M100907, but not by the dopamine blocker haloperidol (74). Furthermore, in keeping with the similarities between acute psychotic states and the syndrome induced by hallucinogens, latent inhibition is also disrupted by LSD and other serotoninergic hallucinogens (75), as it is in acutely ill schizophrenic patients. These effects can be blocked by the putative antipsychotic M100907 (75). Thus, the effects of hallucinogens on habituation, PPI, and latent inhibition in animals have some predictive validity with regard to both specific abnormalities exhibited by patients in the early stages of schizophrenia and the effects of antipsychotics (74). The construct validity of this model is based on compelling evidence that both the symptoms of schizophrenia and the effects of hallucinogens reflect exaggerated responses to sensory and cognitive stimuli, theoretically resulting from failures in normal filtering or gating processes such as habituation, PPI, or latent inhibition (1 ,3 ,9). Accordingly, 5-HT_{2A} antagonism by itself might be effective in the treatment of certain forms of schizophrenia. Indeed, a rather selective 5-HT_{2A} antagonist, M100907, appears to have efficacy as an antipsychotic in some patients with schizophrenia, despite having negligible affinity for dopamine receptors (76). This finding suggests the possibility of a nondopaminergic mechanism for a treatment of subtypes of schizophrenia and provides important support for the predictive validity of the hallucinogen model of psychosis.

Glutamatergic Models

Dysfunctional glutamate neurotransmission has been implicated in schizophrenia, primarily because noncompetitive antagonists of the NMDA subtype of glutamate receptors, including PCP and ketamine, produce a behavioral syndrome in healthy humans that closely resembles symptoms of schizophrenia and is frequently misdiagnosed as acute schizophrenia (77 ,78). The syndrome includes positive symptoms, such as paranoia, agitation, and auditory hallucinations; negative symptoms, such as apathy, poverty of thought, and social withdrawal; and cognitive deficits, such as impaired attention and working memory. The remarkable similarity of PCP-precipitated behaviors with the diverse array of symptoms associated with schizophrenia has prompted the use of PCP (and its analogue ketamine) in pharmacologic models of schizophrenia in both basic and clinical studies. Notably, whereas psychotic episodes are generally associated with prolonged abuse of amphetamine, a single exposure to PCP or ketamine can produce the cognitive deficits and several symptoms listed above in healthy humans. Thus, acute exposure to these compounds is considered a useful pharmacologic tool for producing some aspects of schizophrenic symptomatology in the laboratory animal.

Several interesting aspects of this model distinguish it from monoamine-based models. For example, the behavioral effects of PCP and related compounds are not, for the most part, mediated by increased dopamine transmission and therefore are not blocked by typical antipsychotics (see below). Similarly, in normal human volunteers, the psychotomimetic effects of ketamine are not blocked by typical antipsychotics, but they are reduced significantly by the prototypal atypical antipsychotic clozapine (79). Therefore, this model may be especially useful for testing the effectiveness of atypical and perhaps even novel antischizophrenia drugs. In fact, the first non-monoaminergic ligands (including a glycine-site agonist and a metabotropic glutamate-receptor agonist), which have recently entered clinical trials, have been based on preclinical PCP models (41 ,80). Another attractive aspect of the NMDA antagonist model is that, unlike the dopamine-based models, it has strong construct validity for studying the cognitive and attentional deficits in schizophrenia. In laboratory animals, NMDA antagonists impair working memory, set shifting, and other cognitive functions that are related to schizophrenia (31). More importantly, in clinical studies, direct comparison of schizophrenic patients with healthy volunteers receiving subanesthetic doses of ketamine have indicated no significant difference in scores for thought disorder between the two groups (81).

In rats and monkeys, noncompetitive NMDA antagonists, including PCP and ketamine, produce a range of behavioral abnormalities that have important relationships to schizophrenic symptomatology. These drugs produce both locomotor hyperactivity and stereotyped behaviors. Although they also increase dopamine neurotransmission in limbic regions (82), their motor-activating effects appear to be dopamine-independent (83). At rather low doses, PCP retards habituation of the startle response without affecting startle reactivity (84), a pattern similar to that seen in parallel studies in schizophrenic patients (9). Also as in schizophrenia, PCP-treated rats exhibit marked deficits in social behavior. Although typical antipsychotics have no reliable effect on the PCP-induced disturbance in social behavior in rats, the atypical antipsychotics clozapine, sertindole, and olanzapine appear to reverse the effects partially (22). In terms of sensorimotor gating measures, PPI is reduced or eliminated in rats by psychotomimetic noncompetitive NMDA antagonists, including PCP, dizocilpine (MK-801), and ketamine (14 ,15). As with apomorphine and as in schizophrenia, both intramodal and cross-modal PPI is sensitive to noncompetitive NMDA antagonists (59). In contrast to

the effects of dopamine agonists on PPI, but in keeping with the results of studies of the subjective effects of ketamine in humans, the PPI-disruptive effects of NMDA antagonists are *not* reversed by typical antipsychotics such as haloperidol or selective D1 or D2 antagonists. Importantly, these effects *are* reversed by the atypical antipsychotics clozapine, olanzapine, quetiapine, and remoxipride (14 ,15). These findings raise the possibility that the PCP-induced disruption of PPI may be a useful model for identifying compounds with atypical antipsychotic potential.

In addition to acute dosing with PCP and ketamine, withdrawal from repeated administration of PCP has been proposed to be a useful model for some aspects of schizophrenia (85). Although this model lacks some of the important characteristics of acute models, such as lack of an effect on PPI, it produces an enduring cognitive impairment that is highly relevant to schizophrenic symptomatology.

DEVELOPMENTAL MODELS

Part of "50 - Animal Models Relevant to Schizophrenia Disorders "

The best-characterized animal model in this class is that proposed by Lipska and Weinberger (86 ,87), which involves neonatal excitotoxic lesions of the ventral hippocampus. These lesions produce postpubertal behavioral disturbances, such as increased spontaneous, amphetamine-induced, and NMDA antagonist-induced locomotion. They also produce potentiated apomorphine-induced stereotypies, disruption of PPI, reduced cataleptic response to haloperidol, impaired working memory, and hypersensitivity to stressful stimuli. Furthermore, this manipulation results in alterations in some cellular and molecular markers that may have relevance to schizophrenia (88), such as reduced expression of the glutamate transporter excitatory amino acid transporter 1 ($EAAT_1$) and glutamic acid decarboxylase (GAD67). To the limited extent that they have been tested, dopamine antagonists, including classic and atypical antipsychotic drugs, ameliorate the behavioral abnormalities produced by neonatal ventral hippocampal lesions. It will be important in the future to examine the predictive power of this model for the identification of antipsychotic drugs more thoroughly with measures that are not sensitive to the effects of antipsychotic drugs in sham-lesioned rats.

Other strategies that have been used to disrupt early cortical development include systemic exposure to L-nitroarginine, a nitric oxide synthase inhibitor that disrupts neuronal maturation (89), or the antimitotic agent methyazoxymethanol (54 ,90). These models produce morphologic changes relevant to schizophrenia, such as altered neurogenesis and reduced cortical volume. They also produce some of the behavioral characteristics associated with schizophrenia, such as stereotypy, cognitive impairments, and deficits in PPI. As yet, the predictive validity of this model in terms of sensitivity to antipsychotic treatments remains to be determined.

Isolation rearing of rats has also been used as a manipulation to generate models related to schizophrenia and models of depression and attention-deficit/hyperactivity disorder (ADHD). In the context of schizophrenia, the focus has been on the disruptions of PPI rather than the locomotor hyperactivity observed in isolation-reared rats. Indeed, comparisons among different strains of rats indicate that both effects are strain-dependent but appear in different strains (91 ,92 and 93). Thus, as with a variety of pharmacologic manipulations, locomotor hyperactivity and deficient PPI are readily dissociable behavioral phenomena, even though both have been used in animal models related to schizophrenia. In several laboratories, isolation-reared rats have been shown to exhibit a neuroleptic-reversible deficiency in PPI in comparison with group-reared controls (91 ,94). This effect of isolation rearing appears to be specific to development; similar isolation of adult rats fails to produce the deficit in PPI observed in isolation-reared rats (95). Furthermore, as in the most common form of schizophrenia, the PPI deficits are not evident before puberty but emerge at about that time (96). Converging evidence for an influence of isolation rearing on gating mechanisms in adulthood stem from the observation that the rat analogue of the P50 sensory gating deficit of schizophrenia is also seen in isolation-reared rats (97). Because these deficits in PPI and P50 gating are not associated with concomitant deficits in latent inhibition (95), which occurs only in acutely ill schizophrenic patients (19), it would appear that the isolation-rearing model is more relevant to chronic than to acute schizophrenia. In addition to reversals by typical antipsychotics (haloperidol, raclopride), reversals of the isolation-induced deficits in PPI by clozapine, risperidone, quetiapine, olanzapine, and the putative antipsychotic M100907 have been observed (94 ,98). Thus, PPI deficits in isolation-reared rats may be a valuable paradigm that—like the apomorphine-induced disruption of PPI—is sensitive, but not specific, in its ability to identify compounds with atypical antipsychotic properties. The potential advantage of the isolation-rearing model, as of other models involving developmental perturbations, is that it does not rely on the administration of a drug or the introduction of an artificial lesion to produce the behavior of interest. When the behavior studied in the model is induced by a drug, such as a dopamine or serotonin agonist, the model is typically effective in identifying the corresponding antagonist. Indeed, most of the animal models of schizophrenia have relied on dopaminergic psychostimulants and have proved to be largely limited to the detection of dopamine antagonists. The major message of the fact that clozapine is effective, even at doses that achieve low levels of dopamine receptor occupancy, is that new treatments can be identified for patients with schizophrenia, and that these novel treatments may not involve dopamine antagonism. The isolation-rearing manipulation presumably produces a deficit in PPI by virtue of a substantial reorganization of neural circuits through the course of development.

Hence, such a model has the potential to identify completely novel antipsychotic treatments simply because it does not require the administration of a drug.

GENETIC MODELS

Part of "50 - Animal Models Relevant to Schizophrenia Disorders "

Genetic contributions to schizophrenia have been clearly established in family studies. Although the focus of considerable research, the application of linkage analyses to schizophrenia has not generally proved successful, perhaps because schizophrenia does not represent a single phenotype. Nevertheless, it remains possible that genetic approaches will lead to etiologically based models

Strain Differences

Genetic factors appear to be critical determinants of both sensory and sensorimotor gating in rats. For example, some inbred strains of mice are deficient in gating of the N40 event-related potential (17), which is the rodent analogue of the P50 gating deficit seen in schizophrenia. Indeed, a linkage between the P50 gating deficit in patients with schizophrenia and a specific chromosomal marker associated with the gene for the α_7 subunit of the nicotinic acetylcholine receptor has been demonstrated in a series of elegant studies (16). The potential power of cross-species studies of specific behavioral abnormalities in psychiatric disorders is exemplified by the parallel between these human linkage studies and the observation that the strain of mice that is most deficient in gating of the N40 event-related potential is also the most deficient in α_7-nicotinic receptors (17). Such parallel investigations in patients and animals provide an exemplar for the application of modern molecular biological techniques to the generation and validation of animal models of psychiatric disorders. However, this genetically related deficit in sensory gating does not extend to studies of sensorimotor gating as measured by PPI of the startle response. Thus, mice in which the α_7-nicotinic receptors have been deleted by genetic engineering exhibit normal levels of PPI (99). Nevertheless, other evidence indicates that PPI is regulated by genetic factors. For example, strain-related differences in the dopaminergic modulation of PPI have been reported (100). More relevant to the recent indications that PPI deficits are evident in family members of schizophrenia patients (101), Ellenbroek et al. (102) utilized pharmacogenetic selective breeding to produce strains of rats that were either sensitive (APO-SUS) or insensitive (APO-UNSUS) to the effects of apomorphine on gnawing behavior. Within either a single generation or after many generations of selective breeding, APO-SUS rats and their offspring exhibited significantly less PPI and latent inhibition than did APO-UNSUS rats. Apparently, the physiologic substrates that regulate behavioral sensitivity to apomorphine (presumably some feature related to dopamine-receptor transduction) are associated with substrates that regulate both PPI and latent inhibition, which are transmitted genetically. In another approach, comparisons of several inbred strains of rats identified some strains that exhibit deficits in PPI (103). Because these strains did not exhibit hearing impairments, the genetically determined deficit in PPI likely represents a deficit in sensorimotor gating processes.

Genetically Modified Animals

Other examples of nonpharmacologically based models relevant to schizophrenia are emerging from the field of molecular biology, in which genetic engineering is being used to generate transgenic and knockout animals. In the absence of established candidate genes, the use of mutant animals in models of schizophrenia has focused on the identification of phenotypic differences in behaviors considered relevant to the clinical disorder (i.e., validity has been sought primarily in the dependent measure rather than in the independent variable). For example, schizophrenia-like deficits in PPI of startle have been observed in specific strains of mice (104) and in “knockout” mice in which specific genes have been deleted (105). The focus of genetic engineering in the mouse is beginning to prompt extensions of pharmacologic studies from the rat to the mouse. Although much more such work is needed, it is already abundantly clear that species differences in pharmacologic effects between mice and rats will complicate the application of some schizophrenia-related rat models to mice. For example, in rats, antipsychotic drugs by themselves have minimal effects on PPI, in contrast to their marked suppression of locomotor activity and ability to improve latent inhibition. Hence, rat PPI models can identify antipsychotic effects only if a drug reverses the effects of a disruption in PPI produced by another drug, a lesion, or a developmental manipulation such as isolation rearing. In mice, however, it appears that antipsychotics improve PPI in mice that have not been manipulated (106). This important difference means that it may be easier to detect antipsychotic effects in mice, but also that it will be much more difficult to demonstrate a reversal of a PPI deficit produced by an experimental manipulation.

In an approach that is distinctly different from the candidate gene approach, genetically modified mice have been used to test specific hypotheses of relevance to animal models of schizophrenia. For example, although most pharmacologic evidence in rat had implicated the D2 subtype of the family of dopamine receptors in the PPI-disruptive effects of dopamine agonists, gene knockout mice proved useful in testing this conclusion more definitively. Ralph et al. (60) tested the effects of amphetamine on PPI in D2-, D3-, and D4-receptor knockout mice and corresponding wild-type mice. Only the mice lacking the D2 subtype of receptor failed to show the normal effect of amphetamine on PPI. Although knockout manipulations are confounded

by developmental adaptations, such a study takes advantage of the specificity that represents the fundamental strength of the knockout technology, as no dose of a drug can ever inactivate all of a given receptor without interacting with other receptors at the same time.

Another model with relevance to the etiology and pathophysiology of schizophrenia involves the $NMDAR_1$ (NR_1) "knock-down" line of mutant mice (107). These animals display exaggerated spontaneous locomotion and stereotypy in addition to deficits in social and sexual interactions. Interestingly, preliminary studies indicate that some of these behavioral abnormalities may be ameliorated with a single dose of haloperidol or clozapine. This is an intriguing model that, with further characterization, may advance our understanding of the long-term effects of congenital NMDA-receptor hypofunction. Nevertheless, its relevance to schizophrenia may be questioned by the fact that no evidence has been found in schizophrenia for abnormalities in genes that express subunits of the NMDA receptor (108). Furthermore, these animals, as would be expected, show no behavioral reaction to the NMDA antagonists PCP and dizocilpine. Schizophrenic patients, on the other hand, exhibit a profound exacerbation of preexisting symptoms after exposure to a single dose of PCP (109) or ketamine (110).

CONCLUSIONS

Part of "50 - Animal Models Relevant to Schizophrenia Disorders "

Establishing the construct, etiologic, and predictive validity of animal models relevant to schizophrenia-related disorders is limited by the paucity of rigorous experimental data derived from clinical studies. The major source of validation remains the ability of established antipsychotic drugs to demonstrate efficacy, measured by broadly defined clinical scales in heterogeneous groups of patients. Specific measures of clinical subtypes, clinical course, and symptom-specific treatment effects that can be translated into relevant animal models are needed to overcome the limitations inherent in relying on global assessments of treatment efficacy for validity assessment. The objective study of such measures in translational research is critical for the eventual identification of new antipsychotic treatments. Such treatments not only could help patients who are resistant to treatment with typical antipsychotics but also could help in the treatment of the negative and cognitive symptoms that do not appear to be treated adequately even by the newest generation of atypical antipsychotics.

ACKNOWLEDGMENTS

Part of "50 - Animal Models Relevant to Schizophrenia Disorders "

This work was supported by National Institutes of Health (NIH) research grants DA02925 (MG), MH42228 (MG), MH52885 (MG), MH48404 (BM), and MH 01616 (BM); the VISN 22 Mental Illness Research, Education, and Clinical Center (MG); and the Veterans Administration National Center for Schizophrenia (BM).

REFERENCES

1. Geyer MA, Markou A. Animal models of psychiatric disorders. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:787–798.

2. Matthysse S. Animal models in psychiatric research. *Prog Brain Res* 1986;65:259–270.

3. Segal DS, Geyer MA. Animal models of psychopathology. In: Judd LL, Groves PM, eds. *Psychobiological foundations of clinical psychiatry*. Philadelphia: JB Lippincott Co, 1985:1–14.

4. Creese I. *Stimulants: neurochemical, behavioral, and clinical perspectives*. New York: Raven Press, 1983.

5. Segal DS, Geyer MA, Schuckit A. Stimulant-induced psychosis: an evaluation of animal models. In: Youdim MBH, Lovenberg W, Sharman DF, et al., eds. *Essays in neurochemistry and neuropharmacology*. New York: John Wiley and Sons, 1981:95–130.

6. Snyder S. Amphetamine psychosis: a "model" schizophrenia mediated by catecholamines. *Am J Psychiatry* 1973;130:61–67.

7. Ellenbroek B, Cools A. Animal models with construct validity for schizophrenia. *Behav Pharmacol* 1990;1:469–490.

8. Braff DL, Geyer MA. Sensorimotor gating and schizophrenia: human and animal model studies. *Arch Gen Psychiatry* 1990;47:181–188.

9. Geyer MA, Braff DL. Startle habituation and sensorimotor gating in schizophrenia and related animal models. *Schizophr Bull* 1987;13:643–668.

10. McGhee A, Chapman J. Disorders of attention and perception in early schizophrenia. *Br J Med Psychol* 1961;34:103–116.

11. Bolino F, Di Michele V, Di Cicco L, et al. Sensorimotor gating and habituation evoked by electrocutaneous stimulation in schizophrenia. *Biol Psychiatry* 1994;36:670–679.

12. Taiminen T, Jaaskelainen S, Ilonen T, et al. Habituation of the blink reflex in first-episode schizophrenia, psychotic depression and non-psychotic depression. *Schizophr Res* 2000;44:69–79.

13. Cadenhead K, Geyer M, Braff D. Impaired startle prepulse inhibition and habituation in schizotypal patients. *Am J Psychiatry* 1993;150:1862–1867.

14. Geyer MA, Braff DL, Swerdlow NR. Startle-response measures of information processing in animals: relevance to schizophrenia. In: Haug M, Whalen RE, eds. *Animal models of human emotion and cognition*. Washington, DC: APA Books, 1999:103–116.

15. Swerdlow NR, Geyer MA. Using an animal model of deficient sensorimotor gating to study the pathophysiology and new treatments of schizophrenia. *Schizophr Bull* 1998;24:285–301.

16. Freedman R, Adler LE, Bickford P, et al. Schizophrenia and nicotinic receptors. *Harvard Rev Psychiatry* 1994;2:179–192.

17. Stevens KE, Freedman R, Collins AC, et al. Genetic correlation of hippocampal auditory evoked response and alpha-bungarotoxin binding in inbred mouse strains. *Neuropsychopharmacoly* 1996;15:152–162.

18. Weiner I, Lubow RE, Feldon J. Disruption of latent inhibition by acute administration of low doses of amphetamine. *Pharmacol Biochem Behav* 1988;30:871–878.

19. Gray NS, Pilowsky LS, Gray JA, et al. Latent inhibition in drug-naïve schizophrenics: relationship to duration of illness and dopamine D2 binding using SPET. *Schizophr Res* 1995;17:95–107.

20. Swerdlow NR, Braff DL, Hartston H, et al. Latent inhibition in schizophrenia. *Schizophr Res* 1996;20:91–103.

21. Ellenbroek BA, Cools AR. Animal models of psychotic disturbances. In: den Beor JA, Westenberg HGM, van Praag HM, eds. *Advances in the neurobiology of schizophrenia*. Chichester: John Wiley and Sons, 1995:89–109.

22. Sams-Dodd R. Phencyclidine-induced stereotyped behaviour and social isolation in the rat: a possible animal model of schizophrenia. *Behav Pharmacol* 1996;7:3–23.

23. Green M. What are the functional consequences of neurocognitive deficits in schizophrenia? *Am J Psychiatry* 1996;153:321–330.

24. Goldberg TE, Gold JM, Braff DL. Neuropsychological functioning and time-linked information processing is schizophrenia. *Rev Psychiatry* 1991;10:60–78.

25. Hoff AL, Faustman WO, Wieneke M, et al. The effects of clozapine on symptom reduction, neurocognitive function, and clinical management in treatment-refractory state hospital schizophrenic inpatients. *Neuropsychopharmacology* 1996;15:361–369.

26. Keefe RS, Silva SG, Perkins DO, et al. The effects of atypical antipsychotic drugs on neurocognitive impairment in schizophrenia: a review and meta-analysis. *Schizophr Bull* 1999;25:201–222.

27. Meltzer H, McGurk S. The effects of clozapine, risperidone, and olanzapine on cognitive function in schizophrenia. *Schizophr Bull* 1999;25:233–255.

28. Gevins AS, Bressler S, Cutillo B, et al. Effects of prolonged mental work on functional brain topography. *Electroencephalogr Clin Neurophysiol* 1990;76:339–350.

29. Goldman-Rakic P. Cerebral cortical mechanisms in schizophrenia. *Neuropsychopharmacology* 1994;10:22S–27S.

30. Kesner RP, Hunt ME, Williams JM, et al. Prefrontal cortex and working memory for spatial response, spatial location and visual object information in the rat. *Cereb Cortex* 1996;6:311–318.

31. Aultman JM, Moghaddam B. Distinct contributions of glutamate and dopamine receptors to temporal aspects of rodent working memory using a clinically relevant task. *Psychopharmacology (Berl)* 2001;153:353–364.

32. Robinson AL, Heaton RK. The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. *J Consult Clin Psychol* 1980;48:605–614.

33. Selz M, Reitan RM. Rules for neuropsychological diagnosis: classification of brain function in older children. *J Consult Clin Psychol* 1979;47:258–264.

34. Goldberg TE, Saint-Cyr JA, Weinberger DR. Assessment of procedural learning and problem solving in schizophrenic patients by Tower of Hanoi type tasks. *J Neuropsychiatry Clin Neurosci* 1990;2:165–173.

35. Ragozzino M, Wilcox C, Raso M, et al. Involvement of rodent prefrontal cortex subregions in strategy switching. *Behav Neurosci* 1999;113:32–41.

36. Harvey P, Lenzenweger M, Keefe R, et al. Empirical assessment of the factorial structure of clinical symptoms in schizophrenic patients: formal thought disorder. *Psychiatry Res* 1992;44:141–151.

37. Nestor PG, O'Donnell BF, Niznikiewicz MA, et al. Neuromodulation of attention in schizophrenia. *Psychiatr Ann* 1999;29:633–640.

38. Robbins T, Everitt B, Marston H, et al. Comparative effects of ibotenic acid- and quisqualic acid-induced lesions of the substantia innominata on attentional function in the rat: further implications for the role of the cholinergic neurons of the nucleus basalis in cognitive processes. *Behav Brain Res* 1989;35:221–240.

39. Mirsky A, Rosvold H. The use of psychoactive drugs as a neuropsychological tool in studies of attention in man. In: Uhr I, Miller J, eds. *Drugs and behaviour*. New York: John Wiley and Sons, 1960:375–392.

40. Castner SA, Goldman-Rakic PS. Long-lasting psychotomimetic consequences of repeated low-dose amphetamine exposure in rhesus monkeys. *Neuropsychopharmacology* 1999;20:10–28.

41. Moghaddam B, Adams B. Reversal of phencyclidine effects by a group II metabotropic glutamate receptor agonist in rats. *Science* 1998;281:1349–1352.

42. Robbins TW. Arousal and attention: psychopharmacological and neuropsychological studies in experimental animals. In: Parasuraman R, ed. *The attentive brain*. Cambridge, MA: MIT Press, 1998.

43. Andreasen N, Nasrallah H, Dunn V, et al. Structural abnormalities in the frontal system in schizophrenia: a magnetic resonance imaging study. *Arch Gen Psychiatry* 1986;43:136–144.

44. Akbarian S, Bunney WE Jr, Potkin SG, et al. Altered distribution of nicotinamide-adenine dinucleotide phosphate-diaphorase cells in frontal lobe of schizophrenics implies disturbances of cortical development. *Arch Gen Psychiatry* 1993;50:169–177.

45. Selemon LD, Goldman-Rakic PS. The reduced neuropil hypothesis: a circuit-based model of schizophrenia. *Biol Psychiatry* 1999;45:17–25.

46. Benes FM, McSparren J, Bird ED, et al. Deficits in small interneurons in prefrontal and cingulate cortices of schizophrenic and schizoaffective patients. *Arch Gen Psychiatry* 1991;48:996–1001.

47. Akbarian S, Huntsman MM, Kim JJ, et al. GABAa receptor subunit gene expression in human prefrontal cortex: comparison of schizophrenics and controls. *Cereb Cortex* 1995;5:550–560.

48. Volk DW, Austin MC, Pierri JN, et al. Decreased glutamic acid decarboxylase67 messenger RNA expression in a subset of prefrontal cortical gamma-aminobutyric acid neurons in subjects with schizophrenia. *Arch Gen Psychiatry* 2000;57:237–245.

49. Akil M, Edgar CL, Pierri JN, Casali S, et al. Decreased density of tyrosine hydroxylase-immunoreactive axons in the entorhinal cortex of schizophrenic subjects. *Biol Psychiatry* 2000;47:361–370.

50. Akbarian S, Sucher NJ, Bradley D, et al. Selective alterations in gene expression for NMDA receptor subunits in prefrontal cortex of schizophrenics. *J Neurosci* 1996;16:19–30.

51. Gao XM, Sakai K, Roberts RC, et al. Ionotropic glutamate receptors and expression of *N*-methyl-D-aspartate receptor subunits in subregions of human hippocampus: effects of schizophrenia. *Am J Psychiatry* 2000;157:1141–1149.

52. Cintra L, Granados L, Aguilar A, et al. Effects of prenatal protein malnutrition on mossy fibers of the hippocampal formation in rats of four age groups. *Hippocampus* 1997;7:184–191.

53. Fatemi SH, Emamian ES, Kist D, et al. Defective corticogenesis and reduction in Reelin immunoreactivity in cortex and hippocampus of prenatally infected neonatal mice. *Mol Psychiatry* 1999;4:145–154.

54. Talamini LM, Koch T, Ter Horst GJ, et al. Methylazoxymethanol acetate-induced abnormalities in the entorhinal cortex of the rat: parallels with morphological findings in schizophrenia. *Brain Res* 1998;789:293–306.

55. Laruelle M, Abi-Dargham A, van Dyck C, et al. SPECT imaging of amphetamine-induced dopamine release in drug-free schizophrenic subjects. *Proc Natl Acad Sci U S A* 1996;93:9235–9340.

56. Breier A, Su T, Saunders R, et al. Schizophrenia is associated with elevated amphetamine-induced synaptic dopamine concentrations: evidence from a novel positron emission tomography method. *Proc Natl Acad Sci U S A* 1997;94:2569–2574.

57. Lieberman JA, Sheitman BB, Kinon BJ. Neurochemical sensitization in the pathophysiology of schizophrenia: deficits and dysfunction in neuronal regulation and plasticity. *Neuropsychopharmacology* 1997;17:205–229.

58. Braff D, Grillon C, Geyer M. Gating and habituation of the startle reflex in schizophrenic patients. *Arch Gen Psychiatry* 1992;49:206–215.

59. Geyer MA, Swerdlow NR, Mansbach RS, et al. Startle response models of sensorimotor gating and habituation deficits in schizophrenia. *Brain Res Bull* 1990;25:485–498.

60. Ralph RJ, Varty GB, Kelly MA, et al. The dopamine D2 but not D3 or D4 receptor subtype is essential for the disruption of prepulse inhibition produced by amphetamine in mice. *J Neurosci* 1999;19:4627–4633.

61. Swerdlow NR, Braff DL, Taaid N, et al. Assessing the validity of animal model of deficient sensorimotor gating in schizophrenic patients. *Arch Gen Psychiatry* 1994;51:139–154.

62. Weike AI, Bauer U, Hamm AO. Effective neuroleptic medication removes prepulse inhibition deficits in schizophrenia patients. *Biol Psychiatry* 2000;47:61–70.

63. Kumari V, Soni W, Sharma T. Normalization of information processing deficits in schizophrenia with clozapine. *Am J Psychiatry* 1999;156:1046–1051.

64. Ellinwood E, Sudilovsky A, Nelson L. Evolving behavior in the clinical and experimental amphetamine (model) psychoses. *Am J Psychiatry* 1973;130:1088.

65. Bowers M, Freedman D. “Psychedelic” experiences in acute psychosis. *Arch Gen Psychiatry* 1966;15.

66. Gouzoulis-Mayfrank E, Habermeyer E, Hermle L, et al. Hallucinogenic drug-induced states resemble acute endogenous psychoses: results of an empirical study. *Eur Psychiatry* 1998;13:399–406.

67. Hollister LE. *Chemical psychoses: LSD and related drugs*. Springfield, IL: Charles C Thomas Publisher, 1968.

68. Weil-Malherbe H, Szara SI. *The biochemistry of functional and experimental psychoses*. Springfield, IL: Charles C Thomas Publisher, 1971.

69. Strassman RJ, Qualls CR, Berg LM. Differential tolerance to biological and subjective effects of four closely related spaced doses of *N,N*-dimethyltryptamine in humans. *Biol Psychiatry* 1996;39:784–795.

70. Rosenberg DE, Isbell H, Miner EJ, et al. The effect of *N,N*-dimethyltryptamine in human subjects tolerant to lysergic acid diethylamide. *Psychopharmacologia* 1964;5:217–227.

71. Joyce JN, Shane A, Lexow N, et al. Serotonin uptake sites and serotonin receptors are altered on the limbic system of schizophrenics. *Neuropsychopharmacology* 1993;8:315–336.

72. Meltzer HY, Matsubara S, Lee JC. Classification of typical and atypical antipsychotic drugs on the basis of dopamine D-1, D-2 and serotonin-2 pKi values. *J Pharmacol Exp Ther* 1989;251:238–246.

73. Glennon RA, Titeler M, McKenney JD. Evidence for 5-HT_2 involvement in the mechanism of action of hallucinogenic agents. *Life Sci* 1984;24:2505–2511.

74. Geyer MA. Behavioral studies of hallucinogenic drugs in animals: implications for schizophrenia research. *Pharmacopsychiatry* 1998;2:73–79.

75. Hitchcock JM, Lister S, Fischer TR, et al. Disruption of latent inhibition in the rat by the 5-HT_2 agonist DOI: effects of MDL 100,907, clozapine, risperidone and haloperidol. *Behav Brain Res* 1997;88:43–49.

76. Schmidt CJ, Kehne JH, Carr AA. M100,907: a selective 5-HT_{2A} receptor antagonist for the treatment of schizophrenia. *CNS Drug Rev* 1997;3:49–67.

77. Javitt DC, Zukin SR. Recent advances in the phencyclidine model of schizophrenia. *Am J Psychiatry* 1991;148:1301–1308.

78. Krystal JH, Karper LP, Seibyl JP, et al. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans: psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry* 1994;51:199–214.

79. Malhotra AK, Adler CM, Kennison SD, et al. Clozapine blunts *N*-methyl-D-aspartate antagonist-induced psychosis: a study with ketamine. *Biol Psychiatry* 1997;42:664–668.

80. Javitt DC, Balla A, Sershen H, et al. A. E. Bennett Research Award. Reversal of phencyclidine-induced effects by glycine and glycine transport inhibitors. *Biol Psychiatry* 1999;45:668–679.

81. Adler C, Malhotra A, Elman I, et al. Comparison of ketamine-induced thought disorder in healthy volunteers and thought disorder in schizophrenia. *Am J Psychiatry* 1999;156:1646–1649.

82. Mathe JM, Nomikos CG, Schilstrom B, et al. Non-NMDA excitatory amino acid receptors in the ventral-tegmental area mediate systemic dizocilpine (MK-801)-induced hyperlocomotion and dopamine release in the nucleus accumbens. *J Neurosci Res* 1998;51:583–592.

83. Adams B, Moghaddam B. Corticolimbic dopamine neurotransmission is temporally dissociated from the cognitive and locomotor effects of phencyclidine. *J Neurosci* 1998;18:5545–5554.

84. Geyer MA, Segal DS, Greenberg BD. Increased startle responding in rats treated with phencyclidine. *Neurobehav Toxicol Teratol* 1984;6:1–4.

85. Jentsch JD, Redmond DE Jr, Elsworth JD, et al. Enduring cognitive deficits and cortical dopamine dysfunction in monkeys after long-term administration of phencyclidine. *Science* 1997;277:953–955.

86. Lipska BK, Weinberger DR. Delayed effects of neonatal hippocampal damage on haloperidol-induced catalepsy and apomorphine-induced stereotypic behaviors in the rat. *Brain Res Dev Brain Res* 1993;75:213–222.

87. Lipska BK, Weinberger DR. Genetic variation in vulnerability to the behavioral effects of neonatal hippocampal damage in rats. *Proc Natl Acad Sci U S A* 1995;92:8906–8910.

88. Lipska BK, Weinberger DR. To model a psychiatric disorder in animals: schizophrenia as a reality test. *Neuropsychopharmacology* 2000;23:223–239.

89. Black MD, Selk DE, Hitchcock JM, et al. On the effect of neonatal nitric oxide synthase inhibition in rats: a potential neurodevelopmental model of schizophrenia. *Neuropharmacology* 1999;38:1299–1306.

90. Grace AA, Moore H. Regulation of the information flow in the nucleus accumbens: a model for the pathophysiology of schizophrenia. In: Lenzenweger MF, Dworkin RH, eds. *Origins and development of schizophrenia: advances in experimental psychopathology*. Washington, DC: American Psychological Association, 1998:123–160.

91. Geyer MA, Wilkinson LS, Humby T, et al. Isolation rearing of rats produces a deficit in prepulse inhibition of acoustic startle similar to that in schizophrenia. *Biol Psychiatry* 1993;34:361–372.

92. Paulus MP, Bakshi VP, Geyer MA. Isolation rearing affects sequential organization of motor behavior in post-pubertal but not pre-pubertal Lister and Sprague-Dawley rats. *Behav Brain Res* 1998;94:271–280.

93. Varty GB, Geyer MA. Effects of isolation rearing on startle reactivity, habituation, and prepulse inhibition of male Lewis, Sprague-Dawley, and Fischer F344 rats. *Behav Neurosci* 1998;112:1450–1457.

94. Varty GB, Higgins GA. Examination of drug-induced and isolation-induced disruptions of prepulse inhibition as models to screen antipsychotic drugs. *Psychopharmacology* 1995;122:15–26.

95. Wilkinson LS, Killcross AS, Humby T, et al. Social isolation produces developmentally specific deficits in prepulse inhibition of the acoustic startle response but does not disrupt latent inhibition. *Neuropsychopharmacology* 1994;10:61–72.

96. Bakshi V, Geyer M. Ontogeny of isolation rearing-induced deficits in sensorimotor gating in rats. *Physiol Behav* 1999;67:385–392.

97. Stevens KE, Johnson RG, Rose GM. Rats reared in social isolation show schizophrenia-like changes in auditory gating. *Pharmacol Biochem Behav* 1997;58:1031–1036.

98. Geyer MA, Krebs-Thomson K, Varty GB. The effects of M100907 on pharmacological and developmental animal models of prepulse inhibition deficits in schizophrenia. *Neuropsychopharmacology* 1999;21:S134–S142.

99. Paylor R, Nguyen M, Crawley JN, et al. Alpha7 nicotinic receptor subunits are not necessary for hippocampal-dependent learning or sensorimotor gating: a behavioral characterization of Acra7-deficient mice. *Learning Memory* 1998;5:302–316.

100. Swerdlow NR, Martinez ZA, Hanlon F, et al. Towards understanding the biology of a complex phenotype: rat strain and substrain differences in the sensorimotor gating-disruptive effects of dopamine agonists. *J Neurosci* 2000;20:4325–4336.

101. Cadenhead KS, Swerdlow NR, Shafer KM, et al. Modulation of the startle response and startle laterality in relatives of schizophrenic patients and schizotypal personality disordered subjects: evidence of inhibitory deficits. *Am J Psychiatry* 2000;157:1660–1668.

102. Ellenbroek BA, Geyer MA, Cools AR. The behavior of APO-SUS rats in animal models with construct validity for schizophrenia. *J Neurosci* 1995;15:7604–7611.

103. Palmer AA, Dulawa SC, Mottiwala AA, et al. Prepulse startle deficit in the Brown Norway rat: a potential genetic model. *Behav Neurosci* 2000;114:374–388.

104. Paylor R, Crawley JN. Inbred strain differences in prepulse inhibition of the mouse startle response. *Psychopharmacology* 1997;132:169–180.

105. Lijam N, Paylor R, McDonald MP, et al. Social interaction and sensorimotor gating abnormalities in mice lacking Dvl1. *Cell* 1997;90:895–905.

106. McCaughran JJ, Mahjubi E, Decena E, et al. Genetics, haloperidol-induced catalepsy and haloperidol-induced changes in acoustic startle and prepulse inhibition. *Psychopharmacology* 1997;134:131–139.

107. Mohn AR, Gainetdinov RR, Caron MG, et al. Mice with reduced NMDA receptor expression display behaviors related to schizophrenia. *Cell* 1999;98:427–436.

108. Nishiguchi N, Shirakawa O, Ono H, et al. Novel polymorphism in the gene region encoding the carboxyl-terminal intracellular domain of the NMDA receptor 2B subunit: analysis of association with schizophrenia. *Am J Psychiatry* 2000;157:1329–1331.

109. Luby E, Cohen B, Rosenbaum G, et al. Study of a new schizophrenomimetic drug—sernyl. *Am Med Assoc Arch Neurol Psychiatry* 1959;81:363–369.

110. Lahti AC, Koffel B, LaPorte D, et al. Subanesthetic doses of ketamine stimulate psychosis in schizophrenia. *Neuropsychopharmacology* 1995;13:9–19.

51

Endophenotypes in Studies of the Genetics of Schizophrenia

David L. Braff

Robert Freedman

David L. Braff: Department of Psychiatry, University of California at San Diego, La Jolla, California.

Robert Freedman: Department of Psychiatry and Pharmacology, University of Colorado, Denver, Colorado.

The power and appeal of the molecular biology mantra, “DNA to RNA to protein,” to explicate cell biology comes from its universal appearance and application in all species, from microorganisms to human beings. Based on this mantra, the genomes of viruses, bacteria, fruit flies, and now humans are being mapped and sequenced, so that all the genes and, ultimately, their corresponding biological activity can be identified. As these genes are identified, it is reasonable to ask how this information can be related to the inheritance of risk for psychiatric illness. For a bacterial enzyme, genetic coding of the amino acid sequence of proteins can be closely associated with a functional change in enzymatic activity. For a complex psychiatric illness, as defined by DSM-IV criteria, the relationship is obviously not as straightforward. Psychiatric illnesses such as schizophrenia are generally conceptualized as multifactorial and most likely reflect the combined influence and interactions of both genetic and nongenetic factors. Furthermore, there is no reason to presuppose that only one gene is responsible for a complex psychiatric disorder such as schizophrenia, as there is in some simple mendelian illnesses. Persons who are ill may differ in more than one gene from the rest of the population, and different sets of genes may be associated with illness in different populations. Thus, how best to use the power of molecular genetics to understand the inheritance and pathophysiology of complex genetic psychiatric illnesses remains an enigma that is only now beginning to be solved.

In the simplest and most commonly used strategy of molecular genetics that is applied to complex psychiatric disorders, it is assumed that the distribution of illness in a family represents the effect of a single gene, and techniques of genetic analysis are used to identify that gene. This approach does not necessarily overlook the complexity of psychiatric illness, but it assumes that the effect (i.e., signal) of a single gene will be discerned in a complex, “noisy” genetic background if samples sizes are large enough or if the population is sufficiently homogeneous (e.g., 1). An attractive feature of this approach is that the search for genes is not constrained by preexisting hypotheses about the biology of the illness, which, in the case of schizophrenia, is still unclear. A second commonly applied strategy is, in fact, the opposite approach; an assumption is made about the biology of the illness and then candidate genes associated with that biology are examined to determine if they are mutated. Both approaches have been successful to a limited extent for explicating the genetics of schizophrenia. Replicable linkages for schizophrenia have been obtained at several locations (e.g., chromosomes 1, 6, 8, 13, 15, and 22), but genetic mutations have not as yet been identified at these sites (2). On the other hand, DNA mutations have been found in candidate genes such as *NURR1*, the gene for the receptor for retinoic acid, a pathway critical in neuronal development, but these mutations seem to be found in only a small proportion of schizophrenic patients (3).

This chapter describes a third approach, which attempts to make use of the power of the molecular biology mantra by identifying brain dysfunctions that may be caused by a single genetic abnormality. The rationale comes from the mantra itself. If discrete genetic abnormalities are associated with schizophrenia, then each of them should cause a specific protein change that is reflected in a corresponding discrete functional abnormality. Even if several genes are abnormal, along with additional environmental factors, the functional abnormality resulting from each gene should generally be identifiable. Theoretically, the relationship between these functional abnormalities and genes, discovered either by genetic linkage or by candidate gene analysis, should be stronger than the association to the illness itself because the illness itself results from a mixture of genetic and nongenetic abnormalities that may vary between different individuals and families. As is true for the other approaches described above, this approach has not yet led to the identification

of the genes that are associated with and that may even cause schizophrenia in most cases. Nevertheless, the strategy has been useful for gene discovery in other complex illnesses, such as colon cancer and hemochromatosis. In colon cancer, the formation of multiple polyps, rather than cancer itself, has been found to be the genetically heritable trait (4), and in hemochromatosis, a high serum level of iron, rather than the clinically recognized illness, has been found to be the more penetrant heritable trait (5).

Endophenotype is often used as the descriptive term for these discrete, genetically determined phenotypes that may be part of a complex illness. The search for endophenotypes is not straightforward because no *a priori* criterion can be used to decide if a particular element of schizophrenia or any other psychiatric illness reflects the effect of a single gene. Putative endophenotypes have ranged from clinical characterizations, such as the presence of schizotypy in relatives of schizophrenic patients (6), to the neurophysiologic and neuropsychological measures described in this chapter, to structural measures of specific, functionally important regions of the brain and ventricular size. Because none of these phenotypes has yet led to the identification of a specific molecular deficit, it has not been proved that any one of them is actually linked to a specific genetic abnormality. In this context, even if endophenotypes turn out to be multiple, rather than single, gene phenomena, their genetic architecture, even as complex endophenotypes, may turn out to be simpler than schizophrenia in certain families. The sections below outline the stage of investigation for a number of putative phenotypes, from presence in schizophrenia probands and their relatives to statistically significant genetic linkage to a chromosomal locus.
Several points must be considered in the assessment of endophenotypes. First, because these are putative genetic traits, their biology begins at conception, so that by the time they are measured in adulthood, their expression may have been modified by such factors as development, aging, brain injury, and medication and substance abuse and. Second, most genes expressed in the brain are expressed in many different brain areas, so that their ultimate functional expression may involve much more than the simple phenotype being measured. Third, many genes expressed in the brain are also involved in the development of neurons, so that their most important functional effects may have occurred prenatally. Fourth, according to Mendel's second law, every genetic trait segregates independently in a family, so that if schizophrenia is a multifactorial trait, some siblings should express specific phenotypes independently of other phenotypes. These siblings may be better subjects for characterizing the phenotype than the patients themselves, whose multiple deficits may obscure the unique phenotype. Finally, because the aim of genetics generally is to identify affected individuals who have or do not have a particular genetic abnormality, the measurement of the putative phenotype must clearly separate most affected and unaffected individuals, regardless of whether a quantitative or discrete variable is used. The range of effect sizes for several putative endophenotypes is shown in Table 51.1 , which reflects another point. The measurement of endophenotypes is in itself a complex endeavor in which modest-appearing paradigmatic manipulations lead to significant shifts in the signal of the dependent measure being assessed.

TABLE 51.1. ENDOPHENOTYPES AND THE GENETICS OF SCHIZOPHRENIA: EFFECT SIZE DIFFERENCES BETWEEN SCHIZOPHRENIA SPECTRUM GROUPS AND NORMAL COMPARISON SUBJECTS

Phenotypes	Schizophrenia Patients	Clinically Unaffected Relatives of Schizophrenia Patients	Schizotypal Personality Disorder Patients	References
P50 Suppression	0.92–1.29	0.79	0.79	30, 31, 203
Prepulse Inhibition	0.51–0.85	1.0	1.45	46, 47, 54
Smooth-Pursuit Eye Movement	2.0–3.0	0.29–1.3	0.29	92, 93, 204
Antisaccade	4.88–6.38	1.38–3.75	0.75–1.36	101, 204
Executive Functioning (Wisconsin Card Sorting Test)	0.47–1.97	0.73–1.6	0.72	118, 126, 127, 132
Working Memory (Letter–Number Span)	1.42–2.2	0.42	0.73–1.04	126, 132, 205
Thought Disorder (Thought Disorder Index; Ego Impairment Index)	1.56–2.98	0.34–0.83	1.04–1.28	204, 206–208
Continuous Performance Test	0.45–3.30	0.46–2.97	0.45–0.78	118, 139, 140, 147
Span of Apprehension	0.5–2.5	0.6–1.5		139, 152, 209–212
Visual Backward Masking	0.33–0.65	0.43–0.57	0.45–0.67	168, 173, 213, 214
P300 Event–Related Potential	0.45–1.05	0.17	0.36	192, 215
Reaction Time	0.59–1.05	0.44	0.79–0.99	168, 216, 217

Effect sizes in schizophrenia patients, clinically unaffected relatives of schizophrenia patients, and schizotypal personality disorder patients compared with those in normal subjects. These effect sizes were computed by using the means and standard deviations for the normal comparison subjects and the means of the patient groups. The range of values differs from study to study because different investigators used different patient populations taking different types and amounts of medications; also, the experimental paradigms, although similar, often differed in terms of stimulus parameters. Of course, in some of these studies, multiple conditions were used, some of which were needed to establish floor and ceiling effects. In these cases, we generally cite the most robust effect sizes.

The search for endophenotypes takes advantage of genetic strategies to evaluate the current state of understanding the pathophysiology of schizophrenia. The initial endophenotype was schizotypy, which was proposed to be a pure expression of schizotaxia, the genetic predisposition for schizophrenia. Schizotypy itself does not generally show mendelian segregation, so that the likelihood that it reflects a single genetic trait is now considered small. However, the presence of schizotypy in family members has been related to linkage of schizophrenia at a specific chromosomal locus in a subset of families, so that a reexamination of schizotypy as an endophenotype in some families may once again be productive. Inhibitory interneurons have increasingly become a focus of interest in the biology of schizophrenia. Many of the endophenotypes described below are attempts to demonstrate inhibitory neuronal function by means of psychophysiologic and neurophysiologic techniques. Structural phenotypes have been limited to the measurement of brain volume. As functional brain-imaging techniques become more advanced, so that specific neuronal functions can be demonstrated, it is likely that these techniques will also be used. Magnetic resonance spectroscopy of the amino acids associated with neuronal function, such as *N*-acetylasparate, is an example (7).

CANDIDATE ENDOPHENOTYPES FOR GENETIC STUDIES

Part of "51 - Endophenotypes in Studies of the Genetics of Schizophrenia "

Given the considerations discussed above, it is important to stress that although the DSM-IV diagnostic criteria for schizophrenia may be clinically and administratively useful, they are not likely to be optimally useful as phenotypes in genetic studies (8 ,9 ,10 and 11). The search for and use of new, nondiagnostic, non–DSM-IV-based candidate endophenotypes parallels our search for the corresponding candidate genes in complex human genetic disorders such as schizophrenia. Also, we understand that in accounting for the genetic diathesis or vulnerability to schizophrenia, we are "accounting" for, at most, 50% to 70% of the variance of the disorder; the remaining variability resides in nongenetic "second hits," such as neonatal or *in utero* neural damage to the developing hippocampus (12 ,13 ,14 and 15) or other factors. A plethora of studies indicate that in addition to mutant genes, a second level of environmental or other generalized or specific stressor probably must act as a second hit in the central nervous system. An example of the result of this need for a second hit is illustrated by the fact that clinically "unaffected" relatives of patients with schizophrenia have endophenotypic markers of abnormalities in some or all of the measures listed below but do not have the disorder of schizophrenia. Therefore, it appears clear that some nongenetic contributions (not necessarily reflected by these endophenotypes) are crucially important in the expression of some forms of this elusively heterogeneous and complex disorder of schizophrenia. In searching for non–diagnosis-based "candidate endophenotypes," we are not alone in schizophrenia research because many disorders, from diabetes to hypertension to bipolar disorder, also present the same conundra and difficult conceptual issues.

In schizophrenia research, it seems reasonable to classify candidate endophenotypes into structural and functional abnormalities. Because of limits of chapter length, we do not discuss structural endophenotypes (e.g., widely dispersed, decreased, generalized gray matter; decreased superior temporal gyrus volume; deficits of hippocampal or temporal lobe volume (16); gray matter volume or organizational abnormalities in various subsections of the prefrontal cortex, most significantly the dorsolateral prefrontal cortex) that may be useful in genetic studies (17 ,18). It is important to note that such structural abnormalities may be correlated with some of the functional abnormalities discussed below. In Table 51.1 , some functional endophenotypes are listed, along with estimates of the effect sizes of deficits of each in schizophrenic patients, clinically unaffected relatives of schizophrenic patients, and schizotypal patients in comparison with normal subjects. In addition, reasonable and well-understood neural substrates for these measures are known, as discussed. Table 51.1 summarizes what we have selected as important and representative (but not all-inclusive) candidate endophenotypes for genetic studies in schizophrenia, with an emphasis on the information-processing abnormalities that have assumed a central role in the search for candidate endophenotypes (10 ,11). Identifying these endophenotypes for genetic studies is only a first step; after they have been identified, complex strategies must be employed, as discussed above and below, to conduct linkage, association, and other genetic studies on the candidate endophenotypes so that we can identify the candidate genes that are likely contributing to the endophenotypes (and their neural substrates) present in the complex disorder of schizophrenia. In the next section, we describe possible functional endophenotypes, with an emphasis on gating and oculomotor abnormalities.

Abnormalities of Sensorimotor Gating

The concept of deficits of sensory gating in schizophrenia derives from the clinical observation that patients report failures of information processing characterized by poor sensory gating—an inability to screen out trivial stimuli and focus on salient aspects of the environment so as to process information smoothly and successfully navigate through life (10 ,19 ,20). Gating functions are commonly assessed with P50 suppression and prepulse inhibition (PPI) of the startle response.

P50 Suppression

Initial studies at the University of Colorado have identified P50 suppression as an important candidate endophenotype in studies of schizophrenia (21 ,22 ,23 and 24). In a typical paradigm, P50 suppression occurs when two clicks are presented with a 500-millisecond interval between them. The small P50 event potential wave elicited by the click stimulus can be identified when many trials (e.g., 30 to 100 or more) are performed; this number of trials plus various filtering strategies provides investigators with a robust signal-to-noise ratio for identifying and quantifying the P50 wave. A P50 wave is generated to the first click and another to the second click. Across multiple studies, it has been found that the second P50 wave is normally suppressed; suppression can probably be attributed to the activation of inhibitory processing and circuitry by the first P50 stimulus. In normal subjects, the second P50 wave typically is diminished by 80% in comparison with the first wave (Fig. 51.1).

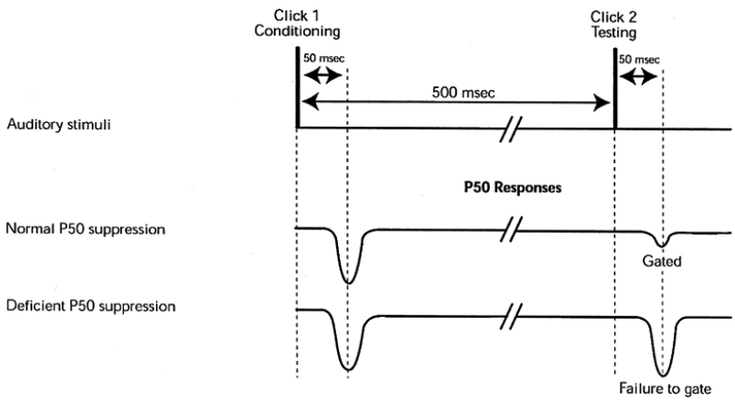


FIGURE 51.1. Pairs of auditory clicks are presented to subjects and EEG is averaged across trials. The P50 component of the auditory event-related potential is measured in response to the first and second clicks.

Initial studies (21 ,22 ,23 and 24) demonstrated an expected failure of suppression in schizophrenic patients, consistent with theories of failed inhibitory function or impaired sensorimotor gating in schizophrenia (25). These studies of deficits in P50 suppression in schizophrenic patients have been widely replicated (26 ,27 ,28 ,29 ,30 and 31). The failure of P50 suppression in schizophrenic patients is not necessarily specific to this one disorder. For example, Franks et al. (32) reported that P50 suppression is also deficient in patients with acute mania but “normalizes” with time, whereas the deficits of P50 suppression are more persistent in schizophrenic patients (32). This finding is consistent with the idea that genetic “diatheses” may be shared between schizophrenia and mania. P50 suppression deficits have also been shown (and replicated) in “clinically unaffected” family members of schizophrenic patients (24 ,31 ,33 ,34 and 35). The P50 suppression abnormalities of these family members normalize following administration of the cholinergic nicotinic receptor stimulant nicotine (36), as do those of schizophrenic patients (37). This finding has raised interest in the critical importance of the cholinergic system in P50 suppression, and some of the cholinergic neurobiological substrates of P50 suppression deficits have been elucidated.

As discussed in the section on PPI , suppression is probably the function of a more wide-ranging neural circuitry prominently involving hippocampal structures (38). The use of P50 suppression as a candidate endophenotype in genetic studies is probably the most advanced of any of the endophenotypes we discuss here; a specific linkage of P50 suppression with a genetic marker at the locus of the α_7 subunit of the nicotinic receptor gene (11) has been identified in the first study linking a candidate endophenotype of information processing in schizophrenia to a specific

chromosomal region. It is important to stress that these types of studies do not identify a “schizophrenia endophenotype,” but rather the linkage of deficits in P50 suppression (characteristic of schizophrenia) to a specific chromosome region. Future studies will have to identify the specific genetic deficit(s) (e.g., specific single-nucleotide polymorphisms) associated with abnormalities of P50 suppression.

In terms of our assessment of candidate endophenotypes and genetic studies, it is important to note that medications have an influence on P50 suppression abnormalities in schizophrenia. It appears that atypical antipsychotic medications may reverse the P50 suppression deficits in schizophrenic patients (39 ,40 ,41 and 42). If these initial results continue to be confirmed, the search for candidate endophenotypes will be complicated by the fact that atypical (and perhaps, in some circumstances, typical) antipsychotic medications are increasingly being utilized as first-line agents in the treatment of schizophrenia. We may thus face the circumstance of examining schizophrenic patients whose P50 suppression deficits have been “normalized” and then conducting family studies in which these deficits may appear in unaffected relatives of schizophrenic patients. Should this occur, genetic statistical strategies will have to be utilized that will allow us to “exclude” the “normalized” schizophrenic patient from analysis and utilize only clinically unaffected family members in genetic (e.g., linkage) studies. Much more information will be generated in the next several years, and the use of what we would term *“null proband” strategies* may be necessary as schizophrenic patients who are not medicated or are neuroleptic-naïve become more difficult to ascertain and are replaced by patients treated with atypical antipsychotic medications. In addition, the use of drug withdrawal strategies to unmask endophenotypic markers has come under increasing criticism (43) and is becoming more difficult to justify ethically in comparison with other promising research strategies (e.g., 44).

Prepulse Inhibition of the Startle Response

Since 1978 (45), PPI deficits of the startle response have been consistently identified in schizophrenic patients. PPI of the startle response occurs as follows. Normally, an intense and powerful sensory stimulus elicits a whole-body startle response in almost all mammals. This rapid, intense sensory stimulus may be sound or light, or it may be tactile (e.g., an air puff). When a weak prestimulus precedes the startling stimulus by approximately 100 milliseconds, PPI occurs. Schizophrenic patients and their relatives show deficits in PPI (45 ,46 ,47 ,48 ,49 ,50 ,51 ,52 ,53 and 54) (Fig. 51.2). This is the second commonly studied form of sensorimotor gating (along with P50 suppression) (see ref. 25 for a discussion of gating abnormalities in schizophrenia). PPI is being increasingly used in schizophrenia research. It is important to distinguish the PPI paradigm described above from a similar but quite distinct paradigm in which attentional allocation to the prepulse (55) is used in an attempt to increase the degree of PPI. The PPI paradigm typically used in schizophrenia research, described above, is a “neutral” or “uninstructed” paradigm that largely taps into involuntary and automatic information processing (56); the instructed paradigm is different because subjects attend to the prepulse and thus the paradigm identifies so-called voluntary attentional deficits. This section describes the more widely studied uninstructed PPI.

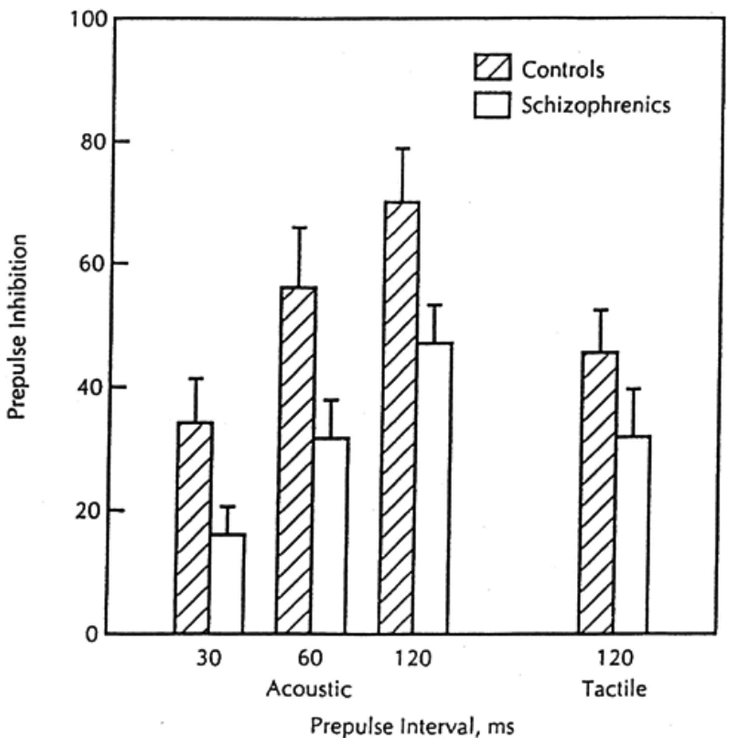


FIGURE 51.2. Across all prepulse-to-pulse intervals tested, the schizophrenic patients showed a loss of gating effect of the prepulse that preceded the startle stimulus. (From Braff DL, Grillon C, Geyer MA. Gating and habituation of the startle reflex in schizophrenic patients. *Arch Gen Psychiatry* 1992;49:206–215, with permission.)

Much like deficits of P50 suppression, PPI deficits are not unique to schizophrenia. PPI deficits are characteristic of a “family” of disorders in which cognitive, sensory, and motor information undergoes a failure of gating. Patients with gating disorders include those with schizophrenia (45 ,46 ,47 ,48 ,49 ,50 ,51 ,52 and 53), obsessive-compulsive disorder (with obsessive and ungated ideas) (57), and Huntington syndrome (58) and Tourette syndrome (59) (with ungated motor activity).

The clinical correlates of PPI in schizophrenia comprise a rich database. PPI deficits have been correlated with distractibility (60), perseverative responses on the Wisconsin Card Sorting Test (61), and most prominently thought disorder (62), especially when PPI and thought disorder are measured at the same time (63). These deficits are also associated with an earlier age of onset (51). Modest correlations have been found with both positive and negative symptoms, and the symptom correlates may be associated with subcortical dopamine hyperactivity and reciprocal frontal dopamine

hypoactivity (47). An initial report has described PPI deficits in clinically unaffected family members of schizophrenic patients (54), and further work is needed to understand the heritability pattern of PPI deficits in family members of schizophrenic patients. Much of what is known about the neural substrate of PPI can be attributed to the extensive work of Swerdlow, Geyer, Braff, and their associates. It appears that PPI is modulated mostly by the ventral cortico-striato-pallido-thalamic (CSPT) circuitry originally described by Swerdlow and Koob (64), based on the pioneering work of Penney, Alexander, and Young on the dorsal loci of the CSPT circuits. The circuitry cannot be described in detail here, but lesion infusion studies and a variety of other strategies have established the animal model of PPI deficits in schizophrenia as a robust area of study (65 ,66 and 67). For example, in rat pups with ventral hippocampal lesions, PPI levels are normal until adolescence, when PPI deficits appear (68 ,69), a finding that supports the neurodevelopmental model of PPI deficits as it applies to an integrated model of schizophrenia. Apomorphine used as a dopamine D2 agonist induces PPI deficits that are reversible with typical or atypical antipsychotic medications. Phencyclidine induces PPI deficits that are differentially reversed by atypical (but not typical) antipsychotic medications. The interested reader is referred to Swerdlow et al. (70) for further discussion of these issues.

Some initial results utilizing between-subjects rather than the more compelling within-subjects designs indicate that PPI deficits in schizophrenic patients may be reversed or “normalized” by antipsychotic medications (51 ,53); however, no linkage studies have utilized PPI, although the genetic contributions to PPI have been elucidated by the fact that PPI levels differ in different rat strains (71), and differential sensitivity to PPI deficits has been observed in these strains (72 ,73 and 74). The increasing use of isolation rearing (75 ,76 ,77 and 78) and knockout mice (79 ,80 ,81 and 82) in PPI studies will undoubtedly yield much more information about the neural and genetic contributions to PPI. In parallel, human linkage studies (already in progress) are being conducted.

Oculomotor Function

Oculomotor function is another important measure that has been used in schizophrenia research. As in gating, two fundamental paradigms have been utilized: eye tracking, or smooth pursuit, and the antisaccade task. Eye-tracking dysfunction in schizophrenic patients was first reported by Diesendorf and Dodge (83), and their work was later extended by numerous investigators, including Holzman and colleagues (84 ,85). Across quantitative and qualitative studies, the eye-tracking deficits seen in schizophrenia have been well documented (86 ,87 ,88 and 89).

Smooth-Pursuit Eye Movement

In a typical study of smooth-pursuit eye movement, a signal is presented to subjects on a computer screen and their ability to track the target smoothly is assessed. Schizophrenic patients do not track the target smoothly; they exhibit fragmented eye movements and resulting eye movement abnormalities, such as “catch-up” saccades. This task has been very useful in family studies; clinically unaffected family members of schizophrenic patients also exhibit eye movement dysfunctions that typically involve an inability to follow a target smoothly, which leads to a variety of abnormalities, including “catch up” saccades (85 ,86 ,88 ,90 ,91 ,92 ,93 and 94). In a series of studies, Siever et al. (95) and others (91 ,93 ,96 ,97) also reported abnormalities of smooth-pursuit eye movement in schizotypal patients.

Antisaccade Task

The antisaccade task has also been widely employed in schizophrenia research as another oculomotor task and as a potential endophenotype. In the antisaccade task, the subject first fixates on a centrally presented visual cue. A target stimulus is then presented to the left or right of the fixation stimulus, and the subject is instructed to look away from the target stimulus; if the stimulus is presented 3 degrees to the left of the fixation point, the subject is expected to look 3 degrees to the right and inhibit the natural tendency to “follow” the target to the left. Voluntary inhibitory functions are utilized to suppress the normal tendency to look at the target stimulus and gaze in the opposite direction. This task, like some of the other measures discussed above, uses inhibition and is largely volitional (like smooth-pursuit eye movement) rather than automatic (like gating). Schizophrenic patients show marked deficits in performing this task; an initial gaze directed toward rather than away from the target stimulus is characteristic (98 ,99 ,100 and 101). The magnitude (i.e., effect size) of the performance deficits in schizophrenic patients and clinically unaffected family members is large (Table 51.1), and the large difference in effect size between probands and normal comparison subjects makes the antisaccade task an excellent candidate endophenotype for genetic studies (101 ,102). Within the schizophrenia spectrum, it is notable that antisaccade deficits occur in family members of schizophrenic patients and in patients with schizotypal personality disorder (96 ,101 ,103 ,104). In this way, the antisaccade deficit meets the second criterion for a candidate endophenotype—that is, a candidate endophenotype should appear in clinically unaffected family members of schizophrenic patients (and perhaps in schizotypal patients).

Neuropsychological Tasks

A plethora of candidate endophenotypes have been derived from the neuropsychological literature. It is well-known that schizophrenic patients exhibit a wide range of neuropsychological deficits (105 ,106) and that these deficits extend to clinically unaffected family members (107). Deficits have

been reported in several important domains: (a) *executive function*, as assessed by the Wisconsin Card Sorting Test (108); (b) *working memory*, as assessed by the Letter–Number Span (109), and (c) *thought disorder*, commonly derived from the processing of stimuli from the Rorschach Test to yield the Thought Disorder Index (110) and the Ego Impairment Index (111). These cognitive dysfunctions are frequently found in family and twin studies in clinically unaffected family members (112 ,113 ,114 ,115 ,116 ,117 ,118 ,119 ,120 and 121), schizotypal patients (122 ,123 ,124 ,125 and 126), and clinically unaffected monozygotic twins discordant for schizophrenia itself (127). The neural substrates of many of these abnormalities are well understood and are being rapidly explicated because these tasks are very well suited to performance during functional brain imaging (e.g., 128). For example, it appears that the Wisconsin Card Sorting Test relies on dorsolateral prefrontal cortex (128 ,129 ,130 and 131) and related distributed circuit structures. Working memory utilizes a complex neural substrate that includes the prefrontal cortex and related structures. It important when working memory is utilized to be clear about whether the test assesses simple delayed recall (transient online storage) or the more complex storage, manipulation, and recall (executive functioning) of visuospatial or verbal memory; these are two distinct neuropsychological processes that probably utilize at least partially distinct neural substrates influenced by at least partially different sets of genes (132). Functional imaging experiments in thought disorder are more preliminary, and the neural substrate of thought disorder is now being explicated.

Continuous Performance Task

The Continuous Performance Task (CPT) is another measure that has been widely applied in the study of schizophrenic patients (133 ,134 ,135 and 136). In the basic form of this task, the subject is presented with a string of stimuli and asked to identify target stimuli from among background or noise stimuli. The CPT is thought to tap into the function of vigilance; schizophrenic patients commonly have significant deficits in the CPT. The CPT also has the advantage that it can be presented in a variety of “degraded” forms in which the signal-to-noise ratio of the stimulus to be identified is attenuated through a variety of parametric manipulations; these make the vigilance and identification task more difficult to perform and may correspondingly increase group separation between schizophrenic patients, their clinically unaffected relatives, and appropriate comparison subjects. The CPT is one of the most thoroughly studied tasks in schizophrenia, in no small degree because of the pioneering work of Nuechterlein et al. (135) and others (137 ,138) who have consistently demonstrated CPT deficits in schizophrenic patients. The fact that unaffected relatives of patients with schizophrenia (139 ,140 ,141 ,142 ,143 ,144 and 145) and schizotypal disorder (146 ,147 and 148) have parallel deficits supports the utility of the task as a candidate endophenotype for genetic studies of information-processing deficits in schizophrenia. These deficits appear to measure stable markers of schizophrenia that may be associated with a genetic vulnerability to the illness and are seen in neuroleptic-naïve, first-break, and neuroleptic-withdrawn schizophrenic patients and their siblings (142).

Span of Apprehension

The utility of Span of Apprehension as a candidate endophenotype, like that of the others measures discussed in this chapter, is supported by a vast amount of literature, only a brief summary of which can be presented here. In its most simplified form, Span of Apprehension refers to the number of items that can be apprehended or attended to and subsequently recalled at one time from an array of stimuli. The interested reader is referred to a particularly scholarly discussion by Asarnow et al. (149). As in the other measures discussed here, the Span of Apprehension has yielded a pattern of interesting results that makes the task another excellent candidate endophenotype in schizophrenia. Additionally, Span of Apprehension deficits have been found in clinically unaffected family members of schizophrenic patients (150 ,151 and 152) and in patients with schizotypal disorder (153 ,154). Recently, in a study of normal twins, Bartfai et al. (155) reported a significant genetic component in the Span of Apprehension task, which further strengthens its utility in genetic studies of schizophrenia.

Visual Backward Masking

In Visual Backward Masking, a simple target stimulus presented with a tachistoscope (156 ,157 and 158) or, more recently, a computer (159 ,160) is followed by a complex, usually powerful masking stimulus of interlocking *X*s that overlap the area of target presentation. A subject reports the target stimulus when it is presented alone. As the masking stimulus appears closer in time to the target (e.g., 100 milliseconds after the target, like the interstimulus interval in PPI experiments), the subject is no longer able identify the target stimulus. Schizophrenic patients are subjected to the effects of the mask on target identification at an interstimulus interval at which normal subjects have little trouble distinguishing the target stimulus (161). Explanations for the masking effect extend from integration of the mask with the target stimulus to interruption of target stimulus processing by the mask (161 ,162 ,163 and 164). Whatever the mechanism, the phenomenon is readily identified as a marker in schizophrenic patients (158 ,165 ,166 ,167 and 168), family members of schizophrenic patients (159 ,160 ,169), and schizotypal patients (170 ,171 ,172 and 173). The specificity of the deficit is unclear, although in one study, manic patients at the height of psychosis showed visual masking deficits that were reversible over time with treatment. In this study, it was reported that the deficits of schizophrenic patients with a good prognosis, who

typically respond to treatment, may also be reversible over time (174). The underlying neural mechanism linked to masking deficits involves the dorsal and ventral information-processing substrates that are supported by magnocellular and parvocellular neurons (175 ,176). Because both clinically unaffected family members and schizotypal patients exhibit deficits of Visual Backward Masking, it may well serve as an important candidate endophenotype in genetic studies.

P300 Event-Related Potential

Since the initial reports of Callaway and colleagues (177 ,178 ,179 and 180) of P300 deficits in schizophrenia, a large number of investigators have identified their topography, lateralization, and neural basis. The results of these studies make the P300 event-related potential an excellent candidate endophenotype for understanding the genetic basis of the deficits.

Schizophrenic patients have long been known to have deficits in the P300 component of the event-related potential (177 ,178 ,181). The P300 wave occurs about 300 milliseconds after stimulus presentation and is commonly thought to reflect the apportionment of attention to a stimulus that is relatively novel or rich in information. Many paradigms utilize a series of rather neutral stimuli and then use an “attention-grabbing” stimulus to elicit a large P300 wave. Although the variability of the latency properties of the P300 event-related potential wave may account for part of the diminution in schizophrenia, repeated studies report that schizophrenic patients show a decreased P300 wave amplitude over time (182 ,183 ,184 and 185). The fact that these deficits are also found in unaffected family members of schizophrenic patients (186 ,187 ,188 ,189 and 190) and in schizotypal patients (191 ,192 and 193) supports the utilization of the P300 wave as a candidate endophenotype. The original work of Callaway et al. (178) and more recent studies by McCarley and associates (182 ,194 ,195 and 196) have contributed to an understanding of the P300 neural circuitry that supports the generation of this wave. It appears that the P300 wave is generated from the temporal lobes, perhaps the superior temporal gyrus of the brain. Along with a diminution of the P300 wave in schizophrenic patients, the volume of the superior temporal gyrus gray matter is also diminished. Lateralization findings indicate that it is probably the left P300 wave that is differentially diminished in schizophrenic patients, matching the volume depletion of the left superior temporal gyrus.

Older studies of reaction time and exciting new and evolving studies of mismatch negativity (197) offer a wide range of potentially useful endophenotypes that may prove to be especially interesting in schizophrenia research.

SUMMARY

Part of "51 - Endophenotypes in Studies of the Genetics of Schizophrenia "

A multitude of interlocking studies, only some of which have been reviewed here, point to information-processing deficits and closely related inhibitory abnormalities as excellent candidate endophenotypes for genetic studies of schizophrenia. The heritability of several of these candidate endophenotypes has already been assessed in genetic studies. In addition, the neuronal mechanisms of many of the endophenotypes are currently being investigated through neurophysiologic studies in both humans (e.g., functional imaging) and related animal models. The ultimate utility of physiologic endophenotypes may be to correlate the wealth of emerging but nonfunctional genetic information with the critically important and functionally significant underlying neurobiology of these endophenotypes. The task of identifying genes that convey a risk for schizophrenia is now under way, generally with the use of either the clinical phenotype of schizophrenia or risk-related endophenotypes. Findings that many of the linkage sites are positive for both schizophrenia and bipolar disorder (e.g., 198) will undoubtedly stimulate a reexamination of what aspect of psychotic psychopathology is being transmitted at each genetic locus, and which nongenetic factors, such as neonatal ventral hippocampal lesions, interact with these genes to produce schizophrenia (e.g., 68) versus bipolar disorder. Additionally, within cohorts of schizophrenic patients, some of these information-processing endophenotypes overlap with each other, both behaviorally and in terms of their underlying neural substrates. This allows for the exciting possibility of constructing “composite” phenotypes consisting of neurologically coherent combinations of more than one of these identified markers (102).

As the molecular mantra states, the fundamental unit of genetic transmission is an abnormality in the structure or expression of a protein. Presumably, most of those protein abnormalities affect neuronal functions that can be measured as changes in physiologic functions, such as the endophenotypes we have described above. The elucidation of how different genetic abnormalities, singly or in combination, contribute to the neuronal pathophysiology of psychosis may help to redefine the nosology of psychotic illnesses and point the way to new treatment approaches. The power and value of endophenotypes is that they illuminate genetically mediated risk/vulnerability factors that often interact with nongenetic factors to produce the syndrome of schizophrenia. Thus, in the pool of genetic strategies and techniques that can be used to understand complex genetic psychiatric disorders (199 ,200 ,201 and 202), endophenotype-based strategies play an important and informative role. In colon cancer, the inherited genetic factor is familial polyposis rather than cancer itself (4). In parallel, it is quite likely that failures of information processing/inhibition are the genetically transmitted risk factors that interact with nongenetic factors to produce the clinical disorder of schizophrenia. Thus, the identification and genetic analysis of these endophenotypes should prove particularly valuable in understanding the genetic basis of schizophrenia. These studies will also facilitate a fuller understanding of how genetic

and nongenetic factors interact to produce this devastating illness and, it is hoped, point the way to more effective treatments.

ACKNOWLEDGMENTS

Part of "51 - Endophenotypes in Studies of the Genetics of Schizophrenia "

This work was supported in part by grants from the National Institute of Mental Health (MH42228) and the Department of Veteran Affairs (VISN 22 MIRECC; Mental Illness Research, Education, and Clinical Center).

REFERENCES

1. Brzustowicz LM, Hodgkinson KA, Chow EW, et al. Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22. *Science* 2000;288:678–682.

2. Pulver AE. Search for schizophrenia susceptibility genes. *Biol Psychiatry* 2000;47:221–230.

3. Buervenich S, Arvidsson M, Carmine A, et al. *NURR1* mutations in cases of schizophrenia and manic depressive disorder. *Neuropsychiatr Genet (in press)*.

4. Leppert M, Burt M, Hughes JP, et al. Genetic analysis of an inherited predisposition to colon cancer in a family with a variable number of adenomatous polyps. *N Engl J Med* 1990;322:904–908.

5. Lalouel JM, Le Mignon L, Simon M, et al. Genetic analysis of idiopathic hemochromatosis using both qualitative (disease status) and quantitative (serum iron) information. *Am J Hum Genet* 1985;37:700–718.

6. Faraone SV, Kremens WS, Lyons MJ, et al. Diagnostic accuracy and linkage analysis: how useful are schizophrenia spectrum phenotypes? *Am J Psychiatry* 1995;152:1286–1290.

7. Callicot JH, Egan MF, Bertolino A, et al. Hippocampal *N*-acetyl aspartate in unaffected siblings of patients with schizophrenia: a possible intermediate neurobiological phenotype. *Biol Psychiatry* 1998;44:941–950.

8. Hyman SE. The NIMH perspective: next steps in schizophrenia research. *Biol Psychiatry* 2000;47:1–7.

9. Weinberger DR. Schizophrenia: new phenes and new genes. *Biol Psychiatry* 1999;46:3–7.

10. Braff DL. Psychophysiological and information processing approaches to schizophrenia. In: Charney DS, Nestler E, Bunney BS, eds. *Neurobiological foundation of mental illness*. New York: Oxford University Press, 1999:258–271.

11. Freedman R, Coon H, Myles-Worsley M, et al. Linkage of a neurophysiological deficit in schizophrenia to a chromosome 15 locus. *Proc Natl Acad Sci U S A* 1997;94:587–592.

12. Weinberger DR. Hippocampal injury and chronic schizophrenia [Letter, Comment]. *Biol Psychiatry* 1991;29:509–511.

13. Weinberger DR. Cell biology of the hippocampal formation in schizophrenia. *Biol Psychiatry* 1999;45:395–402.

14. McNeil TF, Cantor-Graae E, Weinberger DR. Relationship of obstetric complications and differences in size of brain structures in monozygotic twin pairs discordant for schizophrenia. *Am J Psychiatry* 2000;157:203–212.

15. Kinney DK, Levy DL, Yurgelun-Todd DA, et al. Inverse relationship of perinatal complications and eye-tracking dysfunction in relatives of patients with schizophrenia: evidence for a two-factor model. *Am J Psychiatry* 1998;155:976–978.

16. McCarley RW, Wible CG, Frumin M, et al. MRI anatomy of schizophrenia. *Biol Psychiatry* 1999;45:1099–1119.

17. Andreasen NC, Nopoulos P, O'Leary DS, et al. Defining the phenotype of schizophrenia: cognitive dysmetria and its neural mechanisms. *Biol Psychiatry* 1999;46:908–920.

18. Braff DL. Connecting the "dots" of brain dysfunction in schizophrenia: what does the picture look like? *Arch Gen Psychiatry* 1999;56:791–793.

19. McGhie A, Chapman J. Disorders of attention and perception in early schizophrenia. *Br J Med Psychol* 1961;34:103–116.

20. Braff DL, Geyer MA. Sensorimotor gating and schizophrenia: human and animal model studies. *Arch Gen Psychiatry* 1990;47:181–188.

21. Adler LE, Pachtman E, Franks RD, et al. Neurophysiological evidence for a defect in neuronal mechanisms involved in sensory gating in schizophrenia. *Biol Psychiatry* 1982;17:639–654.

22. Freedman R, Adler LE, Waldo MC, et al. Neurophysiological evidence for a defect in inhibitory pathways in schizophrenia: comparison of medicated and drug-free patients. *Biol Psychiatry* 1983;18:537–551.

23. Freedman R, Adler LE, Gerhardt GA, et al. Neurobiological studies of sensory gating in schizophrenia. *Schizophr Bull* 1987;13:669–678.

24. Siegel C, Waldo M, Mizner G, et al. Deficits in sensory gating in schizophrenic patients and their relatives. Evidence obtained with auditory evoked responses. *Arch Gen Psychiatry* 1984;41:607–612.

25. Braff DL, Geyer MA. Sensorimotor gating and schizophrenia: human and animal model studies. *Arch Gen Psychiatry* 1990;47:181–188.

26. Nagamoto HT, Adler LE, Waldo MC, et al. Sensory gating in schizophrenics and normal controls: effects of changing stimulation interval. *Biol Psychiatry* 1989;25:549–561.

27. Nagamoto HT, Adler LE, Waldo MC, et al. Gating of auditory response in schizophrenics and normal controls. Effects of recording site and stimulation interval on the P50 wave. *Schizophr Bull* 1991;4:31–40.

28. Boutros NN, Zouridakis G, Overall J. Replication and extension of P50 findings in schizophrenia. *Clin Electroencephalogr* 1991;22:40–45

29. Judd LL, McAdams L, Budnick B, et al. Sensory gating deficits in schizophrenia: new results. *Am J Psychiatry* 1992;149:488–493.

30. Clementz BA, Geyer MA, Braff DL. P50 suppression among schizophrenia and normal comparison subjects: a methodological analysis. *Biol Psychiatry* 1997;41:1035–1044.

31. Clementz BA, Geyer MA, Braff DL. Poor P50 suppression among schizophrenia patients and their first-degree biological relatives. *Am J Psychiatry* 1998;155:1691–1694.

32. Franks RD, Adler LE, Waldo MC, et al. Neurophysiological studies of sensory gating in mania: comparison with schizophrenia. *Biol Psychiatry* 1983;18:989–1005.

33. Waldo MC, Adler LE, Freedman R. Defects in auditory sensory gating and their apparent compensation in relatives of schizophrenics. *Schizophr Res* 1988;1:19–24.

34. Waldo MC, Carey G, Myles-Worsley M, et al. Co-distribution of a sensory gating deficit and schizophrenia in multi-affected families. *Psychiatry Res* 1991;39:257–268.

35. Waldo M, Myles-Worsley M, Madison A, et al. Sensory gating deficits in parents of schizophrenics. *Am J Med Genet* 1995;60:506–511.

36. Adler LE, Hoffer LJ, Griffith J, et al. Normalization by nicotine of deficient auditory sensory gating in the relatives of schizophrenics. *Biol Psychiatry* 1992;32:607–616

37. Adler LE, Hoffer LD, Wiser A, et al. Normalization of auditory physiology by cigarette smoking in schizophrenic patients. *Am J Psychiatry* 1993;150:1856–1861.

38. Adler LE, Olincy A, Waldo M, et al. Schizophrenia, sensory gating, and nicotinic receptors. *Schizophr Bull* 1998;24:189–202.

39. Nagamoto HT, Adler LE, Hea RA, et al. Gating of auditory P50 in schizophrenics: unique effects of clozapine. *Biol Psychiatry* 1996;40:181–188.

40. Nagamoto HT, Adler LE, McRae KA, et al. Auditory P50 in schizophrenics on clozapine: improved gating parallels clinical improvement and changes in plasma 3-methoxy-4-hydroxyphenylglycol. *Neuropsychobiology* 1999;39:10–17.

41. Yee CM, Nuechterlein KH, Morris SE, et al. P50 suppression in recent-onset schizophrenia: clinical correlates and risperidone effects. *J Abnorm Psychol* 1998;107:691–698.

42. Light GA, Geyer MA, Clementz BA, et al. Normal P50 suppression in schizophrenia patients treated with atypical antipsychotic medications. *Am J Psychiatry* 2000;157:767–771.

43. National Bioethics Advisory Commission. *Research involving persons with mental disorders that may affect decision-making capacity*. Rockville, MD: National Bioethics Advisory Commission, 1998.

44. Davis KL, Braff DL, Weinberger DR. Protecting research subjects and psychiatric research: we can do both [Editorial]. *Biol Psychiatry* 1999;46:727–728.

45. Braff D, Stone C, Callaway E, et al. Prestimulus effects on human startle reflex in normals and schizophrenics. *Psychophysiology* 1978;14:339–343.

46. Braff DL, Grillon C, Geyer MA. Gating and habituation of the startle reflex in schizophrenic patients. *Arch Gen Psychiatry* 1992;49:206–215.

47. Braff DL, Swerdlow NR, Geyer MA. Symptom correlates of prepulse inhibition deficits in male schizophrenic patients. *Am J Psychiatry* 1999;156:596–602.

48. Grillon C, Ameli R, Charney DS, et al. Startle gating deficits occur across prepulse intensities in schizophrenic patients. *Biol Psychiatry* 1992;32:939–943.

49. Bolino F, Di Michele V, Di Cicco L, et al. Sensorimotor gating and habituation evoked by electro-cutaneous stimulation in schizophrenia. *Biol Psychiatry* 1994;36:670–679.

50. Perry W, Braff DL. Information-processing deficits and thought disorder in schizophrenia. *Am J Psychiatry* 1994;151:363–367.

51. Kumari V, Soni W, Sharma T. Normalization of information processing deficits in schizophrenia with clozapine. *Am J Psychiatry* 1999;156:1046–1051.

52. Parwani A, Duncan EJ, Bartlett E, et al. Impaired prepulse inhibition of acoustic startle in schizophrenia. *Biol Psychiatry* 2000;47:662–669.

53. Weike AI, Bauer U, Hamm AO. Effective neuroleptic medication removes prepulse inhibition deficits in schizophrenia patients. *Biol Psychiatry* 2000;47:61–70.

54. Cadenhead KS, Swerdlow NR, Shafer KM, et al. Modulation of the startle response and startle laterality in relatives of schizophrenic patients and schizotypal personality disordered subjects: evidence of inhibitory deficits. *Am J Psychiatry* 2000;157:1660–1668.

55. Filion DL, Dawson ME, Schell AM. Modification of the acoustic startle-reflex eye blink: a tool for investigating early and late attentional processes. *Biol Psychol* 1993;35:185–200.

56. Callaway E, Naghdi S. An information processing model for schizophrenia. *Arch Gen Psychiatry* 1982;39:339–347.

57. Swerdlow NR, Benbow CH, Zisook S, et al. A preliminary assessment of sensorimotor gating in patients with obsessive compulsive disorder. *Biol Psychiatry* 1993;33:298–301.

58. Swerdlow NR, Paulsen J, Braff DL, et al. Impaired prepulse inhibition of acoustic and tactile startle response in patients with Huntington's disease. *J Neurol Neurosurg Psychiatry* 1995;58:192–200.

59. Castellanos FX, Fine EJ, Kaysen D, et al. Sensorimotor gating in boys with Tourette's syndrome and ADHD: preliminary results. *Biol Psychiatry* 1996;39:33–41.

60. Karper LP, Freeman GK, Grillon C, et al. Preliminary evidence of an association between sensorimotor gating and distractibility in psychosis. *J Neuropsychiatry Clin Neurosci* 1996;8:60–66.

61. Butler RW, Jenkins MA, Geyer MA, et al. Wisconsin Card Sorting deficits and diminished sensorimotor gating in a discrete subgroup of schizophrenic patients. In: Tamminga CA, Schulz SC, eds. *Schizophrenia* (*Advances in neuropsychiatry and psychopharmacology*, vol 1). New York: Raven Press, 1991:163–168.

62. Perry W, Braff DL. Information-processing deficits and thought disorder in schizophrenia. *Am J Psychiatry* 1994;151:363–367.

63. Perry W, Geyer MA, Braff DL. Sensorimotor gating and thought disturbance measured in close temporal proximity in schizophrenic patients. *Arch Gen Psychiatry* 1999;56:277–281.

64. Swerdlow NR, Koob GF. Dopamine, schizophrenia, mania, and depression: toward a unified hypothesis of cortico-striato-pallido-thalamic function. *Behav Brain Sci* 1987;10:197–245.

65. Swerdlow NR, Geyer MA. Using an animal model of deficient sensorimotor gating to study the pathophysiology and new treatments of schizophrenia. *Schizophr Bull* 1998;24:285–301.

66. Swerdlow NR, Taaid N, Oostwegel JL, et al. Towards a cross-species pharmacology of sensorimotor gating: effects of amantadine, bromocriptine, pergolide and ropinirole on prepulse inhibition of acoustic startle in rats. *Behav Pharmacol* 1998;9:389–396.

67. Geyer MA, Braff DL, Swerdlow NR. Startle-response measures of information processing in animals: relevance to schizophrenia. In: Haug M, Whalen RE, eds. *Animal models of human emotion and cognition*. Washington, DC: APA Books, 1999:103–116.

68. Lipska BK, Swerdlow NR, Geyer MA, et al. Neonatal excitotoxic hippocampal damage in rats causes post-pubertal changes in prepulse inhibition of startle and its disruption by apomorphine. *Psychopharmacology* 1995;122:35–43.

69. Swerdlow NR, Lipska BK, Weinberger DR, et al. Increased sensitivity to the sensorimotor gating-disruptive effects of apomorphine after lesions of medial prefrontal cortex or ventral hippocampus in adult rats. *Psychopharmacology* 1995;122:27–34.

70. Swerdlow NR, Braff D L, Taaid N, et al. Assessing the validity of an animal model of deficient sensorimotor gating in schizophrenic patients. *Arch Gen Psychiatry* 1994;51:139–154.

71. Swerdlow NR, Martinez ZA, Hanlon F, et al. Toward understanding the biology of a complex phenotype: rat strain and substrain differences in the sensorimotor gating-disruptive effects of dopamine agonists. *J Neurosci* 2000;20:4325–4336.

72. Rigdon G. Differential effects of apomorphine on prepulse inhibition of acoustic startle reflex in two rat strains. *Psychopharmacology (Berl)* 1990;102:419–421.

73. Swerdlow NR, Varty GB, Geyer MA. Discrepant findings of clozapine effects on prepulse inhibition of startle: is it the route or the rat? *Neuropsychopharmacology* 1998;18:50–56.

74. Swerdlow NR, Martinez ZA, Hanlon FM, et al. Towards the genetics of a complex phenotype: strain analyses of drug effects on startle gating. *Biol Psychiatry* 2000;47:121.

75. Bakshi VP, Swerdlow NR, Braff DL, et al. Reversal of isolation rearing-induced deficits in prepulse inhibition by Seroquel and olanzapine. *Biol Psychiatry* 1998;43:436–445.

76. Varty GB, Geyer MA. Effects of isolation rearing on startle reactivity, habituation, and prepulse inhibition in male Lewis, Sprague-Dawley, and Fischer F344 rats. *Behav Neurosci* 1998;112:1450–1457.

77. Bakshi VP, Geyer MA. Ontogeny of isolation rearing-induced deficits in sensorimotor gating in rats. *Physiol Behav* 1999;67:385–392.

78. Varty GB, Braff DL, Geyer MA. Is there a critical developmental “window” for isolation rearing-induced changes in prepulse inhibition of the acoustic startle response? *Behav Brain Res* 1999;100:177–183.

79. Dulawa SC, Hen R, Scearce-Levie K, et al. Serotonin$_{1B}$-receptor modulation of startle reactivity, habituation, and prepulse inhibition in wild-type and serotonin$_{1B}$ knockout mice. *Psychopharmacology* 1997;132:125–134.

80. Dulawa SC, Hen R, Scearce-Levie K, et al. 5-HT$_{1B}$ receptor modulation of prepulse inhibition: recent findings in wild-type and 5-HT$_{1B}$ knockout mice. *Ann N Y Acad Sci* 1998;861:79–84.

81. Ralph RJ, Varty GB, Kelly MA, et al. The dopamine D2, but not D3 or D4, receptor subtype is essential for the disruption of prepulse inhibition produced by amphetamine in mice. *J Neurosci* 1999;19:4627–4633.

82. Geyer MA. Assessing prepulse inhibition of startle in wild-type and knockout mice. *Psychopharmacology* 1999;147:11–13.

83. Diesendorf AR, Dodge R. An experimental study of the ocular reactions of the insane from photographic records. *Brain* 1908;31:451–489.

84. Holzman PS, Proctor LR, Hughes DW. Eye-tracking patterns in schizophrenia. *Science* 1973;181:179–181.

85. Holzman PS, Proctor LR, Levy DL, et al. Eye-tracking dysfunctions in schizophrenic patients and their relatives. *Arch Gen Psychiatry* 1974;31:143–151.

86. Holzman PS, Kringlen E, Levy DL, et al. Deviant eye tracking in twins discordant for psychosis: a replication. *Arch Gen Psychiatry* 1980;37:627–631.

87. Levin S, Jones A, Stark L, et al. Identification of abnormal patterns in eye movements of schizophrenic patients. *Arch Gen Psychiatry* 1982;39:1125–1130.

88. Iacono WG, Moreau M, Beiser M, et al. Smooth-pursuit eye tracking in first-episode psychotic patients and their relatives. *J Abnorm Psychol* 1992;101:104–116.

89. Levy DL, Holzman PS, Matthysse S, et al. Eye tracking and schizophrenia: a selective review. *Schizphr Bull* 1994;20:47–62.

90. Chen Y, Nakayama K, Levy DL, et al. Psychophysical isolation of a motion-processing deficit in schizophrenics and their relatives and its association with impaired smooth pursuit. *Proc Natl Acad Sci U S A* 1999;96:4724–4729.

91. Clementz BA, Sweeney JA, Hirt M, et al. Phenotypic correlations between oculomotor functioning and schizophrenia-related characteristics in relatives of schizophrenic probands. *Psychophysiology* 1991;28:570–578.

92. Clementz BA, Grove WM, Iacono WG, et al. Smooth-pursuit eye movement dysfunction and liability for schizophrenia: implications for genetic modeling. *J Abnorm Psychol* 1992;101:117–129.

93. Thaker GK, Cassady S, Adami H, et al. Eye movements in spectrum personality disorders: comparison of community subjects and relatives of schizophrenic patients. *Am J Psychiatry* 1996;153:362–368.

94. Thaker GK, Ross DE, Cassady SL, et al. Smooth pursuit eye movements to extraretinal motion signals: deficits in relatives of patients with schizophrenia. *Arch Gen Psychiatry* 1998;55:830–836.

95. Siever LJ, Coursey RD, Alterman IS, et al. Clinical, psychophysiological, and neurological characteristics of volunteers with impaired smooth pursuit eye movements. *Biol Psychiatry* 1989;26:35–51.

96. O’Driscoll GA, Lenzenweger MF, Holzman PS. Antisaccades and smooth pursuit eye tracking and schizotypy. *Arch Gen Psychiatry* 1998;55:837–843.

97. Kelley MP, Bakan P. Eye tracking in normals: SPEM asymmetries and association with schizotypy. *Int J Neurosci* 1999;98:27–81.

98. Clementz BA, McDowell JE, Zisook S. Saccadic system functioning among schizophrenia patients and their first-degree biological relatives. *J Abnorm Psychol* 1994;103:277–287.

99. Sereno AB, Holzman, PS. Antisaccades and smooth pursuit eye movements in schizophrenia. *Biol Psychiatry* 1995;37:394–401.

100. Maruff P, Danckert J, Pantelis C, et al. Saccadic and attentional abnormalities in patients with schizophrenia. *Psychol Med* 1998;28:1091–1100.

101. McDowell JE, Myles-Worsley M, Coon H, et al. Measuring liability for schizophrenia using optimized antisaccade stimulus parameters. *Psychophysiology* 1999;36:138–141.

102. Myles-Worsley M, Coon H, McDowell J, et al. Linkage of a composite inhibitory phenotype to a chromosome 22q locus in eight Utah families. *Am J Med Genet* 1999;88:544–550.

103. Crawford TJ, Sharma T, Puri BK, et al. Saccadic eye movements in families multiply affected with schizophrenia: the Maudsley Family Study. *Am J Psychiatry* 1998;155:1703–1710.

104. Gooding DC. Antisaccade task performance in questionnaire-identified schizotypes. *Schizophr Res* 1999;35:157–166.

105. Braff DL, Heaton R, Kuck J, et al. The generalized pattern of neuropsychological deficits in outpatients with chronic schizophrenia with heterogeneous Wisconsin Card Sorting Test results. *Arch Gen Psychiatry* 1991;48:891–898.

106. Heaton R, Paulsen JS, McAdams LA, et al. Neuropsychological deficits in schizophrenics: relationship to age, chronicity, and dementia. *Arch Gen Psychiatry* 1994;51:469–476.

107. Kremen WS, Seidman LJ, Pepple JR, et al. Neuropsychological risk indicators for schizophrenia: a review of family studies. *Schizophr Bull* 1994;20:103–119.

108. Heaton RK. *A manual for the Wisconsin Card Sort Test*. Odessa, FL: Psychological Assessment Resources, 1981.

109. Gold J, Carpenter C, Randolph C, et al. Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Arch Gen Psychiatry* 1997;54:159–165.

110. Holzman PS, Shenton ME, Solovay MR. Quality of thought disorder in differential diagnosis. *Schizophr Bull* 1986;12:360–372

111. Perry W, Viglione D, Braff DL. The Ego Impairment Index and schizophrenia: a validation study. *J Pers Assess* 1992;59:165–175.

112. Franke P, Maier W, Hain C, et al. Wisconsin Card Sorting Test: an indicator of vulnerability to schizophrenia? *Schizophr Res* 1992;6:243–249.

113. Scarone S, Abbruzzese M, Gambini O. The Wisconsin Card Sorting Test discriminates schizophrenic patients and their siblings. *Schizophr Res* 1993;10:103–107.

114. Battaglia M, Abbruzzese M, Ferri S, et al. An assessment of the Wisconsin Card Sorting Test as an indicator of liability to schizophrenia. *Schizophr Res* 1994;14:39–45.

115. Cannon TD, Zorrilla LE, Shtasel D, et al. Neuropsychological functioning in siblings discordant for schizophrenia and healthy volunteers. *Arch Gen Psychiatry* 1994;51:651–661.

116. Keefe RS, Silverman JM, Roitman SE, et al. Performance of nonpsychotic relatives of schizophrenic patients on cognitive tests. *Psychiatry Res* 1994;53:1–12.

117. Kremen WS, Goldstein JM, Seidman LJ, et al. Sex differences in neuropsychological function in non-psychotic relatives of schizophrenic probands. *Psychiatry Res* 1997;66:131–144.

118. Toomey R, Faraone SV, Seidman LJ, et al. Association of neuropsychological vulnerability markers in relatives of schizophrenic patients. *Schizophr Res* 1998;31:89–98.

119. Faraone SV, Seidman LJ, Kremen WS, et al. Neuropsychological functioning among the nonpsychotic relatives of schizophrenic patients: a 4-year follow-up study. *J Abnorm Psychol* 1999;108:176–181.

120. Kinney DK, Holzman PS, Jacobsen B, et al. Thought disorder in schizophrenic and control adoptees and their relatives. *Arch Gen Psychiatry* 1997;54:475–479.

121. Park S, Holzman PS, Goldman-Rakic PS. Spatial working memory deficits in the relatives of schizophrenic patients. *Arch Gen Psychiatry* 1995;52:821–828.

122. Park S, Holzman PS, Lenzenweger MF. Individual differences in spatial working memory in relation to schizotypy. *J Abnorm Psychol* 1995;104:355–363.

123. Trestman RL, Keefe RS, Mitropoulou V, et al. Cognitive function and biological correlates of cognitive performance in schizotypal personality disorder. *Psychiatry Res* 1995;59:127–136.

124. Voglmaier MM, Seidman LJ, Salisbury D, et al. Neuropsychological dysfunction in schizotypal personality disorder: a profile analysis. *Biol Psychiatry* 1997;41:530–540.

125. Suhr JA. Executive functioning deficits in hypothetically psychosis-prone college students. *Schizophr Res* 1997;27:29–35.

126. Cadenhead KS, Perry W, Shafer K, et al. Cognitive functions in schizotypal personality disorder. *Schizophr Res* 1999;37:123–132.

127. Goldberg TE, Torrey EF, Gold JM, et al. Genetic risk of neuropsychological impairment in schizophrenia: a study of monozygotic twins discordant and concordant for the disorder. *Schizophr Res* 1995;17:77–84.

128. Weinberger DR, Berman KF, Zec RF. Physiologic dysfunction of dorsolateral prefrontal cortex in schizophrenia. I. Regional cerebral blood flow evidence. *Arch Gen Psychiatry* 1986;43:114–124.

129. Weinberger DR, Berman KF, Illowsky BP. Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia. III. A new cohort and evidence for a monoaminergic mechanism. *Arch Gen Psychiatry* 1988;45:609–615.

130. Berman KF, Illowsky BP, Weinberger DR. Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia. IV. Further evidence for regional and behavioral specificity. *Arch Gen Psychiatry* 1988;45:616–622.

131. Perry W, Swerdlow N R, McDowell JE, et al. Schizophrenia and frontal lobe functioning: evidence from neuropsychology, cognitive neuroscience, and psychophysiology. In: Miller BL, Cummings JL, eds. *The human frontal lobes: functions and disorders*. New York: The Guilford Press, 1999:509–521.

132. Perry W, Heaton RK, Potterat E, et al. Working memory in schizophrenia: transient “on line” storage versus executive functioning. *Schizophr Bull* 2001;27:157–176.

133. Cornblatt BA, Lenzenweger MF, Erlenmeyer-Kimling L. The continuous performance test, identical pairs version: II. Contrasting attentional profiles in schizophrenic and depressed patients. *Psychiatry Res* 1989;29:65–85.

134. Cornblatt BA, Keilp JG. Impaired attention, genetics, and the pathophysiology of schizophrenia. *Schizophr Bull* 1994;20:31–46.

135. Nuechterlein KH, Dawson ME, Green MF. Information-processing abnormalities as neuropsychological vulnerability indicators for schizophrenia. *Acta Psychiatr Scand Suppl* 1994;384:71–79.

136. Siegel BV, Nuechterlein KH, Abel L, et al. Glucose metabolic correlates of continuous performance test performance in adults with a history of infantile autism, schizophrenics, and controls. *Schizophr Res* 1995;17:85–94.

137. Rutschmann J, Cornblatt B, Erlenmeyer-Kimling L. Sustained attention in children at risk for schizophrenia. Report on a continuous performance test. *Arch Gen Psychiatry* 1977;34:571–575.

138. Walker E, Shaye J. Familial schizophrenia. A predictor of neuromotor and attentional abnormalities in schizophrenia. *Arch Gen Psychiatry* 1982;39:1153–1156.

139. Nuechterlein KH, Asarnow RF, Subotnik KL, et al. Neurocognitive vulnerability factors for schizophrenia: convergence across genetic risk studies and longitudinal trait/state studies. In: Lenzenweger MF, Dworkin RH, eds. *Origins and development of schizophrenia: advances in experimental psychopathology*. Washington, DC: American Psychological Association, 1998:299–327.

140. Franke P, Maier W, Hardt J, et al. Attentional abilities and measures of schizotypy: their variation and covariation in schizophrenic patients, their siblings, and normal control subjects. *Psychiatry Res* 1994;54:259–272.

141. Keefe RS, Silverman JM, Mohs RC, et al. Eye tracking, attention, and schizotypal symptoms in nonpsychotic relatives of patients with schizophrenia. *Arch Gen Psychiatry* 1997;54:169–176.

142. Finkelstein JR, Cannon TD, Gur RE, et al. Attentional dysfunctions in neuroleptic-naïve and neuroleptic-withdrawn schizophrenic patients and their siblings. *J Abnorm Psychol* 1997;106:203–212.

143. D’Amato T, Saoud M, Triboulet P, et al. Vulnerability to schizophrenia. I: Familial nature of neuropsychologic indicators. *Encephale* 1998;24:442–448.

144. Chen WJ, Liu SK, Chang CJ, et al. Sustained attention deficit and schizotypal personality features in nonpsychotic relatives of schizophrenic patients. *Am J Psychiatry* 1998;155:1214–1220.

145. Laurent A, Saoud M, Bougerol T, et al. Attentional deficits in patients with schizophrenia and in their non-psychotic first-degree relatives. *Psychiatry Res* 1999;89:147–159.

146. Lenzenweger MF, Cornblatt BA, Putnick M. Schizotypy and sustained attention. *J Abnorm Psychol* 1991;100:84–89.

147. Roitman SE, Cornblatt BA, Bergman A, et al. Attentional functioning in schizotypal personality disorder. *Am J Psychiatry* 1997;154:655–660.

148. Chen WJ, Hsiao CK, Lin CC. Schizotypy in community samples: the three-factor structure and correlation with sustained attention. *J Abnorm Psychol* 1997;106:649–654.

149. Asarnow RF, Granholm E, Sherman T. Span of apprehension in schizophrenia. In: Steinhauer SR, Gruzelier JH, Zubin J, eds. *Neuropsychology, psychophysiology and information processing* (*Handbook of schizophrenia*, vol 5). Amsterdam: Elsevier Science, 1991.

150. Wagener DK, Hogarty GE, Goldstein MJ, et al. Information processing and communication deviance in schizophrenic patients and their mothers. *Psychiatry Res* 1986;18:365–377.

151. Maier W, Franke P, Hain C, et al. Neuropsychological indicators of the vulnerability to schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 1992;16:703–715.

152. D’Amato T, Saoud M, Triboulet P, et al. Vulnerability to schizophrenia. I: Familial nature of neuropsychologic indicators. *Encephale* 1998;24:442–448.

153. Asarnow RF, Nuechterlein KH, Marder SR. Span of apprehension performance, neuropsychological functioning, and indices of psychosis-proneness. *J Nerv Ment Dis* 1983;171:662–669.

154. Williams LM. Further evidence for a multidimensional personality disposition to schizophrenia in terms of cognitive inhibition. *Br J Clin Psychol* 1995;34:193–213.

155. Bartfai A, Pedersen NL, Asarnow RF, et al. Genetic factors for the Span of Apprehension Test: a study of normal twins. *Psychiatry Res* 1991;38:115–124.

156. Braff DL, Saccuzzo DP. Information processing dysfunction in paranoid schizophrenia: a two factor deficit. *Am J Psychiatry* 1981;138:1051–1056.

157. Braff DL, Saccuzzo DP. Effect of antipsychotic medication on speed of information processing in schizophrenic patients. *Am J Psychiatry* 1982;139:1127–1130.

158. Braff DL, Saccuzzo DP. The time course of information-processing deficits in schizophrenia. *Am J Psychiatry* 1985;142:170–174.

159. Green MF, Nuechterlein KH, Breitmeyer B. Backward masking performance in unaffected siblings of schizophrenic patients. *Arch Gen Psychiatry* 1997;54:465–472.

160. Green MF, Nuechterlein KH. Backward masking performance as an indicator of vulnerability to schizophrenia. *Acta Psychiatr Scand* 1999;395:34–40.

161. Braff DL, Saccuzzo DP, Geyer MA. Information processing dysfunction in schizophrenia: studies of visual backward masking, sensorimotor gating and habituation. In: Zubin J, Steinhauer S, Gruzelier JH, eds. *Neuropsychology, psychophysiology, and information processing* (*Handbook of schizophrenia*, vol 4). Amsterdam: Elsevier Science, 1991:303–334.

162. Turvey MT. On peripheral and central processes in vision: inferences from an information-processing analysis of masking with patterned stimuli. *Psychol Rev* 1973;80:1–52.

163. Green MF, Nuechterlein KH, Mintz J. Backward masking in schizophrenia and mania. II. Specifying a mechanism. *Arch Gen Psychiatry* 1994;51:939–944.

164. Breitmeyer BG, Ganz L. Implications for sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. *Psychol Rev* 1976,83:1–36.

165. Saccuzzo DP, Braff DL. Early information processing deficit in schizophrenia: new findings using RDC schizophrenic subgroups and manic controls. *Arch Gen Psychiatry* 1981;38:175–179.

166. Green M, Walker E. Symptom correlates of vulnerability to backward masking in schizophrenia. *Am J Psychiatry* 1986;143:181–186.

167. Rund BR. Backward-masking performance in chronic and nonchronic schizophrenics, affectively disturbed patients, and normal control subjects. *J Abnorm Psychol* 1993;102:74–81.

168. Cadenhead KS, Geyer MA, Butler RW, et al. Information processing deficits of schizophrenia patients: relationship to clinical ratings, gender and medication status. *Schizophr Res* 1997;28:51–62.

169. Lieb K, Denz E, Hess R, et al. Preattentive information processing as measured by backward masking and text on detection tasks in adolescents at high genetic risk for schizophrenia. *Schizophr Res* 1996;21:171–182.

170. Braff DL. Impaired speed of information processing in nonmedicated schizotypal patients. *Schizophr Bull* 1981;7:499–508.

171. Saccuzzo DP, Braff DL, Sprock J, et al. The schizophrenia spectrum: a study of the relationship among the Rorschach, MMPI, and visual backward masking. *J Clin Psychol* 1984;40:1288–1294.

172. Merritt RD, Balogh DW. Backward masking as a function of spatial frequency. A comparison of MMPI-identified schizotypics and control subjects. *J Nerv Ment Dis* 1990;178:186–193.

173. Cadenhead KS, Perry W, Braff DL. The relationship of information-processing deficits and clinical symptoms in schizotypal personality disorder. *Biol Psychiatry* 1996;40:853–858.

174. Saccuzzo DP, Braff DL. Early information processing deficit in schizophrenia: new findings using RDC schizophrenic subgroups and manic controls. *Arch Gen Psychiatry* 1981;38:175–179.

175. Green MF, Nuechterlein KH, Mintz J. Backward masking in schizophrenia and mania. II. Specifying the visual channels. *Arch Gen Psychiatry* 1994;51:945–951.

176. Cadenhead KS, Serper Y, Braff DL. Transient versus sustained visual channels in the visual backward masking deficits of schizophrenia patients. *Biol Psychiatry* 1998;43:132–138.

177. Callaway E. Averaged evoked responses in psychiatry. *J Nerv Ment Dis* 1966;143:80–91

178. Callaway E, Jones RT, Donchin E. Auditory evoked potential variability in schizophrenia. *Electroencephalogr Clin Neurophysiol* 1970;29:421–428.

179. Jones RT, Callaway E. Auditory evoked responses in schizophrenia—a reassessment. *Biol Psychiatry* 1970;2:291–298.

180. Donchin E, Callaway E, Jones RT. Auditory evoked potential variability in schizophrenia. II. The application of discriminant analysis. *Electroencephalogr Clin Neurophysiol* 1970;29:429–440.

181. Braff DL, Callaway E, Naylor H. Very short-term memory dysfunction in schizophrenia. Defective short time constant information processing in schizophrenia. *Arch Gen Psychiatry* 1977;34:25–30.

182. McCarley RW, Shenton ME, O'Donnell BF, et al. Auditory P300 abnormalities and left posterior superior temporal gyrus volume reduction in schizophrenia. *Arch Gen Psychiatry* 1993;50:190–197.

183. Ford JM, White P, Lim KO, et al. Schizophrenics have fewer and smaller P300s: a single-trial analysis. *Biol Psychiatry* 1994;35:96–103.

184. Strik WK, Dierks T, Franzek E, et al. P300 in schizophrenia: interactions between amplitudes and topography. *Biol Psychiatry* 1994;35:850–856.

185. Turetsky B, Colbath EA, Gur RE. P300 subcomponent abnormalities in schizophrenia: II. Longitudinal stability and relationship to symptom change. *Biol Psychiatry* 1998;43:31–39.

186. Roxborough H, Muir WJ, Blackwood DH, et al. Neuropsychological and P300 abnormalities in schizophrenics and their relatives. *Psychol Med* 1993;23:305–314.

187. Friedman D, Squires-Wheeler E. Event-related potentials (ERPs) as indicators of risk for schizophrenia. *Schizophr Bull* 1994;20:63–74.

188. Schreiber H, Stolz-Born G, Kornhuber HH, et al. Electrophysiologic correlates of selective attention in children and adolescents at increased risk of schizophrenia. *Z Kinder Jugenpsychiatr* 1996;24:282—292.

189. D'Amato T, Karoumi B, Rosenfeld F, et al. Vulnerability to schizophrenia. II: Familial status of auditory evoked potential abnormalities. *Encephale* 1999;25:288–295.

190. Weisbrod M, Hill H, Niethammer R, et al. Genetic influence on auditory information processing in schizophrenia: P300 in monozygotic twins. *Biol Psychiatry* 1999;46:721–725.

191. Salisbury DF, Voglmaier MM, Seidman LJ, et al. Topographic abnormalities of P3 in schizotypal personality disorder. *Biol Psychiatry* 1996;40:165–172.

192. Trestman RL, Horvath T, Kalus O, et al. Event-related potentials in schizotypal personality disorder. *J Neuropsychiatry Clin Neurosci* 1996;8:33–40.

193. Klein C, Berg P, Rockstroh B, et al. Topography of the auditory P300 in schizotypal personality. *Biol Psychiatry* 1999;45:1612–1621.

194. Morstyn R, Duffy FH, McCarley RW. Altered P300 topography in schizophrenia. *Arch Gen Psychiatry* 1983;40:729–734.

195. Salisbury DF, Shenton ME, Sherwood AR, et al. First-episode schizophrenic psychosis differs from first-episode affective psychosis and controls in P300 amplitude over left temporal lobe. *Arch Gen Psychiatry* 1998;55:173–180.

196. Salisbury DF, Shenton ME, McCarley RW. P300 topography differs in schizophrenia and manic psychosis. *Biol Psychiatry* 1999;45:98–106.

197. Javitt DC, Grochowski S, Shelley AM, et al. Impaired mismatch negativity (MMN) generation in schizophrenia as a function of stimulus deviance, probability, and interstimulus/interdeviant interval. *Electroencephalogr Clin Neurophysiol* 1998;108:143–153.

198. Kelsoe JR, Sadovnick AD, Kristbjarnarson H, et al. Possible locus for bipolar disorder near the dopamine transporter on chromosome 5. *Am J Med Genet* 1996;67:533–540.

199. Moldin S. Report of the National Institute of Mental Health's genetics workshop: summary of research. *Biol Psychiatry* 1999;45:573–602.

200. Barondes SH. Report of the National Institute of Mental Health's genetics workshop: introduction. *Biol Psychiatry* 1999;45:559.

201. Freedman R, Adler LE, Leonard S. Alternative phenotypes for the complex genetics of schizophrenia. *Biol Psychiatry* 1999;45:551–558.

202. Hyman SE. Introduction to the complex genetics of mental disorders. *Biol Psychiatry* 1999;45:518–521.

203. Cadenhead KS, Light GA, Geyer MA, et al. Sensory gating deficits assessed by the P50 event-related potential in subjects with schizotypal personality disorder. *Am J Psychiatry* 2000;157:55–59.

204. Holzman PS, Coleman M, Lenzenweger MF, et al. Working memory deficits, antisaccades, and thought disorder in relation to perceptual aberration. In: Raine A, Lencz T, et al., eds. *Schizotypal personality*. New York: Cambridge University Press, 1995:353–381.

205. Cadenhead KS, Shafer K, Perry W, et al. Working memory in the schizophrenia spectrum. *Biol Psychiatry* 1998;43:127S.

206. Kinney DK, Holzman PS, Jacobson B, et al. Thought disorder in schizophrenic and control adoptees and their relatives. *Arch Gen Psychiatry* 1997;475–479.

207. Perry W, Braff DL. Thought disorder and the group of schizophrenia. *Schizophr Res* 1993;9:107.

208. Perry W, Cadenhead KS, Braff DL. Thought disorder in the group of schizophrenias. *Biol Psychiatry* 1992;31:117–118A.

209. Ito M, Kanno M, Mori Y, et al. Attention deficits assessed by Continuous Performance Test and Span of Apprehension Test in Japanese schizophrenic patients. *Schizophr Res* 1997;23:205–211.

210. Buchanan RW, Strauss ME, Breier A, et al. Attentional impairments in deficit and nondeficit forms of schizophrenia. *Am J Psychiatry* 1997;154:363–370.

211. Miller MB, Chapman LJ, Chapman JP, et al. Schizophrenic deficit in span of apprehension. *J Abnorm Psychol* 1990;99:313–316.

212. Elkins IJ, Cromwell RL, Asarnow RF. Span of apprehension in schizophrenic patients as a function of distractor masking and laterality. *J Abnorm Psychol* 1992;101:53–60.

213. Saccuzzo DS, Cadenhead KS, Braff DL. Backward versus forward visual masking deficits in schizophrenic patients: centrally, not peripherally, mediated. *Am J Psychiatry* 1996;153:1564–1570.

214. Cadenhead KS. Unpublished data (personal communication), 2000.

215. Frangou S, Sharma T, Alarcon G, et al. The Maudsley Family Study, II: Endogenous event-related potentials in familial schizophrenia. *Schizophr Res* 1997;23:45–53.

216. Maier W, Franke P, Kopp B, et al. Reaction time paradigms in subjects at risk for schizophrenia. *Schizophr Res* 1994;13:35–43.

217. Sarkin AJ, Dionisio DP, Hillix WA, et al. Positive and negative schizotypal symptoms relate to different aspects of crossover reaction time task performance. *Psychiatry Res* 1998;81:241–249.

52

Neurochemistry of Schizophrenia: Glutamatergic Abnormalities

James H. Meador-Woodruff

Joel E. Kleinman

James H. Meador-Woodruff: Department of Psychiatry, University of Michigan, Ann Arbor, Michigan.

Joel E. Kleinman: Clinical Brain Disorders Branch, National Institutes of Mental Health Neuroscience Center, Washington, DC.

Multiple neurotransmitters have been implicated in schizophrenia. Dopamine is the neurotransmitter most often hypothesized to be associated with the pathophysiology of schizophrenia for two reasons. First, dopaminergic agonists can cause or exacerbate psychotic symptoms. Second, the correlation between antipsychotic efficacy and D2 dopamine-receptor blockade is excellent. For these reasons, a number of postmortem studies have focused on the dopaminergic system in schizophrenic brain. Although the results of these studies have generally been negative, the few positive findings have rarely been replicated, with the notable exception of increased striatal D2-receptor expression, which may be secondary to prior neuroleptic treatment. These studies of dopaminergic abnormalities in postmortem brain in schizophrenia have been recently reviewed (1 ,2).

Given the lack of findings associated with the dopamine system in the brain in schizophrenia, the elucidation of other potential neurotransmitter substrates of this illness has been an area of recent investigation. Glutamatergic dysfunction has been hypothesized to occur in schizophrenia, and this has been one of the most active areas of neurotransmitter research in this illness during the past few years. In this chapter, the glutamate hypothesis of schizophrenia is reviewed, the complexity of the molecules associated with the glutamate synapse is outlined, and postmortem neurochemical data suggesting glutamatergic abnormalities in schizophrenia are presented.

GLUTAMATE AND SCHIZOPHRENIA

Part of "52 - Neurochemistry of Schizophrenia: Glutamatergic Abnormalities "

Several lines of evidence have implicated glutamatergic dysfunction in schizophrenia. Dissociative anesthetics, especially phencyclidine (PCP) and ketamine, can cause psychotic symptoms in normal humans (3 ,4), and worsen these symptoms in persons with schizophrenia (5 ,6 and 7). Unlike catecholamine agonists, PCP can produce both the positive and negative (deficit) symptoms associated with this illness. PCP and related compounds are uncompetitive inhibitors of the *N*-methyl-D-aspartate (NMDA) subtype of glutamate receptor. Hence, this pharmacologic literature has been interpreted as suggesting that schizophrenia may be associated with decreased NMDA-receptor activity (5 ,8).

Several other reasons make a glutamate-receptor hypothesis of schizophrenia attractive. Schizophrenia is believed to have a neurodevelopmental component, and the NMDA receptor is critical in guiding axons to their targets in development (9). Further, NMDA receptors may be important in processes that lead to synaptic pruning seen in adolescence, which has been hypothesized to be abnormal in schizophrenia (10). Cognitive functioning depends on the plasticity mediated in part by NMDA receptors, and schizophrenics often have cognitive deficits (11). Finally, the reduction of gray matter in several brain regions seen in schizophrenia has been suggested to be the result of neurotoxicity mediated by NMDA receptors (12). A constellation of symptoms, findings, and hypotheses of schizophrenia can be parsimoniously explained by NMDA-receptor dysfunction.

The NMDA receptor is one of multiple subtypes of the glutamate receptor, however, and all these subtypes have functional interrelationships. Thus, although NMDA-receptor abnormalities have been hypothesized in schizophrenia, apparent NMDA-receptor dysregulation could be associated with abnormalities of another receptor subtype that interacts with the NMDA receptor, which in turn results in a breakdown of normal glutamatergic transmission in schizophrenia.

GLUTAMATE-RECEPTOR SUBTYPES

Part of "52 - Neurochemistry of Schizophrenia: Glutamatergic Abnormalities "

The four classes of glutamate receptors are functionally and pharmacologically distinct (Fig. 52.1 and Fig. 52.2). The ionotropic

glutamate receptors, AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid), kainate, and NMDA are each composed of four or five subunits that form ligand-gated ion channels. The metabotropic glutamate receptors (mGluRs) are all seven transmembrane-domain, G protein-coupled receptors (13 ,14).

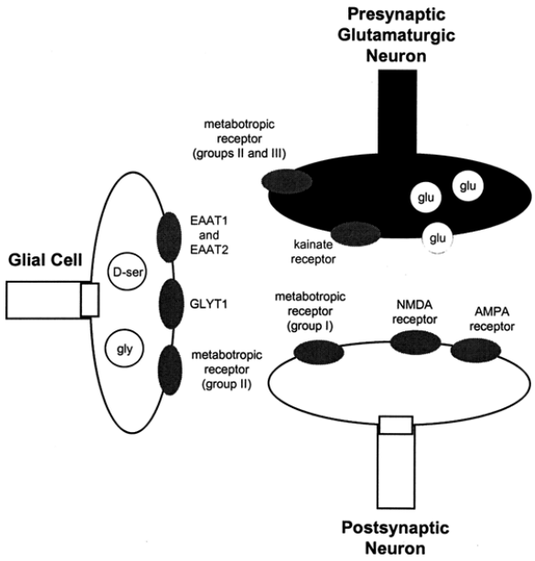


FIGURE 52.1. Diagram of a typical glutamatergic synapse. Recent data suggest that glutamatergic transmission requires three cells: a presynaptic glutamate-releasing cell, a presynaptic glial cell that releases the endogenous agonist for the glycine co-agonist site (recently reported to be D-serine), and a postsynaptic neuron. The various glutamate receptors and transporters are differentially expressed by these three distinct cell populations. The glutamate uptake transporter $EAAT_3$ (excitatory amino acid transporter 3), which is not shown on this figure, appears to be expressed primarily on the cell body and dendrites.

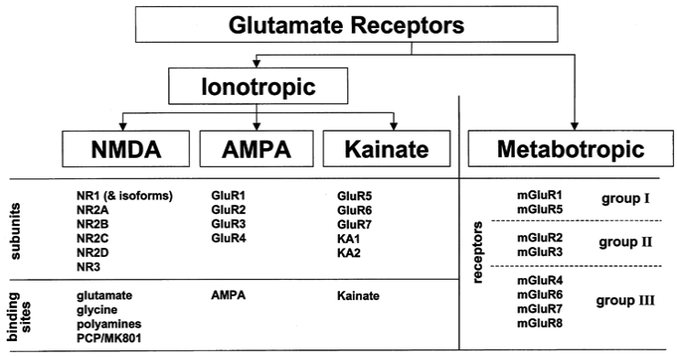


FIGURE 52.2. Subtypes of glutamate receptors. Three families of ionotropic glutamate receptors (*N*-methyl-D-aspartate, AMPA, and kainate) are known, each of which is composed of distinct subunits and has identifiable binding sites. The metabotropic receptors cluster into three groups, members of which share pharmacologic and structural features.

The AMPA-receptor subunits are derived from a family of four genes that have been named *GluR1* through *GluR4*. The transcripts from each of these genes are expressed in one of two isoforms, termed *flip* and *flop*, that result from alternative splicing. In addition, the final subunit protein of the AMPA receptor subunits has amino acids at specific locations in the ion channel that can vary according to RNA editing (13 ,14). Thus, a potential exists for considerable heterogeneity in the final assembled AMPA receptors, based on subunit composition and post-translational modification. The assembled AMPA receptors contain several binding sites: one for glutamate, another at which competitive antagonists such as CNQX (6-cyano-7-nitro-quinoxalindione) act, and yet another where desensitization modulators exert their influence. Subunit composition appears to confer unique pharmacologic properties to the final receptors (15 ,16 ,17 ,18 and 19). For example, decreased calcium influx in AMPA receptors that contain the $GluR_2$ subunit drastically diminishes the electrophysiologic activity of these receptors.

Kainate receptors are also ligand-gated ion channels composed of subunits derived from genes for the low-affinity $GluR_5$ through $GluR_7$ and high-affinity KA_1 through KA_2 subunits (13 ,14). The transcripts associated with these five subunits also undergo alternative splicing and editing. Final assembled kainate receptors may be composed of five identical subunits, or they may be heteromers composed of low- and high-affinity subunits, with pharmacologic properties that differ from those of low-affinity or high-affinity homomers.

The NMDA receptor subunits are encoded by five genes termed *NR1* and *NR2A* through *NR2D* (13 ,14). An *NR3* gene has also been identified, although this subunit appears to be expressed primarily during early development (20 ,21 and 22). NR_1 is expressed as one of eight isoforms because of alternative splicing of exons 5, 21, and 22 (13 ,14 ,23 ,24). As in the case of the AMPA and kainate receptors, transcription of the NR_1 subunit presents an important level for the regulation of the expression of functional NMDA receptors. This regulation can influence certain properties of the final functional NMDA receptors, including the pharmacology of their binding sites.

The pharmacologic regulation of the NMDA receptor depends on the unique combination of binding sites (13 ,14). A primary agonist site exists for the binding of glutamate. A separate glycine co-agonist site must also be occupied before glutamate can activate the ion channel; recent reports suggest that D-serine produced by astrocytes is the endogenous ligand for this site (25 ,26 ,27 and 28). Modulatory binding sites for polyamines, protons, neuropeptides including dynorphin, and zinc have also been identified. Additionally, magnesium ions block the ion channel of the NMDA receptor complex at physiologic concentrations. This blockade is voltage-dependent; partial depolarization of the cell membrane extrudes the magnesium ion. Therefore, presynaptic glutamate release and postsynaptic pre-depolarization are both required for NMDA receptor activity. Finally, a site within the ion channel itself is associated with the binding of uncompetitive antagonists of the NMDA receptor, such as PCP, ketamine, and MK-801. These antagonists are use-dependent (i.e., the ion channel must be opened for these compounds to bind to the receptor), so cooperativity between multiple sites is necessary for occupancy by uncompetitive antagonists.

These binding sites are associated with different subunits, and their affinities can vary depending on subunit composition. NR_1 homomers have been shown to form glycine binding sites, but an NR_2 subunit appears to be required to form both glutamate and MK-801 binding sites (29 ,30 ,31 and 32). Further, receptors containing NR_{2A} subunits have a higher affinity for compounds that bind to the glutamate agonist site, whereas receptors with NR_{2A} or NR_{2B} subunits have higher affinities for MK-801 binding than do receptors with NR_{2C} or NR_{2D} subunits (31). In addition, NMDA receptors containing particular NR_1 splice variants have a higher affinity for MK-801 than do receptors with others, irrespective of NR_2 co-assembly (33). Receptors with NR_{2B} subunits are associated with a higher affinity for polyamine modulators (31 ,34). Therefore, differential subunit combinations confer unique binding properties to the NMDA receptors and probably are associated with subtle electrophysiologic differences within a population of NMDA receptors.

Eight mGluRs have been cloned and are grouped (group I, group II, and group III) based on pharmacology, sequence homology, and linkage to signal transduction pathways (35 ,36 ,37 ,38 ,39 and 40). These mGluRs belong to a unique subset of G protein-coupled receptors with seven transmembrane domains and large, extracellular amine termini. When expressed in heterologous systems, group I mGluRs have been shown to stimulate phospholipase C, phosphoinositide hydrolysis, and the formation of cyclic adenosine monophosphate (cAMP) (41 ,42 ,43 and 44). In heterologous systems, groups II and III mGluRs inhibit forskolin-stimulated cAMP formation and adenylyl cyclase, possibly via a G_i protein (39 ,40 ,45 ,46). The metabotropic receptors have been the target of considerable recent interest because a functional relationship appears to exist between the group II metabotropic and NMDA receptors (47).

Each glutamate receptor subtype appears to have a unique role in glutamatergic neurotransmission. Glutamate receptors interact at multiple levels, as AMPA, kainate, and metabotropic receptors all affect NMDA-receptor activity. Accordingly, although the NMDA receptor is typically hypothesized to be dysregulated in schizophrenia, disturbances of any of the glutamate receptors could result in a condition

that produces the appearance of an abnormally functioning NMDA receptor.

ABNORMALITIES OF GLUTAMATE RECEPTORS IN SCHIZOPHRENIA

Part of "52 - Neurochemistry of Schizophrenia: Glutamatergic Abnormalities "

Given the possibility of glutamate-receptor dysfunction in schizophrenia, the expression of all four families of the glutamate receptor have been studied in schizophrenic brain. As would be expected, these investigations have primarily targeted limbic regions that have been implicated in schizophrenia, particularly limbic cortex, striatal areas, medial temporal lobe structures, and, more recently, the thalamus. These investigations have also targeted multiple levels of gene expression, including subunit messenger RNA (mRNA) and protein levels, and final binding sites have been studied. In the following sections, the studies that have been published for each receptor subtype in postmortem brain in schizophrenia are reviewed.

AMPA Receptors

Of all of the glutamate receptors in schizophrenia, the AMPA receptor has been studied the most, as summarized in Table 52.1 . When the AMPA-associated subunits were first cloned, Harrison et al. (48) examined the expression of the mRNA encoding the $GluR_1$ subunit in medial temporal lobe structures in schizophrenia. A consistent decrease in the expression of this subunit transcript was found in hippocampal regions, an abnormality that was statistically significant only in the CA3 region. These investigators subsequently extended their finding and demonstrated that $GluR_1$-subunit mRNA is decreased in multiple hippocampal subfields (dentate gyrus, CA3, and CA4) and also in the subiculum (49). They also reported that $GluR_2$-subunit mRNA is decreased in the medial temporal lobe in schizophrenia, particularly in the parahippocampal gyrus (49), and continued their examination of AMPA-receptor expression in the medial temporal lobe by determining the patterns of expression of the flip and flop isoforms of the $GluR_1$ and $GluR_2$ subunits. Decreased expression of $GluR_2$-subunit

mRNA was again found in hippocampal structures, and both the flip and flop variants were reduced, the flop to a greater extent (50).

TABLE 52.1. AMPA RECEPTOR BINDING AND SUBUNIT EXPRESSION IN SCHIZOPHRENIA

	Ligand or Subunit	Findings	Brain Regions Studied	Reference
Receptor binding sites	[^{3}H]CNQX		caudate	57
	[^{3}H]CNQX	none	putamen, nucleus accumbens	57
	[^{3}H]AMPA	none	caudate, putamen, nucleus accumbens	55
	[^{3}H]CNQX		CA4, CA3	53
	[^{3}H]AMPA	none	frontal cortex, putamen, nucleus accumbens	58
	[^{3}H]AMPA	none	caudate, putamen, nucleus, accumbens	56
	[^{3}H]AMPA		CA2	54
	[^{3}H]AMPA	none	dentate gyrus, CA1, CA3, subiculum	54
	[^{3}H]AMPA	none	thalamus	61
Subunit protein expression	$GluR_1$		parahippocampal gyrus	50
		none	CA1, CA3, CA4, subiculum	50
	$GluR_{2/3}$		CA4	50
		none	dentate gyrus, CA1, CA3, subiculum parahippocampal gyrus	50
	$GluR_1$	none	hippocampus	52
	$GluR_2$, $GluR_3$	none	cingulate cortex	52
Subunit mRNA expression	$GluR_1$		dentate gyrus, CA3, CA4, subiculum	49
		none	CA1, parahippocampal gyrus	49
	$GluR_2$		dentate gyrus, CA3, CA4, subiculum	49
		none	CA1	49
	$GluR_1$, $GluR_2$, $GluR_3$, $GluR_4$	none	caudate, putamen, nucleus accumbens	55
	$GluR_1$		CA3	48
		none	dentate gyrus, CA1, CA4, subiculum	48
	$GluR_1$, $GluR_2$, $GluR_3$, $GluR_4$	none	caudate, putamen, nucleus accumbens	56
	$GluR_1$, $GluR_3$		thalamus	61
	$GluR_2$, $GluR_4$	none	thalamus	61
	$GluR_1$		frontal cortex	59
	$GluR_1$	none	frontal cortex	59
	$GluR_2$ flip	none	hippocampus	51
	$GluR_2$ flop		hippocampus	51
	flip-flop ratio		hippocampus	51

AMPA, α-amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid; CNQZ, 6-cyano-7-nitro-quinoxalindione

Several studies have examined the expression of the AMPA-subunit proteins in the medial temporal lobe in schizophrenia. Using quantitative immunocytochemical analyses, Eastwood et al. (51) reported decreased expression of the AMPA subunits in medial temporal lobe structures. In particular, $GluR_1$ immunoreactivity was noted to be significantly reduced in the parahippocampal gyrus, and combined $GluR_{2/3}$ immunoreactivity was decreased in the CA4 subfield of the hippocampus. On the other hand, Breese and co-workers (52) found no differences in $GluR_1$, $GluR_2$, or $GluR_3$ immunoreactivity in schizophrenia when they used Western analysis in hippocampal samples.

AMPA-receptor binding has also been studied in medial temporal lobe structures. Using [^{3}H]CNQX to label the AMPA receptor, Kerwin et al. (53) noted decreased binding to the AMPA receptor in the schizophrenic hippocampus, particularly in the CA3 and CA4 subfields. More recently, Gao and colleagues (54) found decreased [^{3}H]AMPA binding in CA2, but not in other hippocampal fields or associated structures. The convergence of these data is that AMPA-receptor expression is decreased in the medial temporal lobe in schizophrenia, a decrease that involves alterations of subunit gene expression in addition to the final binding site.

Although the medial temporal lobe data are the most robust, AMPA-receptor expression has also been examined in other brain regions in schizophrenia. In two studies, none of the AMPA-associated subunit transcripts were changed in striatal subregions (caudate, putamen, and nucleus accumbens) in schizophrenia (55 ,56). To date, subunit protein levels have not been reported in striatal regions. Binding to the AMPA receptor has been determined in striatal regions, but results have not been consistent. Although Noga and colleagues (57) reported an increase in AMPA binding, determined with [^{3}H]CNQX, in caudate, putamen, and accumbens in schizophrenia, no differences in [^{3}H]AMPA binding were found in striatal regions in schizophrenia in three other reports (55 ,58 ,56).

The cortex has also been studied for alterations of AMPA-receptor expression in schizophrenia. In one study, no differences in the expression of any of the AMPA-associated subunit mRNAs were found in prefrontal or occipital cortex in schizophrenia (55), although Sokolov (59), using reverse transcriptase polymerase chain reaction (RT-PCR), reported decreased $GluR_1$ mRNA in superior frontal gyrus. Breese et al. (52) found no differences in $GluR_2$ or $GluR_3$ protein in cingulate cortex as determined by Western analysis. Several groups have studied [^{3}H]AMPA binding in cortical areas in schizophrenia (55 ,60), with generally negative results.

Recently, the neurochemical anatomy of the thalamus has become a subject of interest in schizophrenia research. The AMPA receptor is expressed in multiple nuclei of the human thalamus. In a recent report (61), although [^{3}H]AMPA binding was not different in limbic thalamic nuclei in schizophrenia, the transcripts encoding the $GluR_1$ and $GluR_3$ subunits were both found to be reduced in the face of normal levels of $GluR_2$ and $GluR_4$ mRNA. These results suggest that alterations in the stoichiometry of subunit composition may be associated with the AMPA receptor in the schizophrenic thalamus.

Kainate Receptors

The kainate receptor has been the subject of study in the brain in schizophrenia, as summarized in Table 52.2 . Although the medial temporal lobe has been the best-studied region in the schizophrenic brain for AMPA-receptor expression, fewer studies have systematically focused on the kainate receptor in these structures. Porter and colleagues (62) found decreased expression of $GluR_6$ and KA_2 mRNA in several hippocampal regions, results paralleling similar data for the AMPA subunits in the medial temporal lobe. In this same study, $GluR_6$ mRNA was not found to be changed in the schizophrenic cerebellum. Only one study to date has examined any of the kainate subunit proteins; $GluR_5$ was studied by Western analysis and was not changed in schizophrenic hippocampus (52), although the antisera used in this study cross-reacts with $GluR_6$ and $GluR_7$.

TABLE 52.2. KAINATE RECEPTOR BINDING AND SUBUNIT EXPRESSION IN SCHIZOPHRENIA

	Ligand or Subunit	Findings	Brain Regions Studied	Reference
Receptor binding sites	[^{3}H]kainate	none	caudate, putamen, nucleus accumbens	57
	[^{3}H]kainate	none	thalamus	61
	[^{3}H]kainate		CA4, CA3, CA2, CA1, dentate gyrus, parahippocampal gyrus	53
	[^{3}H]kainate	none	CA3, CA2, CA1, dentate gyrus, subiculum	54
	[^{3}H]kainate	none	prefrontal cortex	65
	[^{3}H]kainate		frontal cortex	64
	[^{3}H]kainate	none	temporal cortex	64
	[^{3}H]kainate		prefrontal cortex	60
	[^{3}H]kainate		prefrontal cortex	63
	[^{3}H]kainate	none	striatum, occipital cortex	63
	[^{3}H]kainate	none	caudate, putamen, nucleus accumbens	56
Subunit protein expression	$GluR_{5,6,7}$	none	hippocampus	52
Subunit mRNA expression	$GluR_6$		dentate gyrus, CA3	62
	$GluR_6$	none	cerebellum	62
	KA_2		dentate gyrus, CA2, CA3	62
	KA_2	none	CA4, CA1	62
	KA_2		thalamus	61
	$GluR_{5,6,7}$, KA_2	none	thalamus	61
	$GluR_7$		prefrontal cortex	63
	KA_2		prefrontal cortex	63
	$GluR_{5,6}$, KA_1	none	prefrontal cortex	63
	$GluR_{5,6,7}$, KA_1, KA_2	none	striatum	63
	$GluR_7$		frontal cortex	59
	$GluR_7$	none	frontal cortex	59
	KA_1		frontal cortex	59
	KA_1	none	frontal cortex	59
	$GluR_{5,6,7}$, KA_1, KA_2	none	caudate, putamen, nucleus accumbens	56

Kainate-receptor expression has been examined in multiple cortical regions. Sokolov (59) has published data suggesting that $GluR_7$- and KA_1-subunit transcripts are decreased in the superior frontal gyrus in schizophrenia, similar to the decreases this investigator noted for some of the subunits associated with the AMPA and NMDA receptors. In a recent study examining transcripts of kainate-receptor subunits in the prefrontal cortex (63), a shift in subunit stoichiometry was found in multiple cytoarchitectural regions of the prefrontal cortex, with increased expression of $GluR_7$ mRNA and decreased expression of KA_2 mRNA in the face of normal expression of the other kainate subunits. In this same study, no changes in transcripts of kainate-receptor subunits were noted in Brodmann area 17.

Several studies have examined the expression of transcripts of the kainate-receptor subunit in subcortical structures. Two reports (56 ,63) noted no alterations of these subunits in multiple striatal regions in schizophrenia. On the other hand, a recent study noted decreased levels of KA_2 mRNA but normal levels of other transcripts of kainate-receptor subunits in limbic thalamic nuclei in schizophrenia.

Kainate-receptor binding has been studied in multiple brain regions in schizophrenia by several independent groups. All these studies have used [^{3}H]kainate to label this receptor. In general, kainate-receptor binding has been reported to be altered in multiple cortical areas in schizophrenia (63 ,64 and 65). Data on the expression of kainate binding sites in medial temporal lobe structures are inconsistent;

one study reported decreased [^{3}H]kainate binding in the hippocampus and parahippocampal gyrus (53), but another found no differences in binding in medial temporal lobe structures (54). Although kainate-receptor binding has been reported to be abnormal in cortical structures, it has not been found to differ in subcortical regions in schizophrenia; [^{3}H]kainate is unchanged in both striatal subregions (56 ,57 ,63 ,65) and limbic thalamic nuclei (61) in this illness.

NMDA Receptors

Although most hypotheses of glutamatergic dysfunction in schizophrenia invoke the NMDA receptor, relatively few studies of this receptor subtype have been carried out to date (Table 52.3). Only several studies have been published that examine the expression of the NMDA subunits in schizophrenic brain, and all these have focused on mRNA levels. In a comprehensive examination of all the NMDA subunits in prefrontal cortex, Akbarian et al. (66) found no absolute differences between controls and schizophrenic patients for any of the NMDA subunits, but the contribution of NR_{2D} to the total pool of NR_2 transcripts was elevated in the schizophrenic patients. Recently, Gao et al. (54) found an altered stoichiometry of NMDA subunits in hippocampus, with decreased NR_1 and increased NR_{2B} mRNA expression but normal NR_{2A} expression, in schizophrenia. Several other studies have been published in which only the NR_1 transcript was measured; in one study, this molecule was reported to be decreased in superior temporal cortex (67), and in another, it was decreased in superior frontal cortex (59).

	Ligand or Subunit	Findings	Brain Regions Studied	Reference
Receptor binding sites	[^{3}H]MK-801	none	caudate, putamen, nucleus accumbens	57
	[^{3}H]MK-801		putamen	68
	[^{3}H]MK-801	none	frontal cortex, entorhinal cortex, hippocampus, amygdala	68
	[^{3}H]L-689, 560		caudate, putamen,	72
	[^{3}H]L-689, 560	none	nucleus accumbens	
	[^{3}H]L-689, 560		temporal cortex	30
	[^{3}H]L-689, 560	none	motor cortex	30
	[^{3}H]CGP39653	none	temporal cortex, motor cortex	30
	[^{3}H]ifenprodil		temporal cortex	30
	[^{3}H]ifenprodil	none	motor cortex	30
	[^{3}H]ifenprodil		thalamus	61
	[^{3}H]MDL105,519		thalamus	61
	[^{3}H]MK-801	none	thalamus	61
	[^{3}H]CGP39653	none	thalamus	
	[^{3}H]ifenprodil	none	caudate, putamen, nucleus accumbens	56
	[^{3}H]MDL105,519			
	[^{3}H]MK-801			
	[^{3}H]CGP39653			
Subunit mRNA expression	NR_{2D}	none	prefrontal cortex	66
	$NR_{1,2A,2B,2C}$	none	prefrontal cortex	66
	$NR_{1,2A,2B,2C,2D}$	none	cerebellum, parietotemporal cortex	66
	$NR_{1,2B,2C}$		thalamus	61
	$NR_{2A,2D}$	none	thalamus	61
	NR_1		temporal cortex	67
	NR_1		frontal cortex	59
	NR_1	none	frontal cortex	59
	NR_1		dentate gyrus, CA3	54
	NR_{2A}	none	hippocampus, subiculum	54
	NR_{2B}		CA2, CA3	54
	$NR_{1,2A,2B,2C,2D}$	none	caudate, putamen, nucleus accumbens	56

TABLE 52.3. NMDA RECEPTOR BINDING AND SUBUNIT EXPRESSION IN SCHIZOPHRENIA

Several recent studies have examined the expression of the NMDA receptor subunits in subcortical structures in schizophrenia. In one study (56), NR_1, NR_{2A}, NR_{2B}, NR_{2C}, and NR_{2D} mRNAs were measured in the caudate, putamen, and nucleus accumbens in schizophrenia; no significant differences were found in comparison with control striata. On the other hand, significant reductions of NR_1, NR_{2B}, and NR_{2C} transcripts (but not of NR_{2A} and NR_{2B} transcripts) were found in dorsomedial and anterior thalamic nuclei in this disorder (61).

Because of the myriad binding domains of the NMDA complex, studies of receptor binding are difficult to interpret and are subject to the selection of radioligand. Further, it has become apparent that certain subunit compositions are associated with specific binding sites, so it is possible that some but not all binding sites on the NMDA receptor are altered in schizophrenia. The best-studied of the

NMDA-associated sites is the ion channel/PCP site. In general, studies in which [^{3}H]MK-801 was used have been relatively unimpressive. In an early study (68), increased [^{3}H]MK-801 binding was reported in the schizophrenic putamen, but no differences were noted in frontal cortex or multiple medial temporal lobe regions, including the hippocampus, amygdala, and entorhinal cortex. A more recent study (57) found no differences in caudate, putamen, or nucleus accumbens. The ion channel site has also been studied with the ligand [^{3}H]TCP, and again minimal changes were noted. In one study (69), no changes were found in multiple cortical areas, putamen, or cerebellum. A subsequent report (70) observed no differences between controls and schizophrenic patients in hippocampus, amygdala, or polar frontal cortex (Brodmann area 10), but increased [^{3}H]TCP binding was noted in orbitofrontal cortex (Brodmann area 11) in the schizophrenic patients.

The other NMDA-associated binding sites have been studied more recently. The primary agonist site for glutamate has been studied with [^{3}H]glutamate in the hippocampus, and no differences have been found in schizophrenia (53 ,54). The glycine co-agonist site has also been studied. Using [^{3}H]glycine, Ishimaru and colleagues (71) reported increased binding in multiple cortical areas in schizophrenia. Recently, the glycine site was studied in striatum with [^{3}H]L-689,560, and increased binding was noted in putamen, but not caudate or accumbens, in schizophrenia (72).

Several comprehensive studies examining multiple binding sites associated with the NMDA receptor complex in subcortical structures have recently been reported. In one of them (56), binding to the glutamate (measured with [^{3}H]CGP39653) and glycine (measured with [^{3}H]MDL105,519) agonist sites, the intrachannel/PCP site ([^{3}H]MK-801), and the polyamine modulatory site ([^{3}H]ifenprodil) were determined in caudate, putamen, and nucleus accumbens in schizophrenia. In this study, no differences were noted between controls and schizophrenic subjects. On the other hand, a study in thalamus from this same group (61), in which the same ligands were used to label the four sites, found decreased expression of binding associated with the glycine and polyamine sites, but not the intrachannel/PCP site or glutamate binding domain in limbic nuclei, in schizophrenia. These changes in some but not all binding sites in the thalamus were also associated

with changes in the stoichiometry of the various NMDA-associated subunit transcripts in these nuclei.

Metabotropic Receptors

Very little has been published about this family of receptors in schizophrenic brain (Table 52.4). In one study, the mRNAs encoding the metabotropic receptors $mGluR_3$ and $mGluR_5$ were measured in prefrontal cortex (73). Although $mGluR_3$ mRNA was not changed in schizophrenia in multiple areas of the prefrontal cortex, $mGluR_5$ was increased in the orbitofrontal cortex (Brodmann area 11), but not in Brodmann areas 9 or 10. Cell-level analysis revealed that this increase was secondary to increased expression of $mGluR_5$ mRNA in pyramidal cells in lamina III of this area of prefrontal cortex. More recently, the expression of the transcripts encoding seven of the eight cloned metabotropic receptors was reported in schizophrenic and control thalamus (74). No differences were found in the expression of the mGluRs in six different thalamic nuclei in schizophrenia in this study.

Receptor	Findings	Brain Regions Studied	Reference
$mGluR_{1,2,3,4,5,7,8}$	none	thalamus	74
$mGluR_3$	none	prefrontal cortex	73
$mGluR_5$		prefrontal cortex	73

TABLE 52.4. METABOTROPIC GLUTAMATE RECEPTOR RNA EXPRESSION IN SCHIZOPHRENIA

GLUTAMATE TRANSPORTERS

Part of "52 - Neurochemistry of Schizophrenia: Glutamatergic Abnormalities "

In addition to the glutamate receptors, other molecules at the glutamate synapse are critical for normal glutamatergic neurotransmission (Fig. 52.1). At least five glutamate uptake transporters, excitatory amino transporter 1 ($EAAT_1$) through $EAAT_5$, are expressed in the glutamate synapse (75). $EAAT_1$ is predominantly expressed in astrocytes of the cerebellum, although expression is also significant in the forebrain. $EAAT_2$ is expressed in both astrocytes and neurons but has a more widespread distribution in the brain (76). Severe neuropathology and epilepsy develop in knockout mice for the *EAAT2* gene, which confirms its importance in normal glutamatergic function. $EAAT_3$ is a neuronal transporter expressed in multiple limbic regions. $EAAT_4$ expression is restricted to Purkinje cells of the cerebellum, and $EAAT_5$ is confined to the retina.

Although glutamate transporters affect the function of all four glutamate receptor subtypes, the glycine transporter family may specifically affect NMDA receptor-mediated activity. Glycine is an NMDA receptor co-agonist, and glycine transporter inhibitors affect normal NMDA-receptor function and reverse PCP-induced behaviors (77 ,78 ,79 ,80 and 81). The two families of glycine transporters are $GLYT_1$ and $GLYT_2$; three isoforms of $GLYT_1$ have overlapping expression in astrocytes throughout the human brain, whereas $GLYT_2$ is restricted to the hindbrain and spinal cord (82 ,83). By altering the availability of glutamate for its receptors, changes in the expression of the transporters may induce profound changes at the level of receptor function. Further, given that the NMDA receptor may depend on glycine as a co-agonist, abnormal synaptic levels of this amino acid may be associated with disturbed function of the NMDA receptor.

Initially, the quantification of glutamate uptake sites in schizophrenia preceded the identification of the EAAT subtypes, and conflicting data have been obtained in schizophrenic prefrontal cortex and basal ganglia with use of the nonselective transporter ligand [^{3}H]D-aspartate (Table 52.5). Early studies found decreases in striatal uptake sites (84 ,85); however, later studies did not replicate these findings (57 ,86). Similarly, increases in frontal cortical uptake sites (64) were not confirmed in follow-up studies (84 ,87). The discrepancies in this literature may be in part a consequence of the nonselectivity of [^{3}H]D-aspartate for the multiple

transporter subtypes; shifts in transporter subtype expression may occur in the absence of changes in total uptake sites. Consistent with this interpretation is the recent demonstration of decreased $EAAT_2$ mRNA levels in prefrontal cortex of schizophrenics (73). This change is in the opposite direction of that noted in previous studies examining radioligand binding to the transporters (64 ,84), which suggests that a shift from $EAAT_2$ to $EAAT_1/EAAT_3$ expression may occur in prefrontal cortex in schizophrenia.

TABLE 52.5. EXCITATORY AMINO ACID BINDING IN SCHIZOPHRENIA

Ligand	Findings	Brain Regions Studied	Reference
[^{3}H]glutamate	none	caudate, putamen, nucleus accumbens	57
[^{3}H]aspartate		frontal cortex	64
[^{3}H]aspartate	none	temporal cortex	64
[^{3}H]aspartate		anterior cingulate gyrus	87
[^{3}H]aspartate	none	hippocampus, temporal cortex	87
[^{3}H]glutamate	none	CA4, CA3, CA2, CA1, dentate gyrus, parahippocampal gyrus	53
[^{3}H]glutamate	none	CA3, CA2, CA1, dentate gyrus, subiculum	54
[^{3}H]aspartate		putamen, globus pallidus caudate,	70
[^{3}H]aspartate	none	nucleus accumbens	70

OTHER NEUROMODULATORS AND MARKERS

Part of "52 - Neurochemistry of Schizophrenia: Glutamatergic Abnormalities "

An alternative mechanism for altering glutamate neurotransmission involves neuropeptide modulators of glutamate-mediated neurotransmission (88 ,89 ,90 and 91). For instance, cholecystokinin (CCK) augments glutamate-mediated neurotransmission (88 ,91). CCK is expressed in subgroups of γ-aminobutyric acid (GABA)- and glutamate-containing neurons in the entorhinal cortex (92 ,93 and 94). Several postmortem studies have found abnormalities in CCK, CCK receptors, and CCK mRNA expression in schizophrenia, both in the frontal and temporal lobes (95 ,96 ,97 and 98). A cell-based silver grain analysis confirmed the involvement of layer VI, finding a reduction in the level of CCK mRNA expression per pyramidal cell (99). This is further supported by other molecular studies involving the measurement of complexin I and complexin II mRNAs, which suggest preferential involvement of excitatory pyramidal neurons in the mesial temporal lobe in schizophrenia (100 ,101).

A second neuropeptide neuromodulator concentrated in glutamate neurons, *N*-acetylaspartylglutamate (NAAG), antagonizes the effects of glutamate at NMDA receptors (102). NAAG is cleaved by glutamate carboxypeptidase II (formerly referred to as *N*-acetyl-α-linked acidic dipeptidase), a membrane-spanning glial enzyme, to yield glutamate and *N*-acetylaspartate (NAA). One study of NAAG and glutamate carboxypeptidase II found decreased glutamate carboxypeptidase II activity in prefrontal cortex and hippocampus and increased NAAG levels in the prefrontal cortex of schizophrenic patients relative to normal controls (103). Moreover, *in vivo* magnetic resonance spectroscopic imaging has revealed selective reductions in NAA in the dorsolateral prefrontal cortex and hippocampal formation of schizophrenic subjects (104 ,105). This suggests that NAA, a marker of neuronal integrity, may be decreased specifically and regionally in schizophrenia secondary to decreases in glutamate carboxypeptidase II.

CONCLUSIONS

Part of "52 - Neurochemistry of Schizophrenia: Glutamatergic Abnormalities "

Converging evidence indicates that abnormalities of glutamatergic neurotransmission occur in specific brain regions in schizophrenia. Although the hippocampus and associated structures have been the best studied, emerging data point to glutamatergic abnormalities in other areas of the brain that are likely to be associated with the pathophysiology of schizophrenia, including limbic cortex, striatal regions, and thalamus. Pharmacologic evidence suggests involvement of the NMDA receptor in schizophrenia, but other studies and theoretic considerations indicate that other molecules associated with glutamatergic transmission are also abnormal in this illness.

Studies in postmortem samples of the molecules associated with the glutamate synapse have not been conducted in a systematic and comprehensive fashion; however, several general principles are emerging from available data. First, although abnormalities of the glutamate synapse have been reported primarily in hippocampal regions, recent data suggest that thalamocortical circuits may also be abnormal. Interestingly, the striatal subregions appear to be less affected than medial temporal lobe and thalamocortical pathways. Second, all four families of glutamate receptors have been reported to be abnormal in brain in schizophrenia, although in region- and circuit-specific patterns. Third, changes are apparent at both transcriptional and translational levels of gene expression. Fourth, the ionotropic glutamate receptors have been studied most, and results thus far reveal changes in ionotropic receptor binding sites in addition to subunit changes suggestive of altered stoichiometry of subunit composition. The metabotropic receptors are just beginning to be studied, but the few available reports do suggest abnormalities of these receptors.

The literature on postmortem neurochemical studies of glutamatergic molecules in schizophrenia supports the hypothesis of abnormal glutamatergic neurotransmission in this illness that particularly involves the ionotropic receptors. These data suggest that novel strategies that permit the modulation of these receptors may prove to be of therapeutic utility in this illness, and may also provide clues about the pathophysiologic substrate of schizophrenia.

REFERENCES

1. Joyce JN, Meador-Woodruff JH. Linking the family of D2 receptors to neuronal circuits in human brain: insights into schizophrenia. *Neuropsychopharmacology* 1997;16:375–384.

2. Meador-Woodruff JH. Novel D2-like dopamine receptors in schizophrenic brain. In: Gattaz WF, Hafner H, eds. *Search for the causes of schizophrenia* (vol 4). *Balance of the century*. Berlin: Springer-Verlag, 1999:251–260.

3. Krystal JH, Karper LP, Seibyl JP, et al. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans: psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry* 1994;51:199–214.

4. Malhotra AK, Pinals DA, Weingartner H, et al. NMDA receptor function and human cognition: the effects of ketamine in healthy volunteers. *Neuropsychopharmacology* 1996;14:301–307.

5. Javitt DC, Zukin SR. Recent advances in the phencyclidine model of schizophrenia. *Am J Psychiatry* 1991;148:1301–1308.

6. Lahti AC, Koffel B, LaPorte D, et al. Subanesthetic doses of ketamine stimulate psychosis in schizophrenia. *Neuropsychopharmacology* 1995;13:9–19.

7. Malhotra AK, Pinals DA, Adler CM, et al. Ketamine-induced exacerbation of psychotic symptoms and cognitive impairment in neuroleptic-free schizophrenics. *Neuropsychopharmacology* 1997;17:141–150.

8. Carlsson M, Carlsson A. Schizophrenia: a subcortical neurotransmitter imbalance syndrome? *Schizophr Bull* 1990;16:425–432.

9. Rakic P, Bourgeois JP, Goldman-Rakic PS. Synaptic development of the cerebral cortex: implications for learning, memory, and mental illness. *Prog Brain Res* 1994;102:227–243.

10. Feinberg I. Cortical pruning and the development of schizophrenia. *Schizophr Bull* 1990;16:567–568.

11. Daw NW, Stein PSG, Fox K. The role of NMDA receptors in information processing. *Annu Rev Neurosci* 1993;16:207–222.

12. Zipursky RB, Lim KO, Sullivan EV, et al. Widespread cerebral grey matter volume deficits in schizophrenia. *Arch Gen Psychiatry* 1992;49:195–205.

13. Hollmann M, Heinemann S. Cloned glutamate receptors. *Annu Rev Neurosci* 1994;17:31–108.

14. Wheal HV, Thomson AM. *Excitatory amino acids and synaptic transmission*, second ed. New York: Academic Press, 1995.

15. Burnashev N, Monyer H, Seeburg PH, et al. Divalent ion permeability of AMPA receptor channels is dominated by the edited form of a single subunit. *Neuron* 1992;8:189–198.

16. Geiger JRP, Melcher T, Koh D-S, et al. Relative abundance of subunit mRNAs determines gating and Ca^{2+} permeability of AMPA receptors in principal neurons and interneurons in rat CNS. *Neuron* 1995;15:193–204.

17. Hollmann M, Hartley M, Heinemann S. Ca^{2+} permeability of KA-AMPA-gated glutamate receptor channels depends on subunit composition. *Science* 1991;252:851–853.

18. Jonas P, Racca C, Sakmann B, et al. Differences in Ca^{2+} permeability of AMPA-type glutamate receptor channels in neocortical neurons caused by differential GluR-B subunit expression. *Neuron* 1994;12:1281–1289.

19. Swanson GT, Kamboj SK, Cull-Candy SG. Single-channel properties of recombinant AMPA receptors depend on RNA editing, splice variation, and subunit composition. *J Neurosci* 1997;17:58–69.

20. Das S, Sasaki YF, Rothe T, et al. Increased NMDA current and spine density in mice lacking the NMDA receptor subunit NR3A. *Nature* 1998;393:377–381.

21. Sucher NJ, Akbarian S, Shi CL, et al. Developmental and regional expression pattern of a novel NMDA receptor-like subunit (NMDAR-L) in the rodent brain. *J Neurosci* 1995;15:6509–6520.

22. Ciabarra AM, Sullivan JM, Gahn LG, et al. Cloning and characterization of NR3A: a developmentally regulated member of a novel class of the ionotropic glutamate receptor family. *J Neurosci* 1995;15:6498–6508.

23. Durand GM, Bennett MVL, Zukin RS. Splice variants of the *N*-methyl-D-aspartate receptor NR1 identify domains involved in regulation of polyamines and protein kinase C. *Proc Natl Acad Sci U S A* 1993;90:6731–6735.

24. Nakanishi S. Molecular diversity of glutamate receptors and implications for brain function. *Science* 1992;258:597–603.

25. Mothet JP, Parent AT, Wolosker H, et al. D-Serine is an endogenous ligand for the glycine site of the *N*-methyl-D-aspartate receptor. *Proc Natl Acad Sci U S A* 2000;97:4926–4931.

26. Wolosker H, Blackshaw S, Snyder SH. Serine racemase: a glial enzyme synthesizing D-serine to regulate glutamate-*N*-methyl-D-aspartate neurotransmission. *Proc Natl Acad Sci U S A* 1999;96:13409–13414.

27. Schell MJ, Brady RO Jr, Molliver ME, et al. D-Serine as a neuromodulator: regional and developmental localizations in rat brain glia resemble NMDA receptors. *J Neurosci* 1997;17:1604–1615.

28. Schell MJ, Molliver ME, Snyder SH. D-Serine, an endogenous synaptic modulator: localization to astrocytes and glutamate-stimulated release. *Proc Natl Acad Sci U S A* 1995;92:3948–3952.

29. Boeckman FA, Aizenman E. Stable transfection of the NMDAR1 subunit in Chinese hamster ovary cells fails to produce a functional *N*-methyl-D-aspartate receptor. *Neurosc Lett* 1994;173:189–192.

30. Grimwood S, LeBourdelles B, Whiting PJ. Recombinant human NMDA homomeric NMDAR1 receptors expressed in mammalian cells form a high-affinity glycine antagonist binding site. *J Neurochem* 1995;64:525–530.

31. Lynch DR, Anegawa NJ, Verdoorn T, et al. *N*-methyl-D-aspartate receptors: different subunit requirements for binding of glutamate antagonists, glycine antagonists, and channel-blocking agents. *Mol Pharmacol* 1994;45:540–545.

32. Monyer H, Sprengel R, Schoepfer R, et al. Heteromeric NMDA receptors: molecular and functional distinction of subtypes. *Science* 1992;256:1217–1220.

33. Rodriguez-Paz JM, Anantharam Y, Treistman SN. Block of the *N*-methyl-D-aspartate receptor by phencyclidine-like drugs is influenced by alternative splicing. *Neurosci Lett* 1995;190:147–150.

34. Gallagher MJ, Huang H, Pritchett DB, et al. Interactions between ifenprodil and the NR2B subunit of the *N*-methyl-D-aspartate receptor. *J Biol Chem* 1996;271:9603–9611.

35. Corti C, Restituito S, Rimland JM, et al. Cloning and characterization of alternative mRNA forms for the rat metabotropic glutamate receptors mGluR7 and mGluR8. *Eur J Neurosci* 1998;10:3629–3641.

36. Nakanishi S. Metabotropic glutamate receptors: synaptic transmission, modulation, and plasticity. *Neuron* 1994;13:1031–1037.

37. Pin J-P, Duvoisin R. The metabotropic glutamate receptors: structure and functions. *Neuropharmacology* 1995;34:1–26.

38. Prezeau L, Carrette J, Helpap B, et al. Pharmacological characterization of metabotropic glutamate receptors in several types of brain cells in primary cultures. *Mol Pharmacol* 1994;45:570–577.

39. Saugstad JA, Kinzie JM, Shinohara MM, et al. Cloning and expression of rat metabotropic glutamate receptor 8 reveals a distinct pharmacological profile. *Mol Pharmacol* 1997;51:119–125.

40. Schoepp DD. Novel functions for subtypes of metabotropic glutamate receptors. *Neurochem Int* 1994;24:439–449.

41. Aramori I, Nakanishi S. Signal transduction and pharmacological characteristics of a metabotropic glutamate receptor, mGluR1, in transfected CHO cells. *Neuron* 1992;8:757–765.

42. Joly C, Gomeza J, Brabet I, et al. Molecular, functional, and pharmacological characterization of the metabotropic glutamate receptor type 5 splice variants: comparison with mGluR1. *J Neurosci* 1995;15:3970–3981.

43. Pickering DS, Thomsen C, Suzdak PD, et al. A comparison of two alternatively spliced forms of a metabotropic glutamate receptor coupled to phosphoinositide turnover. *J Neurochem* 1993;61:85–92.

44. Pin JP, Waeber C, Prezeau L, et al. Alternative splicing generates metabotropic glutamate receptors inducing different patterns of calcium release in *Xenopus* oocytes. *Proc Natl Acad Sci U S A* 1992;89:10331–10335.

45. Conn PJ, Pin J-P. Pharmacology and functions of metabotropic glutamate receptors. *Annu Rev Pharmacol Toxicol* 1997;37:205–237.

46. McCool BA, Pin JP, Brust PF, et al. Functional coupling of rat group II metabotropic glutamate receptors to an omega-conotoxin GVIA-sensitive calcium channel in human embryonic kidney 293 cells. *Mol Pharmacol* 1996;50:912–922.

47. Moghaddam B, Adams BW. Reversal of phencyclidine effects by a group II metabotropic glutamate receptor agonist in rats. *Science* 1998;281:1349–1352.

48. Harrison PJ, McLaughlin D, Kerwin RW. Decreased hippocampal expression of a glutamate receptor gene in schizophrenia. *Lancet* 1991;337:450–452.

49. Eastwood SL, McDonald B, Burnet PWJ, et al. Decreased expression of mRNAs encoding non-NMDA glutamate receptors GluR1 and GluR2 in medial temporal lobe neurons in schizophrenia. *Mol Brain Res* 1995;29:211–223.

50. Eastwood SL, Burnet PWJ, Harrison PJ. GluR2 glutamate receptor subunit flip and flop isoforms are decreased in the hippocampal formation in schizophrenia: a reverse transcriptase-polymerase chain reaction (RT-PCR) study. *Mol Brain Res* 1997;44:92–98.

51. Eastwood SL, Kerwin RW, Harrison PJ. Immunoautoradiographic evidence for a loss of alpha-amino-3-hydroxy-5-methly-4-isoxazole propionate-preferring non-*N*-methyl-D-aspartate glutamate receptors within the medial temporal lobe in schizophrenia. *Biol Psychiatry* 1997;41:636–643.

52. Breese CR, Freedman R, Leonard SS. Glutamate receptor subtype expression in human postmortem brain tissue from schizophrenics and alcohol abusers. *Brain Res* 1995;674:82–90.

53. Kerwin R, Patel S, Meldrum B. Quantitative autoradiographic analysis of glutamate binding sites in the hippocampal formation in normal and schizophrenic brain post mortem. *Neuroscience* 1990;39:25–32.

54. Gao X-M, Sakai K, Roberts RC, et al. Ionotropic glutamate receptors and expression of *N*-methyl-D-aspartate receptor subunits in subregions of human hippocampus: effects of schizophrenia. *Am J Psychiatry* 2000;157:1141–1149.

55. Healy DJ, Haroutunian V, Powchik P, et al. AMPA receptor binding and subunit mRNA expression in prefrontal cortex and striatum of elderly schizophrenics. *Neuropsychopharmacology* 1998;19:278–286.

56. Meador-Woodruff JH, Hogg AJ, Smith RE. Striatal ionotropic glutamate receptor expression in schizophrenia, bipolar disorder, and major depressive disorder. *Brain Res Bull (in press)*.

57. Noga JT, Hyde TM, Herman MM, et al. Glutamate receptors in the postmortem striatum of schizophrenic, suicide, and control brains. *Synapse* 1997;27:168–176.

58. Freed WJ, Dillon-Carter O, Kleinman JE. Properties of [^{3}H]AMPA binding in postmortem human brain from psychotic subjects and controls: increases in caudate nucleus associated with suicide. *Exp Neurol* 1993;121:48–56.

59. Sokolov BP. Expression of NMDAR1, GluR1, GluR7, and KA1 glutamate receptor mRNAs is decreased in frontal cortex of "neuroleptic-free" schizophrenics: evidence of reversible up-regulation of typical neuroleptics. *J Neurochem* 1998;71:2454–2564.

60. Toru M, Kurumaji A, Kumashiro S, et al. Excitatory amino acidergic neurones in chronic schizophrenic brain. *Mol Neuropharmacol* 1992;2:241–243.

61. Ibrahim H, Hogg AJ, Healy DJ, et al. Ionotropic glutamate receptor binding and subunit mRNA expression in thalamic nuclei in schizophrenia. *Am J Psychiatry* 2000;157:1811–1823.

62. Porter RHP, Eastwood SL, Harrison PJ. Distribution of kainate receptor subunit mRNAs in human hippocampus, neocortex and cerebellum, and bilateral reduction of hippocampal GluR6 and KA2 transcripts in schizophrenia. *Brain Res* 1997;751:217–231.

63. Meador-Woodruff JH, Davis KL, Haroutunian V. Abnormal kainate receptor expression in prefrontal cortex in schizophrenia. *Neuropsychopharmacology* 2001;24:545–552.

64. Deakin JFW, Slater P, Simpson MDC, et al. Frontal cortical and left temporal glutamatergic dysfunction in schizophrenia. *J Neurochem* 1989;52:1781–1786.

65. Nishikawa T, Takashima M, Toru M. Increased [^{3}H]kainic acid binding in the prefrontal cortex in schizophrenia. *Neurosci Lett* 1983;40:245–250.

66. Akbarian S, Sucher NJ, Bradley D, et al. Selective alterations in gene expression for NMDA receptor subunits in prefrontal cortex of schizophrenics. *J Neurosci* 1996;16:19–30.

67. Humphries C, Mortimer A, Hirsch S, et al. NMDA receptor mRNA correlation with antemortem cognitive impairment in schizophrenia. *Neuropharmacol Neurotoxicol* 1996;7:2051–2055.

68. Kornhuber J, Mack-Burkhardt F, Riederer P, et al. [^{3}H]MK-801 binding sites in postmortem brain regions of schizophrenic patients. *J Neural Transm* 1989;77:231–236.

69. Weissman AD, Casanova MF, Kleinman JE, et al. Selective loss of cerebral cortical sigma, but not PCP binding sites, in schizophrenia. *Biol Psychiatry* 1991;54:41–54.

70. Simpson MDC, Slater P, Royston MC, et al. Alterations in phencyclidine and sigma binding sites in schizophrenic brains. *Schizophr Res* 1992;6:41–48.

71. Ishimaru M, Kurumaji A, Toru M. Increases in strychnine-insensitive glycine binding sites in cerebral cortex of chronic schizophrenics: evidence for glutamate hypothesis. *Biol Psychiatry* 1994;35:84–95.

72. Aparicio-Legarza MI, Davis B, Hutson PH, et al. Increased density of glutamate/*N*-methyl-D-aspartate receptors in putamen from schizophrenic patients. *Neurosci Lett* 1998;241:143–146.

73. Ohnuma T, Augood SJ, Arai H, et al. Expression of the human excitatory amino acid transporter 2 and metabotropic glutamate receptors 3 and 5 in the prefrontal cortex from normal individuals and patients with schizophrenia. *Mol Brain Res* 1998;56:207–217.

74. Richardson-Burns SM, Haroutunian V, Davis DL, et al. Metabotropic glutamate receptor mRNA expression in the schizophrenic thalamus. *Biol Psychiatry* 2000;47:22–28.

75. Robinson MB. The family of sodium-dependent glutamate transporters: a focus on the GLT-1/EAAT2 subtype. *Neurochemistry* 1999;33:479–491.

76. Rothstein JD, Martin L, Levey AI, et al. Localization of neuronal and glial glutamate transporters. *Neuron* 1994;3:713–725.

77. Berger AJ, Dieudonne S, Ascher P. Glycine uptake governs glycine site occupancy at NMDA receptors of excitatory synapses. *J Neurophysiol* 1998;80:3336–3340.

78. Bergeron R, Meyer TM, Coyle JT, et al. Modulation of *N*-methyl-D-aspartate receptor function by glycine transport. *Proc Natl Acad Sci U S A* 1998;95:15730–15734.

79. Javitt DC, Sershen H, Hashim A, et al. Reversal of phencyclidine-induced hyperactivity by glycine and the glycine uptake inhibitor glycyldodecylamide. *Neuropsychopharmacology* 1997;17:202–204.

80. Supplisson S, Bergman C. Control of NMDA receptor activation by a glycine transporter co-expressed in *Xenopus* oocytes. *J Neurosci* 1997;17:4580–4590.

81. Toth E, Lathja A. Antagonism of phencyclidine-induced hyperactivity by glycine in mice. *Neurochem Res* 1986;11:393–400.

82. Kim KM, Kingsmore SF, Han H, et al. Cloning of the human glycine transporter type 1: molecular and pharmacological characterization of novel isoform variants and chromosomal localization of the gene in the human and mouse genomes. *Mol Pharmacol* 1994;45:608–617.

83. Morrow JA, Collie IT, Dunbar DR, et al. Molecular cloning and functional expression of the human glycine transporter GlyT2 and chromosomal localization of the gene in the human genome. *FEBS Lett* 1998;439:334–340.

84. Aparicio-Legarza MI, Cutts AJ, Davis B, et al. Deficits of [^{3}H]D-aspartate binding to glutamate uptake sites in striatal and accumbens tissue in patients with schizophrenia. *Neurosci Lett* 1997;232:13–16.

85. Simpson MDC, Slater P, Royston MC, et al. Regionally selective deficits in uptake sites for glutamate and gamma-aminobutyric acid in the basal ganglia in schizophrenia. *Psychiatry Res* 1992;42:273–282.

86. Simpson MDC, Slater P, Deakin JFW, et al. Absence of basal ganglia amino acid neuron deficits in schizophrenia in three collections of brains. *Schizophr Res* 1998;31:167–175.

87. Simpson MDC, Slater P, Deakin JFW. Comparison of glutamate and gamma-aminobutyric acid uptake binding sites in frontal and temporal lobes in schizophrenia. *Biol Psychiatry* 1998;44:423–427.

88. Breukel AI, Lopes da Silva FH, Ghijsen WE. Cholecystokinin (CCK-8) modulates vesicular release of excitatory amino acids in rat hippocampal nerve endings. *Neurosci Lett* 1997;234:67–70.

89. Kasparov S, Pawelzik H, Zieglgansberger W. Thyrotropin-releasing hormone enhances excitatory postsynaptic potentials in neocortical neurons of the rat *in vitro. Brain Res* 1994;656:229–235.

90. Koenig ML, Yourick DL, Meyerhoff JL. Thyrotropin-releasing hormone (TRH) attenuates glutamate-stimulated increases in calcium in primary neuronal cultures. *Brain Res* 1996;730:143–149.

91. You ZB, Godukhin O, Goiny M, et al. Cholecystokinin-85 increases dynorphin B, aspartate and glutamate release in the fronto-parietal cortex of the rat via different receptor subtypes. *Naunyn Schmiedebergs Arch Pharmacol* 1997;355:576–581.

92. Fredens K, Stengaard-Pedersen K, Larsson LI. Localization of enkephalin and cholecystokinin immunoreactivities in the perforant path terminal fields of the rat hippocampal formation. *Brain Res* 1984;304:255–263.

93. Kohler C, Chan-Palay V. The distribution of cholecystokinin-like immunoreactive neurons and nerve terminals in the retrohippocampal region in the rat and guinea pig. *J Comp Neurol* 1982;210:136–146.

94. Somogyi P, Hodgson AJ, Smith AD, et al. Different populations of GABAergic neurons in the visual cortex and hippocampus of cat contain somatostatin- or cholecystokinin- immunoreactive material. *J Neurosci* 1984;4:2590–2603.

95. Farmery SM, Owen F, Poulter M, et al. Reduced high affinity cholecystokinin binding in hippocampus and frontal cortex of schizophrenic patients. *Life Sci* 1985;36:473–477.

96. Kerwin R, Robinson P, Stephenson J. Distribution of CCK binding sites in the human hippocampal formation and their alteration in schizophrenia: a post-mortem autoradiographic study. *Psychol Med* 1992;22:37–43.

97. Virgo L, Humphries C, Mortimer A, et al. Cholecystokinin messenger RNA deficit in frontal and temporal cerebral cortex in schizophrenia. *Biol Psychiatry* 1995;37:694–701.

98. Zachrisson O, de Belleroche J, Wendt KR, et al. Cholecystokinin CCK(B) receptor mRNA isoforms: expression in schizophrenic brains. *Neuroreport* 1999;10:3265–3268.

99. Bachus SE, Hyde TM, Herman MM, et al. Abnormal cholecystokinin mRNA levels in entorhinal cortex of schizophrenics. *J Psychiatr Res* 1997;31:233–256.

100. Eastwood SL, Burnet PW, Harrison PJ. Expression of complexin I and II mRNAs and their regulation by antipsychotic drugs in the rat forebrain. *Synapse* 2000;36:167–177.

101. Harrison PJ, Eastwood SL. Preferential involvement of excitatory neurons in medial temporal lobe in schizophrenia. *Lancet* 1998;352:1669–1673.

102. Puttfarcken PS, Handen JS, Montgomery DT, et al. *N*-acetyl-aspartylglutamate modulation of *N*-methyl-D-aspartate-stimulated [^{3}H]norepinephrine release from rat hippocampal slices. *J Pharmacol Exp Ther* 1993;266:796–803.

103. Tsai G, Passani LA, Slusher BS, et al. Abnormal excitatory neurotransmitter metabolism in schizophrenic brains. *Arch Gen Psychiatry* 1995;52:829–836.

104. Bertolino A, Knable MB, Saunders RC, et al. The relationship between dorsolateral prefrontal *N*-acetylasparatate measures and striatal dopamine activity in schizophrenia. *Biol Psychiatry* 1999;45:660–667.

105. Bertolino A, Roffman JL, Lipska BK, et al. Postpubertal emergence of prefrontal neuronal deficits and altered dopaminergic behaviors in rats with neonatal hippocampal lesions. *Soc Neurosci Abst* 1999;520.8(abst).

53

Neural Circuitry Approaches to Understanding the Pathophysiology of Schizophrenia

David A. Lewis

David A. Lewis: University of Pittsburgh, Pittsburgh, Pennsylvania

PAST AND PREVAILING PATHOPHYSIOLOGIC MODELS OF PSYCHIATRIC DISORDERS

Part of "53 - Neural Circuitry Approaches to Understanding the Pathophysiology of Schizophrenia "

Models of the nature of brain dysfunction in major psychiatric disorders have evolved substantially over the past several decades in parallel with the progression in knowledge resulting from investigations of the neurobiological mechanisms that contribute to normal cognition, emotion, and behavior. Some models of psychopathology have emphasized, in their simplest forms, the central contribution of excesses or deficits in the functional activity of a given neurotransmitter to the disease process of interest. In general, these models have been very useful in motivating investigations of the molecular underpinnings and biochemical functions of the neurotransmitter systems of interest, and in spurring the development of novel psychopharmacologic agents that influence these systems. However, in extreme cases, these models tended to view individual psychiatric disorders as the consequences *solely* of the postulated disturbance in the neurotransmitter of interest, and consequently they tended to be conceptually similar to historic views that altered levels of the four classical bodily humors produced different forms of madness. In addition to this limited conceptual perspective, neurotransmitter-based models sometimes seemed to attribute behavioral, emotional, or cognitive functions to neurotransmitters, instead of explicitly recognizing that neurotransmitters have defined actions on receptors, whereas behaviors, emotions, and cognitive abilities represent emergent properties of the integrated activity of large networks of neurons.

Other models of psychopathology have emphasized the critical role of localized disturbances in individual brain regions, an approach that, in extreme cases, has been critiqued as “neophrenology.” Although these models have been useful in stimulating studies of the structure–function relationships of the implicated brain regions, they have been limited in a number of respects, including the inability to account for the array of signs and symptoms that typically constitute the syndromic diagnosis of a psychiatric disorder.

These two types of models were influenced, at least in part, by extrapolations from earlier successes in the study of Parkinson’s disease, which was then viewed as a single neurotransmitter (e.g., dopamine) disease owing to a localized neuropathology (e.g., cell death in the substantia nigra). However, in recent years, these two general approaches have given way to neural circuitry-based models that reflect a fuller appreciation of the fact that neurotransmitters act in an anatomically constrained fashion to produce specific biochemical effects at the cellular level, and that the localization of function(s) is a consequence of the flow of information processing through the neural circuits within a given brain region and those linking that region to other brain areas. These latter types of models (1 ,2 and 3) incorporate the recognition that complex brain functions, such as those that are disturbed in major psychiatric disorders, are subserved by the coordinated activity of distributed ensembles of neurons.

Consequently, the goal of this chapter is to examine the use of a neural circuitry-based approach to understanding the pathophysiology of a psychiatric disorder, using studies of the neurobiological basis for cognitive dysfunction in schizophrenia as an example. Specifically, this chapter: (a) considers the convergent lines of evidence that suggest that the neural circuitry involving the dorsal prefrontal cortex is disturbed in this disorder, (b) reviews the normal organization of this circuitry as revealed through studies in animals, (c) assesses the evidence regarding the integrity of this circuitry in schizophrenia, and (d) discusses new opportunities

for neural circuitry-based studies of the pathophysiology of schizophrenia.

EXAMINING SCHIZOPHRENIA AS A DISORDER OF NEURAL CIRCUITRY

Part of "53 - Neural Circuitry Approaches to Understanding the Pathophysiology of Schizophrenia "

Clues from the Clinical Syndrome

Although schizophrenia and Alzheimer's disease (AD) both received these appellations approximately a century ago, their initial descriptions were based on different types of data. For schizophrenia, the disorder was recognized by the presence of a constellation of clinical features and a particular longitudinal course, whereas the identification of AD was based on clinicopathological correlations. Indeed, the observation of neurofibrillary tangles and neuritic plaques in the cerebral cortex of the victims of AD provided the foundation for the rich array of anatomic, biochemical, and molecular genetic studies in the past two decades that have produced the wealth of current knowledge regarding the pathophysiology of this disease.

In contrast, studies of the pathophysiology of schizophrenia continue to depend, in large part, on clues derived from the clinical syndrome. Although the number and diversity of the clinical features of this illness make this approach daunting, a reasonable case can be made that the disturbances in cognition commonly seen in schizophrenia represent a core feature of the illness. For example, at least some of the other signs and symptoms of schizophrenia may be conceptualized as secondary to the cognitive disturbances (4), cognitive abnormalities can be identified during the prodromal phase of the illness and in those at increased risk for the disorder (5), cognitive symptoms appear to be persistent across the course of the illness (6), and the severity of cognitive impairment may be the best predictor of long-term outcome (7 ,8). Thus, the clinical features of schizophrenia argue for an emphasis on the neural circuitry that normally subserves the types of cognitive functions that are disturbed in the disorder.

Clues from Developmental Features

Although the idea that schizophrenia is a late consequence of an early developmental lesion (9) has many merits, direct evidence for a brain abnormality that could be explained by a mechanism operating prior to or at the onset of the clinical symptoms of the illness, and that could not be attributed to factors associated with having the illness, has proven difficult to obtain. For example, reports of cytoarchitectural disturbances in the entorhinal cortex of schizophrenic subjects (10 ,11) attracted a great deal of attention because the reported findings were strongly suggestive of an abnormality in neuronal migration (12); however, subsequent studies with larger sample sizes have both failed to confirm these reports and have provided likely methodologic explanations for the initial findings (13 ,14 ,15 and 16). The report of an altered distribution of interstitial neurons in the subcortical white matter was also conceptually very attractive because it strongly suggested an early developmental lesion (17); however, the majority of schizophrenic subjects appear to lack such abnormalities (18 ,19).

However, one of the characteristics of schizophrenia is the tendency for clinical symptoms to first appear during late adolescence or early adulthood, and it has been argued that hypotheses of the pathophysiology of schizophrenia must accommodate this age of onset (9 ,20). Although the average age of first hospitalization for patients with schizophrenia is in the early or mid-twenties for men and women, respectively (21 ,22), psychotic symptoms may appear months or even years prior to hospitalization (22 ,23 and 24). In addition, deterioration in other areas, such as scholastic performance and sociability, precedes the onset of the overt symptoms of schizophrenia by some time (5 ,23 ,24), and may represent strong predictors of the subsequent appearance of the disorder. Thus, developmental events occurring during the second decade of life may play a critical role in the appearance of cognitive dysfunction in schizophrenia.

Clues from Distributed Brain Abnormalities

Structural and functional neuroimaging studies in subjects with schizophrenia have implicated a number of brain regions as potential sites of dysfunction or morphologic abnormalities. For example, certain brain regions, such as the medial temporal lobe (including the hippocampus, amygdala, and parahippocampal gyrus), the superior temporal gyrus, the dorsal prefrontal cortex and the thalamus, have all been shown to have reduced total volume in subjects with schizophrenia, although the magnitude of the decrease and its consistency across studies has not been uniform (25). Similarly, functional imaging studies have shown alterations in the activation of some of these brain areas under different conditions, especially when subjects are performing tasks that normally are associated with a change in activation (26). Although some of the reported findings cannot be accommodated in this way, a number of the affected regions share reciprocal connections. Thus, pathophysiologic models that account for the reported abnormalities in two or more of these regions, and the connections that link them, may be particularly promising.

Incorporating These Types of Clues into Research Strategies for Identifying Neural Circuitry Abnormalities

Based on the three types of clues summarized in the preceding, one can ask whether they converge or triangulate on neural circuit(s) that may be preferentially involved in the pathophysiology of schizophrenia. Although this approach

suggests a number of possible candidate circuits for investigation, the dorsal prefrontal cortex (dPFC) may be considered a prototypic nodal point for circuit analysis in schizophrenia for the following reasons.

First, from the perspective of clinical clues, subjects with schizophrenia perform poorly on cognitive tasks that involve working memory, the ability to transiently maintain information in order to guide a subsequent response (27). For example, individuals with schizophrenia exhibit impairments on oculomotor delayed-response tasks (28), a cognitive paradigm on which nonhuman primates with structural or reversible cooling lesions of the dPFC perform poorly (29). Consistent with these observations, subjects with schizophrenia also fail to show normal activation of the dPFC when attempting to perform tasks that tap working memory (30).

Second, from the developmental perspective, the circuitry of the primate dPFC clearly undergoes marked refinements during adolescence, although certainly some other brain regions that have not been as well studied are also likely to show such changes. For example, the number of excitatory synapses in the dPFC declines by 50% during adolescence in both monkeys and humans (31 ,32). In addition, substantial changes occur in markers of excitatory, inhibitory, and modulatory inputs to pyramidal neurons in deep layer 3 of primate dPFC. The apparent laminar specificity of at least some of these changes raises the possibility that circuits involving these pyramidal cells may be preferentially affected in schizophrenia (33).

Third, from the perspective of regional brain analyses, the PFC has been shown to have subtle reductions in gray matter volume in a number of structural imaging studies of schizophrenia (25). The failure of other studies to detect such structural abnormalities in the PFC has been hypothesized to be a consequence of several factors, including a reduction in total PFC volume that approximates the level of sensitivity of MRI and the restriction of volumetric changes to certain regions or gyri of the dPFC (25). Consistent with these interpretations, postmortem observations indicate that cortical thickness in the dPFC may be reduced by 3% to 12% in subjects with schizophrenia (34 ,35 ,36 and 37), although these changes do not always achieve statistical significance. In addition, some (38 ,39 and 40), but not all (41), *in vivo* proton spectroscopy studies indicate that N-acetyl aspartate (NAA), a putative marker of neuronal and/or axonal integrity, is reduced in this brain region. Interestingly, the magnitude of these NAA changes in the dPFC was correlated with the degree of impaired activation in other brain regions during working memory tasks, raising the possibility that a neuronal abnormality in the dPFC could account for distributed functional disturbances in the working memory network (40).

Other lines of evidence suggest that these changes may reflect disturbances in the synaptic connectivity of the dPFC in schizophrenia. For example, the presence of decreased phosphomonoesters and increased phosphodiesters (42 ,43), as measured by P^{31} spectroscopy, in the PFC of schizophrenic subjects has been interpreted to reflect an increased breakdown of membrane phospholipids, and consequently a decreased number of synapses. In addition, a recent gene expression profiling study using cDNA microarrays found that the group of genes encoding proteins that regulate presynaptic secretory machinery were most consistently altered (44). Furthermore, reduced levels of synaptophysin, an integral membrane protein of small synaptic vesicles, have also been observed in the dPFC of subjects with schizophrenia in most (45 ,46 ,47 ,48 and 49), but not all (50), studies.

For these reasons, the next two sections of this chapter focus on a summary of the normal organization of dPFC circuitry and on a review of the evidence suggesting that this circuitry is disturbed in schizophrenia.

OVERVIEW OF DPFC CIRCUITRY IN PRIMATES

Part of "53 - Neural Circuitry Approaches to Understanding the Pathophysiology of Schizophrenia "

Direct studies of the circuitry of the human dPFC have obvious limitations; however, the available cross-species studies indicate that the macaque monkey brain provides an accurate and useful model system for understanding the general organization of the human dPFC. Thus, this section reviews the constituent cell types and the patterns of intrinsic and extrinsic connectivity that characterize the primate dPFC.

Cell Types

Pyramidal Neurons

About 70% of all cortical neurons are pyramidal cells (51). The majority of pyramidal cells have a characteristically shaped cell body that gives rise to a single apical dendrite that is oriented perpendicular to the cortical surface and frequently ends in a terminal tuft (Fig. 53.1). An array of shorter basilar dendrites spread in a radial fashion at the base of the cell body. Both the apical and basilar dendrites are coated with short protrusions or spines, which represent the principal targets of most excitatory synaptic inputs to pyramidal neurons. Because dendritic spines are actively formed or resorbed in response to changes in presynaptic inputs, dendritic spines provide a good estimate of the number of excitatory synapses that pyramidal cells receive (52). Typically, pyramidal neurons possess 6,000 to 10,000 dendritic spines (53). In addition to receiving one excitatory input, some dendritic spines also receive a synapse with the features suggestive of an inhibitory input. Inhibitory terminals also synapse on the dendritic shafts, cell body, and axon initial segment of pyramidal cells. Typically, pyramidal cells receive about 2,000 inhibitory synapses on dendritic shafts, 200 on the cell soma, and 20 on the axon initial segment (53).

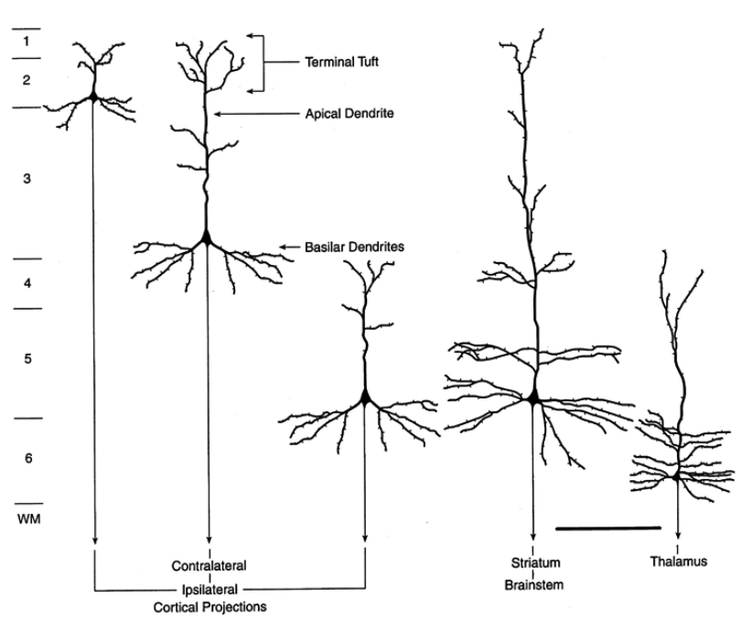


FIGURE 53.1. Schematic drawing illustrating the characteristic morphologic features of pyramidal neurons in different cortical layers. Note that the laminar location of the cell soma tends to be associated with the major projection target of the principal axon. Arabic numbers, cortical layers; WM, white matter. Adapted from Jones EG. Laminar distribution of cortical efferent cells. In: Peters A, Jones EG, eds. *Cerebral cortex*, vol 1. New York: Plenum Press, 1984:521–553.

The axons of pyramidal cells typically give rise to intrinsic collaterals, which travel either horizontally or vertically within the gray matter, and a principal axonal projection, which enters the white matter and travels to another brain region. These axons utilize excitatory amino acids, such as glutamate, as a neurotransmitter and form synapses that have the characteristic morphology associated with excitatory neurotransmission. These so-called Gray's Type I synapses are characterized by the presence of small round vesicles in the axon terminal, and a postsynaptic density that is thick and asymmetric in appearance (54).

Nonpyramidal Neurons

Of the other major type of cortical neurons, nonpyramidal cells, about 90% utilize the inhibitory neurotransmitter γ-aminobutyric acid (GABA). The axons of these generally small, local circuit or interneurons, arborize within the cortical gray matter and form Gray's Type II synapses that are characterized by pleomorphic vesicles in the axon terminal and symmetric pre- and postsynaptic densities. GABA neurons constitute approximately 25% of all neurons in the primate neocortex (55), and are comprised of about 12 distinct subtypes (56 ,57). Although the differences among subtypes can be described on the basis of the morphologic features of the cell body and proximal dendrites (e.g., bipolar, multipolar, bitufted), the most discriminating and functionally meaningful classification system is based on the organization of the axonal arbor and synaptic targets of the axon terminals. In addition, GABA neurons are chemically heterogeneous, and separate subpopulations can be identified by the presence of specific neuropeptides or calcium-binding proteins (58 ,59).

Together, these morphologic and chemical features define subpopulations of GABA neurons that appear to have different biophysical properties and different roles in dPFC circuitry (Fig. 53.2). For example, GABA neurons of the chandelier class, which may also express either the neuropeptide corticotropin-releasing factor (60) or the calcium-binding protein parvalbumin (61 ,62), are found primarily in cortical layers 2 to 5 in monkey dPFC (56). The axon terminals of these neurons, which are arrayed as distinct vertical structures (termed "cartridges"), form Gray's Type II synapses exclusively with the axon initial segment of pyramidal

neurons (63 ,64 ,65 ,66 ,67 and 68), the site of action potential generation in pyramidal cells. Each chandelier cell may contact up to 300 pyramidal neurons within a radius of 100 to 150 µm from its cell body (69). Thus, chandelier cells exert critical inhibitory control over the activity of a localized group of pyramidal neurons. In contrast, the axons of wide arbor (basket) neurons spread horizontally for considerable distances (up to 1.0 mm) within the dPFC (56) and form Gray's Type II synapses with the cell bodies, dendritic shafts, and dendritic spines of pyramidal neurons (70). Wide arbor neurons may be specialized to provide inhibitory constraints over the activity of spatially segregated populations of PFC pyramidal neurons (71 ,72). A third example, double bouquet cells, which contain the calcium-binding proteins calbindin or calretinin (58), have radially restricted axonal arbors that synapse with the dendritic shafts of both pyramidal and local circuit neurons (73).

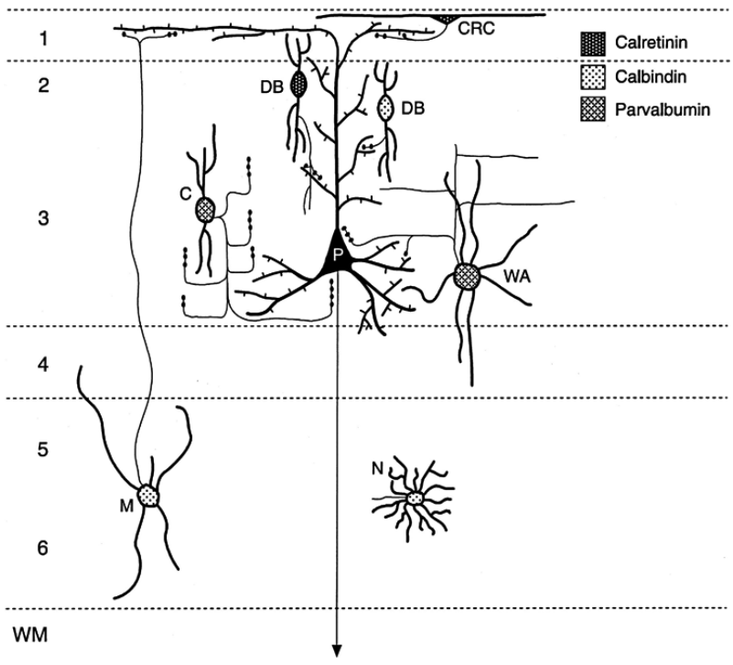


FIGURE 53.2. Schematic drawing of the synaptic interactions between different classes of local circuit neurons and a layer 3 pyramidal neuron (P) in monkey prefrontal cortex. C, parvalbumin (PV)-labeled chandelier neuron; CRC, calretinin (CR)- and/or CB (calbindin)-labeled Cajal-Retzius neuron; DB, CR- or CB-labeled double-bouquet neuron; M, CB-labeled Martinotti cell; N, CB-labeled neurogliaform neuron; WA, PV-labeled wide arbor (basket) neuron. Modified from Condé F, Lund JS, Jacobowitz DM, et al. Local circuit neurons immunoreactive for calretinin, calbindin D-28k, or parvalbumin in monkey prefrontal cortex: distribution and morphology. *J Comp Neurol* 1994,341:95–116.

Laminar Arrangement of Neurons

The dPFC, like other cortical association regions, is composed of six layers that can be distinguished according to the size and packing density of their constituent neurons (Fig. 53.3). Layer 1, which is located just below the pial surface, contains relatively few neurons, but approximately 90% of these neurons utilize GABA. Layers 2 and 4 are thin and densely packed with small "granular" cells. The majority of these neurons are small pyramidal cells, and GABA neurons constitute approximately 30% and 15% of all neurons in layers 2 and 4, respectively (55). Layers 3 and 5, the thickest cortical layers, contain prominent pyramidal neurons with a classic morphology. In both of these layers, the size and packing density of the pyramidal neurons is greater near their borders with layer 4. These patterns make it possible to subdivide layers 3 and 5 into superficial (toward the pial surface) and deep zones. Pyramidal cells in layer 6 have a modified or atypical appearance. In general, GABA cells constitute 20% to 30% of the neurons in layer 3, and about 15% of the neurons in layers 5 and 6 (55).

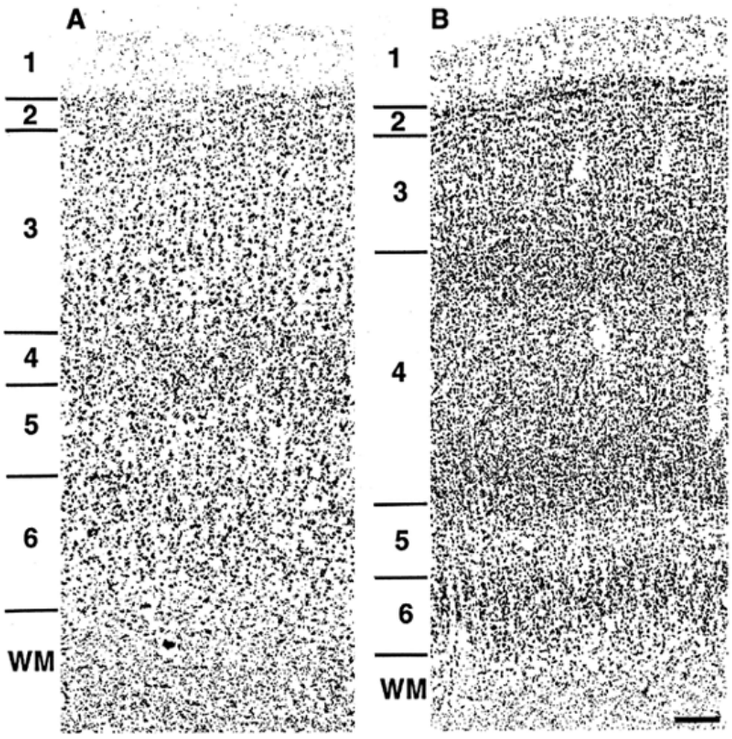


FIGURE 53.3. Nissl-stained sections showing laminar and regional differences in cell size in packing density in area 46 *(A)* of the human prefrontal cortex and area 17, primary visual cortex *(B)*. Calibration bar = 200 µm. Modified from Lewis DA. The organization of cortical circuitry. In: Harrison PJ, Roberts GW, eds. *The neuropathology of schizophrenia: progress and interpretation*. Oxford: Oxford University Press, 2000:235–256.

Different types of axonal projections to the dPFC terminate in certain cortical layers, and projections from the dPFC to other brain regions generally originate from pyramidal neurons in specific lamina. Thus, the laminar specificity of abnormalities in the dPFC in schizophrenia may reveal information about the types of connections that are affected. Nevertheless, owing to the vertical spread of the dendrites of many neurons, alterations in afferents to one layer can influence the function of neurons whose cell body is located in a different layer.

Organization of Projections from the dPFC

Pyramidal neurons located in layers 2 and superficial 3 tend to project to nearby cortical regions in the same hemisphere, whereas those in deep layer 3 frequently project to more distant ipsilateral regions or across the corpus callosum to cortical regions in the other hemisphere (Fig. 53.1). Projections within the same hemisphere are termed associational projections, whereas those to the contralateral hemisphere are termed callosal projections. Pyramidal cells in layers 5 and 6 project subcortically, with outputs from layer 5 directed to the striatum, superior colliculus, and other subcortical structures, and those from layer 6 preferentially directed to the thalamus. The efferent axons of pyramidal neurons tend to project to only one other brain region, although a small percentage of these neurons do give rise to collateral projections (74 ,75). In contrast to these extrinsic projections, the small pyramidal cells in layer 4 project primarily within the gray matter and relatively few send axons into the underlying white matter.

Although this laminar distribution of cortical efferents is generally accurate, many exceptions exist. For example, 25% to 30% of the pyramidal neurons furnishing associational projections to other PFC regions are located in the infragranular layers (layers 5 and 6) (74), and approximately 15% to 20% of neurons projecting to the striatum are located in layer 3 (76). In addition, the nature of the cortical output to a given region may vary with the location of the cell body of origin. For example, neurons in layer 6 provide “modulatory” inputs to cells in higher order thalamic nuclei (such as the mediodorsal thalamic nucleus, the principal source of thalamic projections to the dPFC) as well as inputs to the thalamic reticular nucleus, which regulates thalamocortical interactions. In contrast, thalamic projections originating in layer 5 do not innervate the reticular nucleus and appear to provide “driving” afferent inputs to higher order nuclei (77).

The innervation patterns of the intrinsic axon collaterals of pyramidal cells also tend to differ across cortical layers (71). Pyramidal neurons in layers 2 and 3 furnish local collaterals that arborize in the vicinity of the cell body, as well as horizontal axon projections that spread for considerable distances through the gray matter and then give rise to discrete clusters of axon terminals in the supragranular layers. Although pyramidal neurons in layers 5 and 6 also furnish horizontal intrinsic collaterals, these have a more limited spread and do not terminate in spatially segregated clusters. In contrast, pyramidal cells in layer 4 furnish predominantly vertically oriented axons, which appear to be specialized for interlaminar connections. The intrinsic axonal connections of pyramidal neurons in layers 2 and 3 of the dPFC also appear to be specialized relative to the homologous neurons in at least some other cortical regions. For example, although approximately 95% of the long distance intrinsic and associational axon projections of these neurons target the dendritic spines of other pyramidal cells (78), the synaptic targets of the local axon collaterals of these cells are equally divided between spines and dendritic shafts of GABA neurons (79).

Organization of Projections to the dPFC

The dPFC shares connections with a number of other cortical regions (Fig. 53.4). The inputs from these areas frequently terminate across all cortical layers, although the different layers tend to be preferentially innervated depending upon the source of the inputs. For example, inputs from cortical regions that have a well-developed layer 4 tend to terminate more prominently in layers 4 to 6, whereas those that originate in regions with a poorly developed layer 4 tend to terminate in layers 1 to 3 (80). In some cases, afferents from different regions (e.g., callosal inputs from the contralateral PFC and associational inputs from the posterior parietal cortex) tend be distributed into interdigitated columns (81).

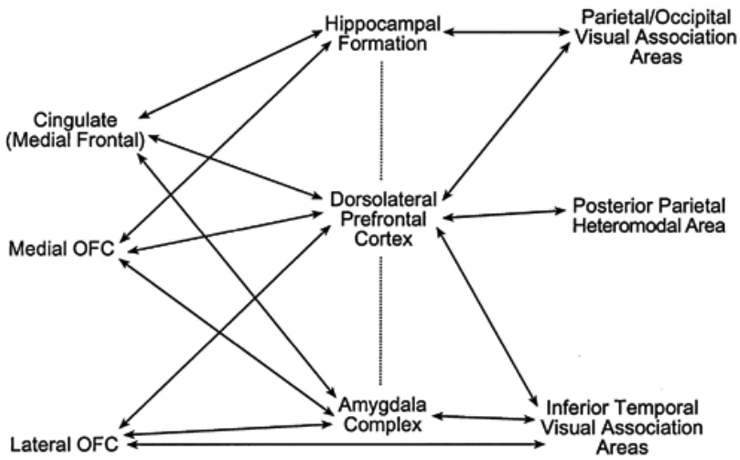


FIGURE 53.4. Schematic diagram of principal corticocortical connections within the frontal lobe regions, including primary limbic connections, and between posterior sensory cortices and frontal lobe regions. Double-head arrows indicate that most connections are reciprocal. Relatively sparse direct connections between dorsolateral prefrontal cortex and limbic structures (hippocampal formation and amygdala nuclei) are depicted with a dashed line. OFC, orbitofrontal cortex. Modified from Kaufer D, Lewis DA. Frontal lobe anatomy and cortical connectivity. In: Miller BL, Cummings JL, eds. *The human frontal lobes: functions and disorders*. New York: Guilford Publications, 1998:27–44.

In contrast to these cortical inputs, afferents from thalamic relay nuclei, such as the medial dorsal nucleus, project dominantly to layers deep 3 and 4, with a minor projection to layer 6 (82). Although afferents from the amygdala project more densely to orbital than dorsal regions of the PFC, these tend to terminate in layers 1 and 6 (83).

Subcortical nuclei containing monoamines or acetylcholine also exhibit distinct laminar patterns of termination in the PFC, along with substantial regional differences in relative

density. Dopamine (DA)-containing axons from the ventral mesencephalon have a bilaminar distribution in the PFC (84), forming a dense band in layers 1 through the most superficial portion of layer 3 and a second band of lower density in layers deep 5 and 6. In more densely innervated regions, such as dorsomedial PFC (area 9), labeled axons are also present in high density in the middle cortical layers, forming a third distinctive band in deep layer 3. The noradrenergic (NA) projection from the locus coeruleus exhibits a different, and in some ways complementary, laminar innervation pattern to that of DA axons (85 ,86). The density of NA axons is substantially greater in the deep cortical layers, especially layer 5, than in the more superficial cortical laminae. In particular, few NA axons are present in layer 1, which receives a dense DA innervation. In contrast, the relatively uniform laminar distribution of cholinergic (87) and serotonergic (88) axons contrasts with the substantial heterogeneity exhibited by both DA and NA axons.

INTEGRITY OF PREFRONTAL CORTICAL CIRCUITRY IN SCHIZOPHRENIA

Part of "53 - Neural Circuitry Approaches to Understanding the Pathophysiology of Schizophrenia "

In this section, we consider how this knowledge about the normal organization of dPFC circuitry can be used to interpret studies of the integrity of different types of neural elements in subjects with schizophrenia. As noted, several lines of evidence support the hypothesis that schizophrenia is associated with a decrease in the synaptic connectivity of the dPFC. However, these abnormalities do not appear to be a consequence of a decreased complement of dPFC neurons, because several postmortem studies (34 ,37 ,89) have reported either a normal or increased cell packing density in the dPFC. In addition, the one study that used unbiased approaches to determine the total number of PFC neurons did not observe a reduction in subjects with schizophrenia (90); however, the approaches used in these studies probably lacked adequate sensitivity to detect reduced numbers of small subpopulations of PFC neurons. Some studies that focused on certain neuronal subpopulations have reported decreases in density of small neurons in layer 2 (91) or of the parvalbumin-containing subpopulation of GABA neurons (92). However, the latter abnormality was not observed in another study (35), and it should be noted that a reduction in neuronal density when using immunocytochemical markers might reflect an alteration in the target protein rather than in the number of cells.

Smaller neuronal cell bodies could also contribute to the observed reduction in PFC gray matter in schizophrenia. Interestingly, two groups have reported that the somal volume of pyramidal cells in deep layer 3 of dPFC area 9 is decreased in subjects with schizophrenia (89 ,93). In addition, this reduction in somal volume may be associated with a decrease in total length of the basilar dendrites of these neurons (94). In contrast, the size of GABA neurons does not appear to be reduced (35 ,95), although less rigorous methods were employed in these studies.

In summary, although a reduction in neuron number cannot be completely excluded, the subtle reduction in dPFC gray matter in schizophrenia may be attributable to a combination of smaller neurons and a decrease in dPFC neuropil, the axon terminals, distal dendrites and dendritic spines that represent the principal components of cortical synapses. Indeed, as described in more detail below, these two factors may be interrelated.

Candidate Sources for Synaptic Reductions

The apparent reduction in synaptic connectivity in the dPFC of subjects with schizophrenia may be attributable to one or more of the following sources of synapses: axon terminals intrinsic to the dPFC, or afferents from other cortical regions, the thalamus, or other subcortical locations. Although none of these sources can be excluded at present, some are more likely to be major contributors than others. For example, one subcortical source, the DA projections from the mesencephalon, may be reduced in number in schizophrenia as evidenced by the report of diminished densities of axons immunoreactive for tyrosine hydroxylase, the rate limiting enzyme in catecholamine synthesis, and the DA membrane transporter in the dPFC (96). However, these reductions appeared to be restricted to the deep cortical layers. Furthermore, an *in vivo* neuroimaging study found a reduced density of DA D1 receptors in the dPFC of subjects with schizophrenia (97); however, DA axons are estimated to contribute less than 1% of cortical synapses (84). Consequently, even the complete loss of DA projections to the dPFC could not, in isolation, account for the observed reductions in gray matter volume or synaptophysin protein levels. The relatively small contributions that NA-, serotonin-, and acetylcholine-containing axons make to the total number of synapses in the dPFC also argues against disturbances in these systems as the principal cause of reduced synaptic connectivity in this brain region.

The fact that layer 3 pyramidal cells, which give rise to a substantial number of intrinsic excitatory synapses, have reduced somal volume in subjects with schizophrenia (89 ,93) may suggest that the synapses furnished by the intrinsic axon collaterals of these neurons are reduced in number, because somal volume tends to be correlated with the size of a neuron’s axonal arbor (98 ,99). Evidence for a disturbance in intrinsic connectivity is supported by recent studies using cDNA microarray profiling of the expression of over 7,000 genes in dPFC area 9 of subjects with schizophrenia (44). Among 250 functional gene groups, the most marked changes in expression were present in the group of genes that encode for proteins involved in the regulation of presynaptic neurotransmitter release. Although these findings very likely indicate a general impairment in the efficacy of synaptic

transmission within the dPFC in schizophrenia, whether they represent a “primary” abnormality intrinsic to the dPFC or a “secondary” response to altered afferent drive to this brain region remains to be determined. Furthermore, because the specific genes in this group that were most altered appeared to differ across subjects, it seems unlikely that these findings can be explained solely by reduction in the number of intrinsic dPFC synapses. Consistent with this view, the expression of synaptophysin mRNA does not appear to be reduced in the dPFC of subjects with schizophrenia (50 ,100), suggesting that the reduction in this synaptic protein marker in the dPFC may have an extrinsic source. Consistent with this interpretation, synaptophysin mRNA levels are reduced in cortical areas that do furnish projections to the dPFC (101 ,102). However, whether these transcriptional changes are present in PFC-projecting neurons, and if so, whether they result in reduced levels of synaptophysin protein in the terminal fields of these neurons, have not been assessed.

Decreased synaptic connectivity within the PFC might also be attributable to altered inputs from the thalamus. For example, some structural MRI studies have revealed a reduction in thalamic volume in subjects with schizophrenia (103 ,104 ,105 and 106). In addition, thalamic volume was correlated with prefrontal white matter volume in schizophrenic subjects (107), suggesting that a reduction in thalamic volume was associated with fewer axonal projections to the PFC. Consistent with these observations, postmortem studies have revealed reductions of 17% to 25% in volume and 27% to 40% in total neuron number of the medial dorsal thalamic nucleus (MDN), the principal source of thalamic projections to the PFC (108 ,109 and 110). Interestingly, the available data suggest that these abnormalities exhibit topographic specificity. For example, reduced cell numbers have been reported in the MDN and the anterior thalamic nuclei (which project to the PFC and anterior cingulate cortex), whereas the ventral posterior medial nucleus, a sensory relay nucleus, appears to be unaffected (109 ,110). In addition, within the MDN, one study indicates that neuron number is significantly decreased in the parvocellular subdivision (which projects principally to the dPFC), but not in the magnocellular subdivision (which projects principally to the ventral PFC) (110). Finally, studies in both subjects with schizophrenia who never received antipsychotic medications (111) and monkeys treated for 1 year with haloperidol (112) suggest that these medications do not account for the reduction in MDN neuron number. However, despite the apparent consistency of the published studies, a deficient number of MDN neurons in schizophrenia must still be considered a preliminary finding given the relatively small sample sizes reported to date, and the fact that potential confounds, such as comorbid conditions (e.g., alcoholism), have not been adequately assessed.

Certainly, a reduction in cell number in the MDN could contribute both to the decrease in synaptic markers in the PFC, and, given the dependence of working memory tasks on the integrity of thalamo–prefrontal connections (29), to the disturbances in working memory observed in schizophrenia. However, in contrast to other species, the primate MDN contains both cortically projecting neurons and local circuit neurons. Thus, it is critical to determine which subpopulation(s) of MDN neurons are affected in schizophrenia. Interestingly, the density of neurons in the anterior thalamic nuclei that contain parvalbumin (113), a calcium-binding protein present in thalamic projection neurons (114), is reduced in schizophrenia; however, whether this reduction represents an actual loss of neurons, as opposed to an activity-dependent decrease in parvalbumin expression, is not known.

Within the dPFC, five other lines of evidence are also consistent with a reduction in inputs from the MDN (Fig. 53.5). First, a preliminary report notes that subjects with schizophrenia, but not those with major depression, have a decreased density of parvalbumin-labeled varicosities (putative axon terminals) in layers deep 3 and 4, the principal termination zone of thalamic projections to the PFC (115). In contrast, parvalbumin-labeled varicosities were not decreased in layers 2 to superficial 3, suggesting that the reduction in layers deep 3 and 4 might not be attributable to changes in the axon terminals of the parvalbumin-containing subset of cortical GABA neurons present in cortical layers 2 to 5 (35). Thus, the laminar-specific reduction of

parvalbumin varicosities in schizophrenia is consistent with a decreased number of MDN terminals in the dPFC, although a laminar-specific reduction in the axon terminals of local circuit neurons cannot be conclusively excluded.

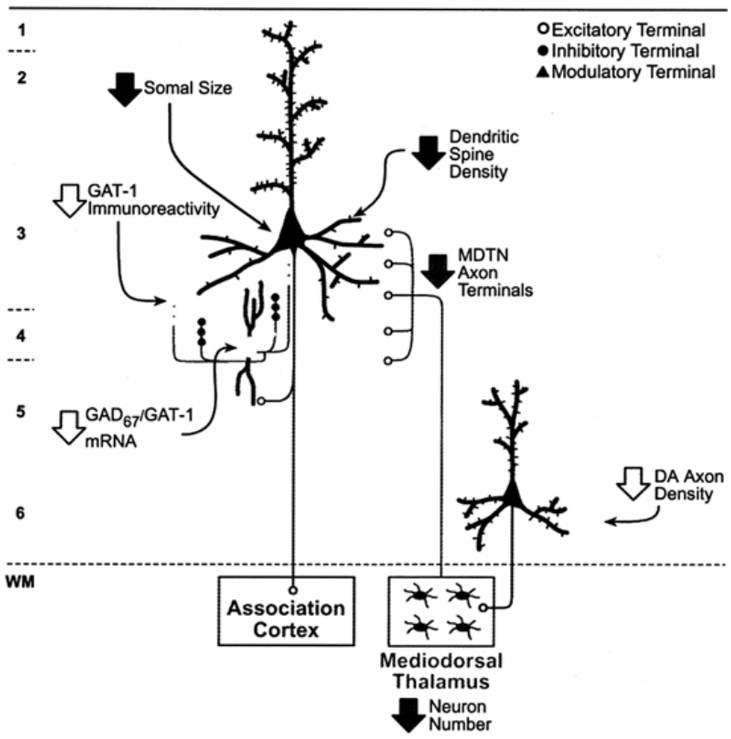


FIGURE 53.5. Schematic diagram summarizing disturbances in the connectivity between the mediodorsal (MD) thalamic nucleus and the dorsal prefrontal cortex (PFC) in schizophrenia. See text for details. Modified from Lewis DA, Lieberman JA. Catching up on schizophrenia: natural history and neurobiology. *Neuron* 2000;28:325=n334.

Second, the thalamic projections to the dPFC principally target the dendritic spines of pyramidal neurons (116). In experimental animals, the elimination of presynaptic axon terminals leads to a resorption of the postsynaptic dendritic spine (117), suggesting that a reduction in MDN projection neurons would be associated with decreased dendritic spine density in the dPFC. The two studies that have examined this issue both found decreased spine density on the basilar dendrites of PFC layer 3 pyramidal neurons (94 ,118), with this decrease most marked for pyramidal neurons located in deep layer 3 (94), those most likely to be targeted by projections from the thalamus. Although the well-documented, remarkable plasticity of dendritic spines must be considered when interpreting these findings, these observations are consistent with a reduction in MDN–dPFC connectivity in schizophrenia. However, the presence of more modest reductions in spine density on pyramidal neurons in cortical layers that do not directly receive MDN input suggest that the observed decrease in deep layer 3 may reflect the combined effect of a deficient number of thalamic and cortical synapses (94).

Third, the size of layer 3 pyramidal neurons is reduced in the dPFC of schizophrenic subjects (89 ,119). Although the possible relationship of these findings to a decrease in MDN inputs is less clear, studies in animals have provided evidence of denervation atrophy of layer 3 pyramidal cells following the loss of other afferent inputs (120).

Fourth, in the primate visual system, monocular deprivation, which results in reduced afferent drive from the thalamus, is associated with a decline in markers of activity in cortical GABA neurons (121), including decreased expression of the mRNA for glutamic acid decarboxylase (GAD_{67}), the synthesizing enzyme for GABA (122). Although the experimental manipulation of the visual system did not involve a partial reduction in thalamic neuron number, if these findings in the visual cortex can be generalized to a deficient number of MDN projections to the dPFC, a reduction in GAD_{67} in the dPFC of schizophrenic subjects might be expected. Consistent with this prediction, both GAD_{67} mRNA and protein levels have been reported to be reduced in the dPFC of schizophrenic subjects (26 ,123 ,124), observations supported by other evidence of reduced GABA neurotransmission in the PFC of schizophrenic subjects. (See ref. 125 for review.)

Finally, the reduction in GAD_{67} mRNA expression in schizophrenia appears to be restricted to a subpopulation of PFC GABA neurons (approximately 25% to 30%), especially those neurons located in the middle cortical layers (124). Consistent with this finding, other studies suggest that the affected subpopulation of GABA neurons includes chandelier cells. The axon terminals of chandelier cells form distinctive vertical arrays (termed cartridges) that synapse exclusively on the axon initial segment of pyramidal neurons (126). Interestingly, expression of the mRNA for GABA membrane transporter (GAT-1) is also undetectable in approximately 25% to 30% of GABA neurons, which have a laminar distribution similar to the neurons with undetectable GAD_{67} mRNA expression (127). In addition, the density of GAT-1 immunoreactive chandelier neuron axon cartridges is decreased in the dPFC of schizophrenic subjects, with the reduction most evident in the middle cortical layers (128 ,129). Thus, given the powerful inhibitory control that chandelier neurons exert over pyramidal cell output, decreased excitatory thalamic drive to the PFC may be partially compensated for by a reduction in chandelier cell-mediated inhibition at the axon initial segment of layer 3 pyramidal cells. This effect could occur via the local axon collaterals of layer 3 pyramidal cells, approximately 50% of which target the dendritic shafts of GABA neurons (79). However, it is important to note that other causes and consequences of the observed alterations in chandelier neurons have not been excluded. (See ref. 124 for a discussion.)

Together, these data are all consistent with the hypothesis that schizophrenia is associated with abnormalities in the projection from the MDN to the dPFC. As in other cortical regions, the connections between the MDN and dPFC are reciprocal, which raises the question of whether the abnormal thalamocortical projection is paralleled by a disturbance in the corticothalamic projection. Studies that have examined PFC neurons in layers 5 and 6, the principal location of corticothalamic projection neurons, have generally not found evidence of a decrease in neuron size or number (34 ,89 ,91), although one study (95) did report decreased neuronal density in layer 6 of PFC (region not specified); however, a reduced density of DA axons was observed selectively in layer 6 of dPFC area 9 in schizophrenic subjects (96). Interestingly, the dendritic shafts and spines of pyramidal cells are the principal synaptic targets of DA axon terminals in layer 6, and DA appears to play a critical role in regulating the influence of other inputs on pyramidal cell activity (130). Thus, a shift in DA neurotransmission in dPFC layer 6 could reflect a change in the modulation of corticothalamic feedback in response to abnormal thalamocortical drive (96).

The significance of these findings depends, in part, on the extent to which they are unique to the diagnosis of schizophrenia, and not a consequence of other factors associated with schizophrenia, such as antipsychotic drug treatment. Some of these findings have been examined for diagnostic specificity, whereas others have not. For example, the reduction in dendritic spine density on deep layer 3 pyramidal neurons was not found in subjects with major depressive disorder (94), and the reduction in density of GAT-1 immunoreactive axon cartridges was not apparent in subjects with nonschizophrenic psychiatric disorders (129). Similarly, to the extent to which it has been examined, these

findings do not appear to be attributable to the abuse of alcohol or other substances, which frequently accompanies the diagnosis of schizophrenia, although further studies in this area certainly are needed.

Similarly, the available evidence suggests that these disturbances in thalamo-prefrontal circuitry are not attributable to treatment with antipsychotic medications. Globally, the increase in cell packing density in the dPFC observed in subjects with schizophrenia was not found in monkeys treated for 6 months with a variety of antipsychotic agents (131). Similarly, treatment of monkeys for 12 months with haloperidol and benztropine at blood levels known to be therapeutic in humans was not associated with a reduction in the size or total neuron number of the MDN (112). The potential influence of antipsychotic drugs on GAT-1–labeled axon cartridges has been examined in several ways with interesting results (129). The density of labeled cartridges was greater in schizophrenic subjects who were on than off antipsychotic medications at the time of death (although both groups showed reduced levels compared to normal controls). In addition, compared to matched control animals, the density of GAT-1-positive cartridges was elevated in monkeys treated for 1 year with haloperidol. Thus, the convergence of these findings suggests both that the pathophysiology of schizophrenia may actually be associated with more marked reductions in GAT-1–immunoreactive cartridge density than those observed in postmortem studies.

OPPORTUNITIES FOR NEURAL CIRCUITRY-BASED STUDIES OF SCHIZOPHRENIA

Part of "53 - Neural Circuitry Approaches to Understanding the Pathophysiology of Schizophrenia "

The data summarized in the preceding section suggest that neural circuitry-based approaches to the study of brain abnormalities in schizophrenia provide: (a) a useful framework to account for the abnormalities observed in individual studies, (b) a platform for the formulation of predictions regarding the outcome of future studies, and (c) the promise of an enhanced ability to understand the neurobiological bases of clinical phenomena. However, a truly neural circuitry-based model of schizophrenia requires an appreciation of the mechanistic relations among the abnormalities observed in different components of the circuitry. Specifically, understanding the pathophysiology of schizophrenia (or any other psychiatric disorder) depends ultimately on knowing how abnormalities in one brain region or circuitry component produce and/or result from disturbances in others, a task that involves a consideration of cause, consequence, and compensation (132). Does a given abnormality represent a primary pathogenetic event (cause), does it reflect a downstream or upstream (given the reciprocal nature of many links between brain regions), secondary, deleterious event (consequence), or does it reveal a homeostatic response intended to restore normal brain function (compensation)? Distinguishing among these three possibilities for each component of a neural network will be necessary for understanding the pathophysiology of the disease as well as for developing novel therapies designed to correct causes and consequences and/or to augment compensatory responses.

Clearly, addressing these types of questions for schizophrenia requires several additional types of investigations. First, it is essential to further assess the normal organization of MDN–dPFC connectivity in nonhuman primates. For example, what patterns of connectivity within the dPFC link inputs from the MDN to the neurons that provide outputs to the MDN or to other brain regions such as the striatum? Second, MDN–dPFC circuitry needs to be further probed in subjects with schizophrenia in ways that inform an understanding of its functional integrity. For example, what are the postsynaptic consequences in pyramidal neurons of the apparent alterations in GABA neurotransmission in chandelier cells? Third, the direction of the pathophysiological changes in MDN–dPFC circuitry in schizophrenia need to be assessed in experimental animal models. For example, can the observations of altered spine density and decreased GAD_{67} mRNA expression in the dPFC be replicated by partial lesions of the MDN in monkeys? Do manipulations of neurotrophin expression in dPFC layer 3 pyramidal cells result in a loss of the MDN neurons that project to the dPFC?

It is also critical to extend these types of investigations to other brain regions that may integrate MDN–dPFC circuitry with broader neural networks. Besides the dPFC, the hippocampal formation is probably the brain region that has been most extensively studied in schizophrenia. Multiple imaging and postmortem studies have documented a slight bilateral reduction in the volume of the hippocampal formation (25 ,133), an observation supported by more recent *in vivo* proton spectroscopy findings of reduced hippocampal N-acetyl aspartate in both unmedicated adult and childhood onset subjects with schizophrenia (134). Although initial reports of hippocampal neuron disarray or misplaced neurons in the superficial layers of the adjacent entorhinal cortex have been widely cited, these observations have not been replicated in other studies. (See ref. 133 for review.) Reduced hippocampal volume also does not appear to be attributable to decreased neuronal number, but several independent studies have found reductions in neuronal cell body size in various subregions of the hippocampus proper (135 ,136 and 137). In addition, there are consistent reports of reductions in the gene products for synaptophysin and related presynaptic markers and in dendritic markers, such as microtubule-associated protein, in certain subdivisions of the hippocampus. (See ref. 12 for review.) Thus, these observations bear some similarity to findings in the dPFC in that disturbances in synaptic connectivity appear to be present in both regions in schizophrenia. How these findings inform our understanding of alterations in the intrinsic connectivity of the hippocampal formation in schizophrenia remains to be

determined, but their possible relation to dPFC abnormalities is suggested by experimental studies in rodents that indicate that dysfunction of the PFC appears postpubertally following perinatal lesions of the hippocampus (138).

As noted in the preceding section, it is also important to consider these findings within the context of the developmental time course of schizophrenia, especially the tendency for prodromal and clinical symptoms to become evident during the second and third decades of life. Although the adolescence-related pruning of excitatory (but not inhibitory) synapses (31 ,32) and their targets, pyramidal cell dendritic spines (139), in the primate dPFC has been well documented, little information exists regarding the extent to which different connections are actually affected by these changes. Interestingly, limited data suggest that the terminals of intrinsic axon collaterals from dPFC layer 3 pyramidal cells may be more extensively pruned than associational cortical projections to the dPFC (140). Knowing whether projections from the MDN are particularly vulnerable to this process might provide critical information for hypotheses regarding the mechanisms underlying disturbances in MDN–dPFC circuitry in schizophrenia.

Another current challenge to the types of neural circuitry-based models of schizophrenia illustrated herein is to understand how the genetic factors that confer susceptibility for the disease contribute to the observed alterations in neural circuitry. In other words, how can molecular genetic and systems neuroscience approaches be integrated in the study of schizophrenia? Schizophrenia appears to be a consequence of multiple interacting genes that individually may have relatively little effect; it is unlikely that all such genes are involved in every individual who meets diagnostic criteria for the disorder. Thus, assessment of the patterns of altered gene expression in the affected brain circuits of subjects with schizophrenia (using cDNA microarray technology or related techniques), and comparison of the chromosomal locations of these genes with regions implicated in schizophrenia through linkage studies (141), may provide convergent approaches to the identification of specific susceptibility genes. For example, as noted, a recent study of gene expression profiling in the dPFC of subjects with schizophrenia revealed that the group of genes encoding proteins involved in the regulation of presynaptic function were most consistently altered (44). In addition, although the subjects with schizophrenia appeared to share a common abnormality in the control of synaptic transmission, they differed in terms of the specific combination of genes that showed reduced expression, a finding that may be consistent with a polygenic model for this disorder. Interestingly, a number of the chromosomal loci that have been implicated in schizophrenia contain genes encoding proteins related to presynaptic function (44). However, this strategy, and the subsequent integration of potential susceptibility genes with neural circuitry models of the illness, rests on the prediction that genes regulating presynaptic function are not homogeneously expressed across classes of neurons and brain regions.

Finally, neural circuitry-based models must also be considered in relation to the clinical heterogeneity of schizophrenia. Is the magnitude of the abnormalities within MDN–dPFC circuitry related to the age of onset or severity of cognitive impairment? Can other clinical features be understood within the context of abnormalities in broader circuits that include connections with MDN and dPFC? Given the limitations involved in making rigorous clinicopathologic correlations in schizophrenia, the answers to these questions may await the development and application of *in vivo* imaging techniques with greater sensitivity, such as those that will permit functional assessments at the levels of individual cortical layers or thalamic nuclei.

REFERENCES

1. Andreasen NC, O'Leary DS, Cizaldo T et al. Schizophrenia and cognitive dysmetria: a positron-emission tomography study of dysfunctional prefrontal-thalamic-cerebellar circuitry. *Proc Natl Acad Sci USA* 1996;93:9985–9990.

2. Jones EG. Cortical development and thalamic pathology in schizophrenia. *Schiz Bull* 1997;23:483–501.

3. Weinberger DR. Anteromedial temporal-prefrontal connectivity: a functional neuroanatomical system implicated in schizophrenia. In: Carroll BJ, Barrett JE, eds. *Psychopathology and the brain*. New York: Raven Press, 1991:25–41.

4. Fuster JM. Frontal lobes. *Curr Opin Neurobiol* 1993;3:160–165.

5. Davidson M, Reichenberg A, Rabinowitz J et al. Behavioral and intellectual markers for schizophrenia in apparently healthy male adolescents. *Am J Psychiatry* 1999;156:1328–1335.

6. Green MF. *Schizophrenia from a neurocognitive perspective: probing the impenetrable darkness*. Boston: Allyn and Bacon, 1998.

7. Weinberger DR, Gallhofer B. Cognitive dysfunction in schizophrenia. *Int Clin Psychopharmacol* 1997;12S:29–36.

8. Green MF. What are the functional consequences of neurocognitive deficits in schizophrenia? *Am J Psychiatry* 1996;153:321–330.

9. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 1987;44:660–669.

10. Jakob H, Beckmann H. Prenatal developmental disturbances in the limbic allocortex in schizophrenics. *J Neural Transm* 1986;65:303–326.

11. Arnold SE, Hyman BT, Van Hoesen GW, et al. Some cytoarchitectural abnormalities of the entorhinal cortex in schizophrenia. *Arch Gen Psychiatry* 1991;48:625–632.

12. Weinberger DR. Cell biology of the hippocampal formation in schizophrenia. *Biol Psychiatry* 1999;45:395–402.

13. Bernstein H-G, Krell D, Baumann B, et al. Morphometric studies of the entorhinal cortex in neuropsychiatric patients and controls: clusters of heterotopically displaced lamina II neurons are not indicative of schizophrenia. *Schiz Res* 1998;33:125–132.

14. Krimer LS, Herman MM, Saunders RC, et al. A qualitative and quantitative analysis of the entorhinal cortex in schizophrenia. *Cerebral Cortex* 1997;7:732–739.

15. Heinsen H, Gössmann E, Rüb U, et al. Variability in the human entorhinal region may confound neuropsychiatric diagnoses. *Acta Anatomica* 1996;157:226–237.

16. Akil M, Lewis DA. The cytoarchitecture of the entorhinal cortex in schizophrenia. *Am J Psychiatry* 1997;154:1010–1012.

17. Akbarian S, Bunney Jr WE, Potkin SG, et al. Altered distribution of nicotinamide-adenine dinucleotide phosphate-diaphorase cells in frontal lobe of schizophrenics implies disturbances of cortical development. *Arch Gen Psychiatry* 1993;50:169–177.

18. Akbarian S, Kim JJ, Potikin SG, et al. Maldistribution of interstitial neurons in prefrontal white matter of the brains of schizophrenic patients. *Arch Gen Psychiatry* 1996;53:425–436.

19. Anderson S, Volk DW, Lewis DA. Increased density of microtubule-associated protein 2-immunoreactive neurons in the prefrontal white matter of schizophrenic subjects. *Schiz Res* 1996;19:111–119.

20. Feinberg I. Schizophrenia: caused by a fault in programmed synaptic elimination during adolescence? *J Psychiatric Res* 1982;17:319–334.

21. Castle DJ, Murray RM. The neurodevelopmental basis of sex differences in schizophrenia. *Psychol Med* 1991;21:565–575.

22. Szymanski S, Lieberman JA, Alvir JM, et al. Gender differences in onset of illness, treatment response, course, and biological indexes in first-episode schizophrenic patients. *Am J Psychiatry* 1995;152:698–703.

23. Haas GL, Sweeney JA. Premorbid and onset features of first-episode schizophrenia. *Schiz Bull* 1992;18:373–386.

24. Larsen TK, McGlashan TH, Johannessen JO, et al. First-episode schizophrenia: II. Premorbid patterns by gender. *Schiz Bull* 1996;22:257–269.

25. McCarley RW, Wible CG, Frumin M, et al. MRI anatomy of schizophrenia. *Biol Psychiatry* 1999;45:1099–1119.

26. Taylor SF. Cerebral blood flow activation and functional lesions in schizophrenia. *Schiz Res* 1996;19:129–140.

27. Goldman-Rakic PS. Working memory dysfunction in schizophrenia. *J Neuropsychiatry* 1994;6:348–357.

28. Park S, Holzman PS. Schizophrenics show spatial working memory deficits. *Arch Gen Psychiatry* 1992;49:975–982.

29. Fuster JM. *The prefrontal cortex: anatomy, physiology, and neuropsychology of the frontal lobe*. Philadelphia: Lippincott-Raven, 1997.

30. Weinberger DR, Berman KF, Zec RF. Physiologic dysfunction of dorsolateral prefrontal cortex in schizophrenia. *Arch Gen Psychiatry* 1986;43:114–124.

31. Huttenlocher PR. Synaptic density in human frontal cortex: developmental changes and effects of aging. *Brain Res* 1979;163:195–205.

32. Bourgeois J-P, Goldman-Rakic PS, Rakic P. Synaptogenesis in the prefrontal cortex of rhesus monkeys. *Cerebral Cortex* 1994;4:78–96.

33. Lewis DA. Development of the prefrontal cortex during adolescence: insights into vulnerable neural circuits in schizophrenia. *Neuropsychopharmacology* 1997;16:385–398.

34. Selemon LD, Rajkowska G, Goldman-Rakic PS. Abnormally high neuronal density in the schizophrenic cortex: a morphometric analysis of prefrontal area 9 and occipital area 17. *Arch Gen Psychiatry* 1995;52:805–818.

35. Woo T-U, Miller JL, Lewis DA. Parvalbumin-containing cortical neurons in schizophrenia. *Am J Psychiatry* 1997;154:1013–1015.

36. Pakkenberg B. Total nerve cell number in neocortex in chronic schizophrenics and controls estimated using optical disectors. *Biol Psychiatry* 1993;34:768–772.

37. Daviss SR, Lewis DA. Local circuit neurons of the prefrontal cortex in schizophrenia: selective increase in the density of calbindin-immunoreactive neurons. *Psychiatry Res* 1995;59:81–96.

38. Bertolino A, Nawroz S, Mattay VS, et al. Regionally specific pattern of neurochemical pathology in schizophrenia as assessed by multislice proton magnetic resonance spectroscopic imaging. *Am J Psychiatry* 1996;153:1554–1563.

39. Deicken RF, Zhou L, Corwin F, et al. Decreased left frontal lobe *N*-acetylaspartate in schizophrenia. *Am J Psychiatry* 1997;154:688–690.

40. Bertolino A, Esposito G, Callicott JH, et al. Specific relationship between prefrontal neuronal N-acetylaspartate and activation of the working memory cortical network in schizophrenia. *Am J Psychiatry* 2000;157:26–33.

41. Stanley JA, Williamson PC, Drost DJ, et al. An *in vivo* proton magnetic resonance spectroscopy study of schizophrenia patients. *Schiz Bull* 1996;22:597–609.

42. Pettegrew JW, Keshavan MS, Panchalingam K, et al. Alterations in brain high-energy phosphate and membrane phospholipid metabolism in first-episode, drug-naive schizophrenics. *Arch Gen Psychiatry* 1991;48:563–568.

43. Stanley JA, Williamson PC, Drost DJ, et al. An *in vivo* study of the prefrontal cortex of schizophrenic patients at different stages of illness via phosphorus magnetic resonance spectroscopy. *Arch Gen Psychiatry* 1995;52:399–406.

44. Mirnics K, Middleton FA, Marquez A, et al. Molecular characterization of schizophrenia viewed by microarray analysis of gene expression in prefrontal cortex. *Neuron* 2000;28:1–20.

45. Karson CN, Mrak RE, Schluterman KO, et al. Alterations in synaptic proteins and their encoding mRNAs in prefrontal cortex in schizophrenia: a possible neurochemical basis for 'hypofrontality.' *Mol Psychiatry* 1999;4:39–45.

46. Honer WG, Falkai P, Chen C, et al. Synaptic and plasticity-associated proteins in anterior frontal cortex in severe mental illness. *Neuroscience* 1999;91:1247–1255.

47. Perrone-Bizzozero NI, Sower AC, Bird ED, et al. Levels of the growth-associated protein GAP-43 are selectively increased in association cortices in schizophrenia. *Proc Natl Acad Sci USA* 1996;93:14182–14187.

48. Davidsson P, Gottfries J, Bogdanovic N, et al. The synaptic-vesicle-specific proteins rab3a and synaptophysin are reduced in thalamus and related cortical brain regions in schizophrenic brains. *Schiz Res* 1999;40:23–29.

49. Glantz LA, Lewis DA. Reduction of synaptophysin immunoreactivity in the prefrontal cortex of subjects with schizophrenia: regional and diagnostic specificity. *Arch Gen Psychiatry* 1997;54:943–952.

50. Eastwood SL, Cairns NJ, Harrison PJ. Synaptophysin gene expression in schizophrenia: Investigation of synaptic pathology in the cerebral cortex. *Br J Psychiatry* 2000;176:236–242.

51. Powell TPS. Certain aspects of the intrinsic organisation of the cerebral cortex. In: Pompeiana O, Marsan CA, eds. *Brain mechanisms and perceptual awareness*. New York: Raven Press, 1981:1–19.

52. Mates SL, Lund JS. Spine formation and maturation of type 1 synapses on spiny stellate neurons in primate visual cortex. *J Comp Neurol* 1983;221:91–97.

53. DeFelipe J, Farinas I. The pyramidal neuron of the cerebral cortex: morphological and chemical characteristics of the synaptic inputs. *Progr Neurobiol* 1992;39:563–607.

54. Peters A, Palay SL, Webster DF. *The fine structure of the nervous system*. New York: Oxford University Press, 1991.

55. Hendry SHC, Schwark HD, Jones EG, et al. Numbers and proportions of GABA-immunoreactive neurons in different areas of monkey cerebral cortex. *J Neurosci* 1987;7:1503–1519.

56. Lund JS, Lewis DA. Local circuit neurons of developing and mature macaque prefrontal cortex: Golgi and immunocytochemical characteristics. *J Comp Neurol* 1993;328:282–312.

57. Fairen A, DeFelipe J, Regidon J. Nonpyramidal neurons, general account. In: Peters A, Jones EG, eds. *Cerebral cortex*, vol. 1. New York: Plenum, 1984:201–245.

58. Condé F, Lund JS, Jacobowitz DM, et al. Local circuit neurons immunoreactive for calretinin, calbindin D-28k, or parvalbumin in monkey prefrontal cortex: distribution and morphology. *J Comp Neurol* 1994;341:95–116.

59. Gabbott PLA, Bacon SJ. Local circuit neurons in the medial prefrontal cortex (areas 24a,b,c, 25 and 32) in the monkey. I. Cell morphology and morphometrics. *J Comp Neurol* 1996;364:567–608.

60. Lewis DA, Foote SL, Cha CI. Corticotropin releasing factor immunoreactivity in monkey neocortex: an immunohistochemical analysis. *J Comp Neurol* 1989;290:599–613.

61. DeFelipe J, Hendry SHC, Jones EG. Visualization of chandelier cell axons by parvalbumin immunoreactivity in monkey cerebral cortex. *Proc Natl Acad Sci USA* 1989;86:2093–2097.

62. Lewis DA, Lund JS. Heterogeneity of chandelier neurons in monkey neocortex: corticotropin-releasing factor and parvalbumin immunoreactive populations. *J Comp Neurol* 1990;293:599–615.

63. Szentagothai J, Arbib M. Conceptual models of neural organization. *Neurosci Res Program Bull* 1974;12:307–510.

64. Jones EG. Varieties and distribution of nonpyramidal cells in the somatic sensory cortex of the squirrel monkey. *J Comp Neurol* 1975;160:205–268.

65. Somogyi P. A specific axo-axonal interneuron in the visual cortex of the rat. *Brain Res* 1977;136:345–350.

66. Fairen A, Valverde F. A specialized type of neuron in the visual cortex of cat: a Golgi and electron microscope study of chandelier cells. *J Comp Neurol* 1980;194:761–779.

67. Peters A, Proskauer CC, Ribak CE. Chandelier neurons in rat visual cortex. *J Comp Neurol* 1982;206:397–416.

68. DeFelipe J, Hendry SHC, Jones EG, et al. Variability in the terminations of GABAergic chandelier cell axons on initial segments of pyramidal cell axons in the monkey sensory-motor cortex. *J Comp Neurol* 1985;231:364–384.

69. Peters A. Chandelier cells. In: Jones EG, Peters A, eds. *Cerebral cortex*, vol 1. New York: Plenum Press, 1984:361–380.

70. Jones EG, Hendry SHC. Basket cells. In: Jones EG, Peters A, eds. *Cerebral cortex*, vol 1. *Cellular components of the cerebral cortex*. New York: Plenum Press, 1984:309–336.

71. Levitt JB, Lewis DA, Yoshioka T et al. Topography of pyramidal neuron intrinsic connections in macaque monkey prefrontal cortex (areas 9 & 46). *J Comp Neurol* 1993;338:360–376.

72. Lund JS, Yoshioka T, Levitt JB. Comparison of intrinsic connectivity in different areas of macaque monkey cerebral cortex. *Cerebral Cortex* 1993;3:148–162.

73. Somogyi P, Cowey A. Double bouquet cells. In: Peters A, Jones EG, eds. *Cerebral cortex*. New York: Plenum Press, 1984:337–360.

74. Schwartz ML, Goldman-Rakic PS. Callosal and intrahemispheric connectivity of the prefrontal association cortex in rhesus monkey: relation between intraparietal and principal sulcal cortex. *J Comp Neurol* 1984;226:403–420.

75. Barbas H. Pattern in the cortical distribution of prefrontally directed neurons with divergent axons in the rhesus monkey. *Cerebral Cortex* 1995;2:158–165.

76. Arikuni T, Kubota K. The organization of prefrontalcaudate projections and their laminar origin in the Macaque monkey: a retrograde study using HRP-gel. *J Comp Neurol* 1986;244:492–510.

77. Guillery RW, Feig SL, Lozsádi DA. Paying attention to the thalamic reticular nucleus. *Trends Neurosci* 1998;21:28–32.

78. Melchitzky DS, Sesack SR, Pucak ML, et al. Synaptic targets of pyramidal neurons providing intrinsic horizontal connections in monkey prefrontal cortex. *J Comp Neurol* 1998;390:211–224.

79. Melchitzky DS, Gonzalez-Burgos G, Barrionuevo G, et al. Synaptic targets of the intrinsic axon collaterals of supragranular pyramidal neurons in monkey prefrontal cortex. *J Comp Neurol* 2001;430:209=n221.

80. Barbas H, Rempel-Clower N. Cortical structure predicts the pattern of corticocortical connections. *Cerebral Cortex* 1997;7:635–646.

81. Goldman-Rakic PS, Schwartz ML. Interdigitation of contralateral and ipsilateral columnar projections to frontal association cortex in primates. *Science* 1982;216:755–757.

82. Giguere M, Goldman-Rakic PS. Mediodorsal nucleus: areal, laminar, and tangential distribution of afferents and efferents in the frontal lobe of rhesus monkeys. *J Comp Neurol* 1988;277:195–213.

83. Amaral DG, Price JL. Amygdalo-cortical projections in the monkey *(Macaca fascicularis). J Comp Neurol* 1984;230:465–496.

84. Lewis DA, Sesack SR. Dopamine systems in the primate brain. In: Bloom FE, Björklund A, Hökfelt T, eds. *Handbook of chemical neuroanatomy*. Amsterdam: Elsevier Science, 1997:261–373.

85. Lewis DA, Morrison JH. The noradrenergic innervation of monkey prefrontal cortex: a dopamine-beta-hydroxylase immunohistochemical study. *J Comp Neurol* 1989;282:317–330.

86. Gaspar P, Berger B, Fabvret A, et al. Catecholamine innervation of the human cerebral cortex as revealed by comparative immunohistochemistry of tyrosine hydroxylase and dopamine-beta-hydroxylase. *J Comp Neurol* 1989;279:249–271.

87. Lewis DA. Distribution of choline acetyltransferase immunoreactive axons in monkey frontal cortex. *Neuroscience* 1991;40:363–374.

88. Lewis DA. The organization of chemically identified neural systems in primate prefrontal cortex: afferent systems. *Progr Neuro-Psychopharmacol Biol Psychiatry* 1990;14:371–377.

89. Rajkowska G, Selemon LD, Goldman-Rakic PS. Neuronal and glial somal size in the prefrontal cortex: a postmortem morphometric study of schizophrenia and Huntington disease. *Arch Gen Psychiatry* 1998;55:215–224.

90. Thune JJ, Hofsten DE, Uylings HBM, et al. Total neuron numbers in the prefrontal cortex in schizophrenia. *Soc Neurosci Abstr* 1998;24:985.

91. Benes FM, McSparren J, Bird ED, et al. Deficits in small interneurons in prefrontal and cingulate cortices of schizophrenic and schizoaffective patients. *Arch Gen Psychiatry* 1991;48:996–1001.

92. Beasley CL, Reynolds GP. Parvalbumin-immunoreactive neurons are reduced in the prefrontal cortex of schizophrenics. *Schiz Res* 1997;24:349–355.

93. Pierri JN, Volk CLE, Auh S, et al. Decreased somal size of deep layer 3 pyramidal neurons in the prefrontal cortex in subjects with schizophrenia. *Arch Gen Psychiatry* 2001;58:466=n473.

94. Glantz LA, Lewis DA. Decreased dendritic spine density on prefrontal cortical pyramidal neurons in schizophrenia. *Arch Gen Psychiatry* 2000;57:65–73.

95. Benes FM, Davidson J, Bird ED. Quantitative cytoarchitectural studies of the cerebral cortex of schizophrenics. *Arch Gen Psychiatry* 1986;43:31–35.

96. Akil M, Pierri JN, Whitehead RE, et al. Lamina-specific alteration in the dopamine innervation of the prefrontal cortex in schizophrenic subjects. *Am J Psychiatry* 1999;156:1580–1589.

97. Okubo Y, Suhara T, Suzuki K, et al. Decreased prefrontal dopamine D1 receptors in schizophrenia revealed by PET. *Nature* 1997;385:634–636.

98. Lund JS, Lund RD, Hendrickson AE, et al. The origin of efferent pathways from the primary visual cortex, area 17, of the macaque monkey as shown by retrograde transport of horseradish peroxidase. *J Comp Neurol* 1975;164:287–304.

99. Gilbert CD, Kelly JP. The projections of cells in different layers of the cat's visual cortex. *J Comp Neurol* 1975;63:81–106.

100. Glantz LA, Austin MC, Lewis DA. Normal cellular levels of synaptophysin mRNA expression in the prefrontal cortex of subjects with schizophrenia. *Biol Psychiatry* 1999;48:389–397.

101. Eastwood SL, Harrison PJ. Hippocampal and cortical growth-associated protein-43 messenger RNA in schizophrenia. *Neuroscience* 1998;86:437–448.

102. Eastwood SL, Burnet PWJ, McDonald B, et al. Synaptophysin gene expression in human brain: A quantitative *in situ* hybridization and immunocytochemical study. *Neuroscience* 1994;59:881–892.

103. Andreasen NC, Arndt S, Swayze V II, et al. Thalamic abnormalities in schizophrenia visualized through magnetic resonance image averaging. *Science* 1994;266:294–298.

104. Buchsbaum MS, Someya T, Teng CY, et al. PET and MRI of the thalamus in never-medicated patients with schizophrenia. *Am J Psychiatry* 1996;153:191–199.

105. Frazier JA, Giedd JN, Hamburger SD, et al. Brain anatomic magnetic resonance imaging in childhood-onset schizophrenia. *Arch Gen Psychiatry* 1996;53:617–624.

106. Wolkin A, Rusinek H, Vaid G, et al. Structural magnetic resonance image averaging in schizophrenia. *Am J Psychiatry* 1998;155:1064–1073.

107. Portas CM, Goldstein JM, Shenton ME, et al. Volumetric evaluation of the thalamus in schizophrenic male patients using magnetic resonance imaging. *Biol Psychiatry* 1998;43:649–659.

108. Pakkenberg B. Pronounced reduction of total neuron number in mediodorsal thalamic nucleus and nucleus accumbens in schizophrenics. *Arch Gen Psychiatry* 1990;47:1023–1028.

109. Young KA, Manaye KF, Liang C-L, et al. Reduced number of mediodorsal and anterior thalamic neurons in schizophrenia. *Biol Psychiatry* 2000;47:944–953.

110. Popken GJ, Bunney Jr. WE, Potkin SG, et al. Subnucleus-specific loss of neurons in medial thalamus of schizophrenics. *Proc Natl Acad Sci USA* 2000;97:9276–9280.

111. Pakkenberg B. The volume of the mediodorsal thalamic nucleus in treated and untreated schizophrenics. *Schiz Res* 1992;7:95–100.

112. Pierri JN, Melchitzky DS, Lewis DA. Volume and neuronal number of the primate mediodorsal thalamic nucleus: effects of chronic haloperidol administration. *Soc Neurosci Abstr* 1999;25:1833.

113. Danos P, Baumann B, Bernstein H-G, et al. Schizophrenia and anteroventral thalamic nucleus: selective decrease of parvalbumin-immunoreactive thalamocortical projection neurons. *Psychiatry Res Neuroimag* 1998;82:1–10.

114. Jones EG, Hendry SHC. Differential calcium binding protein immunoreactivity distinguishes classes of relay neurons in monkey thalamic nuclei. *Eur J Neurosci* 1989;1:222–246.

115. Lewis DA, Cruz DA, Melchitzky DS, et al. Lamina-specific reductions in parvalbumin-immunoreactive axon terminals in the prefrontal cortex of subjects with schizophrenia: evidence for decreased projections from the thalamus. *Am J Psychiatry* 2001;158:1411=n1422.

116. Melchitzky DS, Sesack SR, Lewis DA. Parvalbumin-immunoreactive axon terminals in monkey and human prefrontal cortex: laminar, regional and target specificity of Type I and Type II synapses. *J Comp Neurol* 1999;408:11–22.

117. Parnavelas JG, Lynch G, Brecha N, et al. Spine loss and regrowth in hippocampus following deafferentation. *Nature* 1974;248:71–73.

118. Garey LJ, Ong WY, Patel TS, et al. Reduced dendritic spine density on cerebral cortical pyramidal neurons in schizophrenia. *J Neurol Neurosurg Psychiatry* 1998;65:446–453.

119. Pierri JN, Edgar CL, Lewis DA. Somal size of prefrontal cortical pyramidal neurons in the thalamic recipient zone of subjects with schizophrenia. *Soc Neurosci Abstr* 1998;24:987.

120. Wellman CL, Logue SF, Sengelaub DR. Maze learning and morphology of frontal cortex in adult and aged basal forebrain-lesioned rats. *Behav Neurosci* 1995;109:837–850.

121. Hendry SHC, Jones EG. Activity-dependent regulation of GABA expression in the visual cortex of adult monkeys. *Neuron* 1988;1:701–712.

122. Benson DL, Huntsman MM, Jones EG. Activity-dependent changes in GAD and preprotachykinin mRNAs in visual cortex of adult monkeys. *Cerebral Cortex* 1994;4:40–51.

123. Akbarian S, Kim JJ, Potkin SG, et al. Gene expression for glutamic acid decarboxylase is reduced without loss of neurons in prefrontal cortex of schizophrenics. *Arch Gen Psychiatry* 1995;52:258–266.

124. Volk DW, Austin MC, Pierri JN, et al. Decreased GAD_{67} mRNA expression in a subset of prefrontal cortical GABA neurons in subjects with schizophrenia. *Arch Gen Psychiatry* 2000;57:237–245.

125. Lewis DA, Pierri JN, Volk DW, et al. Altered GABA neurotransmission and prefrontal cortical dysfunction in schizophrenia. *Biol Psychiatry* 1999;46:616–626.

126. Lewis DA. Chandelier cells: shedding light on altered cortical circuitry in schizophrenia. *Mol Psychiatry* 1998;3:468–471.

127. Volk DW, Austin MC, Pierri JN, et al. GABA transporter-1 mRNA in the prefrontal cortex in schizophrenia: decreased expression in a subset of neurons. *Am J Psychiatry* 2001;158:256=n265.

128. Woo T-U, Whitehead RE, Melchitzky DS, et al. A subclass of prefrontal gamma-aminobutyric acid axon terminals are selectively altered in schizophrenia. *Proc Natl Acad Sci USA* 1998;95:5341–5346.

129. Pierri JN, Chaudry AS, Woo T-U, et al. Alterations in chandelier neuron axon terminals in the prefrontal cortex of schizophrenic subjects. *Am J Psychiatry* 1999;156:1709–1719.

130. Goldman-Rakic PS, Leranth C, Williams SM, et al. Dopamine synaptic complex with pyramidal neurons in primate cerebral cortex. *Proc Natl Acad Sci USA* 1989;86:9015–9019.

131. Selemon LD, Lidow MS, Goldman-Rakic PS. Increased volume and glial density in primate prefrontal cortex associated with chronic antipsychotic drug exposure. *Biol Psychiatry* 1999;46:161–172.

132. Lewis DA. Distributed disturbances in brain structure and function in schizophrenia (editorial). *Am J Psychiatry* 1999;157:1=n2.

133. Harrison PJ. The neuropathology of schizophrenia: a critical review of the data and their interpretation. *Brain* 1999;122:593–624.

134. Bertolino A, Kumra S, Callicott JH, et al. Common pattern of cortical pathology in childhood-onset and adult-onset schizophrenia as identified by proton magnetic resonance spectroscopic imaging. *Am J Psychiatry* 1998;155:1376–1383.

135. Benes FM, Sorensen I, Bird ED. Reduced neuronal size in posterior hippocampus of schizophrenic patients. *Schiz Bull* 1991;17:597–608.

136. Arnold SE, Franz BR, Ruben BA, et al. Smaller neuron size in schizophrenia in hippocampal subfields that mediate cortical-hippocampal interactions. *Am J Psychiatry* 1995;152:738–748.

137. Zaidel DW, Esiri MM, Harrison PJ. Size, shape, and orientation of neurons in the left and right hippocampus: investigation of normal asymmetries and alterations in schizophrenia. *Am J Psychiatry* 1997;154:812–818.

138. Weinberger DR, Lipska BK. Cortical maldevelopment, anti-psychotic drugs, and schizophrenia: a search for common ground. *Schiz Res* 1995;16:87–110.

139. Anderson SA, Classey JD, Condé F, et al. Synchronous development of pyramidal neuron dendritic spines and parvalbumin-immunoreactive chandelier neuron axon terminals in layer III of monkey prefrontal cortex. *Neuroscience* 1995;67:7–22.

140. Woo T-U, Pucak ML, Kye CH, et al. Peripubertal refinement of the intrinsic and associational circuitry in monkey prefrontal cortex. *Neuroscience* 1997;80:1149–1158.

141. Pulver AE. Search for schizophrenia susceptibility genes. *Biol Psychiatry* 2000;47:221–230.

142. Jones EG. Laminar distribution of cortical efferent cells. In: Peters A, Jones EG, eds. *Cerebral cortex*, vol 1. New York: Plenum Press, 1984:521–553.

143. Lewis DA. The organization of cortical circuitry. In: Harrison PJ, Roberts GW, eds. *The neuropathology of schizophrenia: progress and interpretation*. Oxford: Oxford University Press, 2000:235–256.

144. Kaufer D, Lewis DA. Frontal lobe anatomy and cortical connectivity. In: Miller BL, Cummings JL, eds. *The human frontal lobes: functions and disorders*. New York: Guilford Publications, 1998;27–44.

145. Lewis DA, Lieberman JA. Catching up on schizophrenia: natural history and neurobiology. *Neuron* 2000;28:325–334.

54

Functional Neuroimaging in Schizophrenia

Karen Faith Berman

Karen Faith Berman: National Institute of Mental Health, Intramural Research Program, Bethesda, Maryland

THEORETICAL PERSPECTIVE

Part of "54 - Functional Neuroimaging in Schizophrenia "

The search for a biological basis of schizophrenia includes a long chapter in which alterations in the characteristics and localization of neural activity are the focus. It is in this arena that functional neuroimaging has had the broadest application and greatest impact in psychiatry. This now extensive body of work has left no doubt that schizophrenia is associated with measurable, objective signs of altered brain function, and clinical and pathophysiologic correlations have begun to emerge. Certain cerebral concomitants of this illness have been consistently demonstrated, although no single pathognomonic neurofunctional abnormality has been delineated yet. Increasingly, it appears that dysfunction of a system of functionally and/or structurally interconnected cortical and limbic brain regions is present to lesser or greater degrees, producing more or less psychopathology in individual patients, and that certain brain regions, such as frontal cortex, may play a special role in this larger picture. Abnormal neurotransmission within these neural circuits appears to be related to core features of schizophrenia, particularly cognitive impairment. Although it is likely that at least some of the functional abnormalities are generative of these features and not simply a response to them, clarification of this “chicken versus egg” issue must be a crucial component of any research program in this area, and the elucidation of the underlying mechanisms is the most important quest of this research at present. Current functional neuroimaging has much to offer in guiding this quest, particularly when combined with new information now available from other fields such as genetics and cognitive science. In this light, the following sections review existing findings, delineate their relationship to other neurobiological and clinical properties of the illness, discuss conceptual issues and controversies, examine methodologic considerations (including technical constraints), summarize new techniques and the new approaches to research design that they allow, and finally describe the most important areas for future research.

HISTORIC PERSPECTIVE

Part of "54 - Functional Neuroimaging in Schizophrenia "

Functional neuroimaging studies utilize the fact that neuronal activation results in regionally increased blood flow and metabolism. This can be measured either by radiotracer methods (e.g., [positron emission tomography] PET regional cerebral blood flow [rCBF] or cerebral glucose metabolic rate) or by a regional effect on the ratio of deoxyhemoglobin to oxyhemoglobin imaged by magnetic resonance techniques (the blood oxygenation level dependent [BOLD] effect). This work began in earnest some 50 years ago with the pioneering studies of Seymour Kety and colleagues who developed the first reproducible, quantitative technique for measuring cerebral blood flow (CBF) as an indicator of neuronal activity, in humans by using nitrous oxide, an inert but soluble gas. When this method was applied to schizophrenia (1), these investigators found no alteration in the overall average CBF level in patients, a result that has largely been confirmed by more recent studies; however, this finding did not rule out the existence of neurophysiologically meaningful changes in specific brain structures.

The next advance came in the 1970s with the establishment of rigorous methods that could differentiate the functional level of specific cortical regions, albeit with only 2-cm anatomic accuracy at best (2). This method, administration of tracer amounts of radioactive xenon, was rapidly applied to the study of schizophrenia. The resulting findings of functional abnormality in the frontal lobe spurred a shift in focus throughout many research domains in the field that remains a prevailing force today. In the 1980s, the advent of tomographic methods, such as single photon emission computed tomography (SPECT) and PET, which both use radioactive compounds as tracers, brought improved interregional spatial resolution on the order of 5 to 6 mm and allowed measurement of subcortical regional function.

In the last decade, functional magnetic resonance imaging (fMRI) has emerged as the premier technique for neuropsychiatric functional neuroimaging. By taking advantage of the differential paramagnetic properties of oxyhemoglobin versus deoxyhemoglobin and the altered ratio between them that occurs when blood volume and blood flow change in response to neural activation, BOLD fMRI uses intrinsic properties of the blood itself rather than an extrinsic contrast or tracer agent, to generate maps of brain function. It is, thus, entirely noninvasive, and measurements can be repeated over time, conferring significant advantage in experiments designed to address important clinical questions in schizophrenia, such as response to medication, correlation with clinical course, differentiation of state versus trait phenomena, determination of the dynamic range of neurophysiological response, and learning. This new methodologic advance brought further improvements in spatial resolution as well as enhanced temporal resolution, which, although still slow (several seconds) compared to neuronal signaling (on the order of 200 ms), improved to the degree that event-related neural activity could be recorded with anatomic precision heretofore unavailable with electrophysiologic methods.

TECHNICAL PERSPECTIVE

Part of "54 - Functional Neuroimaging in Schizophrenia "

As can be seen in the preceding brief history, over the years the sophistication of the questions that could be asked and the hypotheses about schizophrenia that could be tested have paralleled the development of new brain imaging technologies and analytic methods. This parallel development is evident in the evolution of the science from the search for regionally specific pathologic function to that in neural systems, and from measures sensitive only to static pathophysiology to explorations of the dynamic interplay among regions in those neural systems. Therefore, a brief discussion follows of the newest methodologic approaches and latest vistas for research in schizophrenia that they offer.

New Vistas in Data Acquisition

Event-related fMRI is a relatively recent class of experimental design within functional neuroimaging that exploits the superior temporal resolution of fMRI. Unlike previous fMRI and PET approaches that blocked together relatively long (e.g., 20 to 60 sec) periods of similar behavioral trials or conditions and examined average neural response over that time period, event-related fMRI documents the neural response during each individual trial or short behavioral period. This approach has much in common with traditional evoked-potential electrophysiology and offers advantages in experimental design. Because different trial types can be randomly intermixed and then separated for analysis, order effects and habituation can not only be controlled for, but also explicitly investigated. A particular advantage for research in schizophrenia is that neural activity during correct and incorrect trials can be measured separately and compared, allowing more incisive study of the mechanism of cognitive failure and better experimental control of potential confounds based in performance discrepancies that often occur between patient and control groups. Eventrelated fMRI has very recently come into wide use in neuroimaging of cognitive systems in healthy subjects, but as of this writing has had only limited application to the study of schizophrenia.

Multimodal neuroimaging, which seeks to determine the relationship of the neurofunctional abnormalities in patients to other neurobiological features, will have an increasingly important role in delineating the mechanism of schizophrenia. In this approach, blood flow and other measures such as MR spectroscopy, neuroreceptor measurements, and electrophysiology (with MEG or EEG) are determined in the same patients. One example of the richness of the data that can be gleaned is the use of PET or fMRI to measure blood flow in conjunction with EEG or MEG. PET and fMRI allow *localization* of the brain regions that work together during cognition (spatial resolution 3 to 6 mm), but provide relatively little temporal information (temporal resolution in seconds). On the other hand, EEG and MEG have relatively poorer spatial resolution, but provide fine *time resolution* (i.e., milliseconds). Combining these methods, together with the application of the advanced computational cross-registration and source localization techniques that now exist, provides exponentially more information than any of these techniques alone. For example, this allows the determination of the sequence in which various finely localized regions are activated during cognition and the testing of the hypothesis that this sequence of events is altered in schizophrenia. Although this specific multimodal approach has not been applied in schizophrenia, other examples are described in the following.

New Vistas in Data Analysis

The analytic approaches for the two data collection modalities discussed, event-related fMRI and fusion of spatial and temporal neurofunctional data (i.e., PET or fMRI with EEG or MEG), are in their infancy and beyond the scope of this chapter. Another recent set of analytic methods addresses the growing appreciation that the transduction between neural and mental events is not anatomically segregated (i.e., one brain region per one mental or cognitive activity), but rather that even the simplest of cognitive activities requires coordinated actions across widely distributed components of neural systems (interregional integration of neural function). Questions about functional integration and coordinated interregional activity are likely to have particular relevance for schizophrenia, as discussed in later sections. Although the segregational view can be

tested with univariate statistics (multiple *t*-tests or ANOVAs), hypotheses about interregional integration require multivariate approaches. The influence that one brain region exerts or experiences from another is most simply examined via the correlations between brain activity measures for the two anatomic structures; as an operational definition two brain regions can be considered to be functionally coupled if their activities are correlated (3); however, this approach cannot elucidate how other nodes in the network mediate these relationships. To model this mediation requires analysis of the covariance matrix of regions studied in its entirety. Several new methods have been developed. These include structural equation modeling (SEM) and eigenimage analysis. SEM, used in conjunction with acknowledged anatomical models, can characterize and quantify the “functional connectivity” among the multiple components of neural systems. A model of known or hypothesized anatomic pathways is defined first; then a functional model of interest is tested against this model by iterative fitting of the interregional correlational weights. It should be borne in mind that the regional components examined are preselected based on putative pathways and the results from this approach are only as good as the model. In contrast, eigenimage analysis is a data-driven approach that examines patterns of correlation across the entire brain and how they vary in a time-series, revealing distributed brain systems and their temporal dynamics. Single value decomposition or principal component analysis is used to present the percentage of variance accounted for by different patterns of activity or spatial modes, and canonical variates analysis, conceptually similar to factor analysis, can be used to extract connectivity patterns across the entire brain that are most different between the studied groups. Thus, these recently developed methods permit characterization of normal and altered neural connectivity using neuroimaging.

THE FINDINGS

Part of "54 - Functional Neuroimaging in Schizophrenia "

The explosion of functional neuroimaging studies of schizophrenia has resulted in many “findings” and many discrepancies. Nonetheless, several trends spanning the various experimental and methodologic techniques were apparent several years ago (4). First, when brain activity or metabolism is averaged across the entire brain (i.e., “global” function), patients have relatively normal values, unlike in primary degenerative disorders. Second, when scanned during so-called resting conditions (i.e., with no cognitive or motor activities required and no specific sensory input), abnormalities are inconsistent when seen. Although patients may have relatively normal regional patterns of resting brain activity, there appear to be associations between specific resting regional CBF patterns and symptom profiles (5). Changes in lateralization of brain function also have been described (6 ,7). Third, when scanned during cognitive activation, patients tend to be different from normal controls. In particular, they show abnormal prefrontal activity (8) during tests involving working memory (i.e., the system used to hold information in temporary storage to complete a task). They show deficits in cingulate cortex as well as alterations in frontal–temporal and other intracortical functional relationships (12 ,13) during other cognitive tasks, such as cued verbal recall (9) and the Stroop test (10), and in some studies at rest (11). In general, most of these findings have been reproduced in acute, untreated patients, thus excluding a primary role for medication artifacts.

More recent functional brain imaging studies in schizophrenia have focused on: (a) further characterization of the locales and cognitive and behavioral context of neurophysiologic deficits in schizophrenia; (b) delineation of the relationship of the deficits to clinical symptoms and other neurobiological features of the illness; and, most important; (c) attempts to elucidate the pathophysiologic mechanism(s) of the deficits. The following section reviews these areas of observation and others. The most enlightening of these results have emerged from experimental paradigms designed to actively engage neural systems in cognitive or other activities; the most consistent findings concern the function of the prefrontal cortex.

Frontal Lobe Circuits

The frontal lobes have played a prominent role in formulations of schizophrenia since the conceptualization of the illness. In earlier times this role was inferred by clinical analogy with known frontal lobe disorders and findings in neuropsychology and nonhuman primate studies. Substantial indirect but compelling evidence from these multiple domains conferred the status of best studied and most implicated region in schizophrenic pathophysiology on the frontal lobes. Early functional neuroimaging studies substantiated this role, beginning with Ingvar and Franzen’s 1974 seminal finding that patients with schizophrenia had relatively lower blood flow to frontal regions (14). Changes in rCBF in response to cognitive activation were also first observed in these early studies; this approach has been greatly refined and now has largely replaced the inherently ill-defined and variable resting state as the cornerstone of functional brain imaging studies in schizophrenia. These data presaged the body of literature that emerged over the ensuing 15 to 20 years, most of which had similar frontal lobe findings (Table 54.1).

First Author	Publication Year	Reference Number	Imaging Technique	Activation Paradigm
Studies Reporting Decreased Frontal Lobe Activation in Schizophrenia				
Berman KF, et al.	1986	69	Xenon-133 rCBF	WCST/NMT
Weinberger DR, et al.	1986	70	Xenon-133 rCBF	WCST/NMT
Volkow ND, et al.	1987	71	Carbon-11 LCMRglu PET	Eye tracking
Weinberger DR, et al.	1988	45	Xenon-133 rCBF	WCST/NMT
Guich SM, et al.	1989	72	Fluorine-18 LCMRglu PET, E	CPT
Buchsbaum MS, et al.	1990	73	Fluorine-18 LCMRglu PET	CPT
Daniel DG, et al.	1991	44	Xenon-133 SPECT rCBF	WCST
Cantor-Grae E, et al.	1991	15	Xenon-133 rCBF	RPM, VT, WCST
Rubin P, et al.	1991[a]	22	Tc-SPECT rCBF	WCST
Andreasen NC, et al.	1992[a]	24	Xenon-133 rCBF	TOL
Berman KF, et al.	1992	18	Xenon-133 rCBF	WCST/NMT
Buchsbaum MS, et al.	1992[a]	26	Fluorine-18 PET	CPT
Lewis SW, et al.	1992	74	Tc-SPECT rCBF	Verbal fluency
Kawasaki Y, et al.	1993	75	Tc-SPECT rCBF	WCST
Nakashima Y, et al.	1994	76	Oxygen-15 Water PET rCBF	Volitional saccade
Guenther W, et al.	1994	77	Fluorine-18 LCMRglu PET	Complex motor task
Catafau AM, et al.	1994[a]	25	Tc-SPECT rCBF	WCST
Parellada E, et al.	1994[a]	78	SPECT rCBF	WCST
Rubin P, et al.	1994[a]	23	Tc-SPECT rCBF	WCST
Siegel BV, et al.	1995	79	Fluorine-18 LCMRglu PET	CPT
Steinberg JL, et al.	1996[a]	80	Xenon-133 rCBF	WCST/NMT
Yurgelun-Todd D, et al.	1996	81	fMRI	Verbal fluency
Ganguli R, et al.	1997	82	Oxygen-15 Water PET rCBF	Supraspan memory
Shajahan PM, et al.	1997	83	Tc-SPECT rCBF	Oddball task
Volz HP, et al.	1997	84	fMRI	WCST
Gracia Marco R, et al.	1997	85	Tc-SPECT rCBF	WCST
Callicott JH, et al.	1998	86	fMRI	N-back Working memory
Carter CS, et al.	1998	87	Oxygen-15 water PET rCBF	N-back Working memory
Ragland JD, et al.	1998	88	Oxygen-15 water PET rCBF	WCST PART
Curtis VA, et al.	1998	89	fMRI	Word generation
Fletcher PC, et al.	1998	55	Oxygen-15 water PET rCBF	Graded working memory
Parellada E, et al.	1998[a]	90	Tc-SPECT rCBF	Rest, WCST
Stevens AA, et al.	1998	91	fMRI	Tone/auditory working memory
Crespo-Facorro B, et al.	1999	92	PET rCBF	Verbal learning
Volz H, et al.	1999	93	fMRI	CPT
Artiges E, et al.	2000	94	Oxygen-15 water PET rCBF	Random number generation
Higashima M, et al.	2000	95	Tc-SPECT rCBF	Auditory discrimination
Heckers S, et al.	2000	96	Oxygen-15 gas inhalation PET	Visual object recognition
Holcomb HH, et al.	2000	97	Oxygen-15 water PET rCBF	Auditory recognition
Russell TA, et al.	2000	98	fMRI	Mental state attribution
Studies Reporting Increased Frontal Lobe Activation in Schizophrenia				
Heckers S, et al.	1998	48	Oxygen-15 gas inhalation PET	Verbal episodic memory
Manoach DS, et al.	1999	99	fMRI	Sternberg working memory
Manoach DS, et al.	2000	100	fMRI	Sternberg working memory
Callicott JH, et al.	2000	40	fMRI	N-back working memory
Studies Reporting No Between-Group Differences				
Frith CD, et al.	1995	101	Oxygen-15 gas inhalation PET	Verbal fluency
Buckely PF, et al.	1997	102	fMRI	Motor
Spence SA, et al.	1998	103	Oxygen-15 water PET rCBF	Motor
Curtis VA, et al.	1999	104	fMRI	Semantic decision task
Braus DF, et al.	2000	105	fMRI	Motor

[a]Neuroleptic-naive patients.
CPT, continuous performance test; fMRI, functional magnetic resonance imaging; LCMRglu, local cerebral metabolic rate of glucose; NMT, number matching task; PART, paired associate recognition task; PET, positron emission tomography; RPM, Raven's Progressive Matrices; SPECT, single photon emission computed tomography; Tc, technetium99m; TOL, Tower of London; WCST, Wisconsin Card Sorting Test.

TABLE 54.1. FRONTAL LOBE FINDINGS WITH ACTIVATION PARADIGMS IN SCHIZOPHRENIA SINCE 1985

Characterization of Dorsolateral Prefrontal Functional Alterations

The relatively subtle prefrontal functional alterations observed in schizophrenia have been increasingly brought into focus by recent advances in the available neuroimaging armamentarium. Table 54.1 lists activation studies published

over the past 15 years that report frontal lobe results. The overwhelming majority of these investigations have detected abnormal prefrontal responses to a variety of cognitive activities designed to access and/or control frontal neural circuitry, particularly working memory. The prefrontal site most commonly affected is the dorsolateral prefrontal cortex (DLPFC), and, until recently, the physiologic abnormality in this brain region was consistently seen as hyporesponsivity. Indeed, from the time of Ingvar and Franzen's first study and throughout most of the last decade researchers were satisfied that the qualitative nature of the frontal lobe abnormalities was known. However, the relative universality with which the schizophrenic prefrontal cortex had been reported to be hypofunctional, in the past several years, has given way to the notion that the aberrant neural responses in prefrontal cortex are more complex, including *hyper*function under some circumstances. It is noteworthy that abnormally increased prefrontal response has been primarily seen in studies carried out with fMRI, rather than PET, and mainly when the cognitive paradigms take advantage of the temporal properties of fMRI in order to employ shorter blocks of task performance, and/or require task switching. This fact suggests that the anatomic and/or chemical perturbations of the schizophrenic prefrontal cortex can be manifest by inappropriate over recruitment of prefrontal neural circuitry during relatively brief cognitive challenge and failure to sustain this recruitment over longer periods; however, more work is required to understand the implications of these recent observations and to fully characterize them. Regardless of the direction of prefrontal physiologic abnormality, the functional neuroimaging literature leaves little doubt that such abnormality exists. A variety of potential epiphenomenological explanations for this pathophysiology have been considered, and a number of neurobiologically plausible mechanisms have been proposed.

Effect of Neuroleptic Treatment

One potential epiphenomonological confound is the possibility of a causative role of antipsychotic medications, an important consideration because the majority of functional neuroimaging studies have been carried out in patients who were either receiving neuroleptics at the time of study or previously treated and then withdrawn for some prior period. However, taken as a whole, the literature provides little evidence that neuroleptics generate the functional neuroimaging abnormalities. First, although additional longitudinal investigations with newer techniques are necessary, an 18-year follow up (15) showed that prefrontal hypofunction in chronic patients is remarkably stable over time and unaffected by long-term consistent neuroleptic treatment. Second, prefrontal abnormalities like those observed in schizophrenia are not seen in other illnesses in which neuroleptics are used, such as Huntington's disease (16 ,17). Third, a study of monozygotic twins concordant for schizophrenia, but with differing lifetime histories of neuroleptic intake, found that in most pairs the twin with less exposure to neuroleptics was actually the more hypofrontal of the pair during a prefrontally linked task (18), a result opposite to that expected if neuroleptics caused hypofrontality. Fourth, studies examining metabolic or rCBF changes occurring when patients go from the unmedicated to the medicated state are quite inconsistent (4 ,8). Fifth, frontal lobe functional abnormalities have been found in first-degree relatives of patients with schizophrenia who have not been treated with neuroleptics or hospitalized (19 ,20 and 21). Finally and perhaps most conclusively, frontal lobe abnormalities during cognition have been found in a number of studies of young patients who have never received neuroleptics (22 ,23 ,24 ,25 and 26). Thus, in summary, there are little convincing data that neuroleptic treatment is a major factor in frontal lobe functional abnormalities.

The Link with Cognitive Performance

Most prominent among potential epiphenomena and confounds examined has been the effect of poor performance. It is undisputed that patients with schizophrenia perform poorly on the very cognitive tasks that best show prefrontal pathophysiology—the Wisconsin Card Sorting test (WCST), the N-back test, the Tower of London, or others. This is not surprising for the reason that the tasks are chosen because they are thought to reflect core deficits of the illness and they access neural systems relevant to it; however, considerable controversy has arisen around the possibility that the poor performance—or differences in attention or effort and motivation—somehow cause the frontal pathophysiology, rather than the more neurobiologically plausible explanation that underlying pathophysiology is responsible for the poor performance. It is clear that if patients are simply not engaged in a cognitive task during scanning they will not activate relevant brain regions; such results are obviously artifactual. It is less clear what should be predicted if patients (or healthy subjects) are manifestly working at a task, but performing it abnormally.

The relationship between performance level and the degree to which the brain neurophysiologically responds is complex, even in the presumptive absence of pathology in healthy control subjects. Several studies have found significant correlations between prefrontal neural activity and cognitive function, suggesting that these two variables are paradigmatically linked, but both positive and negative (27) relationships have been described (8). The research challenge has been to (a) understand this relationship and (b) tease apart abnormal cognitive performance and abnormal brain activity in patients to determine which is primary. This is a difficult issue to investigate experimentally, and no single study alone can answer the question; however, convergent evidence derived from several different research

directions leads to the conclusion that prefrontal pathophysiology cannot be accounted for as an epiphenomenon (8).

First, studies have been carried out on patient populations who, like schizophrenics, perform poorly on frontal lobe tasks but who have disorders other than schizophrenia. In principal, if the prefrontal physiologic deficit found in patients with schizophrenia is an epiphenomenon of poor performance per se, then other subjects who perform as poorly should have similar prefrontal function. Pathophysiology quite distinct from that characterizing schizophrenia has been reported in Huntington's disease (28) and normal aging (29 ,30) where performance is matched to that in schizophrenia, as well as in Down's syndrome (31) in which performance is worse. These findings indicate that poor performance per se does not necessarily produce the pathophysiologic picture seen in schizophrenia.

A second way to experimentally attack this "chicken and egg" question, and at least on the face of it the most direct way, is to match patients and normal controls for level of performance. However, ensuring good performance in patients (by using different versions of the task for patients and controls or by making the task extremely easy) does not guarantee that the effort involved or the cognitive operations used by the two groups are equated. Moreover, in the absence of neuropsychologic impairment, the neural systems accessed and findings may have very little to do with the illness. Thus, the strategy of employing "easy" tasks that result in some measures of performance being "normal" in patients is not as straightforward an approach to exploring neurophysiologic abnormalities in schizophrenia as it may appear (4 ,8). An alternative to "matching" for good performance is to study normal controls who perform as *poorly* on a given task as the patients. This strategy at least addresses the question of whether normals and patients fail by the same pathophysiologic mechanisms. In a study designed on this premise, patients performing the WCST were found to activate DLPFC less overall than performance-matched normals, but they activated a more anterior area of prefrontal cortex that is not recruited in normals as a group. Moreover, the more a given normal subject activated this "schizophrenic card sort area," the more he or she perseverated (8).

This overactivation of an aberrant area (or of an appropriate area, for that matter) would be difficult to explain on the basis of disengagement from the task and poor performance per se. In a similar example, affected members of monozygotic co-twins discordant for schizophrenia showed hypofunction of prefrontal cortex accompanied by hippocampal hyperfunction during the WCST (32). Analogous results have been described by Friston and colleagues using paced verbal production (33) and have recently been demonstrated with eigenimage analysis in a study of medication-free patients performing the N-back working memory task (34).

Yet a third approach to inform the performance conundrum is to model the abnormal cognition in normal controls. Both underactivation and overactivation have been described in the context of performance difficulties. Goldberg and associates (1998) found that healthy subjects performing a dual task, working memory plus auditory and verbal shadowing, had significant decrements in both performance and DLPFC blood flow (35). Callicott and co-workers (1999) demonstrated that normal controls pushed beyond their working memory capacities also demonstrate reduced DLPFC responses (36). Electrophysiologic recordings in working memory-related neurons of nonhuman primates provide a neurobiological framework for these observations, where failed working memory trials are accompanied by decreased firing rates (37 ,38). Thus, findings of decreased prefrontal response in schizophrenia can be seen as part of an expected curve between working memory load and neural response—a dose–response curve that is shifted in the face of patients' reduced prefrontal capacities. Given the fact that other poorly performing patient populations with pathology that is different than schizophrenia do not show the same prefrontal response one (28 ,30 ,31), this particular shift in the dose–response curve with poor performance is neither inevitable nor the only possible one.

The DLPFC overactivtion response in schizophrenia also has a precedent in the normal neurocognitive imaging literature. Rypma and D'Esposito (1999) found in healthy subjects that the longer the reaction time (an indicator of difficulty with a task), the greater the DLPFC neural response (27). Although this may simply represent increased "time on task," it again provides a context in which to view recent findings of overactivation in patients. Although some clues about the cellular mechanism underlying patients' aberrant prefrontal function are discussed in the following, it remains for future research to determine whether patients and controls share the same source for abnormal prefrontal responses (both underactivation and overactivation) in the context of performance difficulties.

The data summarized in the preceding leave little doubt that, although the patients' poor performance is certainly an integral, primary component of their illness, it is not the cause of the abnormal neural function, rather it is an effect of that pathophysiology. Given this conclusion and the now extensive literature documenting it, the most efficacious use of functional neuroimaging in future schizophrenia research will occur if the poor performance and prefrontal abnormalities are considered as inextricably linked and studies are designed to elucidate the mechanism of these linked phenomena.

Putative Underlying Mechanisms

A number of clues about potential pathophysiologic mechanisms emerge when the neuroimaging findings are considered in light of other neurobiological hallmarks of schizophrenia and findings from different clinical neuroscience research modalities. First, it is likely that cellular pathology in DLPFC contributes to the cognitively linked dysfunction.

The degree of both hyperfunction and hypofunction of DLPFC in patients are predicted by decreased *n*-acetyl-aspartate (NAA), an MR spectroscopy measure of cellular integrity (39 ,40). Second, several compelling lines of evidence suggest a prominent role for dopaminergic dysfunction. It is well documented in nonhuman primates that optimal dopamine function is necessary for maximal working memory and DLPFC physiologic function (41). Similar, albeit less direct, evidence also exists in humans: Pharmacologically altering dopaminergic tone with agents such as amphetamine affects DLPFC activity in both healthy subjects (42 ,43) and patients (44); and a relationship between DLPFC rCBF during the WCST and CSF levels of the dopamine metabolite homovanillic acid has been found in schizophrenia (45). Third, converging data increasingly point to developmental mechanisms. A considerable body of literature now documents that disruption of corticolimbic connectivity in neonatal animals via hippocampal lesions models many features of schizophrenia, including working memory impairment, reduced prefrontal NAA, and dopamine dysregulation (53). Developmental pathology with a genetic basis also appears likely. A recent study links a genetic attribute that affects prefrontal dopamine to both working memory performance and DLPFC activation in patients, their sibs, and unrelated healthy individuals (21). Moreover, an association of schizophrenia with the allele that confers poor prefrontal function (a functional polymorphism in the catechol-o-methyl transferase [COMT] gene that affects enzyme activity) is indicated by several family-based studies. Investigations of this type, which explore the interaction of genetic and neurophysiologic characteristics, hold the greatest promise for elucidating the etiology of the illness and effecting innovative treatments.

Other Frontal Lobe Subregions

Dysfunction, primarily hypofunction, of portions of the frontal lobes other than the DLPFC has also been described. Just as a cognitive foundation for DLPFC dysfunction has been found in working memory and executive function, the dorsal anterior cingulate is classically considered to play an important role in vigilance, attention, and effort. It may, thus, be especially prone to epiphenomenologic effects. In normal subjects the region is activated by a variety of cognitively demanding tasks requiring stimulus-response selection in the face of competing information, such as the Stroop test, complex motor control tasks, verbal fluency, and working memory. These observations have led cognitive neuroscientists to propose more refined cognitive roles such as on line monitoring, conflict monitoring, and error detection (46). Further research is necessary to clarify which of these putative cognitive roles, or which epiphenomena, may be linked to the finding of anterior cingulate underactivation in schizophrenia (9 ,10 and 11). Orbitofrontal cortex, along with the ventral portion of the anterior cingulate, has been most linked to emotion. Few studies have been carried out that formally investigate the function of these regions in schizophrenia.

Lateral And Medial Temporal Lobe

The temporal lobe is of interest in schizophrenia for several reasons. Diseases of the medial temporal lobe can be associated with psychotic symptoms, and some neuropsychological aspects of schizophrenia implicate both lateral and medial temporal lobe. A number of neuroimaging studies have reported functional abnormalities in both lateral and medial temporal lobe structures (47). The data as a whole, however, are less compelling than for frontal lobe, and confounds and potential mechanisms are less well explored. Findings of both hyperfunction and hypofunction have been reported, but the bulk of the evidence leans toward overactivity. Heckers and colleagues (48) reported reduced hippocampal activation during the effort to retrieve poorly encoded material; however, it is of interest that, hippocampal activity appeared to be increased at baseline, again emphasizing the task-dependence of neurofunctional findings in general (48). Several studies point to a role for lateral temporal cortex in hallucinations and other positive symptoms—abnormal functional interactions between temporolimbic and prefrontal structures are also discussed in the following.

Miscellaneous Regional Changes

Functional abnormalities, primarily hypofunction, of many other regions have been reported. Although most are unreplicated, several are worth mentioning. Both increased and decreased basal ganglia activity have been found, but a role for neuroleptic treatment in such findings must be considered. Several investigators have suggested that schizophrenia is characterized by increased posterior cortical activity. In fact, Ingvar and Franzen's seminal 1974 rCBF study of schizophrenia suggested that the hypofrontal pattern represented a redistribution of flow with relatively lower flow to frontal areas as well as relatively higher flow to posterior cortex (14).

Such changes in activity patterns or distributions have received considerable attention. In particular, the notion that schizophrenia may involve disordered functional lateralization has been explored using a variety of methods. Left temporal overactivation was seen in this light in early studies. More recent work suggests that apparent alterations in functional laterality in schizophrenia may not actually reflect abnormal lateralization per se, but rather a failure to organize a lateralized response (6 ,49). For example, Mattay and associates (1997) reported less lateralized and localized lateral premotor area activation in patients during a simple finger movement paradigm (50). This may also be viewed

within the more general context of nonfocalized, less efficient, neurophysiologic responses in schizophrenia.

Interregional Relationships and Functional Connectivity

As the foregoing indicates, a number of brain areas have been shown to function abnormally in schizophrenia, often depending on the cognitive or other conditions under which the scanning is performed. It has been proposed that such multiple, seemingly local changes may be indicators of more ubiquitous dysfunction throughout widely distributed and interactive brain networks (12 ,51), a heuristically appealing pathophysiologic model for schizophrenia given the apparent subtlety of the neurophysiologic abnormalities in the face of the devastating and complex nature of the illness. This conceptualization is consistent with recent trends in viewing higher brain processes as parallel and distributed functions.

Because of the particularly extensive connectivity of the frontal lobe with other cortical and thalamic relay areas and its special role in schizophrenia, it is not surprising that many putative aberrant networks in the illness also involve it. Although the specifics of this network will obviously vary by task, a consistent finding in studies of working memory is a coactivation of prefrontal, anterior cingulate and parietal structures. Consistent with this, Bertolino and colleagues found a tight correlation between DLPFC NAA (indicative of neuronal integrity) and rCBF activation during the WCST, not only in DLPFC, but also with the other nodes in the working memory pathway (39). Because this was not evident in healthy controls, these findings appear to reflect a rate-limiting factor related to the disease process of schizophrenia.

Neurofunctional evidence of abnormal interactions between prefrontal and temporal/limbic areas has accrued for a number of years. Weinberger and associates (52) found in monozygotic twins discordant for schizophrenia an inverse relationship between the volume of the hippocampus (the structural variable that best differentiated well from ill twins) and the degree of dorsolateral prefrontal activation during prefrontal cognition (the physiologic variable that best differentiated the co-twins). This suggests dysfunction of neocortical–limbic connectivity in schizophrenia and is consistent with, if not confirmatory of, a neurodevelopmental mechanism (53). It has been suggested that abnormal development or plasticity of hippocampal connectivity affects the development and function of prefrontal cortex or, alternatively, that both regions are “put at risk” by the same pathologic mechanism (e.g., genetic variation) (54).

The new analytic tools recently developed to search more incisively for evidence of subtle and multidimensional abnormalities across the entire brain (see the foregoing) have provided results that are consistent with and extend the notion of temporohippocampal and prefrontal circuitry failure (13 ,33). During working memory, Meyer-Lindenberg and colleagues (34), using an eigenimage method (discussed in the preceding), uncovered differences between patients and controls in task-independent functional connectivity patterns charcterized by hypofrontality coexisting with increased temporal/hippocampal and cerebellar overactivity in the patients; another, task-related, pattern involving the working memory system (including DLPFC and inferior parietal lobule) was found to be more variable (i.e., showed altered modulation) in the patients, specifically during the working memory condition. Friston and Frith (12), using PET data from a verbal fluency experiment and a method that allowed them to assess patterns of activation most different between normals and patients, found that the prefrontal and temporal coactivations in normals were uncoupled (i.e., did not appear in the same pattern) in the patients. Fletcher and colleagues (55) reported similar results, and Jennings and co-workers (56) using structural equation modeling found altered neural interactions among frontal regions as well as between the frontal and temporal cortices in schizophrenics during a semantic processing task. Disruption of frontal–temporal connectivity has also been found using an EEG coherence measure (57).

Other studies have focused on medial prefrontal and cortical–striatal–thalamic circuit abnormalities (49); Biver and colleagues (58) and Mallet and associates (59), calculating correlations between various regions of glucose metabolic rate in PET, found decreased intrafrontal, as well as frontal–posterior connectivity. Andreasen (60) has advanced a hypothesis implicating compromised connectivity among prefrontal regions, several thalamic nuclei, and the cerebellum as the cause of a fundamental cognitive deficit in schizophrenia. She called the disruption in this circuitry “cognitive dysmetria,” signifying “poor coordination of mental activities” that manifests itself in difficulty in prioritizing, processing, coordinating, and responding to information. This hypothesis is based on a number of studies from her group (61 ,62 and 63) in which the structures enumerated above were found to differ in activation between schizophrenics and controls during several unrelated tasks and in different cohorts, and on the fact that the circuit described is anatomically connected.

In summary, taken together, the functional brain imaging evidence is consistent with the notion that schizophrenia involves dissolution of neuronal interactions and that many features of schizophrenia may best be viewed as dysfunctional interregional circuitry. The details of this circuitry dysfunction differ, depending on the distributed network called into play during the particular behavior, but prefrontal cortex may play a special role.

The Relationship of the Neurofunctional Abnormalities to Clinical Hallmarks

The importance of linking pathophysiologic findings in schizophrenia to clinical aspects of the illness was recognized

early on. Ingvar and Franzen (14) noted that “hypofrontality” was most prominently seen in the most withdrawn, inactive, socially isolative, and “hypointentional” patients, whereas they related the hyperfunction in posterior areas that they observed to a “hypergnostic” component of the illness. The attempt to delineate the clinical and neurobiological implications of the physiologic abnormalities remains an important focus. Studies have primarily searched for neurophysiologic associations with cognitive deficits, symptom clusters, and individual clinical features such as hallucinations. The small sample sizes of some studies and the necessarily phenomenologic nature of research into the neurophysiologic underpinnings of clinical symptoms makes firm conclusions difficult, but some consistent findings have emerged. Frontal lobe dysfunction is consistently linked to negative symptoms and cognitive deficits, particularly working memory and executive function. For example, Goldberg and colleagues (64) used an intra-twin pair difference method in which unaffected co-twins served as individual controls for each patient in the NIMH monozygotic twin sample. Although left hippocampal size predicted a parameter of verbal memory, prefrontal blood flow and perseveration on the WCST were related. These data are part of a growing literature implicating medial temporal and prefrontal regions in symptom expression and some neurocognitive deficits of the illness.

In general, hallucinations are associated with sensory modality-specific activation in brain regions involved in normal sensory processing (65). For example, auditory hallucinations appear linked to language-related regions such as Broca’s area (66) and left superior temporal cortex. Silberswieg and co-workers (67) made similar findings using a technique that quantifies the relationship between brain activity and density of hallucinations during the scanning period.

Specific cerebral blood flow patterns have been associated with distinct syndromes of schizophrenic symptoms (5): psychomotor poverty with decreased activity in DLPFC; disorganization and impaired suppression of inappropriate responses with increased activity in the right anterior cingulate gyrus; and reality distortion, which may arise from disordered internal monitoring, with increased activity medial temporal lobe at a locus activated in normal subjects during internal monitoring of eye movements. Thought disorder has been linked to temporal lobe overactivity. Kaplan and associates (68) found an association of marked psychomotor poverty with superior parietal as well as prefrontal areas, hallucinations and delusions with abnormalities in left temporal cortex, and disorganization with left inferior parietal lobule abnormalities. Further work undoubtedly will refine these interesting clinical and pathophysiologic correlates.

QUESTIONS FOR THE FUTURE

Further characterization of the abnormalities delineated by functional neuroimaging in schizophrenia is a clear goal for the future. One important advance will come from temporal dissection of the abnormal neurophysiologic signals that have now been localized with great anatomic precision. For example, the particularly high degree of both segregation and interaction of the frontal lobe complex appears to be essential for regulating and monitoring the functions it supports via multisynaptic feedback loops modulating posterior brain areas. It is likely that the functional disconnection in schizophrenia described in the preceding includes abnormalities in these feedback loops, which operate on a time scale less than 200 msec. Progress in understanding the etiology of frontal lobe dysfunction in schizophrenia, therefore, requires a methodology that has optimal resolution both spatially (in order to reliably differentiate functionally segregated areas) and temporally (to tap into the time scale in which the feedback loop organization operates). The simultaneous combination of PET or fMRI studies (which afford relatively high spatial resolution) with methods having excellent temporal resolution such as EEG or MEG (which provide temporal resolution in the order of milliseconds) will allow explicit investigation of specific hypotheses about prolongation of the feedback and feed-forward latencies and about disease-related changes in the order in which components of distributed neural systems come into play in schizophrenia.

A second important way in which characterization of the abnormal neurophysiologic signals must advance is further investigation into their relationship to other neurobiological features of the illness. One question that requires closer scrutiny is the relation of the neurophysiologic abnormalities to dopaminergic and other neurochemical parameters. Hypotheses about neurochemical mechanisms can be tested directly with functional brain imaging, both by examining the effects of pharmacologic interventions and direct *in vivo* measurements of various components of neurochemical systems. Also, the relationship of the functional abnormalities to the neurostructural and neurochemical findings described in other chapters must be further elucidated. Not only does such a multimodal approach provide critical cross-validation of the information gleaned from the different technologies and help to rule out epiphenomena, but it is also a means to more closely approach causality and mechanism. For example, links with dopaminergic dysfunction can elucidate putative genetic mechanisms (21). Similarly, the fact that the prefrontal functional abnormalities may relate to structural pathology in other (particularly limbic) areas lends credence to the notion of a neurodevelopmental mechanism, although it does not provide proof; further work, perhaps expanding on insights from animal models (53), will be necessary to test the roles of the temporal lobe and aberrant neural development in the genesis of schizophrenic psychopathology.

Clarifying the mechanism by which the pathophysiology and illness arise is the most important question that can be addressed with functional brain imaging. Considerable

work remains to be done, although some clues have emerged, particularly with regard to the frontal lobe. Longitudinal studies are necessary to differentiate trait from state phenomena. Brain imaging undoubtedly will also continue to impact the current effort on many fronts to uncover the genetic foundations of schizophrenia, by offering new targets for linkage and association studies and providing clues to direct hypothesis-driven genetic investigations, as discussed in this chapter. Such studies provide a unique perspective from which to view brain function, one that offers the possibility of uncovering basic fundamental principals important in the genesis of schizophrenia and that has the potential to lead to direct intervention.

REFERENCES

1. Kety SS, Woodford RB, Harmel MH, et al. Cerebral blood flow and metabolism in schizophrenia. *Am J Psychiatry* 1948;104:765.

2. Lassen NA, Ingvar DH. Radioisotopic assessment of regional cerebral blood flow. *Progr Nucl Med* 1972;1:376.

3. Friston KJ, Frith CD, Liddle PF, et al. Functional connectivity: the principal-component analysis of large (PET) data sets. *J Cereb Blood Flow Metab* 1993;13:5–14.

4. Berman KF, Weinberger DR. Functional brain imaging studies in schizophrenia. In: Charney DS, Nestler EJ, Bunney BS, eds. *Neurobiology of mental illness*. New York: Oxford University Press, 1999:2246–2257.

5. Liddle PF, Friston KJ, Frith CD, et al. Patterns of cerebral blood flow in schizophrenia. *Br J Psychiatry* 1992;160:179–186.

6. Gur RE, Jaggi JL, Shtasel DL, et al. Cerebral blood flow in schizophrenia: effects of memory processing on regional activation. *Biol Psychiatry* 1994;35:3–15.

7. Gur RE, Mozley PD, Resnick SM, et al. Resting cerebral glucose metabolism in first-episode and previously treated patients with schizophrenia relates to clinical features. *Arch Gen Psychiatry* 1995;52:657–667.

8. Weinberger DR, Berman KF. Prefrontal function in schizophrenia: confounds and controversies. *Phil Trans R Soc (Lond)* 1996;351:1495–1503.

9. Fletcher PC, Frith CD, Grasby PM, et al. Local and distributed effects of apomorphine on fronto-temporal function in acute unmedicated schizophrenia. *J Neurosci* 1996;16:7055–7062.

10. Carter CS, Mintun M, Nichols T, et al. Anterior cingulate gyrus dysfunction and selective attention deficits in schizophrenia: [^{15}O]H_2O PET study during single-trial Stroop task performance. *Am J Psychiatry* 1997;154:1670–1675.

11. Haznedar MM, Buchsbaum MS, Luu C, et al. Decreased anterior cingulate gyrus metabolic rate in schizophrenia. *Am J Psychiatry* 1997;154:682–684.

12. Friston KJ, Frith CD. Schizophrenia: a disconnection syndrome? *Clin Neurosci* 1995;3:89–97.

13. Friston KJ. Theoretical neurobiology and schizophrenia. *Br Med Bull* 1996;52:644–655.

14. Ingvar DH, Franzen G. Abnormalities of cerebral blood flow distribution in patients with chronic schizophrenia. *Acta Psychiatr Scand* 1974;50:425-462.

15. Cantor-Graae E, Warkentin S, Franzen G, et al. Aspects of stability of regional cerebral blood flow in chronic schizophrenia: an 18-year followup study. *Psychiatry Res* 1991;40:253–266.

16. Weinberger DR, Berman KF, Iadarola M, et al. Prefrontal cortical blood flow and cognitive function in Huntington's disease. *J Neurol Neurosurg Psychiatry* 1988;51:94–104.

17. Weinberger DR, Berman KF. Speculation on the meaning of metabolic hypofrontality in schizophrenia. *Schizophr Bull* 1988;14:157–168.

18. Berman KF, Torrey EF, Daniel DG, et al. Regional cortical blood flow in monozygotic twins discordant and concordant for schizophrenia. *Arch Gen Psychiatry* 1992;49:927–934.

19. O'Driscoll GA, Benkelfat C, Florencio PS, et al. Neural correlates of eye tracking deficits in first-degree relatives of schizophrenic patients: a positron emission tomography study. *Arch Gen Psychiatry* 1999;56:1127–1134.

20. Blackwood DH, Glabus MF, Dunan J, et al. Altered cerebral perfusion measured by SPECT in relatives of patients with schizophrenia. Correlations with memory and P300. *Br J Psychiatry* 1999;175:357–366.

21. Egan MF, Goldberg TE, Kolachana BS, et al. Effect of COMT Val108/158Met genotype on frontal lobe function and risk for schizophrenia. *Proc Nat Acad of Sci (USA)* 2000;98:6917–6922.

22. Rubin P, Holm S, Friberg L, et al. Altered modulation of prefrontal and subcortical brain activity in newly diagnosed schizophrenia and schizophreniform disorder. A regional cerebral blood flow study. *Arch Gen Psychiatry* 1991;48:987–995.

23. Rubin P, Holm S, Madsen PL, et al. Regional cerebral blood flow distribution in newly diagnosed schizophrenia and schizophreniform disorder. *Psychiatry Res* 1994;53:57–75.

24. Andreasen NC, Rezai K, Alliger R, et al. Hypofrontality in neuroleptic-naive patients and in patients with chronic schizophrenia. Assessment with xenon 133 single-photon emission computed tomography and the Tower of London. *Arch Gen Psychiatry* 1992;49:943–958.

25. Catafau AM, Parellada E, Lomena FJ, et al. Prefrontal and temporal blood flow in schizophrenia: resting and activation technetium-99m-HMPAO SPECT patterns in young neuroleptic-naive patients with acute disease. *J Nucl Med* 1994;35:935–941.

26. Buchsbaum MS, Haier RJ, Potkin SG, et al. Frontostriatal disorder of cerebral metabolism in never-medicated schizophrenics. *Arch Gen Psychiatry* 1992;49:935–942.

27. Rypma B, D'Esposito M. The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. *Proc Natl Acad Sci USA* 1999;96:6558–6563.

28. Goldberg TE, Berman KF, Mohr E, et al. Regional cerebral blood flow andneuropsychology in Huntington's disease and schizophrenia: a comparison of patients matched for performance on a prefrontal type task. *Arch Neurol* 1990;47:418–422.

29. Esposito G, Kirkby JD, Van Horn JL, et al. Impaired Wisconsin card sorting test performance in normal aging and in schizophrenia: PET evidence of different pathophysiological mechanisms for a common cognitive deficit. *Neuroimage* 1996;3:3.

30. Esposito G, Kirkby BS, Van Horn JD, et al. Context-dependent, neural system-specific neurophysiological correlates of aging: mapping PET correlates during cognitive activation. *Brain* 1999;122:963–979.

31. Schapiro MB, Berman KF, Alexander GE, et al. Regional cerebral blood flow in Down syndrome adults during the Wisconsin Card Sorting test: exploring cognitive activation in the context of poor performance. *Biol Psychiatry* 1999;45:1190–1196.

32. Weinberger DR, Berman KF, Ostrem JL, et al. Disorganization of prefrontal-hippocampal connectivity in schizophrenia: a PET study of discordant MZ twins. *Soc Neurosci Abstr* 1993;19:7.

33. Friston KJ, Frith CD, Fletcher P, et al. Functional topography: multidimensional scaling and functional connectivity in the brain. *Cereb Cortex* 1996;6:156–164.

34. Meyer-Lindenberg A, Poline J-B, Kohn P, et al. Altered cortical functional connectivity during working memory in schizophrenia, *Am J Psychiatry* 2001;158.

35. Goldberg TE, Berman KF, Fleming K, et al. Uncoupling cognitive workload and prefrontal cortical physiology: a PET rCBF study. *NeuroImage* 1998;7:296–303.

36. Callicott JH, Mattay VS, Bertolino A, et al. Capacity constraints in working memory revealed by functional MRI. *Cerebral Cortex* 1999;9:20–26.

37. Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 1989;61:331–349.

38. Funahashi S, Bruce CJ, Goldman-Rakic PS. Neuronal activity related to saccadic eye movements in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 1991;65:1464–1483.

39. Bertolino A, Esposito G, Callicott JH, et al. A specific relationship between prefrontal neuronal N-acetyl-aspartate and activation of the working memory cortical network in schizophrenia. *Am J Psychiatry* 2000;157:26–33.

40. Callicott JH, Bertolino A, Mattay VS, et al. Physiological dysfunction of the dorsolateral prefrontal cortex in schizophrenia revisited. *Cereb Cortex* 2000;10:1078–1092.

41. Williams GV, Goldman-Rakic PS. Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* 1995;17:572–575.

42. Mattay VS, Berman KF, Ostrem JL, et al. Dextroamphetamine enhances neural network specific physiological signals: a PET rCBF study. *J Neurosci* 1996;16:4816–4822.

43. Mattay VS, Callicott JC, Bertolino A, et al. Effects of dextroamphetamine on cognitive performance and cortical activation. *Neuroimage* 2000;12:268–275.

44. Daniel DG, Weinberger DR, Jones DW, et al. The effect of amphetamine on regional cerebral blood flow during cognitive activation in schizophrenia. *J Neurosci* 1991;11:1907–1917.

45. Weinberger DR, Berman KF, Illowsky BP. Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia. III. A new cohort and evidence for a monoaminergic mechanism. *Arch Gen Psychiatry* 1988;45:609–615.

46. Bush G, Luu P, Posner MI. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci* 2000;4:215–222.

47. Zakzanis KK, Poulin P, Hansem KT, et al. Searching the schizophrenic brain for temporal lobe deficits: a systematic review and meta-analysis. *Psychol Med* 2000;30:491–504.

48. Heckers S, Rauch SL, Goff D, et al. Impaired recruitment of the hippocampus during conscious recollection in schizophrenia. *Nat Neurosci* 1998;1:318–323.

49. Siegel BV Jr, Buchsbaum MS, Bunney WE Jr, et al. Cortical-striatal-thalamic circuits and brain glucose metabolic activity in 70 unmedicated male schizophrenic patients. *Am J Psychiatry* 1993;150:1325–1336.

50. Mattay VS, Callicott JH, Bertolino A, et al. Abnormal functional lateralization of the sensorimotor cortex in patients with schizophrenia. *Neuroreport* 1997;8:2977–2984.

51. Schroder J, Buchsbaum MS, Siegel BV, et al. Cerebral metabolic activity correlates of subsyndromes in chronic schizophrenia. *Schizophr Res* 1996;19:41–53.

52. Weinberger DR, Berman KF, Suddath R, et al. Evidence of dysfunction of a prefrontal-limbic network in schizophrenia: a magnetic resonance imaging and regional cerebral blood flow study of discordant monozygotic twins. *Am J Psychiatry* 1992;149:890–897.

53. Lipska BK, Weinberger DR. To model a psychiatric disorder in animals. Schizophrenia as a reality test. *Neuropsychopharmacology* 2000;23:223–239.

54. Weinberger DR. Cell biology of the hippocampal formation in schizophrenia. *Biol Psychiatry* 1999;45:395–402.

55. Fletcher PC, McKenna PJ, Frith CD, et al. Brain activations in schizophrenia during a graded memory task studied with functional neuroimaging. *Arch Gen Psychiatry* 1998;55:1001–1018.

56. Jennings JM, McIntosh AR, Kapur S, et al. Functional network differences in schizophrenia: a rCBF study of semantic processing. *Neuroreport* 1998;9:1697–1700.

57. Norman RM, Malla AK, Williamson PC, et al. EEG coherence and syndromes in schizophrenia. *Br J Psychiatry* 1997;170:411–415.

58. Biver F, Goldman S, Luxen A, et al. Altered frontostriatal relationship in unmedicated schizophrenic patients. *Psychiatry Res* 1995;61:161–171.

59. Mallet L, Mazoyer B, Martinot, JL. Functional connectivity in depressive, obsessive-compulsive, and schizophrenic disorders: an explorative correlational analysis of regional cerebral metabolism. *Psychiatry Res* 1998;82:83–93.

60. Andreasen NC. Linking mind and brain in the study of mental illnesses. *Science* 1997;275:1586–1593.

61. Andreasen NC, O'Leary DS, Cizadlo T, et al. Schizophrenia and cognitive dysmetria: a positron-emission tomography study of dysfunctional prefrontal-thalamic-cerebellar circuitry. *Proc Natl Acad Sci USA* 1996;93:9985–9990.

62. Andreasen NC, O'Leary DS, Flaum M, et al. Hypofrontality in schizophrenia: distributed dysfunctional circuits in neuroleptic-naive patients. *Lancet* 1997;349:1730–1734.

63. Andreasen NC, Paradiso S, O'Leary DS. Cognitive dysmetria as an integrative theory of schizophrenia: a dysfunction in cortical-subcortical-cerebellar circuitry? *Schizophr Bull* 1998;24:203–218.

64. Goldberg TE, Torrey EF, Berman KF, et al. Relations between neuropsychological performance and brain morphological and physiological measures in monozygotic twins discordant for schizophrenia. *Psychiatry Res* 1994;55:51–61.

65. Weiss AP, Heckers S. Neuroimaging of hallucinations: a review of the literature. *Psychiatry Res Neuroimag* 1999;92:61–74.

66. McGuire PK, Silbersweig DA, Wright I, et al. The neural correlates of inner speech and auditory verbal imagery in schizophrenia: relationship to auditory verbal hallucinations. *Br J Psychiatry* 1996;169:148–159.

67. Silbersweig DA, Stern E, Frith C, et al. A functional neuroanatomy of hallucinations in schizophrenia. *Nature* 1995;378:176–179.

68. Kaplan RD, Szechtman H, Franco S, et al. Three clinical syndromes of schizophrenia in untreated subjects: relation to brain glucose activity measured by positron emission tomography (PET). *Schizophr Res* 1993;11:47–54.

69. Berman KF, Weinberger DR, Zec RF. Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia. II. Role of neuroleptic treatment, attention, and mental effort. *Arch Gen Psychiatry* 1986;43:126–135.

70. Weinberger DR, Berman KB, Zec RF. Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia. I. Regional cerebral blood flow (rCBF) evidence. *Arch Gen Psychiatry* 1986;43:114–124.

71. Volkow ND, Wolf AP, Van Gelder P, et al. Phenomenological correlates of metabolic activity in 18 patients with chronic schizophrenia. *Am J Psychiatry* 1987;144:151–158.

72. Guich SM, Buchsbaum MS, Burgwald L, et al. Effect of attention on frontal distribution of delta activity and cerebral metabolic rate in schizophrenia. *Schizophr Res* 1989;2:439–448.

73. Buchsbaum MS, Nuechterlein KH, Haier RJ, et al. Glucose metabolic rate in normals and schizophrenics during the Continuous Performance Test assessed by positron emission tomography. *Br J Psychiatry* 1990;156:216–227.

74. Lewis SW, Ford RA, Syed GM, et al. A controlled study of 99mTc-HMPAO single-photon emission imaging in chronic schizophrenia. *Psychol Med* 1992;22:27–35.

75. Kawasaki Y, Maeda Y, Suzuki M, et al. SPECT analysis of regional cerebral blood flow changes in patients with schizophrenia during the Wisconsin Card Test. *Schizophr Res* 1993;10:109–116.

76. Nakashima Y, Momose T, Sano I, et al. Cortical control of saccade in normal and schizophrenic subjects: a PET study using a task-evoked rCBF paradigm. *Schizophr Res* 1994;12:259–264.

77. Guenther W, Brodie JD, Bartlett EJ, et al. Diminished cerebral metabolic response to motor stimulation in schizophrenics: a PET study. *Eur Arch Psychiatry Clin Neurosci* 1994;244:115–125.

78. Parellada E, Catafau AM, Bernardo M, et al. Prefrontal dysfunction in young acute neuroleptic-naive schizophrenic patients: a resting and activation SPECT study. *Psychiatry Res* 1994;55:131–139.

79. Siegel BV Jr, Nuechterlein KH, Abel L, et al. Glucose metabolic correlates of continuous performance test performance in adults with a history of infantile autism, schizophrenics, and controls. *Schizophr Res* 1995;17:85–94.

80. Steinberg JL, Devous MD Sr, Paulman RG. Wisconsin card sorting activated regional cerebral blood flow in first break and chronic schizophrenic patients and normal controls. *Schizophr Res* 1996;19:177–187.

81. Yurgelun-Todd DA, Waternaux CM, Cohen BM, et al. Functional magnetic resonance imaging of schizophrenic patients and comparison subjects during word production. *Am J Psychiatry* 1996;153:200–205.

82. Ganguli R, Carter C, Mintun M, et al. PET brain mapping study of auditory verbal supraspan memory versus visual fixation in schizophrenia. *Biol Psychiatry* 1997;41:33–42.

83. Shajahan PM, Glabus MF, Blackwood DH, et al. Brain activation during an auditory 'oddball task' in schizophrenia measured by single photon emission tomography. *Psychol Med* 1997;27:587–594.

84. Volz HP, Gaser C, Hager F, et al. Brain activation during cognitive stimulation with the Wisconsin Card Sorting Test—a functional MRI study on healthy volunteers and schizophrenics. *Psychiatry Res* 1997;75:145–157.

85. Gracia Marco R, Aguilar Garcia-Iturrospe EJ, Fernandez Lopez L, et al. Hypofrontality in schizophrenia: influence of normalization methods. *Prog Neuropsychopharmacol Biol Psychiatry* 1997;21:1239–1256.

86. Callicott JH, Mattay VS, Ramsey NF, et al. Functional magnetic resonance imaging brain mapping in psychiatry: methodological issues illustrated in a study of working memory in schizophrenia. *Neuropsychopharmacology* 1998;8:186–196.

87. Carter CS, Perlstein W, Ganguli R, et al. Functional hypofrontality and working memory dysfunction in schizophrenia. *Am J Psychiatry* 1998;155:1285–127.

88. Ragland JD, Gur RC, Glahn DC, et al. Frontotemporal cerebral blood flow change during executive and declarative memory tasks in schizophrenia: a positron emission tomography study. *Neuropsychology* 1998;12:399–413.

89. Curtis VA, Bullmore ET, Brammer MJ, et al. Attenuated frontal activation during a verbal fluency task in patients with schizophrenia. *Am J Psychiatry* 1998;155:1056–1063.

90. Parellada E, Catafau AM, Bernardo M, et al. The resting and activation issue of hypofrontality: a single photon emission computed tomography study in neuroleptic-naive and neuroleptic-free schizophrenic female patients. *Biol Psychiatry* 1998;15:787–790.

91. Stevens AA, Goldman-Rakic PS, Gore JC, et al. Cortical dysfunction in schizophrenia during auditory word and tone working memory demonstrated by functional magnetic resonance imaging. *Arch Gen Psychiatry* 1998;55:1097–1103.

92. Crespo-Facorro B, Paradiso S, Andreasen NC, et al. Recalling word lists reveals "cognitive dysmetria" in schizophrenia: a positron emission tomography study. *Am J Psychiatry* 1999;156:386–392.

93. Volz H, Gaser C, Hager F, et al. Decreased frontal activation in schizophrenics during stimulation with the continuous performance test—a functional magnetic resonance imaging study. *Eur Psychiatry* 1999;14:17–24.

94. Artiges E, Salame P, Recasens C, et al. Working memory control in patients with schizophrenia: a PET study during a random number generation task. *Am J Psychiatry* 2000;157:1517–1519.

95. Higashima M, Kawasaki Y, Urata K, et al. Regional cerebral blood flow in male schizophrenic patients performing an auditory discrimination task. *Research* 2000;42:29–39.

96. Heckers S, Curran T, Goff D, et al. Abnormalities in the thalamus and prefrontal cortex during episodic object recognition in schizophrenia. *Biol Psychiatry* 2000;48:651–657.

97. Holcomb HH, Lahti AC, Medoff DR, et al. Brain activation patterns in schizophrenic and comparison volunteers during a matched-performance auditory recognition task. *Am J Psychiatry* 2000;157:1634–1645.

98. Russell TA, Rubia K, Bullmore ET, et al. Exploring the social brain in schizophrenia: left prefrontal underactivation during mental state attribution. *Am J Psychiatry* 2000;157:2040–2042.

99. Manoach DS, Press DZ, Thangaraj V, et al. Schizophrenic subjects activate dorsolateral prefrontal cortex during a working memory task, as measured by fMRI. *Biol Psychiatry* 1999;45:1128–1137.

100. Manoach DS, Gollub RL, Benson ES, et al. Schizophrenic subjects show aberrant fMRI activation of dorsolateral prefrontal cortex and basal ganglia during working memory performance. *Biol Psychiatry* 2000;48:99–109.

101. Frith CD, Friston KJ, Herold S, et al. Regional brain activity in chronic schizophrenic patients during the performance of a verbal fluency task. *J Psychiatry* 1995;167:343–349.

102. Buckley PF, Friedman L, Wu D, et al. Functional magnetic resonance imaging in schizophrenia: initial methodology and evaluation of the motor cortex. *Psychiatry Res* 1997;74:13–23.

103. Spence SA, Hirsch SR, Brooks DJ, et al. Prefrontal cortex activity in people with schizophrenia and control subjects. Evidence from positron emission tomography for remission of 'hypofrontality' with recovery from acute schizophrenia. *Br J Psychiatry* 1998;172:316–323.

104. Curtis VA, Bullmore ET, Morris RG, et al. Attenuated frontal activation in schizophrenia may be task dependent. *Schizophr Res* 1999;37:35–44.

105. Braus DF, Ende G, Hubrich-Ungureanu P, et al. Cortical response to motor stimulation in neuroleptic-naive first episode schizophrenics. *Psychiatry Res* 2000;98:145–154.

55

Structural Magnetic Resonance Imaging Studies in Schizophrenia

Robert W. McCarley

Robert W. McCarley: Harvard Department of Psychiatry, Brockton VAMC and VA Boston Healthcare System, Brockton, Massachusetts

> … We thus come to the conclusion that, in dementia praecox, partial damage to, or destruction of, cells of the cerebral cortex must probably occur, which may be compensated for in some cases, but which mostly brings in its wake a singular, permanent impairment of the inner life (1).

The window on the brain provided by structural imaging has transformed our view of schizophrenia to one that views the very structure of the brain as altered, a view echoing Kraepelin's prescient statement. Beginning with Johnstone's CT findings of enlarged ventricles (which actually confirmed earlier, less systematic pneumoencephalographic studies), subsequent reports using magnetic resonance imaging (MRI) have provided key information detailing volume reductions in particular brain anatomic regions of interest (ROI). These data have provided the major evidence in support of our current view that schizophrenia is a brain disorder with altered brain structure, and consequently involving more than a simple disturbance in neurotransmission.

Section Organization

The next section is a nontechnical introduction to some of the basic concepts of MRI, and may be read independently of the other sections or skipped by those who wish to concentrate on the clinical data. Subsequent sections discuss the application of structural MRI to questions in schizophrenia research, standards for technical quality and reviews of studies, current MRI findings in schizophrenia (limited to studies with defined ROI), and newer technologies.

THE BASIS OF STRUCTURAL MRI AND PULSE SEQUENCES

Part of "55 - Structural Magnetic Resonance Imaging Studies in Schizophrenia "

Addressing the physics of structural MRI is a major topic on its own, and Brown and Smeleka's book (2) is recommended as a good introduction that demands only a limited background in mathematics and physics. The reader is warned that the following brief exposition is highly (over)simplified. Essentially, the tissue characteristics sensed by MRI depend on disruptions of a strong external magnetic field. This external field has aligned the orientation of atomic nuclei by aligning the magnetic field of each nucleus that is associated with its spin direction. Because of its biological ubiquity and good magnetic properties, the most commonly used basis element for MRI is hydrogen, which has a single proton in its nucleus. The reader may want to think of hydrogen nuclei as analogous to a large set of spinning tops or gyroscopes. In a state without an external magnetic field, their direction of spin is random, and so is the net magnetization (a vector), because each rotating proton has a magnetic field that is parallel to its axis of rotation. Applying a strong external magnetic field can be thought of as aligning this set of spinning tops in a uniform direction of spin, snapping them to attention, as it were. The resultant population net magnetization can be thought of as vector aligned with the z-axis (vertical axis in our example), perpendicular to the x-y plane, and in the direction of the external magnetic field. The magnetic field strength is described in units of tesla (T) and most current clinical imagers use an external field of 1.5 T.

This vertically aligned population of protons (vertical magnetization vector) is then perturbed by applied radiofrequency (rf) pulses, which can be thought of as having the effect of moving the magnetization vector away from the vertical axis (z-axis); for example, in our analogy, moving the tops from their average vertical orientation to a "tilt." This applied change of the magnetization vector then decays, and, during the course of the decay, give off energy in the form of radiofrequency emissions. It is this emitted energy that provides the key information for MRI scans.

There are two main kinds of information about tissue characteristics derived from this perturbation decay, often referred to as a "relaxation." The T1 relaxation time is the

time constant describing the time course followed by the “tilted” magnetization vector in returning to its original orientation. The T1 relaxation time is the time required for this vector to return to 63% of its original vertical orientation value following an rf excitation pulse. Again, in our analogy to spinning tops subjected to a tilt, this T1 is the time required to return to about two-thirds of its original vertical orientation.

The second measure, T2 relaxation time, requires us to think about each nucleus (spinning top) individually. Again and again crudely, one can visualize a group of spinning tops oriented upward (z-axis) and then simultaneously tipped away from this vertical orientation. As everyone who has spun tops or played with a gyroscope knows, if one tilts a top from a vertical orientation, the top will not only tilt but will rotate about the vertical (z-axis), a wobble technically called precessing. Figure 55.1 illustrates this process. In the beginning all the individual tops will wobble (precess) about the z-axis with the same frequency, but gradually they will lose their coherence and wobble at different frequencies, leading to progressively less net magnetization in the x-y plane. T2 is the time required for the coherence to decrease to 37% of its original value. This “dephasing time” or T2 relaxation is always less than or equal to T1. Often the term T2* is used to take into account the observed variations in relaxation time owing to inhomogeneities in the tissue being imaged and in the applied magnetic field. The web site, http://ej.rsna.org/ej3/0095-98.fin/index.htm has a nice animated illustration of T2 relaxation (on the menu page, select Diffusion and Magnetic Resonance). Of relevance to this description, Pfefferbaum and colleagues (3) found T2 relaxation times were longer in schizophrenic patients than in controls in both gray and white matter, suggesting possible differences in fundamental tissue organization in schizophrenia.

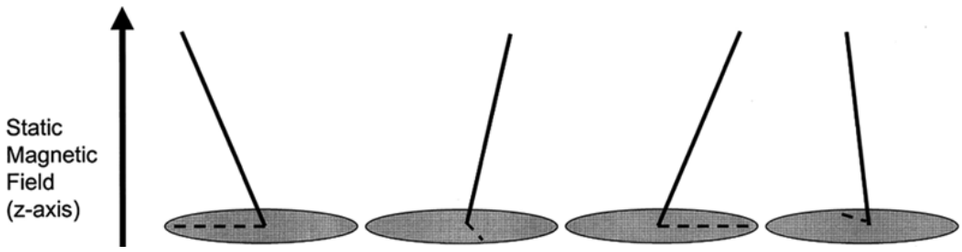


FIGURE 55.1. Illustration of effects of applied magnetic field on hydrogen nuclei (protons, *images 1 to 4*). With only a static magnetic field *(left arrow)* present, all nuclei have the same vertically aligned spin directions parallel to the static magnetic field and the z axis (this state is not illustrated). Application of an rf pulse “tilts” the orientation so there is a transverse plane component *(broken line)* in Image 1. Initially all protons precess uniformly, so that images 1, 2, 3, and 4 can be thought of as successive snapshots (successive moments in time) of the counterclockwise precession (rotation) of the net magnetization vector about the z-axis. Over time the protons dephase and show different precession frequencies; as an illustration of this case, images 1 to 4 should be thought of as a single snapshot of individual protons at the same instant in time. There is no net transverse plane magnetic vector because the individual protons show no uniformity of phase.

In terms of T1, an rf pulse may “tilt” the net magnetization (spin) vector, but usually a second pulse is applied before there is a full return to the vertical orientation, and subsequent rf pulse repetitions lead to a steady-state orientation prior to each new pulse. This new vector depends on a number of values; two are of particular relevance: the T1 relaxation time (how efficiently the protons give up their energy) and number of protons per unit of tissue, *proton density*.

Spatial Localization

The resonance frequency of protons, the frequency at which energy is maximally absorbed by protons, is dependent on the strength of the magnetic field. By applying small magnetic field gradients (typically less than 1% of the total field strength) for short periods of time it is possible to spatially localize the signals resulting from the applied rf pulses. In the presence of a magnetic gradient field each proton will resonate at a unique frequency that depends on its exact position within the gradient field. The MR image is a frequency and phase map of the protons at each point or picture element (pixel) throughout the image. The pixel intensity is proportional to: the number of protons present in the volume represented by the pixel weighted by the T1 and T2 relaxation times. *Different sequences of rf pulses will produce images that mainly reflect one of these variables, and these images are often referred to as proton density-, T1-, or T2-weighted images*. Operationally, the initial step in spatial localization is localization of the rf excitation to a region of space (slice) by the slice selection gradient. When images are viewed, the slice selection direction is always perpendicular to the surface. A second spatial direction is determined by a phase encoding gradient, which differentially alters the precessional frequency of protons at different positions in

the phase encoding direction, thereby enabling spatial localization. In MRI the signal is always detected in the presence of a readout gradient, perpendicular to the slice selection and phase encoding gradient, and producing the third dimension of the image. The readout gradient detects differences arising from both the slice selection gradient and the phase encoding gradient, with the latter varying in amplitude with each repetition of the slice selection and readout gradient.

Pulse Sequences

Appendix A describes commonly used pulse sequences in terms of our knowledge about the relaxation processes for the reader wishing insight into the terminology and rationale of pulse sequences.

STRUCTURAL MRI: WHAT CAN IT TELL US ABOUT SCHIZOPHRENIA?

Part of "55 - Structural Magnetic Resonance Imaging Studies in Schizophrenia "

Structural MRI

In 1984 Smith and co-workers (4) performed the first MRI study of schizophrenia. The capability of structural MRI to provide information about gray and white matter parenchyma of the brain and CSF-filled spaces is new with MRI studies; it represents an important advance over CT studies that poorly visualize parenchyma and can not differentiate gray and white matter. This gray–white differentiation is important for schizophrenia studies, because abnormal tissue classes (tumors, infarcted areas, etc.), which may be detected by CT, have not been found to characterize schizophrenia. The term “structural MRI” is used to differentiate it from “functional MRI,” where indices of short-duration change (e.g., blood oxygenation) are used; this topic is treated in the second part of this chapter.

Our use of the term schizophrenia is in the sense of a syndrome and not a single disease entity. The current major questions about the schizophrenia syndrome include:

1. What are the brain changes in this disorder? Which areas of the brain are affected?
2. What is the cause of the brain changes?
3. At what life stage do brain abnormalities occur and are they static or progressive? Are they developmental (prenatal and perinatal) and/or progressive?
4. How are brain abnormalities related to clinical symptom abnormalities?
5. Are brain findings in schizophrenia distinct from those in affective psychosis?
6. What are the most effective treatments? Is treatment neuroprotective?
7. Are there structural endophenotypes that will help us in the genetic analysis of the disorder?

Structural MRI studies of schizophrenia have the potential of addressing all of these questions, although space restrictions confine our focus primarily on the first question, which has also been the major focus of empirical studies.

WHAT ARE THE DESIRABLE FEATURES OF A STRUCTURAL MRI STUDY?

Part of "55 - Structural Magnetic Resonance Imaging Studies in Schizophrenia "

We here briefly summarize the features.

1. *Thinner is better*. Smaller units of volume analysis (called voxels, for volume element) allow for more precise determination of the irregular contours of brain regions, by reducing the voxel mixing of the desired region with neighboring structures in the voxel. This mixing is called partial voluming. Many earlier studies used MRI acquisitions with “gaps” between slices, with interpolation used to estimate the volume in the “gap”; this obviously limits precision of measurement. Thus studies with thinner slices (1.5 mm is the current standard), and no gaps between slices, will likely lead to more precise MR morphometric volume measures.
2. *Quantitative versus qualitative analysis*. Early studies relied on subjective, visual ratings of abnormalities. There is now general agreement that computation of volumes of the ROI examined is essential. When raters are used, as is generally the case, inter-rater reliability is important, and should be $r \geq 0.85$. Moreover, the ROI should be objectively and clearly defined, so that others can measure the same entity. Such objectively defined criteria should include detailed specification of the internal landmarks used to define each ROI.
3. *Segmentation*. Segmentation involves sorting the tissue classes into gray matter, white matter, or cerebrospinal fluid (CSF). It seems to us that all studies of cortical gyri should, whenever possible, separate gray and white matter in the analysis, because this is a fundamental distinction in brain tissue; however, not all studies distinguish between gray and white matter, making comparisons with studies that do segment gray and white matter problematic. Finally, segmentation is often automated or semiautomated; unfortunately, there is no agreed-on gold standard for the quality of segmentation, because “phantoms” with known composition do not reflect the complexity of the outlines of brain gray or white matter, and postmortem estimates of tissue and fluid volumes may not exactly parallel those *in vivo*. (See ref. 5 for discussion.) Figure 55.2 provides an example of a segmented image with ROI tracing and three-dimensional (3D) reconstruction.

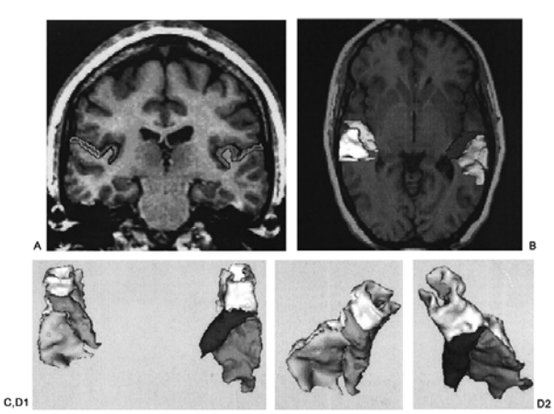


FIGURE 55.2. A: Coronal slice (1.5 mm) through the temporal lobe of a normal control subject. This is a SPGR proton density-weighted image. The regions of interest for the structures are outlined: the gray matter of Heschl’s gyrus (HG) is red on subject left and green on subject right. The gray matter of planum temporale (PT) is labeled yellow on subject left and blue on subject right. **B:** Top-down view of the three-dimensional (3D) reconstruction of HG and PT placed atop an axial magnetic resonance slice. This axial slice has been constructed by reformatting the coronal images. Anterior is top. HG is red on subject left and green on subject right, and PT is blue on subject left and yellow on subject right. **C:** 3D reconstruction of the left and right ROI (color-coded as in *B* but from a slightly different angle of rotation than *B*). Note the tubular structure of the gray matter of the STG, most clearly seen anteriorly, where gray codes non-HG, non-PT portions of STG. **D:** 3D reconstructions viewed from a different angle than *C*. (D1 is subject right and D2 is subject left). (Reproduced from Hirayasu Y, McCarley RW, Salisbury DF, et al. Planum temporale and Heschl’s gyrus volume reduction in schizophrenia: an MRI study of first-episode patients. *Arch Gen Psychiatry* 2000;57:692–699.)

4. *Quality of imager and postacquisition processing*. The quality of the MR scanner is also important and should include technical assessments such as the homogeneity of the magnetic field, which greatly influences the postprocessing segmentation of tissue into different tissue components. Day-to-day assessment of inhomogeneities in the magnetic field is thus a critical quality assurance feature for the quality

of the MR scans, and consequently also the quality of the postprocessing of MR images. Most modern imagers have magnetic fields of 1.5 T or greater, which is important for signal-to-noise ratio. Additionally, postacquisition filtering may improve signal-to-noise ratio (6).

5. *Variability and reliability of MRI findings*. An important question is the variability in MRI findings that is to be expected, not only from variation in measurement techniques, but also from physiologic changes. This becomes increasingly important because MR imaging is done at different time points in the disorder. Unfortunately, there are currently very few studies that have examined the extent of changes in MR over multiple measurements, and more careful controlled studies would be useful. Our laboratory found less than 1% variation over gray–white–CSF segmentation values in one subject who had an MR scan on two different days (5). Kikinis and colleagues (7) presented data from a female subject who received 23 MR scans during 1 year. The variance of the intracranial cavity was only 1.2% over the course of the 23 MR scans. Also, there is little evidence of the idea of a physiologic variation in gray or white matter volume throughout the brain in studies with high-resolution MRI techniques. For example, Gur and associates (8) found changes in volume over 2.5 years in whole frontal lobe in schizophrenia without finding changes in volumes of whole-brain and CSF, a finding difficult to reconcile with the idea

of “whole brain variability” in gray or white matter, or CSF caused by hydration or other factors affecting all brain tissue, for example.

HOW SHOULD LITERATURE FINDINGS BE REPORTED IN REVIEWS? EXPERT OPINION VERSUS COUNTING VERSUS METAANALYSIS

Part of "55 - Structural Magnetic Resonance Imaging Studies in Schizophrenia "

The reader trying to shape an informed opinion on the current state of the field (not only structural MRI but any field) relies on reviews with essentially three main approaches, which can be combined.

Expert Opinion

One approach for reviewers is to survey the literature, and then provide an informed opinion as to the summary trends and findings, with specific citations to drive home the points. Many reviews, in fact, adopt this approach of “trust the reviewer.” The disadvantage is that the reader cannot form a judgment of the *accuracy of the review's conclusions based on the data presented* without reading the literature.

Counting

This approach tabulates the number of studies with findings supporting or not supporting an abnormality in a particular region. The disadvantage of this approach is that the subject *N* of each study and the effect size are not taken into account.

Metaanalysis

Metaanalysis essentially involves weighting each individual study by a function of its *N* and effect size, and then using this information to produce an estimate of the combined effect size (9 ,10 and 11). *If all studies used equivalent technology and subject populations, this would be the method of choice. However, they do not*, and the reader should realize that: (a) MR scanner technology in the past decade has been changing rapidly; therefore, studies are not quantitatively comparable; (b) the extent of detail varies in anatomically based ROI information used in the measurement of images; (c) there is a wide difference in moderator variables of subject gender, chronicity (age of onset), medication, parental SES, etc.; and (d) metaanalysis, especially of MRI studies, is beset with the difficulty of estimating the number of studies with negative findings that did not get published—Rosenthal’s “file drawer” problem (11). Practically, this means one cannot do a review that is both comprehensive and metaanalytically valid, unless the items (a) through (c) had remained constant. Another disadvantage is that many studies must be rejected for reasons of subject or methodologic consistency; this omission has a potential effect of distorting results by omitting technically good studies that could not be included; for example, the metaanalytic study of Nelson and colleagues (12) omitted 45% of the studies on medial temporal lobe found in the literature.

Our summary relies on a combination of these approaches. The conclusions are congruent with our subjective opinions and we provide tabulation of more than a decade’s results of MRI studies; however, we are sympathetic to the need to provide more than a simple “box score” of positive and negative results. Accordingly, we have followed a suggestion of Rosenthal (11), and computed the probability of the observed number of positive and negative findings for each region. This is simply done by using a two-tailed alpha level of $p < .05$ for each study finding positive results, and then using the binomial theorem to calculate the overall probability of finding the observed number of positive studies. The resulting overall probability does not assume a normal distribution, but does assume comparability of the studies. (*Caution:* This assumption does not necessarily hold.) Whatever the degree of comparability, this statistic does have the distinct advantage of providing the reader with a sense of the weight of current evidence for each region of interest beyond that of a simple “majority vote,” which neglects the odds against an individual study’s finding positive results at the $p < .05$ level. We see this as an improvement on simply saying “some studies find… but others…” and providing the reader a backdrop of estimating the weight of evidence from peer-reviewed studies based on a probability of .05, assuming the prerequisites of metaanalysis hold, and then allowing a subjective “dilution factor” for all of the problems of a metaanalytic analysis.

CURRENT STRUCTURAL FINDINGS IN SCHIZOPHRENIA

Part of "55 - Structural Magnetic Resonance Imaging Studies in Schizophrenia "

A complete review is not appropriate here because the explicit focus of this volume is on summary of main trends and new developments. However, because it seemed essential to us that the reader have at least an overview of structural results in schizophrenia (with references), we have drawn on data from a recent comprehensive review (13). This covered all peer-reviewed schizophrenia studies with control groups during the time period 1987 to May 1998, and to our knowledge is the most recently published comprehensive review.

We summarize here the results from this review, whereas Table 55.1 presents the results and references in tabular form. (Table 2 in the original article summarized the subject *N* and characteristics, and should be consulted for further detail [13].)

Brain Region	% +	% −	Total N	N +	N −	References[b]
Whole brain	19	81	31	6	25	+ [Andreasen 1990,1994b; Gur 1994,1998; Jernigan 1991, Nasrallah 1990] − [Barta 1990; Blackwood 1991; Breier 1992; Buchanan 1993; Colombo 1993; DeLisi 1991; DeLisi 1992; Douphinais 1990; Flaum 1995; Harvey 1993; Hirayasu 1998a; Johnstone 1989; Kawasaki 1993; Kelsoe 1988; Marsh 1994; Nopoulos 1995; Reite 1997; Rossi 1990b,1994b; Shenton 1992; Schlaepfer 1994; Sullivan 1998; Vita 1995; Zipursky 1992,1997]
Lateral ventricles	77	23	43	33	10	+ [Andreasen 1990,1994b; Barr 1997; Becker 1990; Bogerts 1990, Bornstein 1992; Buchsbaum 1997; Corey-Bloom 1995; Dauphinais 1990; Degreef 1990,1992a; DeLisi 1991,1992; Egan 1994, Flaum 1995; Gur 1994; Harvey 1993; Johnstone 1989; Kawasaki 1993; Kelsoe 1988; Lauriello 1997; Lim 1996; Marsh 1994,1997; Nasrallah 1990; Nopoulos 1995; Rossi 1988; Stratta 1989; Suddath 1989,1990; Sullivan 1998; Vita 1995; Zipursky 1992] − [Blackwood 1991; Colombo 1993; Hoff 1992; Jernigan 1991; Rossi 1990b; Rossi 1994b; Schwarz 1992; Schwarzkopf 1990; Shenton 1991,1992]
Third ventricle	67	33	24	16	8	+ [Bornstein 1992; Becker 1996; Dauphinais 1990; Degreef 1990,1992a; Egan 1994; Flaum 1995; Kelsoe 1988; Lim 1996; Marsh 1994,1997; Nasrallah 1990, Rossi 1994b; Schwarzkopf 1990; Woodruff 1997a; Sullivan 1998] − [Andreasen 1990; Barta 1990; Colombo 1993; DeLisi 1991; Schwarzkopf 1992; Shenton 1992; Suddath 1990; Zipursky 1992]
Fourth ventricle	0	100	3	0	3	− [Rossi 1988; Shenton 1992; Stratta 1989]
Temporal lobe (TL)						
Whole TL	62	38	37	23	14	+ [Andreasen 1994b; Barta 1990; Becker 1996; Bogerts 1990; Dauphinais 1990; DeLisi 1991; DiMichele1992; Egan 1994; Gur1998; Harvey 1993; Jernigan 1991; Johnstone 1989; Marsh 1997; Rossi 1988,1989b,1990b,1991; Suddath 1989,1990; Sullivan 1998; Woodruff 1997a; Woods 1996; Zipursky 1992] − [Becker 1990; Bilder 1994asym; Blackwood 1991; Colombo 1993; DeLisi 1992; Flaum 1995; Hoff 1992; Kawasaki 1993; Kelsoe 1988; Nopoulos 1995; Raine 1992; Shenton 1992; Swayze 1992; Vita 1995]
Medial TL	77	23	31	24	7	+ [Barta 1990,1997b; Becker 1990,1996; Blackwood 1991; Bogerts 1990,1993; Breier 1992; Buchanan 1993; Dauphinais 1990; DeLisi 1988; Egan 1994; Flaum 1995; Fukuzako 1996b; Hirayasu 1998a; Jernigan 1991; Kawasaki 1993; Marsh 1994; Ohnuma 1997; Rossi 1994b; Shenton 1992; Suddath 1989,1990; Woodruff 1997a] − [Colombo 1993; Corey-Bloom 1995; DeLisi 1991; Harvey 1993; Swayze 1992; Marsh 1997; Zipursky 1994]
Superior temporal	81	19	16	13	3	
Gyrus						
All studies						
Gray matter	100	0	7	7	0	+ [Hajek 1997; Hirayasu 1998a; Menon 1995; Schlaepfer 1994; Shenton 1992; Sullivan 1998; Zipursky 1994] − None
Gray and white matter (combined)	67	33	9	6	3	+ [Barta 1990; Barta 1997b; Flaum 1995; Marsh 1997; Reite 1997; Tune 1996] − [Kulynych 1996; Vita 1995; Woodruff 1997a]
Planum Temporale	63	37	8	5	3	+ [Barta 1997a; DeLisi 1994; Kwon 1999; Petty 1995; Rossi 1992] − [Kleinschmidt 1994; Kulynych 1995; Rossi 1994a]
Frontal lobe	55	45	33	18	15	+ [Andreasen 1994b; Bilder 1994asym; Breier 1992; Buchanan 1993; Gur 1998; Harvey 1993; Jernigan 1991; Nopoulos 1995; Ohnuma 1997; Raine 1992; Rossi 1988; Schlaepfer 1994; Stratta 1989; Sullivan 1998; Woodruff 1997a; Woods 1996; Zipursky 1992,1994] − [Andreasen 1990; Blackwood 1991; Corey-Bloom 1995; DeLisi 1991; Egan 1994; Kawasaki 1993; Kelsoe 1988; Kikinis 1994; Nasrallah 1990; Rossi 1990b; Shenton 1992; Suddath 1989,1990; Vita 1995; Wible 1995]
Parietal lobe	44	56	9	4	5	+ [Andreasen 1994b; Bilder 1994asym; Schlaepfer 1994; Zipursky 1994] − [Egan 1994; Jernigan 1991; Nopoulos 1995; Sullivan 1998; Zipursky 1992]
Occipital lobe	43	57	7	3	4	+ [Andreasen 1994b; Bilder 1994asym; Zipursky 1992] − [Jernigan 1991; Nopoulos 1995; Schlaepfer 1994; Sullivan 1998]
Localized (L) vs. Diffuse (D) Changes in Cortex	**L %**	**D %**	**Total #**	**#L**	**#D**	
Gray matter	86	14	7	6	1	L [Jernigan 1991; Schlaepfer 1994; Shenton 1992; Suddath 1990; Sullivan 1998; Zipursky 1997] D [Zipursky 1992]
Gray and white matter (combined)	50	50	4	2	2	L [Nopoulos 1995; Bilder 1994asym] D [Andreasen 1994b; Buchanan 1993]
Subcortical Structures	**% +**	**% −**	**Total N**	**N +**	**N −**	
Thalamus	67	33	6	4	2	+ [Andreasen 1990,1994a; Buchsbaum 1996; Flaum 1995] − [Corey-Bloom 1995; Portas 1998]
Corpus callosum	67	33	18	12	6	+ [Casanova 1990; DeLisi 1997; DeQuardo 1996; Gunther 1991; Hoff 1994; Lewine 1990; Raine 1990; Rossi 1988,1989a; Stratta 1989; Woodruff 1993; Uematsu 1988] − [Blackwood 1991; Colombo 1994; Hauser 1989; Kawasaki 1993; Kelsoe 1988; Woodruff 1997b]
Basal ganglia	65	35	17	11	6	+ [Breier 1992; Buchanan 1993; Chakos 1994,1995; Elkashef 1994; Hokama 1995; Jernigan 1991; Keshavan 1995; Mion 1991; Ohnuma 1997; Swayze 1992] − [Blackwood 1991; Corey-Bloom 1995; DeLisi 1991; Flaum 1995; Kelsoe 1988; Rossi 1994b]
Cerebellum	33	66	6	2	4	+ [Andreasen 1994b; Breier 1992] − [Coffman 1989; Flaum 1995; Uematsu 1989; Rossi 1993]
Cavum septi pellucidi	91	9	11	10	1	+ [Degreef 1992b,1992c; DeLisi 1993; Kwon 1998; Jurjus 1993; Nopoulos 1996,1997; Uematsu 1989; Scott 1993; Shioiri 1996] − [Fukuzako 1996a]

Note: asym, finding asymmetry difference only. (Planum Temporale citations are mainly of asymmetry differences and do not use the "asym" qualifier.)
[a]Each study is cited by first author and year.
[b]Full references are listed at the end of this chapter.

TABLE 55.1. SUMMARY OF MRI STUDIES REPORTING POSITIVE AND NEGATIVE FINDINGS IN SCHIZOPHRENIA FROM 1987 TO MAY 1998 (FROM McCARLEY ET AL., 1999)

Most studies (81% of 31) did not find abnormalities

of whole brain/intracranial contents, but lateral ventricle enlargement was reported in 77% of 43 studies, third ventricle enlargement in 67% of 24 studies, whereas none of the three studies evaluating the fourth ventricle found abnormalities.

The temporal lobe was the brain parenchymal region with the most consistently documented abnormalities, with 62% of 37 studies finding whole lobe volume decreases. Of all cortical areas surveyed, the superior temporal gyrus most consistently showed volume reduction (81% of 16 studies) and, if the gray matter of this structure was evaluated separately from white matter, all seven studies showed a volume reduction. Fully 77% of the 30 studies of the medial temporal lobe reported abnormalities in one or more of its constituent structures (hippocampus, amygdala, or parahippocampal gyrus). Neuropathologic studies in general support the presence of temporal lobe limbic system abnormalities in schizophrenia (14 ,15), although some do not (16). Unfortunately, there is a lack of quantitative postmortem studies of temporal lobe neocortex.

Despite the presence of functional abnormalities, frontal lobe structural MRI investigations did not consistently find abnormalities, with 55% of the 33 studies describing volume reduction. In a postmortem quantitative study, Selemon and associates (17) found only a small (8%) reduction in prefrontal cortical thickness, a reduction that was not statistically significant, although noteworthy abnormalities in density of various cell types were present in schizophrenia. This and the MRI findings suggest that frontal lobe volume reductions may be small, and near the threshold for MRI detection. The parietal and occipital lobes have been much less studied, and there are about the same percentage of positive and negative findings in each. Most of the seven studies of cortical gray matter (86%) find that volume reductions are not diffuse, but are more pronounced in certain areas, as might be anticipated from the preceding statistics on individual regions of interest.

About two-thirds of the studies of subcortical structures report positive findings, including the six studies of the thalamus, 18 studies of the corpus callosum, and 17 studies of the basal ganglia. Basal ganglia tended to show increased volume when patients were on typical but not on atypical neuroleptics. The cerebellum had mainly negative findings (in four of the six studies), but was not studied with the same volumetric precision as other ROI. Almost all (91%) of the 11 studies of cavum septi pellucidi (CSP) showed that schizophrenics have less fusion of the septum, a developmental abnormality probably linked to limbic system pathology.

Statistics

The binomial theorem computation (using $p < .05$ for a positive study) shows that *all ROI* surveyed in Table 55.1 show a two-tailed $p < .05$ for the number of positive studies, except for the fourth ventricle and cerebellum ($p = .66$), and all ROI had p's $\leq .002$ except for the occipital lobe (.004). Again, caution should be exercised on the strength of these probability estimates because comparability is not strict and an (unknown) percentage of studies with negative results may not have been published.

Clinical Symptom Implications

Clinical scale data suggest that positive symptoms (and disorganization in a three-factor model) may be most closely related to temporal lobe volumetric abnormalities, whereas more limited evidence supports, in some studies, a relationship between negative symptoms and frontal lobe MRI abnormalities. Finally, there now appears to be growing evidence that MRI abnormalities differ in affective (bipolar) psychosis and schizophrenia, with reductions in neocortical gray matter, especially in temporal and prefrontal neocortex being especially prominent in schizophrenia. (See McCarley and co-workers 1999 for a more complete review and Sheline [this volume] for a review of structural MRI in affective disorder [13].)

The presence of CSP and sulco-gyral abnormalities (for the latter, see Kikinis and colleagues) (18) and abnormalities in first-episode patients all suggest a possible developmental origin. However, there are growing (although still limited) data pointing to progression of volumetric abnormalities over time. This suggests that both developmental and progressive features may be present in schizophrenia; these are consistent with, we hypothesize, a “two-hit” model of schizophrenia.

AUTOMATING STRUCTURAL MRI ANALYSIS: BRAIN WARPING AND VOXEL-BASED ANALYSIS

Part of "55 - Structural Magnetic Resonance Imaging Studies in Schizophrenia "

The tedious task of manual definition of regions of interest by tracing outlines—even if assisted by automated segmentation—has prompted interest in using automated methods of MRI analysis. The two (closely related) major classes of methods are: (a) brain warping, using a standard or “atlas brain” to compare and define features on subject brains, and (b) voxel-based analyses. Because these techniques, although promising, are new and thus far have limited data on validity, we have not included the studies in the summary table of ROI findings. This section concludes with a brief discussion of shape analysis, often based on the brain warping techniques described in the first part of this section.

Brain Warping

In one use of this technique the ROI definitions and anatomic features of an index template (atlas brain image) are warped (mapped) onto a new case (target image). In this

context, the atlas brain can be compared to a rubber brain, which is stretched and compressed nonlinearly in order to match the contours of the new brain. At the end of the registration, all the structures previously defined for the atlas brain are also defined for the new brain image. In general, the first step in matching the atlas brain and subject (patient or object brain) is *linear registration* to correct for the differences in size, rotation, and translation between the two brain images. This step is illustrated in the first panel of Fig. 55.3 . (Parenthetically, linear transforms use scaling, translation and rotation uniformly for each element [voxel], whereas nonlinear transforms use different and more complicated transforms for different voxels.)

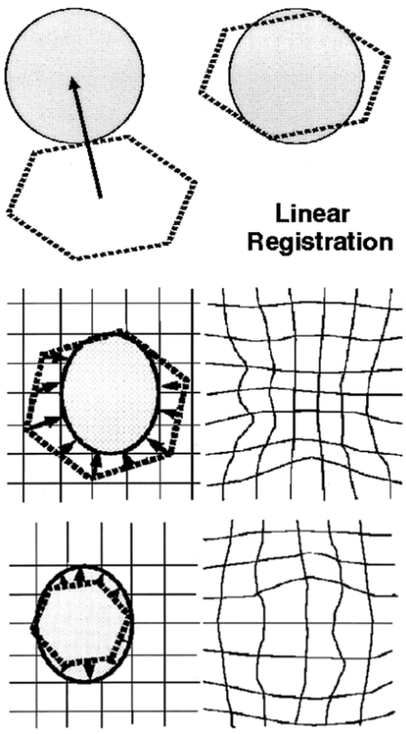


FIGURE 55.3. Schematic of brain image warping. An "atlas" image *(red hexagon)* is warped (mapped) onto a new "patient" image *(yellow oval)*. **Top:** A simple uniform linear transformation (translation, rotation or scaling) does not work. Instead, nonlinear transformations are used to "warp" the "atlas" image onto the "patient" image *(middle and bottom panels)*. This warping resulted in a "vector field" *(blue arrows)*. The nonuniform displacement of each pixel is then represented in the right field of Fig. 55.1, by means of the deformation of the rectangular grid (the elastic membrane). Examples are presented for atlas contraction *(middle panel)* and dilatation *(bottom panel)*.

Nonlinear Elastic Matching

The atlas information can then be projected into other MRI scans by applying an elastic match (i.e., warping the atlas into the shape of the new brain image). The global registration technique used by our lab (19) to match an anatomic MR atlas with defined ROI onto new-segmented MR images, was based on the theory of elastic membranes (20 ,21), and was similar to Grenander and Miller's (22) approach. The elastic membrane model can be intuitively understood as the deformations occurring when a set of points on the membrane is stretched. The goal of the elastic matching algorithm was to find a 3D *vector deformation field* that transformed the source data set (atlas) so that it matched the target data set (patient) with the greatest fidelity, that is, maximized the local similarity between the two data sets under the constraint of using a restricted set of "stretches" of the membrane. The middle and bottom panels of Fig. 55.3 show, in a cartoon, how contraction and expansion of particular voxels lead to a match between the atlas brain and the target brain. The right-hand part of these panels graphically illustrate the voxel-by-voxel deformation (stretching or compression) resulting from the *vector deformation field* used to transform the atlas image. We note that other choices of an "index brain" include the probabilistic atlas used by the Montreal Group (23) and interpolated data from an anatomic atlas, as used by Gee and colleagues (24).

Brain Warping: How Good Is It?

Any new technique must be validated, and brain warping is no exception. The current gold standard is manual ROI definition. The technique used by Iosifescu and co-workers (19 and unpublished data) was compared with manual ROI definition for volumes. Agreement on volumes over 28 subjects (half schizophrenia patients and half controls) was 97% for whole brain volume, 97% for whole white matter, 91% for whole gray matter, 96% for thalamus (both sides), 93% for putamen, 91% for caudate nucleus, and 76% for globus pallidus. A more rigorous and better method of comparison is the extent of overlap of voxels in the manually defined ROI with those in the automated definitions. In measurements of 20 brain cortical and subcortical brain structures on one brain image the extent of voxel-by-voxel correspondence, was defined as:

$$\frac{(\#\text{ voxels in manual ROI also in automated ROI})}{(\#\text{ voxels in manual ROI})}$$

This averaged 90% for subcortical structures and 98% for total gray and white matter volumes; however, for cortical gyri the overlap averaged only 60%. The automated computer algorithm assumed the neuroanatomic variability among subjects to be a topologic invariant. However, cerebral gyri frequently split in two in some subjects, whereas they remain one single structure in others. These differences could not be taken into account by the automated registration in its present form. Taken together, these data suggest:

1. Each automated warping procedure should be compared with the results using manual ROI definitions.
2. Accuracy may be good for subcortical structures (because of their relatively small variability in shape) and total brain gray and white matter.
3. Accuracy is questionable for the neocortex, because of the irregularity of sulco-gyral patterns.

A recent use of the technique of warping is to use the *vector deformation field* to provide a statistical test of whether each voxel is significantly displaced or not (25). (See Fig. 55.3 for a description of extent and direction of warp.) This methodology does not attempt to map gyrally defined ROI, but rather looks at changes in gray matter on a global or regional basis, often using Talairach space.

Gaser and colleagues (26) compared the 3D vector deformation fields required to warp each voxel of an index brain (source not specified, presumably that of the Talairach atlas) onto spatially normalized brains of a large group of schizophrenic patients (n = 85) and controls (n = 75). They then computed the statistical significance of the difference between the schizophrenic patients' and controls' deformation fields, finding volume reduction bilaterally in Talairach spatial locations corresponding to thalamus and superior temporal gyrus and unilateral reductions in the superior and middle frontal gyrus, precentral gyrus, lingual gyrus, and cerebellum. This study forms a good transition to the next section because the spatial normalization techniques and the nonlinear registration method, those of SPM99 and Ashburner and Friston (27), respectively, are described in the following section.

Voxel-Based Morphometry

Ashburner and Friston (27) define this technique as "a voxel-wise comparison of the local concentration of gray matter between two groups of subjects," and have provided a detailed description of this methodology, closely related to that of SPM99. As a first step, this method takes all subject images and normalizes them to the same stereotaxic space, using procedures similar to those used in SPM for fMRI and PET data. This procedure involves an initial linear (affine) match (similar to that described for brain warping) followed by a nonlinear registration using smooth spatial basis functions. These authors emphasize that this spatial normalization "does not attempt to match every cortical feature exactly, but merely corrects for global brain shape differences," thereby differentiating it from more exact attempts at a match, as discussed in the brain warping section. The second step involves segmentation of the normalized images into gray matter, white matter, and CSF. The third step is smoothing using a convolution with a Gaussian kernel, which leads to each voxel being the mean of gray matter density for it and, to a spatially progressively lesser degree, its neighbors. The last step is statistical analysis using the general linear model to identify regions of gray matter concentration that are significantly related to the variable under study (if normality is not present a nonparametric statistical analysis is used).

Compared with manually drawn ROI, this technique has the following clear advantages: (a) enabling of regional comparisons throughout the whole brain without the restrictions of a few selected areas used in the typical manually drawn region of interest methodology; (b) the reduction of labor; and (c) the ability to use large samples with an attendant increase in statistical power as a corollary to (b).

Unfortunately, however, there as been an absence of work comparing the spatial specificity and sensitivity of voxel-based analysis with manual ROI analysis, the current standard, and thus the question of validity has been incompletely addressed.

Wright and co-workers (28) have undertaken some comparisons with manual ROI. They performed manual area measurements of the head of the caudate in the transverse slice 12 mm superior to the intercommissural plane in the untransformed data. They then compared these with voxel values in the transformed data at coordinates corresponding to the center of the caudate in Talairach space for each of 20 the subjects. They found Pearson product-moment correlations between the area measurement and the voxel gray matter values for the transformed data for the 20 subjects to be about r = 0.8. These data do not, unfortunately, provide information on spatial specificity in terms of a measurement of the boundaries of the caudate in the untransformed data for the 20 subjects and the transformed data. Nor do these data, taken from the center point of a regular structure, provide any clear information on how well the transformation would work on the much more irregular cerebral cortex. Because one of the findings with transformed data was decreased gray matter in the schizophrenic group in the voxels corresponding to the right amygdala, one would have liked to see a comparison with manually drawn ROI in this structure as a way of validating the voxel analysis (and/or a comparison in the other regions found to be abnormal, the temporal pole/insula, and left dorsolateral prefrontal cortex). Wright and associates did find that voxel analysis could detect artificial "lesions," created by setting gray matter content to zero in a group of voxels, including a 4- × 4-mm bar and a 12- × 25-mm grid. They did not try more realistic "lesions" with parametric variation of degrees of lesser gray matter content; nor did they quantify

the spatial specificity. In concluding the discussion of this technique, Wright and colleagues voiced the important caveats that voxel based morphometry may not detect “very small gray matter reductions, gray matter reduction in areas of high variability in gray matter volume or gray matter reductions with an inconsistent location.”

A direct comparison of manual ROI and voxel-based analysis would seem to be a high priority, because some estimate of the specificity and sensitivity of voxel analysis for various brain regions and ROI could be formed. Until such validation procedures are done, any results with voxel-based morphometry (VBM) will, of necessity, be viewed by many workers in the field as tentative. Because of the importance of the validity question, our laboratory has recently begun to compare SPM99 VBM results with traditional ROI analysis (Kubicki and colleagues, unpublished data). For VBM applied to whole brain, only the left posterior superior temporal gyrus region was significantly different between schizophrenic and control groups, a finding consistent with our ROI analysis. In a less statistically less stringent analysis (taking into account peak z values and voxel cluster extent), there was significance bilaterally in the anterior cingulate gyri and insula (regions not examined with ROI), but not in medial temporal lobe where ROI analysis showed differences.

Taken together, these data suggest the following methodologic conclusions:

1. Each VBM study should be compared with manual ROI definitions until validity is established.
2. VBM may be useful for generating hypotheses to be validated with traditional ROI analyses.
3. Much work remains to be done in comparing the validity of VBM and ROI analysis, and formulating reasons for any differences.

Shape Analysis

It is readily apparent that ROI shape as well as volume may carry information about pathology. Casanova and colleagues (29) used 3D Fourier techniques to characterize shape of temporal lobe regions, finding schizophrenics and controls differed. However, Fourier techniques cannot pinpoint where in the shape the abnormality occurs, as can brain warping and other methodologies. Csernansky and associates (30), using a variant (based on Grenander’s work) of the brain warping techniques described in the preceding, found that maximal differences between controls and the schizophrenia subjects were localized to the lateral aspect of the head of the hippocampus and medial aspect of the body, where the subiculum is found. The study of shape using a number of different algorithms is a current area of very active interest in MR schizophrenia research, especially in the study of the corpus callosum.

DIFFUSION TENSOR MR IMAGING

Part of "55 - Structural Magnetic Resonance Imaging Studies in Schizophrenia "

This is a new MRI technology that is able to provide information on the orientation and integrity of fiber tracts. In diffusion tensor imaging (DTI), a tensor describing local water diffusion is acquired for each voxel; crudely, this tensor can be thought of as a mathematical description of the direction and velocity component of diffusion relative to the orientation of the chosen coordinate system, or “basis.” Diffusion may be “isotropic,” equal in all directions, as occurs in CSF, and the diffusion volume (3D representation of diffusion pathways) has a spherical geometry in this case. Or diffusion may be “anisotropic” (e.g., not isotropic) and greater in one direction, in which case there is an ellipsoid shape. The limiting case for maximal anisotropy is an infinitely long and thin cylinder. In white matter fiber tracts diffusion is mainly in the direction of the fibers. Factors that affect the shape of the apparent diffusion tensor (shape of the diffusion ellipsoid) in the white matter include the density of fibers, degree of myelination, average fiber diameter, and directional similarity of the fibers in the voxel. For example, the DTI-measured diffusion coefficients are larger when measured along (parallel to) white matter fibers (in the range of 1.0×10^{-3} mm^2/sec) than across the fibers (in the range of 0.6×10^{-3} mm^2/sec). The geometric nature of the measured diffusion tensor within a voxel is thus a meaningful measure of fiber tract organization.

The degree of anisotropy in schizophrenia has been investigated in two recent studies. Using DTI, Buchsbaum and associates (31) reported evidence of lower diffusion anisotropy in some inferior portions of prefrontal white matter in patients with schizophrenia than in controls. Lim and co-workers (32) found that abnormally low white matter anisotropy in patients with schizophrenia was present in both hemispheres and was widespread, extending from frontal to occipital brain regions. For group statistics, Lim and co-workers used the median value of voxel anisotropy (measured as fractional anisotropy; 1 is maximal and 0 minimal) in each slice within the white matter regions of interest in the control and schizophrenia groups. These studies raised the important question of whether white matter connectivity is disturbed in schizophrenia, although Lim and colleagues caution that the proper statistical measures for DTI are still being worked out.

A recent technical advance in DTI has been line scan diffusion imaging (LSDI). This method, in contrast to the commonly used diffusion-sensitized, ultrafast, echo-planar imaging (EPI) technique, is less sensitive to gross motion and cardiovascular pulsations. LSDI also has higher resolution, exhibits minimal image distortion, and does not require cardiac gating, head restraints, or post-processing image correction. It also can be implemented without specialized hardware on all standard MRI scanners.

Recent work has focused on measurements extending beyond the scalar measurement of the degree of anisotropy

in a voxel to characterizing the spatial trajectory and orientation of fiber tracts. Although the individual axons and the surrounding myelin sheaths cannot be revealed with the limited spatial resolution of *in vivo* imaging, distinct bands of white matter fibers with parallel orientation may be distinguished from others running in different directions if MRI techniques are sensitized to water diffusion and the preferred direction of diffusion is determined. Figure 55.4 shows the degree to which orientation of fiber tracts can be quantified and displayed using color-coding. An important point for summarizing data made by Westin and associates (33) is that remaining within the tensor domain when processing is useful, as contrasted with operating on scalars and vectors to produce summary statistics. In processing DTI images, it is important to note that averaging of a diffusion tensor field and then deriving a scalar measure from the averaged field is not the same as averaging a scalar field derived from the original field. By using geometrically defined diffusion measures on locally averaged tensors local directionality consistency can be determined (e.g., existence of larger fiber tracts). This averaging approach can be used to derive a tensor field that will describe macrostructural features in the tensor diffusion data. For example, a measure of linearity derived from the averaged tensor field can be used for quantitative evaluation of fiber tract organization.

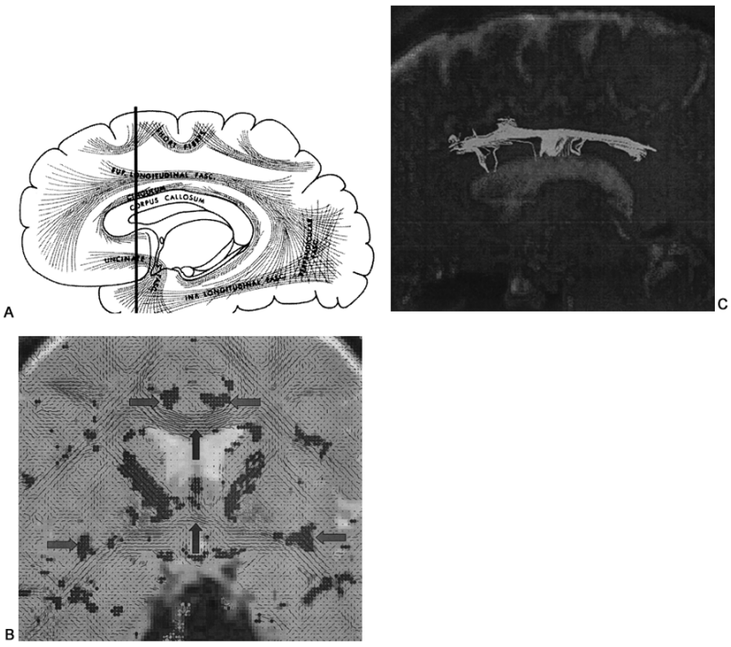


FIGURE 55.4. A: Sagittal schematic of brain fiber tracks. The vertical line shows the approximate plane of the coronal diffusion tensor image to the right. (Adapted from Gray H, Bannister LH, eds. *Gray's anatomy: the anatomical basis of medicine and surgery*, thirty-eighth ed. London: Churchill Livingstone, 1995.) **B:** In this diffusion tensor imaging (DTI) image, white matter tracts that are within the coronal plane are color-coded blue. Note the corpus callosum *(top blue arrow)* and anterior commissure *(bottom blue arrow)*. White matter tracts perpendicular to the plane are coded red-orange. Note the cingulum bundle *(top arrows)*, the white matter tract within the cingulate gyrus, and the uncinate fasiculus *(bottom arrows)*, the tract connecting anterior temporal lobe with inferior frontal lobe. (Unpublished image from our laboratory, Shenton et al., 2000; technique and application discussed in Kubicki M, Maier SE, McCarley RW, et al. *Uncinate fasciculus in schizophrenia: a diffusion tensor study*. American Psychiatric Association New Research Abstracts, 2000.) **C:** Parasagittal image showing anterior–posterior course of the cingulate bundle as constructed from DTI. (Unpublished image courtesy of Stephan Maier and Carl-Fredrik Westin, Surgical Planning Laboratory, Brigham and Women's Hospital).

Still another promising application of DTI is tracking white matter tracts. The operation begins with a seed point in a voxel element and then generates a tracking sequence

if the adjacent elements have similar linear orientation. This similarity is at the voxel level, and does not, of course, permit tracking of individual fibers; rather, it tracks groups of fibers.

CONCLUSION

Part of "55 - Structural Magnetic Resonance Imaging Studies in Schizophrenia "

A clear current and positive trend is to use as much automation as possible in structural MRI analysis because of the labor involved in traditional ROI analysis. Currently, however, the field is still in a state of flux with respect to the validity of the new techniques, such as VBM and brain warping, because of the absence of detailed comparisons with ROI analysis and formulation of the reasons for differences. Validity evaluation for new technologies is thus a high priority item.

Another clear and positive trend is employ new technologies, and to use multimodal imaging, with diffusion tensor imaging as the prime example within the structural field. Similarly, as discussed in another chapter in this volume by Dr. Berman, “functional” imaging is becoming increasingly multimodal and a desideratum is the combination of structural and functional approaches, just as anatomy and physiology are inextricably linked in basic neuroscience studies. The reader will likely notice a certain “mismatch” in the brain regions emphasized in the functional imaging chapter (frontal lobe) and in this structural imaging chapter (temporal lobe). It is clear that functional studies have defined more prominent abnormalities in frontal lobe than in temporal lobe, whereas structural studies have tended to show a greater degree of abnormality in temporal lobe. The mismatch may arise, in part, because the frontal lobe receives input from and communicates with virtually all cortical (and many subcortical) areas. Functional neuroimaging “activation” in a region primarily represents postsynaptic potentials; these and not action potentials constitute the major metabolic and energetic load and hence the main signals used in functional analysis. It is consequently often very difficult to disambiguate abnormalities in input to frontal lobe from intrinsic abnormalities. Similarly, although temporal lobe gray matter volume changes appear quantitatively larger than those in frontal cortex, no brain region acts on its own and interconnections and abnormalities of interconnections, as well as intrinsic volume changes must be considered in the explanation of the features of schizophrenia. This task of seamlessly integrating information from multiple technologies is both one of the most exciting and also the most challenging for work in the next few years.

APPENDIX A: STRUCTURAL MRI PULSE SEQUENCES

STRUCTURAL MRI PULSE SEQUENCES

Spin echo pulse sequences use at least two pulses. The first is an initial excitation pulse (tilting the magnetization vector 90 degrees from the steady-state field orientation), followed by one or more refocusing pulses, and directed 180 degrees from the orientation of the steady-state field. These refocusing pulses reintroduce phase coherence (again, the web site http://ej.rsna.org/ej3/0095-98.fin/index.htm provides a useful animated illustration of this). The reformation of phase coherence induces another signal known as a “spin echo,” which does not have the potential confounds of magnet and tissue inhomogeneity (they remain constant over pulses), and thus this signal provides a better measure of T2. In spin echo pulse sequences the repetition time (TR) is the time between excitation pulses, whereas the echo time (TE) is the time from the excitation pulse to the echo maximum.

Relatively short TR and TE standard single echo sequences produce T1-weighted images. Multiecho sequences produce proton density weighted images at short TE (less than 30 ms) and T2-weighted images at long TE (more than 80 ms) when TR is long enough to allow for nearly compete T1 relaxation (more than 2,000 ms for most tissues). Fast spin echo sequences are a variant of multiecho sequences that maximize efficiency of data collection and shorten acquisition time. They are commonly used to produce T2-weighted images. Inversion recovery pulse sequences are still another variation of the spin echo sequence, in which an additional 180-degree pulse is applied before the excitation pulse, thereby increasing T1 weighting (commonly used for improving contrast between different tissues).

Gradient Echo Pulse Sequences

These sequences do not use 180-degree refocusing pulses. The most commonly used pulse sequence in volumetric work is a spoiled gradient echo sequence, called spoiled GRASS (SPGR) in GE imagers and FLASH in Siemens imagers. This pulse sequence uses a “spoiling scheme” to dephase the transverse (x-y plane) magnetization following signal detection, commonly using “spoiler” (also called “crusher”) gradient pulses that have the same duration and magnitude as the first excitation pulse, but the opposite polarity. This has as a consequence that, at the time of the next excitation, only the longitudinal direction (vertical direction in our analogy) has any remaining coherence. If the first pulse has a low excitation angle (small “tilt” of the tops in our analogy) this allows shorter repetition times to be used, speeding acquisition.

Signal Intensity of Tissue Elements and T1 and T2 Weighting

SPGR pulses lead to proton density-weighted images, because the small “tilt” and short TR diminishes any T1 or T2 effects. In a proton density image produced by the SPGR sequence most commonly used in schizophrenia research, CSF appears dark, gray matter is gray, and white matter

has the most signal (is brightest). Spin-echo sequences can produce proton density, T2- or T1-weighted images. The normal CSF T1 relaxation time is 3,000 ms at 1.5 T, whereas that of fat is 200 to 250 ms; gray matter has a longer T1 relaxation time than white matter and thus shows a brighter signal with sequences allowing longer T1 relaxation times. Because the ability to capture relatively complete T1 relaxation depends on longer TRs, longer TRs thus give brighter CSF and gray matter brighter than white matter. The tissue intensity in T2-weighted images depends on the TE in spin echo sequences. CSF has longer T2 values than other brain tissues and shows up as a bright signal in T2-weighted acquisitions with the long TE values commonly used in schizophrenia volumetric studies. In general, a long TR allows more time for T1 relaxation and produces more signal from tissues with long T1 values, whereas a long TE allows more time for T2 relaxation and produces more signal from tissues with long T2 values.

ACKNOWLEDGMENTS

Supported in part by VA Medical Research Service, Department of Veterans Affairs Center for Clinical and Basic Neuroscience Studies of Schizophrenia, NIMH 40977 and 52807 (RWM). Parts of the introduction are adapted from a previous review: McCarley RW, Wible C, Frumin M, et al. MRI anatomy of schizophrenia. *Biol Psychiatry* 1999;45:1099–1119.

REFERENCES

1. Kraepelin E (Barclay E, Barclay S, Trans.). *Dementia praecox*. New York: New York: Churchill Livingstone, 1971 (German edition published in 1899).

2. Brown MA, Semelka RC. *MRI: basic principles and applications*, second ed. New York: Wiley-Liss, 1999.

3. Pfefferbaum A, Sullivan EV, Hedehus M, et al. Brain gray and white matter transverse relaxation time in schizophrenia. *Psychiatry Res* 1999;91:93–100.

4. Smith RC, Calderon M, Ravichandran GK, et al. Nuclear magnetic resonance in schizophrenia: A preliminary study. *Psychiatry Res* 1984;12:137–147.

5. Kikinis R, Shenton ME, Jolesz FA, et al. Routine analysis of brain and cerebrospinal fluid spaces with MR imaging. *J Magn Reson Imaging* 1992;2:619–629.

6. Gerig G, Kikinis R, Kubler O. Significant improvement of MR image data quality using anisotropic diffusion filtering. *Technical report BIWI-TR-124*. Zurich: ETH, 1990.

7. Shenton ME, Kikinis R, McCarley RW, et al. Application of automated MRI volumetric measurement techniques to the ventricular system in schizophrenics and normals. *Schizophr Res* 1991;5:103–113.

8. Gur RE, Cowell P, Turetsky BI, et al. A follow-up magnetic resonance imaging study of schizophrenia: relationship of neuroanatomical changes to clinical and neurobehavioral measures. *Arch Gen Psychiatry* 1998;55:145–152.

9. Hunter JE, Schmidt FL. *Methods of metaanalysis*. Newbury Park, CA: Sage, 1990.

10. Petitti DB. *Metaanalysis, decision analysis, and cost-effectiveness analysis*. New York: Oxford University Press, 1994.

11. Rosenthal R. *Judgement studies. Design, analysis and metaanalysis*. Cambridge: Cambridge University Press, 1987.

12. Nelson MD, Saykin AJ, Flashman LA, et al. Hippocampal volume reduction in schizophrenia as assessed by magnetic resonance imaging: a meta-analytic study. *Arch Gen Psychiatry* 1998;55:433–440.

13. McCarley RW, Wible C, Frumin M, et al. MRI anatomy of schizophrenia. *Biol Psychiatry* 1999;45:1099–1119.

14. Bogerts B. The neuropathology of schizophrenic diseases: historical aspects and present knowledge. *Eur Arch Psychiatry Clin Neurosci* 1999;249(Suppl 4):2–13.

15. Harrison PJ. The neuropathology of schizophrenia. A critical review of the data and their interpretation. *Brain* 1999;122:593–624.

16. Weinberger DR. Cell biology of the hippocampal formation in schizophrenia. *Biol Psychiatry* 1999;45:395–402.

17. Selemon LD, Rajkowska G, Goldman-Rakic PS. Elevated neuronal density in prefrontal area 46 in brains from schizophrenic patients: application of a three-dimensional stereologic counting method. *J Comp Neurol* 1998;392:402–412.

18. Kikinis R, Shenton ME, Gerig G, et al. Temporal lobe sulco-gyral pattern anomalies in schizophrenia: an *in vivo* MR three-dimensional surface rendering study. *Neurosci Lett* 1994;183:7–12.

19. Iosifescu DV, Shenton ME, Warfield SK, et al. An automated registration algorithm for measuring MRI subcortical brain structures. *Neuroimage* 1997;6:13–25.

20. Schmidt M, Dengler J. Adapting multi-grid methods to the class of elliptic partial differential equation appearing in the estimation of displacement vector fields. In: Cantoni V, Creutzburg R, Levialdi S, et al, eds. *Recent issues in pattern analysis and recognition*. Berlin: Springer, 1989:266—274.

21. Warfield S, Dengler J, Zaers J, et al. Automatic identification of grey matter structures from MRI to improve the segmentation of white matter lesions. *Proceedings of Medical Robotics and Computer Assisted Surgery (MRCAS)*, November, 1995;140–47.

22. Grenander U, Miller MI. Representation of knowledge in complex systems. *J R Stat Soc B* 1994;56:3.

23. Collins DL, Peters TM, Dai W, et al. Model based segmentation of individual brain structures from MRI data. *SPIE, Visual Biomed Comput* 1992;1808:10–23.

24. Gee JC, Reivich M, Bajcsy R. Elastically deforming 3D atlas to match anatomical brain images. *J Comput Assist Tomogr* 1993;17:225–236.

25. Thompson PM, Toga AW. Brain warping. In: Toga AW, ed. *Brain warping*. San Diego: Academic Press, 1999:311–336.

26. Gaser C, Volz H-P, Kiebel S, et al. Detecting structural changes in whole brain based on nonlinear deformations: application to schizophrenia research. *NeuroImage* 1999;10:107–113.

27. Ashburner J, Friston KJ. Voxel-based morphometry: the methods. *NeuroImage* 2000;11:805–821.

28. Wright IC, Ellison ZR, Sharma T, et al. Mapping of grey matter changes in schizophrenia. *Schiz Res* 1999;35:1–14.

29. Casanova MF, Goldberg TE, Suddath RL, et al. Quantitative shape analysis of the temporal and prefrontal lobes of schizophrenic patients: a magnetic resonance image study. *J Neuropsychiatry Clin Neurosci* 1990;2:363–72.

30. Csernansky JG, Joshi S, Wang L, et al. Hippocampal morphometry in schizophrenia by high dimensional brain mapping. *Proc Natl Acad Sci USA* 1998;95:11406–11411.

31. Buchsbaum MS, Tang CY, Peled S, et al. MRI white matter diffusion anisotropy and PET metabolic rate in schizophrenia. *Neuroreport* 1998;9:425–430.

32. Lim KO, Hedehus M, Moseley M, et al. Compromised white matter tract integrity in schizophrenia inferred from diffusion tensor imaging. *Arch Gen Psychiatry* 1999;56:367–374.

33. Westin CF, Maier SE, Khidhir B, et al. Image processing for diffusion tensor magnetic resonance imaging. In: *Proceedings of Second International Conference on Medical Image Computing and Computer-Assisted Interventions*, Cambridge, UK: 1999:441–452.

34. Kubicki M, Maier SE, McCarley RW, et al. *Uncinate fasciculus in schizophrenia: a diffusion tensor study*. American Psychiatric Association New Research Abstracts, 2000.

Table 55.1 References

1. **Andreasen NC, Ehrhardt JC, Swayze II VW, et al**. Magnetic resonance imaging of the brain in schizophrenia: the pathophysiologic significance of structural abnormalities. *Arch Gen Psychiatry* 1990;47:35–44.
2. **Andreasen NC, Arndt S, Swayze II VW, et al**. Thalamic abnormalities in schizophrenia visualized through magnetic resonance image averaging. *Science* 1994a;266:294–298.
3. **Andreasen NC, Flashman L, Flaum M, et al**. Regional brain abnormalities in schizophrenia measured with magnetic resonance imaging. *JAMA* 1994b;272:1763–1769.
4. **Barr WB, Ashtari M, Bilder RM, et al**. Brain morphometric comparison of first episode schizophrenia and temporal lobe epilepsy. *Br J Psychiatry* 1997;170:515–519.
5. **Barta PE, Pearlson GD, Powers RE, et al**. Auditory hallucinations and smaller superior temporal gyral volume in schizophrenia. *Am J Psychiatry* 1990;147:1457–1462.
6. **Barta PE, Pearlson GD, Brill LB, et al**. Planum temporale asymmetry reversal in schizophrenia: replication and relationship to gray matter abnormalities. *Am J Psychiatry* 1997a;154:661–667.
7. **Barta PE, Powers RE, Aylward EH, et al**. Quantitative MRI volume changes in late onset schizophrenia and Alzheimer's disease compared to normal controls. *Psychiatry Res* 1997b;68:65–75.
8. **Becker T, Elmer K, Mechela B, et al**. MRI findings in the medial temporal lobe structures in schizophrenia. *Eur Neuropsychopharmacol* 1990;1:83–186.
9. **Becker T, Elmer K, Schneider F, et al**. Confirmation of reduced temporal limbic structure volume on magnetic resonance imaging in male patients with schizophrenia. *Psychiatry Res* 1996;67:135–143.
10. **Bilder RM, Wu H, Bogerts B, et al**. Absence of regional hemispheric volume asymmetries in first-episode schizophrenia. *Am J Psychiatry* 1994;151:1437–1447.
11. **Blackwood DHR, Young AH, McQueen JK, et al**. Magnetic resonance imaging in schizophrenia: Altered brain morphology associated with P300 abnormalities and eye tracking dysfunction. *Biol Psychiatry* 1991;30:753–769.
12. **Bogerts B, Ashtari M, Degreef G, et al**. Reduced temporal limbic structure volumes on magnetic resonance images in first episode schizophrenia. *Psychiatry Res Neuroimag* 1990;35:1–13.
13. **Bogerts B, Lieberman JA, Ashtari M, et al**. Hippocampus-amygdala volumes and psychopathology in chronic schizophrenia. *Biol Psychiatry* 1993;33:236–246.
14. **Bornstein RA, Schwarzkopf SB, Olson SC, et al**. Third-ventricle enlargement and neuropsychological deficit in schizophrenia. *Biol Psychiatry* 1992;31:954–961.
15. **Breier A, Buchanan RW, Elkashef A, et al**. Brain morphology and schizophrenia: a magnetic resonance imaging study of limbic, prefrontal cortex, and caudate structures. *Arch Gen Psychiatry* 1992;49:921–926.
16. **Buchanan RW, Breier A, Kirkpatrick B, et al**. Structural abnormalities in deficit and nondeficit schizophrenia. *Am J Psychiatry* 1993;150:59–65.
17. **Buchsbaum MS, Someya T, Teng CY, et al**. PET and MRI of the thalamus in never-medicated patients with schizophrenia. *Am J Psychiatry* 1996;153:191–199.
18. **Buchsbaum MS, Yang S, Hazlett E, et al**. Ventricular volume and asymmetry in schizotypal personality disorder and schizophrenia assessed with magnetic resonance imaging. *Schizophr Res* 1997;27:45–53.
19. **Casanova MF, Zito M, Goldberg T, et al**. Shape distortion of the corpus callosum of monozygotic twins discordant for schizophrenia. *Schizophr Res* 1990;3:155–156.
20. **Chakos MH, Lieberman JA, Bilder RM, et al**. Increase in caudate nuclei volumes of first-episode schizophrenic patients taking antipsychotic drugs. *Am J Psychiatry* 1994;151:1430–1436.
21. **Chakos MH, Lieberman JA, Alvir J, et al**. Caudate nuclei volumes in schizophrenic patients treated with typical antipsychotics or clozapine. *Lancet* 1995;345:456–457.
22. **Colombo C, Abbruzzese M, Livian S, et al**. Memory functions and temporal limbic morphology in schizophrenia. *Psychiatry Res Neuroimag* 1993;50:45–56.
23. **Colombo C, Bonfanti A, Scarone S**. Anatomical characteristics of the corpus callosum and clinical correlates in schizophrenia. *Eur Arch Psychiatry Clin Neurosci* 1994;243:244–248.
24. **Coffman JA, Schwarzkopf SB, Olson SC, et al**. Midsagittal cerebral anatomy by magnetic resonance imaging: The importance of slice position and thickness. *Schizophr Res* 1989;2:287–294.
25. **Corey-Bloom J, Jernigan T, Archibald S, et al**. Quantitative magnetic resonance imaging of the brain in late-life schizophrenia. *Am J Psychiatry* 1995;152:447–449.
26. **Dauphinais D, DeLisi LE, Crow TJ, et al**. Reduction in temporal lobe size in siblings with schizophrenia: a magnetic resonance imaging study. *Psychiatry Res Neuroimag* 1990;35:137–147.
27. **Degreef G, Bogerts B, Ashtari M, et al**. Ventricular system morphology in first episode schizophrenia: a volumetric study of ventricular subdivisions on MRI. *Schizophr Res* 1990;3:18.
28. **Degreef G, Ashtari M, Bogerts B, et al**. Volumes of ventricular system subdivisions measured from magnetic resonance images in first-episode schizophrenic patients. *Arch Gen Psychiatry* 1992a;49:531–537.
29. **Degreef G, Bogerts B, Falkai P, et al**. Increased prevalence of the cavum septum pellucidum in magnetic resonance scans and post-mortem brains of schizophrenic patients. *Psychiatry Res: Neuroimaging* 1992b;45:193–199.
30. **Degreef G, Lantos G, Bogerts B, et al**. Abnormalities of the septum pellucidum on MR scans in first-episode schizophrenic patients. *Am J Neuroradiology* 1992c;13:835–840.
31. **DeLisi LE, Hoff AL, Schwartz JE, et al**. Brain morphology in first-episode schizophrenic-like psychotic patients: a quantitative magnetic resonance imaging study. *Biol Psychiatry* 1991;29:159–175.
32. **DeLisi LE, Strizke P, Riordan H, et al**. The timing of brain morphological changes in schizophrenia and their relationship to clinical outcome. *Biol Psychiatry* 1992;31:241–254.
33. **DeLisi LE, Hoff AL, Kushner M, et al**. Increased prevalence of cavum septum pellucidum in schizophrenia. *Psychiatry Res: Neuroimag* 1993;50:193–199.
34. **DeLisi LE, Hoff AL, Neale C, et al**. Asymmetries in the superior temporal lobe in male and female first-episode schizophrenic patients: measures of the planum temporale and superior temporal gyrus by MRI. *Schizophr Res* 1994;12:19–28.
35. **DeLisi LE, Sakuma M, Tew W, et al**. Schizophrenia as a chronic active brain process: a study of progressive brain structural change subsequent to the onset of schizophrenia. *Psychiatry Res* 1997;74:129–140.
36. **DeQuardo JR, Bookstein FL, Green WD, et al**. Spatial relationships of neuroanatomic landmarks in schizophrenia. *Psychiatry Res* 1996;67:81–95.

37. **Di Michele V, Rossi A, Stratta P, et al**. Neuropsychological and clinical correlates of temporal lobe anatomy in schizophrenia. *Acta Psychiatr Scand* 1992;85:484–488.
38. **Egan MF, Duncan CC, Suddath RL, et al**. Event-related potential abnormalities correlate with structural brain alterations and clinical features in patients with chronic schizophrenia. *Schizophr Res* 1994;11:259–271.
39. **Elkashef AM, Buchanan RW, Gellad F, et al**. Basal ganglia pathology in schizophrenia and tardive dyskinesia: an MRI quantitative study. *Am J Psychiatry* 1994;151:752–755.
40. **Flaum M, Swayze II VW, O'Leary DS, et al**. Effects of diagnosis and gender on brain morphology in schizophrenia. *Am J Psychiatry* 1995b;152:704–714.
41. **Fukuzako T, Fukuzako H, Kodama S, et al**. Cavum septum pellucidum in schizophrenia: a magnetic resonance imaging study. *Psychiatry Clin Neurol* 1996a,50:125–128.
42. **Fukuzako H, Fukuzako T, Hashiguchi T, et al**. Reduction in hippocampal formation volume is caused mainly by its shortening in chronic schizophrenia: assessment by MRI. *Biol Psychiatry* 1996b;39:938–945.
43. **Gunther W, Petsch R, Steinberg R, et al**. Brain dysfunction during motor activation and corpus callosum alterations in schizophrenia measured by cerebral blood flow and magnetic resonance imaging. *Biol Psychiatry* 1991;29:535–555.
44. **Gur RE, Mozley PD, Shtasel DL, et al**. Clinical subtypes of schizophrenia: differences in brain and CSF volume. *Am J Psychiatry* 1994;151:343–350.
45. **Gur RE, Cowell P, Turetsky BI, et al**. A follow-up magnetic resonance imaging study of schizophrenia: relationship of neuroanatomical changes to clinical and neurobehavioral measures. *Arch Gen Psychiatry* 1998;55:145–152.
46. **Hajek M, Huonker R, Boehle C, et al**. Abnormalities of auditory evoked magnetic fields and structural changes in the left hemisphere of male schizophrenics—a magnetoencephalographic-magnetic resonance imaging study. *Biol Psychiatry* 1997;42:609–616.
47. **Harvey J, Ron MA, Boulay GD, et al**. Reduction of cortical volume in schizophrenia on magnetic resonance imaging. *Psychol Med* 1993;23:591–604.
48. **Hauser PI, Dauphinais D, Berrettini W, et al**. Corpus callosum dimensions measured by magnetic resonance imaging in bipolar affective disorder and schizophrenia. *Biol Psychiatry* 1989;26:659–668.
49. **Hirayasu Y, Shenton ME, Salisbury DF, et al**. Lower left temporal lobe MRI volumes in patients with first-episode schizophrenia compared with psychotic patients with first-episode affective disorder and normal subjects. *Am J Psychiatry* 1998a;155:1384–1391.
50. **Hoff AL, Riordan H, O'Donnell D, et al**. Anomalous lateral sulcus and cognitive function in first-episode schizophrenia. *Schizophr Bull* 1992;18:257–272.
51. **Hoff AL, Neal C, Kushner M, et al**. Gender differences in corpus callosum size in first-episode schizophrenics. *Biol Psychiatry* 1994;35:913–919.
52. **Hokama H, Shenton ME, Nestor PG, et al**. Caudate, putamen, and globus pallidus volume in schizophrenia: a quantitative MRI study. *Psychiatry Res Neuroimag* 1995;61:209–229.
53. **Jernigan TL, Zisook S, Heaton RK, et al**. Magnetic resonance imaging abnormalities in lenticular nuclei and cerebral cortex in schizophrenia. *Arch Gen Psychiatry* 1991;48:881–890.
54. **Johnstone EC, Owens DGC, Crow TJ, et al**. Temporal lobe structure as determined by nuclear magnetic resonance in schizophrenia and bipolar affective disorder. *J Neurol Neurosurg Psychiatry* 1989;52:736–741.
55. **Jurjus GJ, Nasrallah HA, Olson SC, et al**. Cavum septum pellucidum in schizophrenia, affective disorder, and healthy controls: a magnetic resonance imaging study. *Psychol Med* 1993;23:319–322.
56. **Kawasaki Y, Maeda Y, Urata K, et al**. A quantitative magnetic resonance imaging study of patients with schizophrenia. *Eur Arch Psychol Clin Neurosci* 1993;242:268–272.
57. **Kelsoe JR, Cadet JL, Pickar D, et al**. Quantitative neuroanatomy in schizophrenia: a controlled magnetic resonance imaging study. *Arch Gen Psychiatry* 1988;45:533–541.
58. **Keshavan MS, Bagwell WW, Haas GL, et al**. Does caudate volume increase during follow up in first-episode psychosis? *Schizophr Res* 1995;15:87.
59. **Kikinis R, Shenton ME, Gerig G, et al**. Temporal lobe sulco-gyral pattern anomalies in schizophrenia: an in vivo MR three-dimensional surface rendering study. *Neuroscience Lett* 1994;182:7–12.
60. **Kleinschmidt P, Falkai P, Huang Y, et al**. In vivo morphometry of planum temporale asymmetry in first-episode schizophrenia. *Schizophr Res* 1994;12:9–18.
61. **Kulynych JJ, Vladar K, Fantie BD, et al**. Normal asymmetry of the planum temporale in patients with schizophrenia: three-dimensional cortical morphometry with MRI. *Br J Psychiatry* 1995;166:742–749.
62. **Kulynych JJ, Vladar K, Jones DW, et al**. Superior temporal gyrus volume in schizophrenia: a study using MRI morphometry assisted by surface rendering. *Am J Psychiatry* 1996;153:50–56.
63. **Kwon JS, Shenton ME, Hirayasu Y, et al**. MRI study of cavum septi pellucidi in schizophrenia, affective disorder, and schizotypal personality disorder. *Am J Psychiatry* 1998;155:509–515.
64. **Kwon JS, McCarley RW, Hirayasu Y, et al**. Left planum temporale volume reduction in schizophrenia. *Arch Gen Psychiatry* 1999;56:142–148.
65. **Lauriello J, Hoff A, Wieneke MH, et al**. Similar extent of brain dysmorphology in severely ill women and men with schizophrenia. *Am J Psychiatry* 1997;154:819–825.
66. **Lewine RR, Gulleu LR, Risch SC, et al**. Sexual dimorphism, brain morphology, and schizophrenia. *Schizophr Bul* 1990;16:195–203.
67. **Lim KO, Tew W, Kushner M, et al**. Cortical gray matter volume deficit in patients with first-episode schizophrenia. *Am J Psychiatry* 1996;153:1548–1553.
68. **Marsh L, Suddath RL, Higgins N, et al**. Medial temporal lobe structures in schizophrenia: relationship of size to duration of illness. *Schizophr Res* 1994;11:225–238.
69. **Marsh L, Harris D, Lim KO, et al**. Structural magnetic resonance imaging abnormalities in men with severe chronic schizophrenia and an early age at clinical onset. *Arch Gen Psychiatry* 1997;54:1104–1112.
70. **Menon RR, Barta PE, Aylward EH, et al**. Posterior superior temporal gyrus in schizophrenia: grey matter changes and clinical correlates *Schizophr Res* 1995;16:127–135.
71. **Mion CC, Andreasen NC, Arndt S, et al**. MRI abnormalities in tardive dyskinesia. *Psychiatry Res Neuroimag* 1991;40:157–166.
72. **Nasrallah HA, Schwarzkopf SB, et al**. Gender differences in schizophrenia on MRI brain scans. *Schizophr Bull* 1990;16:205–210.
73. **Nopoulos P, Torres I, Flaum M, et al**. Brain morphology in first-episode schizophrenia. *Am J Psychiatry* 1995;152:1721–1723.
74. **Nopoulos P, Swayze V, Andreasen NC**. Pattern of brain morphology in patients with schizophrenia and large cavum septi pellucidi. *J Neuropsychiatry Clin Neurosci* 1996;8:147–152.
75. **Nopoulos P, Swayze V, Flaum M, et al**. Cavum septi pellucidi in normals and patients with schizophrenia as detected by magnetic resonance imaging. *Biol Psychiatry* 1997;41:1102–1108.
76. **Ohnuma T, Kimura N, Takahashi T, et al**. A magnetic resonance imaging study in first episode disorganized-type patients with schizophrenia. *Psychiatry Clin Neurosci* 1997;51:9–15.
77. **Petty RG, Barta PE, Pearlson GD, et al**. Reversal of asymmetry of the planum temporale in schizophrenia. *Am J Psychiatry* 1995;152:715–721.
78. **Portas CM, Goldstein JM, Shenton ME, et al**. Volumetric evaluation of the thalamus in schizophrenic male patients using magnetic resonance imaging. *Biol Psychiatry* 1998;43:649–659.

79. **Raine A, Harrison GN, Reynolds GP, et al**. Structural and functional characteristics of the corpus callosum in schizophrenics. *Arch Gen Psychiatry* 1990;47:1060–1064.
80. **Raine A, Lencz T, Reynolds GP, et al**. An evaluation of structural and functional prefrontal deficits in schizophrenia: MRI and neuropsychological measures. *Psychiatry Res Neuroimag* 1992;45:123–137.
81. **Reite M, Sheeder J, Teale P, et al**. Magnetic source imaging evidence of sex differences in cerebral lateralization in schizophrenia. *Arch Gen Psychiatry* 1997;54:433–440.
82. **Rossi A, Stratta P, Gallucci M, et al**. Standardized magnetic resonance image intensity study in schizophrenia. *Psychiatry Res* 1988;25:223–231.
83. **Rossi A, Stratta P, Galluci M, et al**. Quantification of corpus callosum and ventricles in schizophrenics with nuclear magnetic resonance imaging: a pilot study. *Am J Psychiatry* 1989a;46:99–101.
84. **Rossi A, Stratta P, D'Albenzio L, et al**. Reduced temporal lobe area in schizophrenia by magnetic resonance imaging: preliminary evidence. *Psychiatry Res* 1989b;29:261–263.
85. **Rossi A, Stratta P, D'Albenzio L, et al**. Reduced temporal lobe areas in schizophrenia: preliminary evidences from a controlled multiplanar magnetic resonance imaging study. *Biol Psychiatry* 1990b;27:61–68.
86. **Rossi A, Stratta A, Michele VD, et al**. Temporal lobe structure by magnetic resonance in bipolar affective disorders and schizophrenia. *J Affective Dis* 1991;21:19–22.
87. **Rossi A, Stratta P, Mattei P, et al**. Planum temporale in schizophrenia: a magnetic resonance study. *Schizophr Res* 1992;7:19–22.
88. **Rossi A, Stratta P, Mancini F, et al**. Cerebellar vermal size in schizophrenia: a male effect. *Biol Psychiatry* 1993;33:354–357.
89. **Rossi A, Serio A, Stratta P, et al**. Planum temporale asymmetry and thought disorder in schizophrenia. *Schizophr Res* 1994a;12:1–7.
90. **Rossi A, Stratta P, Mancini F, et al**. Magnetic resonance imaging findings of amygdala-anterior hippocampus shrinkage in male patients with schizophrenia. *Psychiatry Res* 1994b;52:43–53.
91. **Schlaepfer TE, Harris GJ, Tien AY, et al**. Decreased regional cortical gray matter volume in schizophrenia. *Am J Psychiatry* 1994;151:842–848.
92. **Schwartz JM, Aylward E, Barta PE, et al**. Sylvian fissure size in schizophrenia measured with the magnetic resonance imaging rating protocol of the consortium to establish a registry for Alzheimer's disease. *Am J Psychiatry* 1992;149:1195–1198.
93. **Schwarzkopf SB, Olson SC, Coffman JA, et al**. Third and lateral ventricular volumes in schizophrenia: support for progressive enlargement of both structures. *Psychopharm Bull* 1990;26:385–391.
94. **Scott TF, Price TP, George MS, et al**. Midline cerebral malformations and schizophrenia. *J Neuropsychiatry Clin Neurosci* 1993;5:287–293.
95. **Shenton ME, Kikinis R, McCarley RW, et al**. Application of automated MRI volumetric measurement techniques to the ventricular system in schizophrenics and normal controls. *Schizophr Res* 1991;5:103–113.
96. **Shenton ME, Kikinis R, Jolesz FA, et al**. Abnormalities of the left temporal lobe and thought disorder in schizophrenia: a quantitative magnetic resonance imaging study. *N Engl J Med* 1992;327:604–612.
97. **Shioiri T, Oshitani Y, Kato T, et al**. Prevalence of cavum septum pellucidum detected by MRI in patients with bipolar disorder, major depression and schizophrenia. *Psychol Med* 1996;26:431–434.
98. **Stratta P, Rossi A, Gallucci M, et al**. Hemispheric asymmetries and schizophrenia: a preliminary magnetic resonance imaging study. *Biol Psychiatry* 1989;25:275–284.
99. **Suddath RL, Casanova MF, Goldberg TE, et al**. Temporal lobe pathology in schizophrenia: a quantitative magnetic resonance imaging study. *Am J Psychiatry* 1989;146:464–472.
100. **Suddath RL, Christison GW, Torrey EF, et al**. Anatomical abnormalities in the brains of monozygotic twins discordant for schizophrenia. *N Engl J Med* 1990;332:789–794.
101. **Sullivan EV, Mathalon DH, Lim KO, et al**. Patterns of regional cortical dysmorphology distinguishing schizophrenia and chronic alcoholism. *Biol Psychiatry* 1998;43:118–131.
102. **Swayze II VW, Andreasen NC, Alliger RJ, et al**. Subcortical and temporal structures in affective disorder and schizophrenia: a magnetic resonance imaging study. *Biol Psychiatry* 1992;31:221–240.
103. **Tune L, Barta P, Wong D, et al**. Striatal dopamine D2 receptor quantification and superior temporal gyrus: volume determination in 14 chronic schizophrenic subjects. *Psychiatry Res* 1996;67:155–158.
104. **Uematsu M, Kaiya H**. The morphology of the corpus callosum in schizophrenia: an MRI study. *Schizophr Res* 1988;1:391–398.
105. **Uematsu M, Kaiya H**. Midsagittal cortical pathomorphology of schizophrenia: a magnetic resonance imaging study. *Psychiatry Res* 1989;30:11–20.
106. **Vita A, Dieci M, Giobbio GM, et al**. Language and thought disorder in schizophrenia: brain morphological correlates. *Schizophr Res* 1995;15:243–251.
107. **Wible CG, Shenton ME, Hokama H, et al**. Prefrontal cortex and schizophrenia: a quantitative magnetic resonance imaging study. *Arch Gen Psychiatry* 1995;52:279–288.
108. **Woodruff PW, Pearlson GD, Geer MJ, et al**. A computerized magnetic resonance imaging study of corpus callosum morphology in schizophrenia. *Psychol Med* 1993;23:45–56.
109. **Woodruff PW, Wright IC, Shuriquie N, et al**. Structural brain abnormalities in male schizophrenics reflect fronto-temporal dissociation. *Psychol Med* 1997a;27:1257–1266.
110. **Woodruff PW, Philips ML, Rushe T, et al**. Corpus callosum size and inter-hemispheric function in schizophrenia. *Schizophr Res* 1997b;23:189–196.
111. **Woods BT, Yurgelun-Todd D, Goldstein JM, et al**. MRI Brain abnormalities in chronic schizophrenia: one process or more? *Biol Psychiatry* 1996;40:585–596.
112. **Zipursky RB, Lim DO, Sullivan EV, et al**. Widespread cerebral gray matter volume deficits in schizophrenia. *Arch Gen Psychiatry* 1992;49:195–205.
113. **Zipursky RB, Marsh L, Lim KO, et al**. Volumetric MRI assessment of temporal lobe structures in schizophrenia. *Biol Psychiatry* 1994;35:501–516.
114. **Zipursky RB, Seeman MV, Bury A, et al**. Deficits in gray matter volume represent in schizophrenia but not bipolar disorder. *Schizophr Res* 1997;26:85–92.

56
Therapeutics of Schizophrenia

Seiya Miyamoto

Gary E. Duncan

Donald C. Goff

Jeffrey A. Lieberman

Seiya Miyamoto, Gary E. Duncan, and Jeffrey A. Lieberman: Departments of Psychiatry and Pharmacology, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina.

Donald C. Goff: Psychotic Disorders Program of the Massachusetts General Hospital, and Consolidated Department of Psychiatry, Harvard Medical School, Boston, Massachusetts.

The introduction of the first antipsychotic medication in the early 1950s revolutionized mental health care strategies and led to the era of deinstitutionalization, a period in which patients with schizophrenia and related psychotic disorders were released from state hospitals in large numbers to be cared for in the community (1). Nonetheless, with the growing understanding that a significant percentage of patients responds poorly to conventional antipsychotics, as well as the recognition of discouraging long-term outcomes for schizophrenia, the need to develop new therapeutic agents that work rapidly, potently, broadly, and with fewer side effects has become increasingly appreciated. The reintroduction of clozapine heralded the second generation of atypical antipsychotic drugs and a new pharmacotherapy of schizophrenia. To date, the greater benefits of the atypical antipsychotic drugs in many outcome domains have been demonstrated (2), and novel medications are replacing the conventional antipsychotics as treatments of choice. The development of additional novel strategies to obtain potentially new antipsychotic compounds possessing unique pharmacologic profiles with few side effects is being pursued based on specific hypotheses (3). This chapter provides a review and critique of currently available pharmacologic and psychosocial treatments in schizophrenia, and focuses on investigational treatments and potential strategies for future pharmacotherapy.

HISTORY OF ANTIPSYCHOTIC DRUG DEVELOPMENT

Part of "56 - Therapeutics of Schizophrenia "

Since the discovery of the prototypical antipsychotic chlorpromazine in the early 1950s, a number of neuroleptics were developed based on the hypothesis that schizophrenia reflected a disorder of hyperdopaminergic activity, with the dopamine D_2 receptor most strongly associated with antipsychotic response (4). In many patients with schizophrenia, the widely used conventional antipsychotic drugs (e.g., chlorpromazine and haloperidol) are effective in the treatment of the positive symptoms of schizophrenia, and also in preventing psychotic relapse (5); however, there are crucial limitations in the use of these agents. As many as 25% to 60% of patients treated with conventional antipsychotics remain symptomatic and are labeled either treatment-refractory, or partially responsive (3). In addition, these drugs at best only modestly improve negative symptoms of the deficit syndrome and a range of cognitive impairments, which may be fundamental to the disease (6). Further, conventional antipsychotics cause a variety of side effects both acutely (e.g., extrapyramidal side effects [EPS]) and with long-term exposure (e.g., tardive dyskinesia [TD]) (7 ,8). Such adverse effects may reduce compliance and represent a major drawback of these drugs.

For a number of years, there was a widely held view that any compound that was an effective antipsychotic agent must also induce EPS. The availability of clozapine and other newer atypical antipsychotic agents, however, have disproved this notion. The development of atypical antipsychotic drugs was aimed at increasing the ratio between doses that produce therapeutic effects and those that produce side effects, as well as improving efficacy (e.g., against a broader spectrum of psychopathologic symptoms and the treatment-resistant aspects of the disorder) (1). Although there is currently no uniform definition of the term “atypical,” in its broadest sense it is used to refer to drugs that have at least equal antipsychotic efficacy compared to conventional drugs, without producing EPS or prolactin elevation (1). A more restrictive definition would require that atypical drugs also have superior antipsychotic efficacy (i.e., they are effective in treatment resistant schizophrenic patients, and against negative symptoms and/or neurocognitive deficits).

Although agents like thioridazine were first suggested to

have atypical characteristics, it now is generally accepted that clozapine, first synthesized in 1958, is the prototypical “atypical” antipsychotic (9). Clozapine underwent extensive clinical testing in the 1970s, but its development was halted in the United States, and limited in other countries, because of a relatively high incidence of a potential fatal side effect, agranulocytosis. Nevertheless, its superior outcomes ultimately led to further development and eventual reintroduction beginning in 1990 (10). The renaissance of clozapine was based on several advantages: It appears to be more effective than typical neuroleptic drugs (e.g., chlorpromazine and haloperidol) in treatment refractory schizophrenia (11); it can ameliorate some of the negative as well as positive symptoms of schizophrenia (12); it can reduce relapse; it may improve certain cognitive functions; it may alleviate mood symptoms associated with schizophrenia and reduce the likelihood of suicidal behavior; it has very low liability for EPS and TD; and it does not induce sustained hyperprolactinemia (10). The reintroduction of clozapine represented a breakthrough in the treatment of schizophrenia. In recent years, concerted research and development efforts have been made to produce a second generation of “atypical” antipsychotic drugs, including risperidone, olanzapine, quetiapine, and ziprasidone, with the therapeutic advantages of clozapine, without the properties contributing to its serious side effects (13). Ongoing clinical evaluation of the new “atypical” antipsychotic drugs will eventually allow comprehensive assessment of their efficacy and safety.

REVIEW AND CRITIQUE OF CURRENT SCHIZOPHRENIA PHARMACOTHERAPY

Part of "56 - Therapeutics of Schizophrenia "

Conventional Antipsychotic Drugs

Pharmacology

Conventional or typical antipsychotic drugs can be classified as high, intermediate, or low potency based on their affinity for dopamine D_2 receptors and the average therapeutic dose, compared with a 100-mg dose of chlorpromazine (14). Haloperidol, the prototypical high-potency typical antipsychotic, has relatively high affinity for D_2 receptors and a dose of 2 to 4 mg of haloperidol is equivalent to approximately 100 mg of chlorpromazine. Low-potency drugs (e.g., thioridazine) have a chlorpromazine equivalent dose of more than 40 mg. There is a good correlation between antipsychotic potency and D_2 affinity for conventional antipsychotics of several chemical classes (4). Conventional drugs have various interactions with serotonin receptors, ranging from slight (e.g., haloperidol) to moderate (e.g., chlorpromazine).

Positron emission tomography (PET) and single photon emission computed tomography (SPECT) studies have further elucidated the importance of dopamine receptor occupancy as a predictor of antipsychotic response and adverse effects. Prospective studies have demonstrated that antipsychotic effects require a striatal D_2 receptor occupancy of 65% to 70% (15 ,16 ,17 and 18), and D_2 occupancy greater than 80% significantly increases the risk of EPS (15). Thus, a threshold between 65% and 80% D_2 occupancy appears to represent the optimal therapeutic range to minimize the risk of EPS for typical antipsychotic drugs (18 ,19 and 20). It should be noted, however, that despite adequate D_2 occupancy, many patients do not respond to medication (17). Moreover, results of studies with atypical drugs such as olanzapine indicate that receptor occupancy levels above 80% are not invariably associated with the occurrence of EPS, thus casting some doubt over the generalizability of the D_2 occupancy model with regard to atypical antipsychotics (21).

In preclinical studies, acute treatment with conventional antipsychotics (e.g., haloperidol and fluphenazine) increases the expression of c-fos mRNA or Fos protein in the dorsolateral striatum, as well as the shell of nucleus accumbens in rats (22 ,23 ,24 and 25). Neuroleptic-induced expression of Fos in the nucleus accumbens has been postulated to relate to the antipsychotic activity of both conventional and atypical drugs (26 ,27). The Fos expression in the dorsolateral striatum, which is not induced by clozapine, has been proposed to be predictive of a liability to induce EPS (23 ,27). More recently, it has been reported that haloperidol, but not clozapine, increased the immediate-early gene, arc (activity-regulated cytoskeleton-associated gene) mRNA levels in the rat striatum (28). After chronic treatment, haloperidol also induces an increase in D_2 receptor density and D_{2L} receptor mRNA in the striatum (29 ,30 and 31). Interestingly, several investigators have reported striatal enlargement after chronic treatment with conventional antipsychotics, but not atypical drugs, in both schizophrenic patients (32 ,33) and rats (34). Thus, available data suggest that conventional antipsychotic drugs may induce long-term plastic changes that lead to morphologic alterations in the striatum, and that the efficacy and side-effect profile of typical antipsychotics relate to antagonistic actions at D_2 dopamine receptors.

Efficacy

Although typical neuroleptics vary in side-effect profile and hence tolerability, there is little evidence for differences in efficacy between these drugs (3). However, in rare cases, patients failing a trial of one class may respond to the other. Although conventional neuroleptic drugs are effective for alleviating positive symptoms of schizophrenia, and preventing their recurrence in many patients, they have serious limitations. Approximately 30% of patients with acutely exacerbated psychotic symptoms have little or no response to conventional antipsychotics, and up to 50% of patients have only partial response to medication (5 ,7). Negative symptoms, mood symptoms, and cognitive deficits are marginally responsive to conventional neuroleptics. In particular, primary negative symptoms are very resistant to the

typical drugs (7 ,35). The presence of negative symptoms and cognitive impairment often leads to poor social and vocational function (36 ,37). Thus, in the absence of a clinical response at acute phase of the illness, clinicians often switch to a newer atypical agent (38).

Safety

Most conventional antipsychotics are associated with a wide range and a variable degree of undesirable acute and long-term adverse effects, including EPS; sedation; anticholinergic, autonomic, and cardiovascular effects; weight gain; sexual dysfunction; hyperprolactinemia; and neuroleptic malignant syndrome, a condition that is potentially life threatening (7 ,39). Up to 70% of patients given recommended therapeutic dosages of conventional antipsychotics develop acute EPS (40). The most troublesome neurologic side effect, tardive dyskinesia (TD), can be irreversible, and incidence rates have been estimated at about 5% per year in the nonelderly and as high as 30% per year in the elderly (41). Further, the anticholinergic drugs that are often used to reduce EPS, can also produce serious side effects (e.g., dry mouth, constipation, delirium and memory deficits) (42). All these adverse effects can contribute to treatment noncompliance, and hence increase rates of relapse and rehospitalization during the course of the chronic illness (7 ,39).

Effectiveness

Treatment with typical antipsychotics may result in poorer clinical and quality of life outcomes than with atypical antipsychotics (6). The mean first-year relapse rate during continuing maintenance treatment with conventional antipsychotics is approximately 26% in schizophrenic patients with first or multiple episodes (43). Even under the best conditions, when patients are maintained on therapeutic doses of depot conventional antipsychotics, approximately 30% of discharged patients with schizophrenia will be rehospitalized within 1 year (44). Hospital readmission rates are higher for conventional antipsychotics than for atypical antipsychotics (45). The monthly relapse rate of compliant patients taking optimal doses of a depot neuroleptic is estimated to be 3.5% per month, and the rate for patients who have discontinued their medication is 11.0% per month (44).

In terms of relapse prevention, higher doses of conventional antipsychotics may help stability, yet the patient's quality of life will be reduced because of increased side effects. Often, when considering the best dose of a conventional antipsychotic, there is a trade-off between maximizing relapse prevention and optimizing comfort (46). Although there has been substantial progress in understanding maintenance dosing, for most patients with schizophrenia, this unfortunate trade-off is inevitable with conventional antipsychotic treatment (46).

Atypical Antipsychotic Drugs

A series of atypical compounds has been developed since the introduction of clozapine. These include risperidone, olanzapine, quetiapine, and ziprasidone, which were approved by the FDA in 2000, and aripiprazole and iloperidone, which are in late Phase III development.

Pharmacology

The pharmacologic properties that confer the unique therapeutic properties of atypical antipsychotic drugs are poorly understood despite intensive research efforts. Defining the role of the individual complex actions of clozapine responsible for its unique therapeutic profile (Table 56.1) is necessary for the rational design of new and improved atypical (clozapine-like) antipsychotics because this drug is the prototype atypical drug.

TABLE 56.1. AFFINITY OF ANTIPSYCHOTIC DRUGS FOR HUMAN NEUROTRANSMITTER RECEPTORS (Ki, nM)[a]

Receptor	Clozapine	Risperidone	Olanzapine	Quetiapine	Ziprasidone	Aripiprazole	Iloperidone	Haloperidol
D_1	290	580	52	1,300	130	410[c]	320	120
D_2	130	2.2	20	180	3.1	0.52[c]	6.3	1.4
D_3	240	9.6	50	940	7.2	9.1[c]	7.1	2.5
D_4	47	8.5	50	2,200	32	260[c]	25	3.3
5-HT_{1A}	140	210	2,100	230	2.5		93	3,600
5-HT_{1D}[a,b]	1,700	170	530	>5,100	2			>5,000
5-HT_{2A}	8.9	0.29	3.3	220	0.39	20[d]	5.6	120
5-HT_{2C}	17	10	10	1,400	0.72		43	4,700
5-HT_6	11	2,000	10	4,100	76	160[e]	63	6,000
5-HT_7	66	3	250	1,800	9.3	15[e]	110	1,100
α_1	4	1.4	54	15	13	57[d]	1.4[d]	4.7
α_2	33	5.1	170	1,000	310		160	1,200
H_1	1.8	19	2.8	8.7	47		470[d]	440
m_1	1.8	2,800	4.7	100	5,100			1,600

[a] Values are geometric means of at least three determinations.
[b] Bovine.
[c] CHO cells.
[d] Rat.
[e] HEK cells.
From Duncan GE, Zorn S, Lieberman JA. Mechanisms of typical and atypical antipsychotic drug action in relation to dopamine and NMDA receptor hypofunction hypotheses of schizophrenia. *Mol Psychiatry* 1999;4:418–428. Lawler CP, Prioleau C, Lewis MM, et al. Interactions of the novel antipsychotic aripiprazole (OPC-14597) with dopamine and serotonin receptor subtypes. *Neuropsychopharmacology* 1999;20:612–627. Kongsamut S, Roehr JE, Cai J, et al. Iloperidone binding to human and rat dopamine and 5-HT receptors. *Eur J Pharmacol* 1996;317:417–423.

A distinguishing feature of clozapine in comparison to conventional antipsychotics is the relatively high affinity of clozapine for the 5-HT_{2A} receptor. Meltzer and associates (47) provided evidence that combined 5-HT_{2A}/D_2 antagonistic actions, with greater relative potency at the 5-HT_{2A} receptor, may be critical to atypicality, in terms of enhanced efficacy and reduced EPS liability. Based on this theoretic model, risperidone was developed to mimic the relative 5-HT_{2A}/D_2 affinities of clozapine, although risperidone has substantially higher affinity for both receptors than clozapine (Table 56.1). The reduced EPS side effects associated with low-dose risperidone treatment (4 to 6 mg per day), even at high levels of D_2 receptor occupancy, may be owing to the 5-HT_{2A} antagonistic properties of the drug (47 ,48). However, at higher doses, risperidone produces EPS, indicating that 5-HT_{2A} receptor antagonism alone cannot completely eliminate EPS associated with high D_2 receptor blockade. The potential role of 5-HT_{2A} receptor antagonism in therapeutic responses to atypical antipsychotic drugs may become more apparent when data from clinical trials are available for the selective 5-HT_{2A} antagonist M-100907. However, the results to date support the hypothesis that some degree of D_2 antagonism is still required to achieve antipsychotic effects. Moreover, at this point it is unclear what clinical effects 5-HT_{2A} antagonism confers, in addition to mitigating the adverse effect of striatal D_2 antagonism, and propensity to cause EPS (21).

Risperidone, like clozapine, has relatively high affinity for α_1- and α_2-adrenergic receptors (Table 56.1), but the potential therapeutic significance of the adrenergic receptor blocking properties of clozapine and risperidone is uncertain. Addition of the α_2-antagonist idazoxan to the regime of patients treated with the typical neuroleptic fluphenazine resulted in improved treatment responses in patients refractory

to treatment with fluphenazine alone (49). However, there has been no subsequent confirmation of the effects of α_2 antagonists as adjuncts to typical neuroleptic treatment, and it has been suggested that α_2 agonists may actually be useful for treating cognitive deficits of the disease (50).

Olanzapine is a closely related in chemical structure to clozapine, and the two drugs have many common receptor binding characteristics. Primary considerations in selection of olanzapine for development were the drug's relatively potent antagonistic effects at both D_2 and 5-HT_{2A}receptors (51 ,52). Olanzapine is more potent at 5-HT_{2A} than D_2 receptors (Table 56.1), similar to clozapine and risperidone. In addition, receptor binding characteristics of olanzapine in regard to other dopaminergic, serotonergic, cholinergic, and adrenergic receptor subtypes are similar to clozapine, but there are also some notable distinctions between the two drugs. For example, clozapine has substantially higher affinity for 5-HT_{1A} and 5-HT_7 receptors in comparison to olanzapine (Table 56.1).

Quetiapine is another drug with greater relative affinity for 5-HT_{2A} than for D_2 receptors, but also some affinity for α_1-adrenergic and H_1 receptors (53) (Table 56.1). Interestingly, quetiapine produces only transiently high striatal D_2 occupancy in schizophrenic patients, although the study has clinical and technical limitations (54). Ziprasidone has potent 5-HT_{2A} and D_2 affinities, and like clozapine, it shows 5-HT_{1A} agonist properties that could potentially act as protective effects on the development of EPS. Ziprasidone also has significant affinity for 5-HT_{1D} and 5-HT_{2C}, as well as H_1 and α_1-adrenergic receptors (55) (Table 56.1). Iloperidone has in addition to affinity for 5-HT_{2A} and $_{1A}$ and $D_{2,3}$ receptors, also affinity for the α_1- and α_{2c}-adrenergic receptors. Aripiprazole is distinct from the other atypical antipsychotic drugs because it is selective for the dopamine system and acts through partial agonism.

PET studies showing that therapeutic doses of risperidone and olanzapine produce greater than 70% occupancy of D_2 receptors suggest that D_2 receptor antagonism could be a predominant mechanism of action of these atypical drugs (56 ,57). Clozapine, however, does not exhibit high levels of D_2 receptor occupancy at therapeutically effective dose (15 ,57 ,58), suggesting that D_2 receptor antagonism alone cannot explain the greater therapeutic efficacy of clozapine (13). The low occupancy of striatal D_2 receptors by clozapine could account for its low EPS liability (20 ,58 ,59).

Clozapine, risperidone, and olanzapine occupy more than 80% of 5-HT_{2A} receptors in the therapeutic dose rage in humans (15 ,56 ,57 and 58 ,60). Although 5-HT_{2A} receptor antagonism is likely to be associated with the low EPS liability of risperidone and olanzapine, the role of this molecular action in the superior therapeutic responses to clozapine is unclear (13).

Efficacy

Although the proportion of patients who improve and the magnitude of therapeutic effects vary greatly, atypical antipsychotics

are at least as effective for psychotic symptoms as conventional drugs (3). Well-controlled double-blind studies of atypical antipsychotics suggest that clozapine, risperidone, and olanzapine may be superior to haloperidol for controlling psychotic symptoms (61). At selected doses, risperidone appears to be more effective than haloperidol in treating positive and negative symptoms (53). Olanzapine has been demonstrated to be effective for positive, negative, and depressive symptoms (62), and in some studies the drug was superior to haloperidol and risperidone in terms of negative symptoms and long-term efficacy (63 ,64). However, in a recent large double-blind study (that has only been preliminarily reported), risperidone demonstrated significantly greater efficacy than olanzapine in reducing anxiety/depression and positive symptoms (65). Quetiapine appears to be comparable to chlorpromazine and haloperidol in treating both positive and secondary negative symptoms (61). Similarly, ziprasidone appears to be as effective as haloperidol in alleviating positive and negative signs in an acute treatment study (66), whereas a 52-week placebo-controlled maintenance study found primary and secondary negative symptom efficacy for ziprasidone (67).

To date, clozapine is the only drug that has proven efficacy in treatment-refractory schizophrenia (68 ,69). The efficacy rates for clozapine in treatment-refractory patients vary from 20% to more than 70% (11 ,70 ,71). In some studies, risperidone does not appear to be as effective as clozapine in treatment-resistant schizophrenic patients (72 ,73 and 74); however, Bondolfi and associates (75) found no difference between risperidone and clozapine in treatment-resistant patients. In this latter study, certain methodologic issues may have led to an overestimation of the efficacy of both clozapine and risperidone, and there are questions as to whether the patient population studied represented “truly resistant” patients (69). Further investigation is necessary to adequately compare the relative efficacy of risperidone and clozapine in treatment-resistant patients. Olanzapine was found to be more effective than haloperidol (74 ,76), but not chlorpromazine (77), in treatment-refractory patients. In a recent randomized double-blind study of treatment-resistant schizophrenia, olanzapine and clozapine had similar antipsychotic efficacy (74). Additional studies are needed to reach definitive conclusions regarding efficacy of the newer atypical antipsychotics in treatment-resistant schizophrenia. Results of studies investigating the effects of atypical antipsychotics in treatment-resistant patients are discussed elsewhere in this chapter.

The efficacy of atypical antipsychotics in treating primary negative symptoms has not been clearly demonstrated (61). Thus, the choice of atypical drugs for patients with predominantly negative symptoms is less clear (8). In addition, the effects of atypical antipsychotics on cognitive impairment have not yet been clearly proved. A metaanalysis of 15 studies (only three of which were double-blind) of atypical antipsychotics and cognitive impairment in patients with schizophrenia suggests that they may improve attention and executive function (37). Available results, however, are relatively inconsistent and modest in effect size. Furthermore, there are statistical limitations and a lack of standard conventions in the studies of cognition (78). It appears that there could be significant differences among the atypical drugs in terms of what types of cognition they improve.

Atypical antipsychotics have been associated with a reduction in the incidence of suicidality, which may be relevant to antidepressant effects of these agents, at least in part (6). Clozapine, risperidone, and olanzapine, in particular, appear to have beneficial effects on the depressive component of schizophrenia (6 ,65) (Table 56.2).

	Clozapine	Risperidone	Olanzapine	Quetiapine	Ziprasidone
Clinical effect					
Psychotic symptoms	+++	+++?	+++?	++	++
Negative symptoms	+	+	+	+	+
Cognitive symptoms	++?	++?	++?	?	?
Mood symptoms	+++	++	+++?	++?	++?
Refractory symptoms	+++	+++?	+++?	++?	++?
Side effect					
EPS	—	++[a]	+[a]	—	+[a]
TD	—	+	?	?	?
Prolactin elevation	—	+++	—	—	—

[a] Dose dependent.
EPS, extrapyramidal side effects; TD, tardive dyskinesia; + to +++, weakly (for clinical effect) or active (for side effect) to strongly active; – to —, weak to little activity; ?, questionable to unknown activity.
From Dawkins K, Lieberman JA, Lebowitz BD, et al. Antipsychotics: past and future. National Institute of Mental Health, Division of Services and Intervention Research Workshop, July 14, 1998. *Schizophr Bull* 1999;25:395–404.

TABLE 56.2. CLINICAL AND SIDE-EFFECT PROFILE OF ATYPICAL ANTIPSYCHOTIC DRUGS

Although atypical drugs have shown some instances of superior efficacy in comparison with conventional drugs, they are not effective in all patients and against all symptom dimensions of psychotic disorders (Table 56.2). It is clear

that the atypical drugs are unable to fully reverse already-established impairment in cognition, negative symptoms, and social disability in many patients (79). Thus, the possible use of these agents in the prodromal period of schizophrenia, before the emergence of psychosis, is an important issue to address in the next decade (79).

Safety

Although atypical antipsychotics were developed to improve on the shortcomings of conventional drugs it has already become apparent that they also have significant limitations in terms of side effects in the relatively brief period that they have been in general clinical use (3). As a class, and with some variation between the individual drugs (Table 56.2), they have a much more favorable side-effect profile, particularly in terms of EPS and TD. They do, however, produce side effects, including sedation, hypotension, dry mouth, constipation, sedation, and some types of sexual dysfunction (3). Neuroleptic malignant syndrome has also been reported with atypical antipsychotics such as clozapine, risperidone, and olanzapine (80). Weight gain is the most worrisome and potentially serious side effect that appears to be class wide, except perhaps for ziprasidone and drugs that have not yet been approved for marketing by the FDA, including aripiprazole and iloperidone (81). In particular, weight gain and sedation are common reasons for drug discontinuation for adolescent patients (78). In addition, the atypical antipsychotics have been associated with new onset type II diabetes mellitus (82). It is unclear whether these effects are secondary to weight gain, independent, or causative. Atypical drugs are also associated with increases in cholesterol and lipids, the long-term medical consequences of which are largely unknown (78). It appears prudent to monitor fasting blood sugar and lipid levels in patients treated with these agents. The new atypical drugs also have their own individual and idiosyncratic side-effect profiles (Table 56.2). Thus, each new drug should be evaluated individually in terms of side effects and safety (39).

Clozapine is associated with a very low propensity for EPS and little or no incidence of TD; thus, it is a valuable option for patients who experience EPS (11). However, clozapine can cause serious side effects that impose substantial limitations on its use. Not only must initial dose titration be quite gradual, but also there is a significant occurrence (around 0.9%) of agranulocytosis (83) and seizures, as well as sedation, hypotension, hypersalivation, and weight gain (8). The frequency of agranulocytosis with clozapine is such that regular white blood count monitoring is required (8).

Risperidone has a favorable side effect profile in comparison to haloperidol (84). Risperidone can produce dose-related EPS (≥6 mg per day), but the rate of TD is low (0.6%) for dose currently used (2 to 8 mg per day) (84 ,85). Risperidone is associated with prolactin elevation, hypotension, somnolence, insomnia, and agitation (39 ,86).

The incidence of EPS with olanzapine is not significantly different from that with placebo, and the incidence of olanzapine-related TD is low (1%) (87). There is a risk of mild sedation and mild anticholinergic side effects, and the risk of weight gain appears greater than with risperidone, but comparable to clozapine (78).

Quetiapine is associated with very low levels of EPS and its prolactin level elevation is indistinguishable from that of placebo (88). The incidence of TD with quetiapine is reportedly low or virtually nonexistent, although this remains to be demonstrated prospectively. There is a potential risk of lenticular opacities that were associated in one preclinical study in beagles (89), but have not been found in nonhuman primates or patients, yet monitoring is recommended until additional data are available. The risk of weight gain with quetiapine appears to be less than that with olanzapine and clozapine (78). Although quetiapine has virtually no cholinergic activity, tachycardia is a possible side effect, perhaps secondary to its adrenergic effects on blood pressure (39). There are several other side effects with quetiapine such as decrease in T3 and T4, orthostatic hypotension, and sedation, necessitating gradual dose titration (39).

Ziprasidone has a risk of EPS that is not significantly different from that with placebo (90). The risk of TD is not known. Ziprasidone is associated with mild dyspepsia, nausea, dizziness, and transient somnolence (90). Ziprasidone treatment has been associated with minimal weight gain, which could distinguish it among other atypical agents (80). The FDA delayed ziprasidone approval because of concern about its ability to prolong the Q-T interval (90), but an FDA Advisory Committee recommended its approval for the treatment of schizophrenia in July 2000, and the FDA issued an approval letter in September 2000.

Effectiveness

Considerable evidence indicates that relapse and rehospitalized rates are substantially better with the group of atypical antipsychotics than with conventional antipsychotics for patients who are compliant with their maintenance antipsychotic regimen (46). The decreased EPS liability of the atypical drugs will make it easier to prescribe more effective doses of antipsychotic that can maximize relapse prevention, without simultaneously interfering with the patient's quality of life or motor functioning (46). Patient-based measures of quality of life show improvement with the atypical drugs over the conventional neuroleptics (45).

In one randomized controlled trial comparing clozapine with standard neuroleptic therapy for treatment-resistant schizophrenic inpatients, the actual hospital discharge rates at 1 year were 27% for clozapine and 29% for standard care (91). The clozapine group, however, had decreased readmission rates within the first 6 months compared with the neuroleptic group (3% versus 29%) (91).

In another randomized double-blind comparative study of clozapine and haloperidol in patients with refractory schizophrenia over 1 year, clozapine-treated patients showed significant quality-of-life improvements when compared with haloperidol-treated patients (53% versus 37%) (92). The patients assigned to clozapine had significantly fewer mean days of hospitalization for psychiatric reasons than patients assigned to haloperidol (144 versus 168 days) and used more outpatient services (134 versus 98 units of service) (92).

Several studies have examined the impact of risperidone on health care utilization in the 2 years before and after risperidone treatment in small groups of schizophrenic patients. Decreases of 20% to 31% in the number of hospitalization days were reported (93 ,94), but Viale and colleagues (95) observed an increase of 12% in hospitalization days in the first year of risperidone therapy.

Extensive controlled studies have proven olanzapine to be significantly superior to haloperidol in long-term maintenance of response (62 ,96). The estimated 1-year risk of relapse was 19.7% with olanzapine and 28% with haloperidol (97). Furthermore, a significantly greater proportion of the olanzapine- than risperidone-treated responders maintained their improvement in the extended follow-up after 28 weeks of therapy (63). It is not clear whether the lower relapse rates are owing to increased prophylactic efficacy or better treatment compliance because of better tolerability. To date, there have been no definitive prospective random-assignment studies on compliance rates for atypical antipsychotics (46).

Cost-Effectiveness in Comparison with Conventional Drugs

Atypical antipsychotic drugs are approximately 10 to 40 times more expensive than conventional drugs (98). In the past few years, a number of studies comparing the cost-effectiveness of the atypical antipsychotics with that of the typical drugs have been published. However, many of these studies have frequently been criticized because of limitations in experimental design; thus, the cost-effectiveness of atypical antipsychotics has not yet been fully established (98 ,99). Most of the available cost-effectiveness evidence is from retrospective studies or economic computer models, which have considerable methodologic limitations (98).

Perhaps the best study of the cost-effectiveness of clozapine published to date in terms of its methodology is a randomized controlled trial conducted by Rosenheck and associates (92), that compared clozapine with haloperidol in patients with treatment-refractory schizophrenia over 1 year. After 1 year of treatment, the clozapine group had lower inpatient but higher outpatient costs. The total medical costs (including inpatient hospital costs, outpatient medical costs, and medication costs) of the clozapine group ($58,000) were not significantly lower than the haloperidol group ($61,000). Overall, clozapine was concluded to be cost neutral, although it demonstrated improved clinical outcomes, suggesting that it may be cost-effective (92).

The higher price of olanzapine compared with classic neuroleptics may be offset by reductions in the use of inpatient and outpatient services (45 ,100). For example, Hamilton and colleagues (100) compared the cost-effectiveness of olanzapine to those of haloperidol for the treatment of schizophrenia, in a randomized clinical trial, for 6 weeks (acute phase) and up to 1 year (maintenance phase). The medication costs for olanzapine were about 22 times larger than those for haloperidol after 6 weeks of treatment; however, patients treated with olanzapine had significantly lower inpatient and outpatient medical expenses than patients treated with haloperidol. Overall, mean total medical costs during the acute phase for the olanzapine patients were significantly lower (US$388/6 weeks) than those for the haloperidol patients. As was seen in the acute phase, these total medical cost differences were sustained (US$636 lower per patient for olanzapine over 46 weeks) during the maintenance phase (100). Glazer and Johnstone (99) also reported that the total health care costs for olanzapine treatment for 6 weeks and up to 1 year were lower than those for haloperidol treatment ($431/month lower and $345/month lower, respectively).

Palmer and associates (101) used a decision analytic model to estimate the total medical costs and effectiveness outcomes of olanzapine, haloperidol, and risperidone over 5 years for schizophrenia treatment in the United States. The estimated 5-year total medical cost of olanzapine, haloperidol, and risperidone was US$92,593, $94,132, and $94,468, respectively. The estimated disability-free years of these agents were 3.19 (olanzapine), 2.62 (haloperidol), and 3.15 (risperidone). The quality-adjusted life years (QALYs) were 3.15 (olanzapine), 2.95 (haloperidol), and 3.12 (risperidone). These data suggest a modest cost-effectiveness advantage for olanzapine over haloperidol and risperidone (101), whereas the decision-modeling approach appears to be subjective to imprecision and possible bias (45). There have been no published randomized controlled studies of the cost-effectiveness of risperidone. In addition, so far, no prospective randomized studies have been completed that compare the cost-effectiveness of the atypical antipsychotics to each other for the treatment of schizophrenia. Furthermore, the other atypical drugs are too new to have had their cost-effectiveness evaluated to any significant extent. Additional prospective randomized clinical trials with larger sample sizes and long-term assessment should be conducted in order to evaluate the cost-effectiveness of atypical antipsychotics adequately (45).

First-Episode Patients

Pharmacotherapy

First-episode patients as a group may differ from chronic patients in several aspects of pharmacologic responsiveness.

Relatively high response rates of positive and negative symptoms have been reported in first-episode samples; for example, Lieberman and colleagues (102) reported remission rates of 83% after 1 year of treatment with conventional antipsychotic agents in 70 first-episode patients. Surprisingly, remission did not occur until a median of 11 and mean of 36 weeks of treatment. Despite the apparent heightened responsiveness of first-episode patients, residual cognitive deficits and poor psychosocial adjustment are common (103 ,104). First-episode patients may also require a lower mean dose of antipsychotic medication and may be more sensitive to drug side effects compared to more chronic patients (105). Kopala and colleagues (106) treated 22 first-episode patients openly with risperidone for a mean of 7 weeks and observed a 91% response rate in patients who received risperidone 2 to 4 mg per day compared to a 27% response rate in patients who received a dose of 5 to 8 mg per day. The lower-dose group exhibited no EPS, whereas 32% of the higher-dose group developed akathisia or parkinsonism. However, because this was not a fixed-dose design, conclusions regarding dose–response relationships must be considered preliminary. In a different approach, Sanger and colleagues (107) analyzed results from the 83 first-episode patients (out of a total of 1,996 subjects) who participated in a double-blind, 6-week comparison of olanzapine and haloperidol. First-episode patients who received olanzapine had significantly better clinical response and fewer EPS than the haloperidol group. Of particular interest, first-episode patients treated with olanzapine achieved a significantly higher response rate than chronic patients treated with olanzapine. In addition, chronic patients treated with haloperidol developed significantly fewer EPS than first-episode patients treated with haloperidol. Mean doses of haloperidol and olanzapine were similar between first-episode and chronic patient groups (10.8 versus 11.0 mg per day and 11.6 versus 12.0 mg per day, respectively). Although these findings suggest that the relative benefits of olanzapine (and perhaps of other atypical agents) compared to conventionals may be greater in first-episode patients than chronic patients, issues of nonequivalent dosing between drugs may be of particular concern in light of recent work indicating that optimal D_2 receptor blockade may be achieved in first-episode patients with haloperidol 0.25 to 2 mg per day (18). Two other double-blind controlled studies have been preliminarily reported that address the question of whether first-episode patients respond better to atypical antipsychotic drugs. The first is a 52-week study of clozapine versus chlorpromazine in 164 first-episode treatment naive schizophrenia patients in China (108). The cumulative response rates of patients at 12 and 52 weeks, respectively, were 81.2% and 96.3% for clozapine (mean dose 292 mg per day), and 68.3% and 97.7% for chlorpromazine (mean dose 319 mg per day). The first-episode patients treated with clozapine had more rapid response, fewer EPS, and higher treatment retention and relapse prevention than the chlorpromazine group (108). The second is a comparison between olanzapine and haloperidol in 262 patients with first-episode psychotic disorder (109). At 12 weeks, the patients treated with olanzapine (mean dose 9.1 mg per day) demonstrated a higher response rate (55% versus 46%) and greater cognitive improvement than the patients treated with haloperidol (mean dose 4.4 mg per day).

Response of first-episode patients has also received renewed attention because of the widely held belief that early intervention may favorably affect the course of the illness. This hypothesis, which often invokes “neurotoxicity of untreated psychosis” as a mechanism, is largely based on one naturalistic study reported by Loebel and colleagues (110). Other naturalistic studies have failed to find a relationship between duration of initial untreated illness and outcome (111 ,112 and 113). Prospective controlled trials are needed to determine whether early intervention with specific antipsychotic agents improves the early course of the illness.

Psychosocial Interventions

Psychosocial interventions potentially may have the greatest impact on first-episode patients and their families. Preliminary studies have looked at stress-reduction approaches for patients identified as “premorbid” or at risk for schizophrenia, combining cognitive therapy or stress reduction interventions alone or in combination with medication (114 ,115 and 116). Preliminary studies have indicated that cognitive-behavioral therapy (CBT) approaches that have been developed for patients with treatment-resistant psychosis can be successfully modified for first-episode patients (117). Psychoeducation, family support, and interventions to enhance compliance are also expected to play important roles early in the course of the illness. However, two studies of first-episode patients in Norway failed to find benefit from the addition of behavioral family management (BFM), which emphasizes communications skills, to a basic psychoeducation program (118 ,119). The authors concluded that families of first-episode patients may be in greatest need of information and support, rather than the intensive communication skills training offered by BFM.

Maintenance Treatment

Pharmacotherapy

Maintenance treatment with conventional and atypical antipsychotic medications has consistently demonstrated prophylactic efficacy against relapse. Hogarty (120) reviewed the literature on maintenance treatment with conventional antipsychotic agents and found that the average relapse rate during the first year after hospitalization was 41% with active medication compared to 68% with placebo. Among patients who survived the first year, annual relapse rates with medication dropped to 15%, whereas relapse rates on

placebo remained constant at 65%. This pattern suggests that maintenance treatment is relatively ineffective for a substantial proportion of patients; only after this poorly responsive subgroup is removed from the sample does the benefit of medication become fully apparent. Consistent with this view are the results of a low-dose maintenance treatment trial with depot fluphenazine in which a dose–response relationship only emerged during the second year of follow-up (121 ,122). Depot preparations have significantly lowered relapse rates by an average of 15% compared to oral neuroleptics in six double-blind, randomized trials (123). The advantage of depot administration may be understated in these trials, however, because research subjects were probably poorly representative of typical clinical samples and most trials did not extend beyond 1 year. Research comparing low and standard-dose maintenance with depot neuroleptics has demonstrated a trade-off between adverse effects with higher doses, including neurologic side effects and dysphoria, versus increased relapse rates with lower doses (122 ,124). “Intermittent” maintenance treatment was associated with an unacceptable rate of hospitalizations, whereas relapses associated with low-dose depot medication generally were responsive to rescue with brief augmentation with oral neuroleptic or benzodiazepine; hospitalization rates were not elevated with low compared to standard doses (122 ,124). Carpenter and colleagues (125) reported that administration of diazepam at the earliest sign of exacerbation in medication-free patients was more effective than placebo and comparable to fluphenazine in preventing relapse. This work suggests that lower doses of depot neuroleptic may provide acceptable protection against relapse if accompanied by close monitoring and rapid psychosocial and pharmacologic intervention at the first sign of relapse. These measures presumably will also enhance maintenance treatment with atypical agents, although dose-limiting side effects are not as problematic.

Growing evidence suggests that maintenance treatment with atypical agents provides greater protection against relapse compared to conventional oral agents. In a large, open trial, Essock and colleagues (126) found that chronically hospitalized patients randomized to clozapine were not more likely to be discharged than patients receiving treatment as usual, but once discharged, relapse rates were significantly lower with clozapine. Pooled results from three double-blind extension studies revealed that relapse rates were significantly lower with olanzapine (20%) compared to haloperidol (28%) in patients with schizophrenia and related psychoses (97). Until depot preparations of atypical agents are available for study, it will be difficult to determine whether the advantage of certain atypical agents is primarily the result of enhanced compliance versus a direct modulatory effect on symptom exacerbation. It is clear from depot neuroleptic studies that large numbers of patients relapse despite adequate compliance; relapse in medication-compliant patients is often associated with depression and resolves spontaneously without change in medication (127). Whether all atypical agents are equally effective in preventing relapse is also unknown. In a naturalistic study, Conley and colleagues (128) found that relapse rates were quite similar during the first year after discharge in patients treated with clozapine versus risperidone. During the second year, no additional relapses occurred on clozapine, whereas the rate of relapse on risperidone increased from roughly 13% to 34%. In the only published comparison between risperidone and olanzapine, rates of exacerbation (increase in PANSS score by 20%) were significantly higher at 28 weeks in patients who had responded to risperidone (mean dose 7 mg per day) compared to olanzapine (mean dose 17 mg per day) (63). It will be important to determine whether specific drugs differ in prophylactic efficacy against relapse when compliance is controlled and issues of dosing equivalence are addressed. It is possible that clozapine and perhaps other atypical agents are more effective in suppressing relapse; this effect may be relatively independent of antipsychotic efficacy and mediated by different neurotransmitter systems. Continued development of psychosocial interventions to improve compliance and monitor and respond to early signs of relapse will be equally important.

Psychosocial Interventions

A diverse range of psychosocial interventions has been shown to reduce relapse rates. In over 20 controlled trials, family therapies emphasizing psychoeducation and support have reduced relapse rates for schizophrenia patients who have regular contact with family members (129 ,130). Although differences in theoretical orientations and intensity of treatment have not produced consistent differences in efficacy, recent evidence has suggested that multiple-family psychoeducation groups may be particularly effective (131). Several controlled trials have also indicated that relapse rates can be reduced by assertive community treatment programs (PACT) or similar outreach programs that provide intensive monitoring, skills training, and case management in the community, usually with continuous availability of staff (132 ,133). Social skills training improves role functioning of patients with schizophrenia, but has not substantially reduced symptoms or reduced relapse rates compared to control conditions in most studies (134). In an illuminating study, Herz and colleagues (135) found that a relatively simple, weekly monitoring of schizophrenia patients in psychoeducation groups in conjunction with the availability of rapid pharmacologic and psychosocial interventions at the first sign of decompensation substantially reduced relapse rates, by approximately fourfold, compared to treatment as usual.

Noncompliance

Pharmacotherapy

Cramer and Rosenheck (136) surveyed the literature on antipsychotic medication and found that compliance rates

averaged 42%. Similar surveys have not been conducted looking specifically at atypical agents, although it is generally believed that reduced relapse rates reported with olanzapine and clozapine may reflect, in part, improved compliance (97 ,126). Factors contributing to noncompliance are complex and probably involve the patient's perception of benefits and side effects of medication, as well as the patient's level of insight. Compliance can be compromised by psychosis, agitation, and comorbid substance abuse (137 ,138). Van Putten (139) studied compliance in 85 schizophrenia patients chronically treated with conventional neuroleptics and determined that 46% took less antipsychotic medication than prescribed. Medication refusal was associated with an early dysphoric response, which Van Putten attributed to subtle akathisia. Analysis of responses by 150 schizophrenia patients to a “Drug Attitude Inventory” revealed that, based on responses to 10 items, 89% of patients could be correctly assigned to compliant versus noncompliant categories as determined by clinician assessment of compliance (140). The strongest predictor of compliance was a positive experience with medication—this factor accounted for 60% of the total variance, whereas the factor representing a negative subjective experience accounted for 12%. Factors representing attitudes and beliefs about medication had minimal predictive power. Other studies have also found that a patient's perception of benefit from medication is the strongest predictor of compliance (141). Whereas many clinicians expect atypical agents to achieve higher levels of compliance by virtue of reduced or absent EPS, this view may seriously underestimate the impact of other side effects. Two studies have found that clinicians tend to misjudge the relative distress produced by different medication side effects (142 ,143). Side effects associated with certain atypical agents, such as sedation, patients rated weight gain, drooling, and sexual dysfunction as more distressing than EPS in these surveys (142 ,143 and 144). The advantage of atypical agents in terms of compliance may stem less from their reduced EPS and more from their improved efficacy for symptoms of anxiety, depression, and tension. Whether targeting cognitive deficits and impairment in insight will improve compliance remains to be seen.

Psychosocial Interventions

Most approaches to noncompliance involve psychoeducation, supervision, and supportive therapy in which the benefits of treatment are emphasized, whereas barriers to adherence and medication side effects are minimized (145). Family therapy and social skills training may also exert a positive impact on compliance. Cognitive behavioral approaches have recently been applied to noncompliance by Kemp and colleagues (146 ,147), who developed “compliance therapy,” a four- to six-session intervention based on motivational interviewing techniques that targets attitudes towards medication and discharge planning during acute hospitalizations. In a randomized, controlled trial, compliance therapy was found to improve insight and observer-rated adherence to treatment over an 18-month treatment period (147). Patients in the compliance therapy group also displayed significantly greater improvement in social functioning and lower relapse rates than the control group (147). In addition to educational and skills training approaches, Cramer and Rosenheck (148) demonstrated that interventions that assist patients in remembering to take medications, such as placing microchip schedulers on pill bottles, can also substantially improve compliance.

Treatment Resistance

Estimates of the incidence of treatment resistance have varied with changes in the diagnostic classification of schizophrenia and definitions of treatment response (149), which have tended to obscure potential improvements in outcome associated with advances in pharmacologic and psychosocial treatments. For example, Hegarty and colleagues (150) reviewed results of 320 clinical trials and found that, since the introduction of modern antipsychotics in the mid-twentieth century, about 50% of patients were improved at follow-up, whereas the rate of improvement dropped to 35% in the decade ending in 1994. A narrowing of the diagnostic criteria is believed to account for this decline in response rates. Rates of response have tended to be higher in first-episode psychosis, although dropout rates have been high in this population, particularly with conventional agents (102 ,107). Persistence of psychotic symptoms is more common in drug trials involving chronic patients, presumably reflecting progression of the illness as well as a possible selection bias favoring participation by more refractory patients. If the definition of treatment resistance is broadened to include persistence of negative symptoms, cognitive deficits, or failure to achieve premorbid levels of functioning, treatment resistance can be considered the rule rather than the exception.

Psychotic Symptoms

Antipsychotic Monotherapy

Response of psychotic symptoms to conventional antipsychotics, risperidone, and olanzapine has been associated with D_2 receptor occupancy in excess of 65% (18 ,57), although persistence of psychotic symptoms has been shown to occur despite adequate D_2 blockade in a subgroup of refractory patients (151). As noted, only clozapine has consistently demonstrated efficacy for psychotic symptoms in treatment of refractory patients; the mechanism responsible for this therapeutic advantage remains uncertain. In a sample of 268 patients prospectively established to be neuroleptic resistant, 30% in the clozapine group met criteria for response at 6 weeks compared to 7% treated with chlorpromazine (11). Response rates as high as 60% have been reported

after 6 months in open trials with clozapine in patients less rigorously defined as treatment refractory (152). The extent to which a prolonged trial is necessary to determine efficacy of clozapine and other atypical agents is the subject of debate (153 ,154).

The relative efficacy of atypical agents other than clozapine in patients who have failed conventional neuroleptic therapy is less clear. Marder and colleagues (155) found that schizophrenia patients presumed to be treatment-resistant on the basis of having been hospitalized for 6 months or longer at the time of study entry did not respond to haloperidol 20 mg per day but significantly improved with risperidone 6 mg per day or 16 mg per day compared to placebo. Similarly, analysis of a subgroup of 526 patients from a larger trial identified retrospectively as having had a poor response to at least one prior antipsychotic, revealed greater response of psychotic symptoms to olanzapine (mean dose 11 mg per day) than haloperidol (mean dose 10 mg per day); this difference was significant in the intent-to-treat analysis but not in a comparison of completers (76). Trials specifically designed to study treatment-resistant patients have provided less consistent support for efficacy of risperidone and olanzapine. In 67 schizophrenia patients with histories of neuroleptic resistance, risperidone 6 mg per day significantly improved total BPRS scores compared to haloperidol 15 mg per day at 4 weeks, but response did not differ between groups at 8 weeks (156). In contrast, risperidone produced significantly higher response rates than haloperidol in a large, randomized open trial involving 184 schizophrenia patients with a history of poor response (157). Relative response of psychotic symptoms to risperidone increased over time and reached a maximum improvement compared to haloperidol at the final 12-month assessment. In a 6-week trial designed to mirror the landmark Clozapine Collaborative Trial (11), only 7% of patients prospectively determined to be treatment resistant to haloperidol responded to olanzapine 25 mg per day, a response rate that did not differ from chlorpromazine (77). The same group reported that 41% of 44 patients identified as unresponsive to olanzapine in the preceding study or in an open trial subsequently exhibited a response to clozapine (158). In addition, open trials in which patients have been switched from clozapine to olanzapine or risperidone have reported a high incidence of clinical deterioration, casting doubt on claims for therapeutic equivalence between clozapine and the second-generation agents, at least at the doses tested (159 ,160). Of interest, two controlled trials have found comparable efficacy for risperidone and clozapine. However, in one 4-week trial, the 59 participants were not screened for treatment resistance at baseline and, despite equivalence in outcomes between groups using an LOCF analysis, 25% of the risperidone group dropped out owing to lack of efficacy compared to only 5% in the clozapine group (161).

The evidence is strongest in support of clozapine monotherapy as an intervention for neuroleptic-resistant patients; serum levels of 350 ng/mL or greater have been associated with maximal likelihood of response (162). Given the risk of agranulocytosis, the burden of side effects, and the requirement of white blood cell monitoring, the second-generation agents (risperidone, olanzapine, and quetiapine) are commonly tried before proceeding to clozapine. The appropriate first choice among these agents is unclear; two controlled studies that compared olanzapine and risperidone have produced divergent results, probably reflecting differences in dosing of the two agents and the use of intent-to-treat versus completer analyses (63 ,163). The focus of this research has been on comparisons of mean responses between groups; predictors of response have not been identified, nor have subgroups of patients that may exhibit preferential response to one agent of the class. Many clinicians express the impression that certain patients do respond preferentially to a single agent of this class. Sequential controlled trials of the newer agents in treatment-resistant patients will be necessary to fully examine this issue.

Combinations of Antipsychotics

The practice of combination therapy is gaining widespread popularity in the absence of controlled data in its support (164). In part based on empirical experience and the demonstration that clozapine at optimal doses achieves relatively low degrees of D_2 occupancy, European clinicians commonly add low-doses of neuroleptics to clozapine in partially responsive patients (165). Uncontrolled trials and case reports have described benefits associated with the addition of risperidone (4 mg per day) (159 ,166) and pimozide (167) to clozapine in partially responsive patients. In a small, placebo-controlled trial, addition of sulpiride 600 mg per day to clozapine significantly improved positive and negative symptoms at the end of 10 weeks in 28 subjects (168). Other combinations, most notably olanzapine plus risperidone, are also increasingly employed, often because clinicians perceive improved response during the cross-tapering phase of switching from one to the other. A theoretical rationale for this combination is less apparent, given that each agent produces maximal D_2 and 5-HT_2 occupancy when appropriately dosed (57). If combined treatment with olanzapine and risperidone is found in suitably controlled study designs to offer advantages over optimal monotherapy with either agent, such a finding would argue in favor of the existence of additional contributory receptor actions unique to each drug.

Adjunctive Treatments

A diverse range of adjunctive treatments has been proposed for antipsychotic-resistant schizophrenia, although therapeutic effects generally have been small or inconsistent in controlled trials. Very little data are available from controlled trials augmenting clozapine in partial responders (169). Lithium augmentation frequently has been cited as

the best-established intervention based on positive results from three small studies (170 ,171 and 172); however, two recent placebo-controlled studies found no benefit when well-characterized neuroleptic-resistant patients were treated with lithium (approximately 1.0 mEq/L) added to haloperidol or fluphenazine decanoate (173 ,174). Augmentation with lithium may enhance response of some patients, particularly in the presence of affective symptoms or excitement (175 ,176). Carbamazepine augmentation of conventional neuroleptics has been associated with modest reductions in persistent symptoms, including tension and paranoia, in several controlled trials (177 ,178 and 179), particularly in patients with abnormal EEGs or violence. However, induction of hepatic microsomal enzymes by carbamazepine can substantially lower blood levels of certain antipsychotic agents (180) and in one report, resulted in clinical deterioration (181). Valproate does not significantly affect serum concentrations of most antipsychotic drugs, but results from two small controlled augmentation trials have been inconsistent. Wassef and colleagues (182) reported efficacy for negative symptoms and global psychopathology associated with addition of divalproex to haloperidol in a placebo-controlled 12-week trial in 12 schizophrenia patients hospitalized for acute exacerbation. In contrast, Ko and colleagues (183) found no effect when valproic acid was added to conventional neuroleptics in six treatment-resistant patients in a placebo-controlled crossover design. Augmentation with benzodiazepines also has been advocated, in part, because of the potential role of GABAergic agents in modulating dopamine transmission, although the evidence for efficacy is not compelling (184). Short-term, acute treatment with high-dose benzodiazepines may reduce agitation and psychotic symptoms in as many as 50% of patients (185 ,186), but early reports of benefit of longer-term treatment with benzodiazepines have not been replicated consistently by controlled trials (186 ,187).

Electroconvulsant Therapy and Transcranial Magnetic Stimulation

The most consistent evidence for efficacy in neuroleptic-resistant patients can be found in the literature describing electroconvulsant therapy (ECT) (188). Response rates between 50% and 80% were observed when ECT or the convulsant, Metrozole, were administered unblinded in neuroleptic-naive patients prior to the introduction of antipsychotic medication (189 ,190 and 191). Three double-blind randomized trials comparing neuroleptic plus ECT versus neuroleptic plus sham-ECT have demonstrated a significantly greater and more rapid reduction in psychotic symptoms (delusions) with the combination treatment during 2- to 4-week trials (192 ,193 and 194). Benefits of ECT were lost, however, at follow-up 10 to 28 weeks after treatment. Predictors of a positive response to ECT include acute onset and brief duration of illness (188 ,195 ,196 ,197 and 198). Mood symptoms in schizophrenia patients have tended to be relatively unresponsive to ECT and a diagnosis of schizoaffective disorder did not predict a favorable response (192 ,193 ,194 and 195). Cases describing the successful combination of ECT with clozapine in refractory patients have also been reported, suggesting that augmentation of atypical agents with ECT warrants further investigation (199 ,200). Recently, interest has focused on the potential use of transcranial magnetic stimulation (TMS) as an alternative to ECT in schizophrenia. TMS has shown promising efficacy in depression (201 ,202 and 203). In a preliminary, sham-TMS controlled crossover study in 12 medication-resistant schizophrenia patients, the frequency and severity of auditory hallucinations were significantly reduced following 12 to 16 minutes of stimulation (204). Improvement of auditory hallucinations persisted for a mean of 14 days (range 1 to 60 days). This is an intriguing area for future research, both as a tool to explore the neural circuits underlying symptoms of schizophrenia as well as a potential treatment option in medication-resistant cases.

Psychosocial Interventions

A particularly promising psychosocial approach to medication-resistant psychotic symptoms is cognitive-behavioral therapy (CBT) (205). CBT for psychosis generally consists of alliance formation, examination, and challenge of psychotic beliefs, and the teaching of self-monitoring and coping skills. Four randomized trials, all performed in the United Kingdom, demonstrated superior efficacy for CBT compared to active control treatments on measures of global psychopathology and positive symptoms among chronic, medicated patients (206 ,207 ,208 and 209). A recent metaanalysis determined that the between-groups effect size was .65, favoring CBT over comparison treatments for the response of psychotic symptoms; delusions were generally more responsive than hallucinations (210). Improvements in ratings of psychotic symptoms have been found to persist at follow-up, 1 year after completion of CBT (209). Although therapeutic effects have been impressive, only about half of subjects have displayed improvement in controlled trials (205). Preliminary evidence suggests that patients who exhibit a capacity to entertain alternative explanations for psychotic beliefs at baseline are more likely to respond to CBT (205).

Negative Symptoms

Antipsychotic Monotherapy

Although atypical antipsychotics have generally demonstrated superior efficacy for negative symptoms compared to high-potency conventional agents, the degree of improvement is usually quite modest, leaving substantial levels of residual negative symptoms. For example, across several studies, the effect size of risperidone 6 mg per day compared to placebo on negative symptoms was small (.27) (211). Path analysis has suggested that both risperidone and olanzapine exert direct effects on negative symptoms independent of differences in psychotic, depressive, or extrapyramidal

symptoms (212 ,213). Recently, Volavka and colleagues (74) preliminarily reported a prospective double-blind randomized study, comparing the effects of clozapine, olanzapine, risperidone, and haloperidol, for 14 weeks in 157 treatment-resistant inpatients. Clozapine (mean dose 527 mg per day) and olanzapine (mean dose 30 mg per day), but not risperidone (mean dose 12 mg per day), demonstrated significantly greater efficacy than haloperidol (mean dose 26 mg per day) in reducing negative symptoms (74). However, it is debated whether clozapine's established efficacy for negative symptoms extends to the treatment of primary negative symptoms of the deficit syndrome (153 ,154 ,214). Few data are available from controlled trials to guide treatment of negative symptoms that persist despite optimal treatment with atypical agents (215). Clinicians commonly employ augmentation strategies, but evidence supporting this practice is derived mostly from an older literature describing combinations of augmenting agents added to conventional agents.

Adjunctive Agents

Following clozapine's example as an antagonist of D_2 and 5-HT_2 receptors, investigators combined haloperidol with ritanserin, a relatively selective 5-HT_{2A} and 5-HT_{1C} antagonist (216). In a 6-week, placebo-controlled trial, addition of ritanserin to haloperidol produced significant reductions in negative symptoms (primarily affective expression and social withdrawal) and depressed mood. Addition of 5-HT_2 blockade may improve negative symptoms by enhancing mesocortical dopamine release. Svensson and colleagues demonstrated that 5-HT_2 blockade increases firing of midbrain dopamine neurons and reverses the effects of *N*-methyl-D-aspartate (NMDA) antagonism (217) and hypofrontality (218) on A10 dopamine neuronal firing. Because the available atypical agents achieve maximal occupation of 5-HT_2 receptors at usual therapeutic doses (57), it is unlikely that augmentation with 5-HT_2 antagonists (e.g., nefazodone) will further improve response of negative symptoms.

Another serotonergic augmentation strategy has involved addition of selective serotonin reuptake inhibitors (SSRIs) to conventional neuroleptics, based largely on early empirical observation (219). Fluoxetine and fluvoxamine significantly improved negative symptoms when added to conventional neuroleptics in three of four controlled trials, producing generally modest effects (220). In one study, fluoxetine 20 mg per day added to depot neuroleptics decreased ratings of negative symptoms by 23% compared to a 12% reduction with placebo; this improvement occurred despite a mean 20% elevation in haloperidol serum concentrations and a 65% increase in fluphenazine levels (221). However, addition of sertraline 50 mg per day to haloperidol produced no symptomatic change in an 8-week, placebo-controlled trial in 36 chronic inpatients with schizophrenia (222). In the only reported controlled trial of SSRI augmentation of an atypical agent, fluoxetine at a mean dose of 49 mg per day produced no improvement in negative symptoms when added to clozapine in 33 patients (223).

Anticholinergic agents are commonly added to conventional antipsychotics for control of EPS (224). The atypical agents vary substantially in their muscarinic anticholinergic activity; clozapine is strongly anticholinergic, whereas quetiapine and risperidone exhibit very low affinity for muscarinic receptors (Table 56.1). Addition of anticholinergic agents to conventional agents was associated with reductions in negative symptoms in one study (225) but not others (176 ,226 ,227 and 228). Whether primary negative symptoms are improved by anticholinergics, as suggested by Tandon and colleagues (229), cannot be answered by studies in which subjects are treated with conventional agents; by attenuating psychomotor side effects of the neuroleptic, the anticholinergic may be improving secondary negative symptoms only. To address this issue, two small placebo-controlled trials have administered anticholinergic agents to medication-free patients. Negative symptoms were improved by biperiden in one study (230) and were unchanged with trihexyphenidyl in the other (231). Although the efficacy of augmentation with muscarinic anticholinergic agents for negative symptoms remains poorly established, the potential cognitive impairment that these agents can produce is well described (232 ,233).

Dopamine agonists have also been studied as augmenting agents for negative symptoms. Three of four placebo-controlled trials demonstrated improvement of negative symptoms following a single dose of amphetamine given orally or intravenously (234 ,235 ,236 and 237); in one study efficacy for negative symptoms was not affected by coadministration with pimozide (236). However, Casey and colleagues (238) found no clinical benefit in an extended, 20-week placebo-controlled trial of amphetamine augmentation of chlorpromazine. Augmentation trials of psychostimulants added to atypical agents have not been reported.

As discussed elsewhere in this chapter, augmentation strategies for negative symptoms have recently targeted glutamatergic receptors, in part based on the NMDA antagonist model for schizophrenia and the observation that clozapine differs from conventional agents in its effects on NMDA receptor activity (239). Significant improvements in negative symptoms consistently have been produced in placebo-controlled trials by the addition to conventional antipsychotics of agonists at the glycine site of the NMDA receptor. D-cycloserine, a partial agonist at the glycine site, produced a selective, 23% mean improvement of negative symptoms at 6 weeks that, compared to placebo (7% reduction), represented a large effect size (.80) (240). The full agonist, glycine, at a dose of 60 g per day produced a 30% mean reduction in negative symptoms and also improved a qualitative measure of cognitive functioning (241). Augmentation with another endogenous full agonist, D-serine 30 mg per kg per day, was associated with significant improvements

in negative, positive, and cognitive symptoms when added to conventional agents and to risperidone in an 8-week trial (242). Consistent with evidence that clozapine differs from conventional agents in its effects on NMDA receptor responsiveness, glycine, D-cycloserine, and D-serine did not improve negative symptoms when added to clozapine (242 ,243 ,244 and 245). Whether strategies that enhance NMDA receptor activation will improve response to other atypical agents remains uncertain, although both olanzapine and quetiapine resemble clozapine in certain models of NMDA receptor responsivity.

Psychosocial Treatments

Existing psychosocial approaches have not achieved notable success in the treatment of negative symptoms. Negative symptoms are substantially less responsive to CBT than are psychotic symptoms and patients with prominent negative symptoms are generally poor candidates for CBT (205). Similarly, in a pilot study, Kopelowicz and colleagues (246) found that patients meeting criteria for the deficit syndrome were relatively less likely to benefit from a program of psychoeducation and social skills training than patients without prominent negative symptoms. The presence of negative symptoms also predicts poor outcome in vocational rehabilitation programs for patients with schizophrenia (247). Although most forms of outreach and involvement of deficit syndrome patients in psychosocial programs may improve their quality of life by reducing social isolation and countering apathy, negative symptoms constitute a serious obstacle to participation in such programs and are unlikely to improve with psychosocial treatment.

Mood Symptoms

Antipsychotic Monotherapy

Depressive symptoms are common during all stages of schizophrenia and are associated with poor outcome, including relapse and suicide (248 ,249 and 250). It is not uncommon for patients to present initially with depression during the prodromal stage, prior to the appearance of psychotic symptoms (251). Approximately 25% of first-episode patients exhibit depression, although estimates of the incidence of comorbid depression vary widely according to choice of diagnostic criteria (251 ,252 and 253). The prevalence of depression as defined by moderate scores on depression rating scales ranges between 25% and 50% in chronic patients (252 ,254). Although considerable overlap exists between symptoms of depression and certain negative symptoms (e.g., anhedonia, poor concentration, psychomotor retardation), dysphoria appears to discriminate between the two (255 ,256).

Conventional antipsychotics tend to have little effect on comorbid depression, although anxiety and depression associated with acute psychotic exacerbation frequently respond to neuroleptic monotherapy (257 ,258). However, dysphoric reactions to high-potency conventional agents, although generally not meeting criteria for major depression, can closely resemble the depressive symptoms often associated with the illness (254 ,259 ,260). Clozapine, olanzapine, and risperidone have all demonstrated significantly greater efficacy for depressive symptoms compared to conventional neuroleptics in large, double-blind trials (64 ,211 ,261). Path analysis suggested that 57% of the superior response of depressive symptoms to olanzapine compared to haloperidol was a direct effect, whereas effects on negative symptoms accounted for only 21% and reductions in EPS accounted for 13% of the difference in depressive symptom response (64). Antidepressant activity of the atypical agents may have important clinical consequences because perceived improvement in anxiety and depression is a strong predictor of compliance and emergence of depressive symptoms often accompanies relapse.

Adjunctive Agents

In a placebo-controlled trial reported in 1989, Kramer (258) found that addition of desipramine or amitriptyline 5 weeks after initiating haloperidol to acutely decompensated patients with schizophrenia and depression was associated with poorer antipsychotic response and did not improve depressive symptoms. Subsequently, Siris and colleagues (262 ,263) demonstrated that imipramine added to conventional agents in stable outpatients significantly improved depression without adversely affecting psychotic symptoms. In a carefully controlled trial, imipramine 200 mg per day was associated with substantial improvement in depressive symptoms in 42% of patients compared to 12% with placebo. Hogarty and colleagues (176) found that desipramine improved symptoms of depression, anxiety, and psychosis when added to fluphenazine decanoate in a placebo-controlled trial. Benefits of desipramine were only significant in female patients and did not achieve significance until week 12. The investigators noted that improvement of psychotic symptoms might have resulted from successful prophylaxis against depressive episodes, which were associated with worsening of psychosis. Several trials of tricyclic antidepressants added to conventional agents have been reported; this literature generally supports their use for acute and maintenance treatment of depressive symptoms in stable patients (264 ,265). Augmentation with selective serotonin reuptake inhibitors has been studied primarily as a treatment for negative symptoms—use of these agents in schizophrenia patients with depression is not well studied. Similarly, addition of antidepressants to atypical agents has not been reported in schizophrenia patients with comorbid depression.

Cognitive Symptoms

Antipsychotic Monotherapy

A wide range of cognitive deficits are usually present at the time of the first psychotic episode (266) and remain stable

or only slowly progressive during the course of the illness, independent of psychotic symptoms (267 ,268 and 269). Cognitive deficits are particularly prominent in patients meeting criteria for the deficit syndrome (270) and in patients with tardive dyskinesia (271). The latter association may indicate that cognitive deficits are a risk factor for tardive dyskinesia, or alternatively, that the neurotoxic mechanism responsible for irreversible motoric deficits also compromises cognitive functioning. Targeting cognitive impairments is now a major focus of drug development because cognitive deficits are powerful determinants of vocational and social functioning and may influence quality of life (36) more than psychotic symptoms.

The conventional neuroleptics produce small and inconsistent effects on cognitive functioning; sustained attention improved in some studies, whereas motor control (finger tapping) worsened and memory and executive functioning were minimally affected (272). Recent evidence in monkeys indicates that chronic neuroleptic exposure results in decreased prefrontal cortical D_1 receptor density after 6 months (273); treatment with a D_1 agonist reversed neuroleptic-associated deficits in working memory (274). In normal subjects, clozapine administered as a single 50-mg dose worsened attention, concentration, and motor functioning (275), presumably reflecting sedative and anticholinergic properties. Studies in patients with schizophrenia have found either no effect following a switch to clozapine (276), or improvements in a wide range of cognitive functions, including verbal fluency, attention, and reaction time (37 ,277). In general, clozapine, olanzapine, and risperidone have demonstrated superior efficacy compared to conventional agents on tests of verbal fluency, digit-symbol substitution, fine motor function, and executive function (37 ,277). Atypical agents least affected measures of learning and memory (37). Enhanced performance with atypical agents could result, in part, from reduced parkinsonian side effects because these tests all measure performance during a timed trial (37). Methodologic issues limit comparisons between atypical agents, however, preliminary evidence suggests that risperidone may be more effective for visual and working memory than clozapine (277). In a 12-month, double-blind trial involving 55 schizophrenia patients randomly assigned to olanzapine (mean dose 11 mg per day), risperidone (mean dose 6 mg per day), or haloperidol (mean dose 10 mg per day), risperidone and olanzapine produced significantly greater improvement in verbal fluency compared to haloperidol, and olanzapine was superior to both haloperidol and risperidone in effects on motor skills, nonverbal fluency, and immediate recall (278). However, this finding is complicated by the high incidence of anticholinergic administration prior to the final cognitive assessment; anticholinergics were prescribed to 73% in the haloperidol group, 45% in the risperidone group, and 15% in the olanzapine group. As in efficacy studies for negative symptoms, dose equivalency is an important factor in trials comparing cognitive effects of atypical agents, particularly because excessive dosing can impair performance on time-sensitive tasks and increase anticholinergic exposure.

Adjunctive Agents

Augmentation with glutamatergic agents has shown promise for cognitive deficits in schizophrenia (279). As noted, glycine and D-serine improved ratings of cognitive functioning when added to conventional neuroleptics (241 ,280). Both agents improved the “cognitive subscale” of the PANSS compared to placebo, and D-serine was also associated with improved performance on the Wisconsin Card Sort. These findings are of interest given that NMDA antagonists produce in normal subjects deficits in attention and memory similar to those found in schizophrenia (281 ,282). The partial agonist, D-cycloserine, did not improve cognitive functioning when added to conventional agents in a study that utilized formal cognitive testing, however (240). Positive modulators of the glutamatergic AMPA receptor are also under investigation, as these agents improve performance in tests of learning and memory in animal studies (283). In a preliminary 4-week, placebo-controlled trial involving 19 schizophrenia patients, CX-516, a positive modulator of the glutamatergic AMPA receptor, improved performance on tests of memory and attention when added to clozapine (284). Effect sizes favoring CX-516 over placebo were moderate to large (.5 to 1.2) on tests of cognitive performance.

Psychosocial Treatments

Although cognitive remediation treatments have long been used for brain-injured individuals, similar treatment approaches targeting cognitive deficits in schizophrenia are relatively recent. In small studies in which schizophrenia patients practiced graduated cognitive exercises, performance on laboratory measures of attention and memory function improved, although the functional benefits of these gains are not clear (285 ,286). Brenner and colleagues (287) developed integrated psychological therapy (IPT), a cognitive remediation program in which cognitive exercises are provided in a group format stressing the integration of cognitive skills with social functioning. In a 6-month randomized trial in which patients received IPT or supportive treatment in addition to comprehensive psychiatric rehabilitation, the IPT group displayed greater improvement on the primary outcome measure of interpersonal problem solving and on a laboratory measure of attentional processing (288). This study was conducted prior to the introduction of atypical antipsychotics. Following another approach, Hogarty and Flesher (289) recently developed cognitive enhancement therapy (CET), which combines interactive software and social group exercises to improve socially and behaviorally relevant cognitive functioning. This approach is based on a neurodevelopmental model for cognitive deficits in schizophrenia (290). Preliminary results

from a controlled 1-year trial of CET have also been encouraging (289).

EXPERIMENTAL TREATMENTS AND STRATEGIES

Part of "56 - Therapeutics of Schizophrenia "

Selective Dopamine Antagonists

There are several lines of evidence suggesting that selective dopamine D_4 receptor antagonists may be potential novel antipsychotic drugs. Clozapine has a relatively higher affinity for the D_4 versus D_2 or D_3 receptors (291) (Table 56.1). Not only clozapine, but also a number of clinically efficacious antipsychotics have relatively high affinity for this receptor site (Table 56.1). In addition, an increase in D_4 receptors has been reported in the brains of patients with schizophrenia (292). Furthermore, the D_4 receptor, enriched in the prefrontal cortex and hippocampus, is located in dopamine terminal fields potentially associated with emotion and cognition, but not with movement, underscoring the potential of this receptor as a target. The selective D_4 antagonist, sonepiprazole (U-101387) increases dopamine release in the frontal cortex, but decreases dopamine release in the nucleus accumbens in rats (293). Sonepiprazole attenuates apomorphine-induced impairment of prepulse inhibition in rats (294). It also antagonized the decrease in c-fos expression in the medial prefrontal cortex and neurotensin mRNA in the nucleus accumbens produced by repetitive amphetamine administration in rats, suggesting possible antipsychotic action of the agent (295). Sonepiprazole is currently in Phase II clinical trials in patients with schizophrenia (293). An initial clinical trial with another highly selective D_4 antagonist, L-745,870, failed to demonstrate any antipsychotic activity in the treatment of schizophrenia (296 ,297). Although the single dose tested makes it difficult to draw firm conclusions regarding the potential efficacy of D_4 antagonists as antipsychotic agents (298), this drug actually caused a worsening of symptoms (297). Similarly, NGD-94-1 also did not show clinical efficacy in limited trials in schizophrenics (293). More extensive testing of D_4 antagonists in patients with schizophrenia will be necessary to adequately assess the therapeutic potential of such drugs.

Dopamine Partial Agonists

Partial dopamine agonists are agents with good affinity for one or more dopamine receptors, but with intrinsic activity less than dopamine (3). Thus, such drugs may antagonize the actions of dopamine, yet by agonistic actions, activate other dopamine-related functions (299). It has been proposed that some D_2-like dopamine agonists have a greater affinity for autoreceptors than for heteroreceptors. The action of these agonists at autoreceptors would induce a receptor-mediated inhibition of both the synthesis and release of dopamine from nerve terminals, without producing significant activation of heteroreceptors on target cells (300). Such partial dopamine agonists are therefore proposed to act as dopaminergic “buffers,” reducing dopaminergic transmission without completely blocking it when dopaminergic activity is excessive, or conversely, stimulating it when it is reduced (7 ,299).

Despite the numerous compounds that were developed as partial agonists, none has proved to be sufficiently effective to warrant its full development and introduction for clinical use. The first of this class to show consistent and robust efficacy comparable to clinically used antipsychotic drugs, both conventional and atypical, is aripiprazole (301). Aripiprazole (OPC-14597) is a dual dopamine autoreceptor partial agonist and postsynaptic D_2 receptor antagonist (302 ,303). It has a modest affinity for 5-HT_2 receptors, but no appreciable affinity for D_1 receptors (304) (Table 56.1). Aripiprazole decreased striatal dopamine release (303), and inhibited the activity of dopamine neurons when applied locally to the ventral tegmental area in rats (305). Animal behavioral studies showed that the compound exhibited weak cataleptogenic effects compared to haloperidol and chlorpromazine despite the fact it has almost identical D_2 receptor antagonistic activity (302). The potency of aripiprazole to up-regulate striatal D_2 receptors in response to chronic treatment was much smaller than that of haloperidol, suggesting lower potential for EPS, including tardive dyskinesia (31). Aripiprazole is currently going through worldwide Phase III development. Preliminary clinical studies have shown its efficacy in alleviating both positive and negative symptoms of schizophrenia. Although current dogma suggests that such a D_2-selective agent would cause profound EPS and high sustained prolactin elevation, neither side effect has been seen clinically (306 ,307 and 308). Based on available data, it would appear that aripiprazole is the first compound with partial D_2 agonist properties to be a clinically effective antipsychotic agent. It has been proposed that aripiprazole induces “functionally selective” activation of D_2 receptors coupled to diverse G proteins (and hence different functions), thereby explaining its unique clinical effects (304).

CI-1007 is a new dopamine autoreceptor agonist and partial dopamine D_2/D_3 receptor agonist that is currently under development for the treatment of schizophrenia (309 ,310). In preclinical studies, CI-1007 demonstrated that it inhibited the firing of dopamine neurons and reduced the synthesis, metabolism, utilization, and release of dopamine in the brain (310). In addition, it produced behavioral effects predictive of antipsychotic efficacy and indicated a low liability for EPS and TD (311).

5-HT Agents

The 5-HT_{2A} receptor subtype has received considerable attention because of its potential roles in the therapeutic action of atypical antipsychotic drugs (21 ,312); it is involved

in perception, mood regulation, and motor control (313). Available evidence indicates that 5-HT_{2A} receptor stimulation plays a role in promoting the synthesis and release of dopamine, either by effects on firing rates of dopamine neurons, or via heteroreceptors on dopamine nerve terminals, or both (312 ,313 ,314 and 315). 5-HT_{2A} receptor blockade may therefore contribute to "normalizing" levels of dopamine release (316) and theoretically possess antipsychotic activity.

M-100907 (formerly MDL-100,907) is a selective 5-HT_{2A} receptor antagonist devoid of affinity to dopamine receptors (21). Like the atypical antipsychotics, it decreases the firing rate of A10, but not A9, neurons after chronic treatment (317). M-100907 inhibited the behavioral response not only to amphetamine and cocaine (316 ,317 and 318), but also to NMDA receptor antagonists at doses that did not affect spontaneous activity given alone in rodents (319 ,320 and 321). M-100907, like clozapine, markedly increases dopamine release in the medial prefrontal cortex in rats (322), suggesting that the agent may have efficacy for negative symptoms. In contrast, it attenuates dopamine release in the nucleus accumbens induced by the NMDA receptor antagonist MK-801 (323). M-100907 also antagonized MK-801-induced prepulse inhibition deficit in rats (324). Further, in electrophysiologic studies, it prevented phencyclidine (PCP)-induced blockade of NMDA responses (325). These preclinical results suggest that M-100907 can attenuate variable responses to NMDA receptor antagonists *in vivo* and modulate NMDA receptor-mediated neurotransmission. M-100907, however, exhibited lower antipsychotic efficacy compared with haloperidol in Phase III clinical trials (326). Insufficient data are currently published to adequately judge the efficacy of the drug.

It has been suggested that the partial agonist activity of clozapine at 5-HT_{1A} receptors may contribute to its therapeutic action (313 ,327). Preclinical studies have suggested that serotonin 5-HT_{1A} agonists may potentiate the antipsychotic activity of dopaminergic antagonists (328). Activation of inhibitory 5-HT_{1A} autoreceptors may also counteract the induction of EPS owing to striatal D_2 receptor blockade (329). Further, in schizophrenic patients, increased 5-HT_{1A} receptor binding was seen in the prefrontal cortex (330 ,331). Based on these preclinical data, compounds that act as serotonin 5-HT_{1A} agonists are being developed as potential antipsychotic compounds.

S-16924 is a novel, potential antipsychotic agent with high affinity for dopamine $D_{2/4}$, α_1-adrenergic, and serotonin 5-HT_{2A} receptors, similar to that of clozapine, in addition to being a potent partial 5-HT_{1A} agonist (332). Reflecting its partial agonist actions at 5-HT_{1A} receptors, it attenuates cerebral serotonergic transmission, and preferentially facilitates dopaminergic transmission in mesocortical as compared to mesolimbic and nigrostriatal pathways (333 ,334). S-16924 exhibited a profile of potential antipsychotic activity and low EPS liability in animal behavioral models, similar to clozapine (332).

Muscarinic Agents

In patients with Alzheimer's disease (AD), cholinesterase inhibitors (e.g., physostigmine) have been shown to not only improve cognition, but also reduce hallucinations, delusions, suspiciousness, and other behavioral disturbances sometime associated with the illness (335 ,336 ,337 and 338). Similar positive effects on cognitive and psychotic-like symptoms in AD have been observed after treatment with the direct muscarinic agonist, xanomeline (339). In addition, high doses of some muscarinic antagonists produce psychotic-like symptoms and memory loss (340). Thus, it has been proposed that muscarinic agonists could be novel potential treatments for positive and cognitive symptoms of schizophrenia (341).

Recent findings that partial agonists of m_2/m_4 muscarinic receptors are active in animal models that predict antipsychotic activity suggest potential usefulness of muscarinic agonists in the treatment of schizophrenia (342). The drug (5R,6R) 6-(3-propylthio-1,2,5-thiadiazol-4-yl)-1-azabicyclo[3.2.1]octane (PTAC) is a muscarinic partial agonist at muscarinic m_2 and m_4 receptor subtypes (342). PTAC acts as a functional dopamine antagonist in many paradigms (consistent with known dopamine-acetylcholine interactions), although it has minimal or no affinity for central dopamine receptors. PTAC attenuates apomorphine induced climbing (341), inhibits the effects of D_1 and D_2 dopamine receptors agonists in 6-hydroxydopaminelesioned rats, and antagonizes amphetamine-induced Fos induction and hyperactivity (343). In addition, after chronic administration, PTAC reduced the number of spontaneously active dopamine cells in the ventral tegmental area, but not the substantia nigra (343). Such selective effects on the mesocorticolimbic dopamine projection neurons are similar to those observed for clozapine and olanzapine (344 ,345). The notable preclinical data of the effects of PTAC provide strong encouragement to examine the potential therapeutic effects of M_2/M_4 muscarinic agonists in schizophrenic patients. Among the agents that have been developed for the treatment of AD that are being examined in schizophrenia are donepezil, metrifonate, galantamine, and xanomeline.

Glutamatergic Agents

The NMDA Receptor Hypofunction Hypothesis of Schizophrenia

Since the late 1950s, the anesthetics phencyclidine (PCP) and ketamine have been known to induce "emergence reactions" in 40% to 50% of individuals during the recovery from anesthesia, that resembles some features of schizophrenia (346). Recent work has confirmed and extended the early clinical studies and has demonstrated that subanesthetic doses of ketamine can induce positive, negative, and cognitive schizophrenia-like symptoms in normal humans (281 ,347).

In chronic stabilized schizophrenic patients, subanesthetic doses of ketamine can also exacerbate cognitive impairment and in some cases reproduce specific hallucinations and delusional ideation remarkably similar to those experienced during active phases of the patients' illness (282 ,348 ,349). Both ketamine and PCP are potent noncompetitive NMDA receptor antagonists. These drugs bind to a site within the calcium channel of the NMDA receptor complex, and thereby interfere with calcium flux through the channel. Competitive NMDA receptor antagonists (i.e., drugs that inhibit binding to the glutamate recognition site) are also psychotomimetic (350). The ability of NMDA antagonists to induce a spectrum of schizophrenia-like symptoms has led to the hypothesis that hypofunction of NMDA receptors is involved in the pathophysiology of schizophrenia (346 ,351 ,352 and 353).

Antipsychotic Drug Actions in Relation to the NMDA Receptor Hypofunction

The well-documented psychotomimetic effects of NMDA antagonist in human suggest that effects of the drugs in experimental animals could present useful pharmacologic models of schizophrenia. In our recent studies, striking effects of subanesthetic doses of ketamine were observed on regional brain patterns of ^{14}C-2-deoxyglucose (2-DG) uptake in both rats (354 ,355) and mice (356). Ketamine induces robust and neuroanatomically selective patterns of brain metabolic activation, with especially large effects observed in the hippocampus, nucleus accumbens, and medial prefrontal cortex (354 ,355). Pretreatment of rats with clozapine or olanzapine can completely block these effects of ketamine (357 ,358). However, the typical antipsychotic haloperidol failed to antagonize the brain metabolic activation induced by ketamine (357). Similarly, clozapine and olanzapine, but not haloperidol, effectively block NMDA antagonist-induced electrophysiologic responses (325), deficits in prepulse inhibition (359 ,360), and deficits in social interactions (361). Thus, in a wide range of experimental paradigms, atypical antipsychotic drugs selectively antagonize the consequences of experimentally induced NMDA receptor hypofunction, raising the possibility that the therapeutic effects of these agents may be associated with a similar neurochemical action (362).

Therapeutic Potential of Glycine Site Agonists

If reduced NMDA receptor function is involved in the pathophysiology of schizophrenia, then drugs that enhance NMDA receptor function could be therapeutic agents and potentially improve upon, or supplement, current antipsychotic treatments (13). Direct agonists of the NMDA receptor may not be feasible candidates in this regard, because of the propensity of such drugs to produce excessive excitation and seizures.

Glycine is a positive allosteric modulator and obligatory coagonist at the NMDA receptor (363) and this allosteric regulatory site represents a potential target for drugs to augment NMDA-mediated neurotransmission. Preclinical studies have demonstrated that glycine-site agonists reverse the effects of noncompetitive NMDA receptor antagonists (364). There have been several clinical studies to test effects of different glycine site agonists in patients with schizophrenia. The earliest studies in this regard used glycine in doses of 5 to 15 g per day and obtained inconsistent results (365 ,366). In more recent work with glycine, higher doses were administered (30 to 60 g per day) and more robust and consistent effects were found, primarily in the improvement of negative symptoms (241 ,367 ,368).

D-cycloserine is a partial agonist at the glycine regulatory site on the NMDA receptor. Thus, at low dose of the amino acid, stimulatory responses are observed, but at higher doses, D-cycloserine blocks the effects of endogenous glycine. D-cycloserine has been tested in patients with schizophrenia, and in a very narrow dose range, the agent was shown to improve negative symptoms when administered alone (369), and when added to conventional antipsychotic treatment regimes (240 , 370). The “inverted U”-shaped dose response may result from the partial agonist properties of D-cycloserine, because antagonism of the actions of endogenous glycine would be predicted at higher doses of the drug. Interestingly, when D-cycloserine was administered in conjunction with clozapine, the negative symptoms of the patients worsened (244 ,371). A ready explanation for these effects is not available, but understanding the mechanisms involved in the worsening of negative symptoms after administration of D-cycloserine to clozapine-treated patients may be an important clue in understanding the actions of both of these drugs. The poor penetration of the blood–brain barrier by glycine, and the partial agonistic properties of D-cycloserine, appear to make these agents less than optimal for providing pharmacologic agonism of the glycine regulatory site on the NMDA receptor (13).

D-serine is a full agonist on the strychnine-insensitive glycine site of NMDA receptor (372) and is more permeable than glycine at the blood–brain barrier, thus requiring a lower dosage. In a recent clinical trial, D-serine (30 mg per kg per day) added to neuroleptic treatment in treatment-resistant patients with schizophrenia demonstrated significant improvements not only in negative and cognitive symptoms but also positive symptoms, which is different from glycine (280). These data, together with the results of the clinical investigations with glycine and D-cycloserine (346), offer promise for the therapeutic potential of enhancing NMDA receptor function as a strategy for the pharmacotherapy of schizophrenia. Recently, Wolosker and colleagues (373) purified an enzyme from Type II astrocytes that converts L-serine to D-serine. It may be that effectors of this enzyme (directly or through possible receptor-mediated regulation) can provide a mechanism to modulate NMDA

function. Examining the effects of synthetic compounds with greater potency and full agonistic activity at the glycine regulatory site could be an intriguing line of future research. There are, however, no such compounds available for testing at present.

Potentiation of NMDA Receptor Function by Inhibition of Glycine Uptake

Glycine transporters have been identified on both neuronal and glial cells in the central nervous system. A function of these transporters has been suggested to control the extracellular glycine concentration (374). Although there is some controversy as to whether the glycine regulatory site on the NMDA receptor is saturated under physiologic conditions, recent data demonstrate that inhibition of glycine transport by glycine transporter type 1 antagonist can potentiate electrophysiologic effects of NMDA (374 ,375). Furthermore, the glycine uptake inhibitor glycyldodecylamide attenuated PCP-induced hyperactivity more potently than glycine (364 ,376). These preclinical data suggest that inhibition of glycine uptake could represent a feasible approach to potentiate NMDA receptor-mediated neurotransmission and, possibly, treat schizophrenic patients.

Glutamate Release-Inhibiting Drugs

A number of studies have indicated that administration of relatively low (subanesthetic) doses of NMDA antagonists induces behavioral and brain metabolic activation in experimental animals and humans (362). Consistent with these data, NMDA antagonists increase glutamate release in rats (377). In contrast to the increase in glutamate release by subanesthetic doses of ketamine, anesthetic doses of the drug decreased glutamate levels (377). The effect of different doses of ketamine on glutamate levels is consistent with our observations of increased 2-DG uptake in response to a subanesthetic dose, and reduction in 2-DG uptake in response to an anesthetic dose of ketamine (354).

The stimulatory effect of NMDA receptor antagonism presumably results from disinhibitory actions, perhaps by reducing excitatory input to inhibitory interneurons (362). In hippocampal formation, GABAergic interneurons are more sensitive to the effects of NMDA antagonists than the glutamate-containing pyramidal cells (378), providing support for the hypothesis that NMDA antagonism could result in excitatory effects by disrupting recurrent inhibitory circuits (362).

If behavioral activation induced by NMDA antagonists is related to increased glutamate release, pharmacologic agents that decrease glutamate release should block the effects of the drugs. Glutamate release can be inhibited by Na^{+}-channel blockers, Ca^{2+}-channel blockers, K^{+}-decreasing agents, toxins that prevent fusion of vesicles with the presynaptic membrane, and presynaptic group II metabotropic glutamate autoreceptor agonists (379 ,380 and 381).

Administration of LY-354740, a group II metabotropic glutamate receptor agonist, blocked both behavioral activation and increased glutamate release induced by PCP in rats (382). In humans, Anand and co-workers (381) found that lamotrigine, a new anticonvulsant agent that inhibits glutamate release, can reduce the ketamine-induced neuropsychiatric effects. These data suggest the possibility that glutamate release-inhibiting drugs (e.g., LY-354740 and lamotrigine) could be useful in the treatment of schizophrenia.

AMPA/Kainate Receptor Antagonists

The increased release of glutamate observed in response to NMDA antagonist could mediate some of the behavioral actions of the drugs by activation of non-NMDA receptors, including α-amino-3-hydroxy-5-methy-isoxazole-4-propionic acid (AMPA) and kainate receptors (377). In support of the hypothesis that behavioral effects of NMDA antagonists relate to increased glutamate release, administration of an AMPA/kainate receptor antagonist, LY-293558, partially reversed impairment of working memory induced by subanesthetic doses of ketamine in rats (377). Furthermore, AMPA/kainate receptor antagonists reduce NMDA antagonist-induced hyperlocomotion (383 ,384 and 385) and neurodegeneration (386). These data suggest that AMPA/kainate receptor antagonists may have utility for treatment of cognitive deficits in which NMDA receptor hypofunction is suspect (377).

Potential of Positive Modulators of AMPA Receptors

In apparent contrast to the postulated utility of AMPA/kainate receptor antagonists as antipsychotics, ampakines, a class of compounds that allosterically enhance AMPA receptor function, have also been suggested to represent potential adjunctive treatments for schizophrenia. Ampakines enhance excitatory (glutamatergic) transmission, facilitate long-term potentiation, learning, and memory in rodents (387 ,388), and have synergistic effects with typical and atypical antipsychotics on blocking behavioral effects of methamphetamine (389). In addition, preliminary results suggest that chronic administration of an ampakine (CX-516) can improve negative and cognitive symptoms in schizophrenia patients that also receive clozapine (284). Thus, such findings are paradoxical with regard to the foregoing discussion of the hypothesis that excessive glutamate release may be involved in behavioral effects of reduced NMDA receptor function. Further clinical experience with the effects of positive and negative modulators of non-NMDA glutamate receptors will be needed to clarify the potential of these compounds for treatment of schizophrenia (3).

Protein Kinase C Inhibitors

Accumulating evidence from Manji and colleagues has identified the family of protein kinase C (PKC) isozymes as a common target in the brain for the long-term action of the two structurally highly dissimilar antimanic agents, lithium and valproate (390). Chronic treatment of rats with lithium or valproate induces a reduction in the levels of two PKC isozymes, α and ε, in the frontal cortex and hippocampus, as well as a reduction in the expression of a major PKC substrate, myristoylated alanine-rich C kinase substrate (MARCKS), which has been implicated in long-term neuroplastic events in the developing and adult brain (391). In view of the critical role of the PKC signaling pathway in the regulation of neuronal excitability, neurotransmitter release, and long-term synaptic events, Manji and associates postulated that the attenuation of PKC activity might have antimanic efficacy. In a pilot study, they found marked antimanic efficacy of a potent PKC inhibitor tamoxifen, which is also a synthetic nonsteroidal antiestrogen, in the treatment of acute mania (392). Their heuristic preliminary data suggest that PKC inhibitors may represent a novel class of antimanic agents for the treatment of bipolar disorder, and deserve further study in psychotic syndromes.

Steroidal Agents

Estrogen

The gender effect of delayed onset (by approximately 2 to 5 years) and relatively reduced symptom severity in females has been consistently observed in schizophrenia (393 ,394 and 395). Some, but not all, researchers have found an additional smaller peak of onset of schizophrenia for women at age 40 to 45 years, which is a time of decreasing levels of estrogen associated with menopause (395 ,396). The inverse relationship between estradiol levels and specific psychopathology, especially positive symptoms, was also observed over the menstrual cycle in premenopausal women with schizophrenia (397 ,398). The indirect clinical evidence suggests a potential role for estrogen in delaying the onset or attenuating the severity of psychotic symptoms associated with schizophrenia (393 ,395). In animal behavioral studies, estrogen reduces amphetamine- and apomorphine-induced stereotypy, as well as enhances neuroleptic-induced catalepsy (399). In addition, preclinical biochemical studies have shown that estrogen can alter dopamine D_2 receptor density and affinity in the brain (399), whereas the effect is dependent on the time course of the administration (395). These findings suggest a neuroleptic-like effect of estrogen, and may have important implications for the prevention and therapy of schizophrenia. To date, there have been few treatment studies examining the effect of estrogen in patients with schizophrenia. Lindamer and associates (395) presented a case report of a postmenopausal woman with schizophrenia who had an improvement in positive symptoms with estrogen augmentation of neuroleptic medication. Long-term larger double-blind trials are crucially needed to evaluate the efficacy of estrogen in conjunction with neuroleptic treatment on psychotic symptoms in women with schizophrenia.

Dehydroepiandrosterone

Dehydroepiandrosterone (DHEA) and its sulfate derivative (DHEA-S) are neuroactive neurosteroids that represent steroid hormones synthesized *de novo* in the brain and acting locally on nerve cells (400). Although DHEA and DHEA-S are the most abundant circulating steroid hormones in humans, their precise physiologic roles remain to be elucidated. In humans, DHEA levels in blood rise dramatically at puberty and sustain a monotonic decline with age, reaching very low levels in late life. *In vitro* data suggest that DHEA and DHEA-S enhance neuronal and glial survival and differentiation in mouse embryonic brain tissue cultures (401 ,402 and 403). In addition, DHEA-S shows marked neuroprotective ability against the glutamate-induced toxicity (404) and oxidative stress (405). In rodents, DHEA has been demonstrated to be a positive modulator of the NMDA receptor. In both the adult rat brain and developing mouse brain, DHEA-S was shown to potentiate substantially physiologic responses to NMDA (403 ,406 ,407). The enhancement of physiologic response to NMDA by DHEA has been suggested to result from agonistic actions at s_1 receptors in the brain (407). Consistent with a positive modulatory action of DHEA at the NMDA receptor, the neurosteroid has been demonstrated to enhance memory in mice (408 ,409 ,410 and 411). Moreover, DHEA-S attenuates NMDA receptor antagonist MK-801-induced learning impairment via an interaction with s_1-receptors in mice (412). These preclinical studies provide the neurobiological rationale for the clinical studies to explore the potential utility of DHEA to treat the NMDA receptor hypofunction postulated to occur in schizophrenia. In chronic schizophrenics, significantly lower morning levels of plasma DHEA were observed (413). Further, there are a number of earlier case reports suggesting that DHEA may be useful in the treatment of schizophrenia, especially for negative symptoms (414 ,415 and 416), although these trials were not well controlled. A recent double-blind study of patients with major depression suggests that DHEA has antidepressant effects (417). Although the mechanism of action of DHEA and DHEA-S has to be further characterized, the possibility that these compounds may have efficiency in schizophrenia should be explored.

Phospholipid Compounds

Membrane Phospholipid Hypothesis of Schizophrenia

The membrane phospholipid hypothesis of schizophrenia originated with suggestion by Horrobin (418) that schizophrenia

might be caused by a prostaglandin (PG) deficiency. The proposal was based on several clinical observations of a relationship between pyrexia and the transient dramatic remission of psychosis, the relative resistance to PG-mediated pain and inflammation and reduced rate of rheumatoid arthritis in patients with schizophrenia, and the observation that PGE_1 injected into the CSF of mammals could produce catalepsy (419). Because PGs are derived from membrane essential fatty acid (EFA), Horrobin and colleagues (420) hypothesized that schizophrenia involves a failure to produce PGE_1 from EFA precursors. Interestingly, over two decades ago, it was suggested that the structure and pharmacologic actions of clozapine are consistent with its being a PGE analogue (420). PGEs are potent stimulators of cAMP formation, and cAMP inhibits phospholipase A_2 (PLA_2). In fact, clozapine treatment induced a dramatic rise in erythrocyte membrane concentrations of the major cerebral fatty acids, arachidonic acid (AA) and docosahexaenoic acid (DHA) (421). Thus, a generally unrecognized mechanism of action of clozapine may be on membrane phospholipid composition, in addition to its receptor-blocking profile (421).

The specific EFA content of synaptic membrane plays a significant role in modifying neuronal function. The changes in membrane EFA concentrations alter the biophysical microenvironment and hence, structure and function of membrane proteins, including neurotransmitter receptors, ion channels, and enzymes (419). EFAs also contribute to cellular regulation by acting as a source of precursors for second messengers in intracellular and intercellular signal transduction (419).

In rat models, changes in brain fatty acid concentrations produced by chronic dietary omega-3 fatty acid deficiency alter dopaminergic and serotonergic neurotransmission (422) and induce a decrease in D_2 and increase in 5-HT_2 receptor density in the frontal cortex (423). Impaired behavioral performance and learning are observed in omega-3 deficient rats (424) and have been hypothesized to reflect changes in attention, motivation and reactivity consistent with a deficit in the function of prefrontal dopamine pathways (419).

The phospholipid hypothesis of schizophrenia has been supported by the accumulating consistent clinical findings in schizophrenic patients that indicate reduced levels of erythrocyte membrane EFA, elevated serum and platelet PLA_2 activity (probably owing to accelerated breakdown of membrane phospholipids), and 31-phosphorus cerebral magnetic resonance spectroscopy (MRS) evidence of decreased synthesis and increased breakdown of phospholipids in the prefrontal cortex (419). Furthermore, phospholipid hypotheses are consistent with both dysfunction of multiple neurotransmitter systems and neurodevelopmental abnormalities associated with aberrant cell remodeling, apoptosis, or migration (425).

Omega-3 Fatty Acid

The membrane EFA or PG deficiency hypotheses have provided the rationale for attempts to treat symptoms of schizophrenia with supplementation of PG precursors, including omega-6 and omega-3 fatty acids and PGE_1. Among the studies of these compounds conducted to date, omega-3 EFA treatment has consistently yielded positive results. Two small open trials and a single double-blind trial suggest supplementation with omega-3 eicosapentaenoic acid (EPA) may improve residual symptoms and tardive dyskinesia when added to standard neuroleptic treatment in schizophrenic patients (419). Surprisingly, the more recent case report by Puri and colleagues (426) demonstrates a dramatic and sustained efficacy of omega-3 EPA on both positive and negative symptoms of schizophrenia in a drug-naive patient without any adverse side effects. In addition, cerebral atrophy, observed before omega-3 EPA treatment, was reversed by 6 months of EPA treatment; however, small trials and a single case report make it difficult to draw firm conclusions regarding the potential efficacy of omega-3 EPA. A recent double-blind placebo controlled study of omega-3 EPA as an adjunctive treatment to antipsychotic drugs found no difference between placebo and omega-3 EPA (427).

Trophic Factors

There is converging evidence that an abnormal neurodevelopmental process is accountable for at least a proportion of the pathophysiology of schizophrenia (428). The neurotrophic factors such as nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), and neurotrophin (NT)-3/4/5 play a decisive role in a neurodevelopmental process, including neuronal and glial differentiation, migration, proliferation, and regeneration (429). They are not only active during embryogenesis and organogenesis, but also influence the synaptic organization and synthesis of neurotransmitters in the adult brain, and are therefore involved in the maintenance of neural plasticity (429). Thus, pathologic alterations of the neurotrophic factor system may lead to neural maldevelopment, migration deficits, and disconnections, which are proposed to be the characteristic pathogenetic features of the maldevelopmental hypothesis of schizophrenia (429). A more recent pathophysiologic theory of schizophrenia suggests that it is involved in a limited neurodegenerative process reflected by the progressive and deteriorating clinical course of the illness (430). If neurotrophic factors salvage degenerating neurons, facilitate desirable synaptic connections, and hence, halt the progression of neurodegenerative process of schizophrenia, drugs that selectively stimulate the production of neurotrophic factors could represent a new approach to forestall the progression of schizophrenia and prevent morbidity from increasing (431). However, the lack of consistent evidence supportive

of pathophysiologic progression in schizophrenia has been a weakness of this hypothesis (430). Recently, Riva and associates (432) found that acute or chronic administration of clozapine increased basic fibroblast growth factor (FGF-2) mRNA and protein in the rat striatum, suggesting neuroprotective activity of clozapine. It has been proposed that small molecules that boost the endogenous levels of BDNF or NT-3 might be useful for treating temporally protracted and severe forms of neurodegenerative disease, such as AD or Parkinson's disease (433). Although neurotrophic factors are unable to cross the blood–brain barrier, potential alterations to administration of these factors are transplantation of neurotrophic factor-producing cells, direct transfection of neurotrophic factor gene, and development of compounds that modulate endogenous neurotrophic factor homeostasis and/or the influence their signal transduction mechanisms (429). The augmentation therapy with neurotrophic factors suggests novel and innovative pharmacotherapeutic, but as yet unproved strategies for schizophrenia.

CONCLUSION

Part of "56 - Therapeutics of Schizophrenia "

The therapeutic armamentarium for the treatment of schizophrenia has become rich and varied in the half century since the inception of the pharmacologic era marked by the introduction of chlorpromazine. We now have the capacity to control many of the symptoms of the disorder and restore the lives of patients. Much remains to be done in terms of drug discovery of new and novel agents and the determination of their optimal use in conjunction with psychosocial and adjunctive therapies; however, there is reason to be optimistic that future progress will be relatively swift.

ACKNOWLEDGMENTS

Part of "56 - Therapeutics of Schizophrenia "

Dr. Goff received research support from Cortex Pharmaceuticals, Eli Lilly & Company, Janssen Pharmaceuticals, and Pfizer, Inc. In addition, he has received honoraria and/or served on an advisory board for Eli Lilly, Janssen, and Pfizer, Inc. Dr. Lieberman has served as a consultant for a number of companies including: Janssen, Lilly, Astra Zeneca, Pfizer, Bristol Myers Squibb, Protarga, and Wyeth Ayerst.

REFERENCES

1. Lieberman JA. Atypical antipsychotic drugs as a first-line treatment of schizophrenia: a rationale and hypothesis. *J Clin Psychiatry* 1996;57(Suppl 11):68–71.

2. Leucht S, Pitschel-Walz G, Abraham D, et al. Efficacy and extrapyramidal side-effects of the new antipsychotics olanzapine, quetiapine, risperidone, and sertindole compared to conventional antipsychotics and placebo. A metaanalysis of randomized controlled trials. *Schizophr Res* 1999;35:51–68.

3. Miyamoto S, Duncan GE, Mailman RB, et al. Developing novel antipsychotic drugs: strategies and goals. *Curr Opin CPNS Invest Drugs* 2000;2:25–39.

4. Seeman P. Dopamine receptor sequences. Therapeutic levels of neuroleptics occupy D2 receptors, clozapine occupies D4. *Neuropsychopharmacology* 1992;7:261–284.

5. Kane JM. The current status of neuroleptic therapy. *J Clin Psychiatry* 1989;50:322–328.

6. Meltzer HY. Outcome in schizophrenia: beyond symptom reduction. *J Clin Psychiatry* 1999;60(Suppl 3):3–7.

7. Fleischhacker WW. New drugs for the treatment of schizophrenic patients. *Acta Psychiatr Scand (Suppl)* 1995;388:24–30.

8. Campbell M, Young PI, Bateman DN, et al. The use of atypical antipsychotics in the management of schizophrenia. *Br J Clin Pharmacol* 1999;47:13–22.

9. Meyer JM, Simpson GM. From chlorpromazine to olanzapine: a brief history of antipsychotics. *Psychiatr Serv* 1997;48:1137–1139.

10. Lieberman JA, Kane JM, Johns CA. Clozapine: guidelines for clinical management. *J Clin Psychiatry* 1989;50:329–338.

11. Kane J, Honigfeld G, Singer J, et al. Clozapine for the treatment-resistant schizophrenic. A double-blind comparison with chlorpromazine. *Arch Gen Psychiatry* 1988;45:789–796.

12. Lieberman JA. Understanding the mechanism of action of atypical antipsychotic drugs: a review of compounds in use and development. *Br J Psychiatry* 1993;163:7–18.

13. Duncan GE, Zorn S, Lieberman JA. Mechanisms of typical and atypical antipsychotic drug action in relation to dopamine and NMDA receptor hypofunction hypotheses of schizophrenia. *Mol Psychiatry* 1999;4:418–428.

14. Baldessarini RJ, Katz B, Cotton P. Dissimilar dosing with high-potency and low-potency neuroleptics. *Am J Psychiatry* 1984;141:748–752.

15. Farde L, Nordstrom AL, Wiesel FA, et al. Positron emission tomographic analysis of central D1 and D2 dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine. Relation to extrapyramidal side effects. *Arch Gen Psychiatry* 1992;49:538–544.

16. Kapur S, Remington G, Jones C, et al. High levels of dopamine D2 receptor occupancy with low-dose haloperidol treatment: a PET study. *Am J Psychiatry* 1996;153:948–950.

17. Nordstrom AL, Farde L, Wiesel FA, et al. Central D2-dopamine receptor occupancy in relation to antipsychotic drug effects: a double-blind PET study of schizophrenic patients. *Biol Psychiatry* 1993;33:227–235.

18. Kapur S, Zipursky R, Jones C, et al. Relationship between dopamine D_2 occupancy, clinical response, and side effects: a double-blind PET study of first-episode schizophrenia. *Am J Psychiatry* 2000;157:514–520.

19. Remington G, Kapur S. D2 and 5-HT2 receptor effects of antipsychotics: bridging basic and clinical findings using PET. *J Clin Psychiatry* 1999;60(Suppl 10):15–19.

20. Nyberg S, Farde L. Non-equipotent doses partly explain differences among antipsychotics—implications of PET studies. *Psychopharmacology* 2000;148:22–23.

21. Lieberman JA, Mailman RB, Duncan G, et al. Serotonergic basis of antipsychotic drug effects in schizophrenia. *Biol Psychiatry* 1998;44:1099–1117.

22. Dragunow M, Robertson GS, Faull RL, et al. D2 dopamine receptor antagonists induce fos and related proteins in rat striatal neurons. *Neuroscience* 1990;37:287–294.

23. Robertson GS, Fibiger HC. Neuroleptics increase c-fos expression in the forebrain: contrasting effects of haloperidol and clozapine. *Neuroscience* 1992;46:315–328.

24. Merchant KM, Dorsa DM. Differential induction of neurotensin and c-fos gene expression by typical versus atypical antipsychotics. *Proc Natl Acad Sci USA* 1993;90:3447–3451.

25. Semba J, Sakai M, Miyoshi R, et al. Differential expression of c-fos mRNA in rat prefrontal cortex, striatum, *N. accumbens*, and lateral septum after typical and atypical antipsychotics: an in situ hybridization study. *Neurochem Int* 1996;29:435–442.

26. Deutch AY, Lee MC, Iadarola MJ. Regionally specific effects of atypical antipsychotic drugs on striatal Fos expression: the nucleus accumbens shell as a locus of antipsychotic action. *Mol Cell Neurosci* 1992;3:332–341.

27. Robertson GS, Matsumura H, Fibiger HC. Induction patterns of neuroleptic-induced Fos-like immunoreactivity as predictors of atypical antipsychotic activity. *J Pharmacol Exp Ther* 1994;271:1058–1066.

28. Nakahara T, Kuroki T, Hashimoto K, et al. Effect of atypical antipsychotics on phencyclidine-induced expression of arc in rat brain. *Neuroreport* 2000;11:551–555.

29. Buckland PR, O'Donovan MC, McGuffin P. Both splicing variants of the dopamine D2 receptor mRNA are up-regulated by antipsychotic drugs. *Neurosci Lett* 1993;150:25–28.

30. Rogue P, Hanauer A, Zwiller J, et al. Up-regulation of dopamine D2 receptor mRNA in rat striatum by chronic neuroleptic treatment. *Eur J Pharmacol* 1991;207:165–168.

31. Inoue A, Miki S, Seto M, et al. Aripiprazole, a novel antipsychotic drug, inhibits quinpirole-evoked GTPase activity but does not up-regulate dopamine D2 receptor following repeated treatment in the rat striatum. *Eur J Pharmacol* 1997;321:105–111.

32. Chakos MH, Lieberman JA, Bilder RM, et al. Increase in caudate nuclei volumes of first-episode schizophrenic patients taking antipsychotic drugs. *Am J Psychiatry* 1994;151:1430–1436.

33. Corson PW, Nopoulos P, Miller DD, et al. Change in basal ganglia volume over 2 years in patients with schizophrenia: typical versus atypical neuroleptics. *Am J Psychiatry* 1999;156:1200–1204.

34. Chakos MH, Shirakawa O, Lieberman J, et al. Striatal enlargement in rats chronically treated with neuroleptic. *Biol Psychiatry* 1998;44:675–684.

35. Hawkins KA, Mohamed S, Woods SW. Will the novel antipsychotics significantly ameliorate neuropsychological deficits and improve adaptive functioning in schizophrenia? *Psychol Med* 1999;29:1–8.

36. Green MF. What are the functional consequences of neurocognitive deficits in schizophrenia? *Am J Psychiatry* 1996;153:321–330.

37. Keefe RSE, Silva SG, Perkins DO, et al. The effects of atypical antipsychotic drugs on neurocognitive impairment in schizophrenia: a review and meta-analysis. *Schizophr Bull* 1999;25:201–222.

38. Sharif ZA. Common treatment goals of antipsychotics: acute treatment. *J Clin Psychiatry* 1998;59(Suppl 19):5–8.

39. Barnes TRE, McPhillips MA. Critical analysis and comparison of the side-effect and safety profiles of the new antipsychotics. *Br J Psychiatry* 1999;174(Suppl 38):34–43.

40. Chakos MH, Mayerhoff DI, Loebel AD, et al. Incidence and correlates of acute extrapyramidal symptoms in first-episode of schizophrenia. *Psychopharmacol Bull* 1992;28:81–86.

41. Kane J. Olanzapine in the long-term treatment of schizophrenia. *Br J Psychiatry* 1999;174:26–29.

42. Fleischhacker WW, Hummer M. Drug treatment of schizophrenia in the 1990s. Achievements and future possibilities in optimising outcomes. *Drugs* 1997;53:915–929.

43. Ayuso-Gutierrez JL, del R, V. Factors influencing relapse in the long-term course of schizophrenia. *Schizophr Res* 1997;28:199–206.

44. Weiden PJ, Olfson M. Cost of relapse in schizophrenia. *Schizophr Bull* 1995;21:419–429.

45. Revicki DA. Pharmacoeconomic studies of atypical antipsychotic drugs for the treatment of schizophrenia. *Schizophr Res* 1999;35:S101–S109

46. Weiden P, Aquila R, Standard J. Atypical antipsychotic drugs and long-term outcome in schizophrenia. *J Clin Psychiatry* 1996;57(Suppl 11):53–60.

47. Meltzer HY, Matsubara S, Lee JC. Classification of typical and atypical antipsychotic drugs on the basis of dopamine D1, D2 and Serotonin2 pKi values. *J Pharmacol Exp Ther* 1989;251:238–246.

48. Gerlach J. New antipsychotics: classification, efficacy, and adverse effects. *Schizophr Bull* 1991;17:289–309.

49. Litman RE, Su TP, Potter WZ, et al. Idazoxan and response to typical neuroleptics in treatment-resistant schizophrenia. Comparison with the atypical neuroleptic, clozapine. *Br J Psychiatry* 1996;168:571–579.

50. Friedman JI, Temporini H, Davis KL. Pharmacologic strategies for augmenting cognitive performance in schizophrenia. *Biol Psychiatry* 1999;45:1–16.

51. Bymaster FP, Calligaro DO, Falcone JF, et al. Radioreceptor binding profile of the atypical antipsychotic olanzapine. *Neuropsychopharmacology* 1996;14:87–96.

52. Bymaster FP, Rasmussen K, Calligaro DO, et al. In vitro and in vivo biochemistry of olanzapine: a novel, atypical antipsychotic drug. *J Clin Psychiatry* 1997;58(Suppl 10):28–36.

53. Markowitz JS, Brown CS, Moore TR. Atypical antipsychotics. Part I: pharmacology, pharmacokinetics, and efficacy. *Ann Pharmacother* 1999;33:73–85.

54. Kapur S, Zipursky R, Jones C, et al. A positron emission tomography study of quetiapine in schizophrenia: a preliminary finding of an antipsychotic effect with only transiently high dopamine D2 receptor occupancy. *Arch Gen Psychiatry* 2000;57:553–559.

55. Arnt J, Skarsfeldt T. Do novel antipsychotics have similar pharmacologic characteristics? A review of the evidence. *Neuropsychopharmacology* 1998;18:63–101.

56. Kapur S, Zipursky RB, Remington G, et al. 5-HT2 and D2 receptor occupancy of olanzapine in schizophrenia: a PET investigation. *Am J Psychiatry* 1998;155:921–928.

57. Kapur S, Zipursky RB, Remington G. Clinical and theoretical implications of 5-HT_2 and D_2 receptor occupancy of clozapine, risperidone, and olanzapine in schizophrenia. *Am J Psychiatry* 1999;156:286–293.

58. Nordstrom AL, Farde L, Nyberg S, et al. D1, D2, and 5-HT2 receptor occupancy in relation to clozapine serum concentration: a PET study of schizophrenic patients. *Am J Psychiatry* 1995;152:1444–1449.

59. Seeman P, Tallerico T. Antipsychotic drugs which elicit little or no parkinsonism bind more loosely than dopamine to brain D2 receptors, yet occupy high levels of these receptors. *Mol Psychiatry* 1998;3:123–134.

60. Farde L, Nyberg S, Oxenstierna G, et al. Positron emission tomography studies on D2 and 5-HT2 receptor binding in risperidone-treated schizophrenic patients. *J Clin Psychopharmacol* 1995;15:19S–23S.

61. Remington G, Kapur S. Atypical antipsychotics: are some more atypical than others? *Psychopharmacology* 2000;148:3–15.

62. Tollefson GD, Kuntz AJ. Review of recent clinical studies with olanzapine. *Br J Psychiatry* 1999;174:30–35.

63. Tran PV, Hamilton SH, Kuntz AJ, et al. Double-blind comparison of olanzapine versus risperidone in the treatment of schizophrenia and other psychotic disorders. *J Clin Psychopharmacol* 1997;17:407–418.

64. Tollefson GD, Sanger TM, Lu Y, et al. Depressive signs and symptoms in schizophrenia: a prospective blinded trial of olanzapine and haloperidol. *Arch Gen Psychiatry* 1998;55:250–258.

65. Conley RR, Mahmoud R, the Risperidone Study Group. Efficacy of risperidone vs olanzapine in the treatment of patients with schizophrenia or schizoaffective disorder. *Int J Neuropsychopharmacol* 2000;3(Suppl 1):S151(P.01.219).

66. Goff DC, Posever T, Herz L, et al. An exploratory haloperidol-controlled dose-finding study of ziprasidone in hospitalized patients with schizophrenia or schizoaffective disorder. *J Clin Psychopharmacol* 1998;18:296–304.

67. Arato M, O'Connor R, Meltzer H, et al. Ziprasidone: efficacy in the prevention of relapse and in the long-term treatment of negative symptoms of chronic schizophrenia. 10th Annual Meeting of the ECNP. Austria, 1997.

68. Pickar D, Owen RR, Litman RE, et al. Clinical and biologic response to clozapine in patients with schizophrenia. Crossover comparison with fluphenazine. *Arch Gen Psychiatry* 1992;49:345–353.

69. Bradford DW, Chakos MH, Sheitman BB, et al. Atypical antipsychotic drugs in treatment-refractory schizophrenia. *Psychiatr Ann* 1998;28:618.

70. Lieberman JA, Safferman AZ, Pollack S, et al. Clinical effects of clozapine in chronic schizophrenia: response to treatment and predictors of outcome. *Am J Psychiatry* 1994;151:1744–1752.

71. Kurz M, Hummer M, Kurzthaler I, et al. Efficacy of medium-dose clozapine for treatment-resistant schizophrenia. *Am J Psychiatry* 1995;152:1690–1691.

72. Flynn SW, MacEwan GW, Altman S, et al. An open comparison of clozapine and risperidone in treatment-resistant schizophrenia. *Pharmacopsychiatry* 1998;31:25–29.

73. Breier AF, Malhotra AK, Su TP, et al. Clozapine and risperidone in chronic schizophrenia: effects on symptoms, parkinsonian side effects, and neuroendocrine response. *Am J Psychiatry* 1999;156:294–298.

74. Volavka J, Czobor P, Sheitman B, et al. Clozapine, olanzapine, risperidone, and haloperidol in treatment-resistant patients with schizophrenia and schizoaffective disorder. 38th ACNP Annual Meeting. Acapulco, Mexico, 1999.

75. Bondolfi G, Dufour H, Patris M, et al. Risperidone versus clozapine in treatment-resistant chronic schizophrenia: a randomized double-blind study. The Risperidone Study Group. *Am J Psychiatry* 1998;155:499–504.

76. Breier A, Hamilton SH. Comparative efficacy of olanzapine and haloperidol for patients with treatment-resistant schizophrenia. *Biol Psychiatry* 1999;45:403–411.

77. Conley RR, Tamminga CA, Bartko JJ, et al. Olanzapine compared with chlorpromazine in treatment-resistant schizophrenia. *Am J Psychiatry* 1998;155:914–920.

78. Dawkins K, Lieberman JA, Lebowitz BD, et al. Antipsychotics: Past and future—National Institute of Mental Health Division of Services and Intervention Research Workshop, July 14, 1998. *Schizophr Bull* 1999;25:395–404.

79. Meltzer HY. Treatment of schizophrenia and spectrum disorders: pharmacotherapy, psychosocial treatments, and neurotransmitter interactions. *Biol Psychiatry* 1999;46:1321–1327.

80. Hummer M, Fleischhacker WW. Non-motor side effects of novel antipsychotics. *Curr Opin CPNS Invest Drugs* 2000;2:45–51.

81. Wirshing DA, Wirshing WC, Kysar L, et al. Novel antipsychotics: comparison of weight gain liabilities. *J Clin Psychiatry* 1999;60:358–363.

82. Wirshing DA, Spellberg BJ, Erhart SM, et al. Novel antipsychotics and new onset diabetes. *Biol Psychiatry* 1998;44:778–783.

83. Alvir JM, Lieberman JA, Safferman AZ, et al. Clozapineinduced agranulocytosis. Incidence and risk factors in the United States. *N Engl J Med* 1993;329:162–167.

84. Csernansky J, Okamoto A. Risperidone vs haloperidol for prevention of relapse in schizophrenia and schizoaffective disorders: a long-term double-blind comparison. The 10th Biennial Winter Workshop on Schizophrenia. Davos, Switzerland, 2000.

85. Chouinard G. Effects of risperidone in tardive dyskinesia: an analysis of the Canadian multicenter risperidone study. *J Clin Psychopharmacol* 1995;15:36S–44S.

86. Conley RR, Mahmoud R, the Risperidone Study Group. Risperidone vs olanzapine in the treatment of patients with schizophrenia or schizoaffective disorder: safety comparisons. *Int J Neuropsychopharmacol* 2000;3(Suppl 1):S151 (P.01.218).

87. Tollefson GD, Beasley CMJ, Tamura RN, et al. Blind, controlled, long-term study of the comparative incidence of treatment-emergent tardive dyskinesia with olanzapine or haloperidol. *Am J Psychiatry* 1997;154:1248–1254.

88. Small JG, Hirsch SR, Arvanitis LA, et al. Quetiapine in patients with schizophrenia. A high- and low-dose double-blind comparison with placebo. Seroquel Study Group. *Arch Gen Psychiatry* 1997;54:549–557.

89. Stip E, Boisjoly H. Quetiapine: are we overreacting in our concern about cataracts (the beagle effect)? Can J Psychiatry 1999;44:503.

90. Ferris P. Ziprasidone. *Curr Opin CPNS Invest Drugs* 2000;2:58–70.

91. Essock SM, Hargreaves WA, Dohm FA, et al. Clozapine eligibility among state hospital patients. *Schizophr Bull* 1996;22:15–25.

92. Rosenheck R, Cramer J, Xu W, et al. A comparison of clozapine and haloperidol in hospitalized patients with refractory schizophrenia. Department of Veterans Affairs Cooperative Study Group on Clozapine in Refractory Schizophrenia. *N Engl J Med* 1997;337:809–815.

93. Addington DE, Jones B, Bloom D, et al. Reduction of hospital days in chronic schizophrenic patients treated with risperidone: a retrospective study. *Clin Ther* 1993;15:917–926.

94. Guest JF, Hart WM, Cookson RF, et al. Pharmacoeconomic evaluation of long-term treatment with risperidone for patients with chronic schizophrenia. *Br J Med Econ* 1996;10:59–67.

95. Viale G, Mechling L, Maislin G, et al. Impact of risperidone on the use of mental health care resources. *Psychiatr Serv* 1997;48:1153–1159.

96. Beasley CMJ, Tollefson GD, Tran PV. Efficacy of olanzapine: an overview of pivotal clinical trials. *J Clin Psychiatry* 1997;58(Suppl 10):7–12.

97. Tran PV, Dellva MA, Tollefson GD, et al. Oral olanzapine versus oral haloperidol in the maintenance treatment of schizophrenia and related psychoses. *Br J Psychiatry* 1998;172:499–505.

98. Amin S. Cost-effectiveness of atypical antipsychotics in chronic schizophrenia. *Hosp Med* 1999;60:410–413.

99. Glazer WM, Johnstone BM. Pharmacoeconomic evaluation of antipsychotic therapy for schizophrenia. *J Clin Psychiatry* 1997;58(Suppl 10):50–54.

100. Hamilton SH, Revicki DA, Edgell ET, et al. Clinical and economic outcomes of olanzapine compared with haloperidol for schizophrenia. Results from a randomised clinical trial. *Pharmacoeconomics* 1999;15:469–480.

101. Palmer CS, Revicki DA, Genduso LA, et al. A cost-effectiveness clinical decision analysis model for schizophrenia. *Am J Manag Care* 1998;4:345–355.

102. Lieberman J, Jody D, Geisler S, et al. Time course and biologic correlates of treatment response in first-episode schizophrenia. *Arch Gen Psychiatry* 1993;50:369–376.

103. Bilder RM, Goldman RS, Robinson D, et al. Neuropsychology of first-episode schizophrenia: initial characterization and clinical correlates. *Am J Psychiatry* 2000;157:549–559.

104. Gupta S, Andreasen NC, Arndt S, et al. The Iowa Longitudinal Study of Recent Onset Psychosis: one-year follow-up of first-episode patients. *Schizophr Res* 1997;23:1–13.

105. McEvoy JP, Hogarty GE, Steingard S. Optimal dose of neuroleptic in acute schizophrenia. A controlled study of the neuroleptic threshold and higher haloperidol dose. *Arch Gen Psychiatry* 1991;48:739–745.

106. Kopala LC, Good KP, Honer WG. Extrapyramidal signs and clinical symptoms in first-episode schizophrenia: response to low-dose risperidone. *J Clin Psychopharmacol* 1997;17:308–313.

107. Sanger TM, Lieberman JA, Tohen M, et al. Olanzapine versus haloperidol treatment in first-episode psychosis. *Am J Psychiatry* 1999;156:79–87.

108. Lieberman JA, Phillips M, Kong L, et al. Efficacy and safety of clozapine versus chlorpromazine in first-episode psychosis: results of a 52 week randomized double blind trial. 39th ACNP Annual Meeting. San Juan, Puerto Rico, 2000.

109. Lieberman J, Tohen M, McEvoy J, et al. Olanzapine versus haloperidol in the treatment of first-episode psychosis. 39th ACNP Annual Meeting. San Juan, Puerto Rico, 2000.

110. Loebel AD, Lieberman JA, Alvir JMJ, et al. Duration of psychosis and outcome in first-episode schizophrenia. *Am J Psychiatry* 1992;149:1183–1188.

111. Craig TJ, Bromet EJ, Fennig S, et al. Is there an association between duration of untreated psychosis and 24-month clinical outcome in a first-admission series? *Am J Psychiatry* 2000;157:60–66.

112. Ho BC, Andreasen NC, Flaum M, et al. Untreated initial psychosis: its relation to quality of life and symptom remission in first-episode schizophrenia. *Am J Psychiatry* 2000;157:808–815.

113. Robinson DG, Woerner MG, Alvir JM, et al. Predictors of treatment response from a first-episode of schizophrenia or schizoaffective disorder. *Am J Psychiatry* 1999;156:544–549.

114. Falloon IR, Kydd RR, Coverdale JH, et al. Early detection and intervention for initial episodes of schizophrenia. *Schizophr Bull* 1996;22:271–282.

115. McGorry PD, Edwards J, Mihalopoulos C, et al. EPPIC: an evolving system of early detection and optimal management. *Schizophr Bull* 1996;22:305–326.

116. Yung AR, McGorry PD, McFarlane CA, et al. Monitoring and care of young people at incipient risk of psychosis. *Schizophr Bull* 1996;22:283–303.

117. Haddock G, Morrison AP, Hopkins R, et al. Individual cognitive-behavioural interventions in early psychosis. *Br J Psychiatry (Suppl)* 1998;172:101–106.

118. Linszen D, Dingemans P, Van der Does JW, et al. Treatment, expressed emotion and relapse in recent onset schizophrenic disorders. *Psychol Med* 1996;26:333–342.

119. Nugter A, Dingemans P, Van der Does JW, et al. Family treatment, expressed emotion and relapse in recent onset schizophrenia. *Psychiatr Res* 1997;72:23–31.

120. Hogarty GE. Depot neuroleptics: the relevance of psychosocial factors—a United States perspective. *J Clin Psychiatry* 1984;45:36–42.

121. Marder SR, Hubbard JW, Van Putten T, et al. Pharmacokinetics of long-acting injectable neuroleptic drugs: clinical implications. *Psychopharmacology* 1989;98:433–439.

122. Marder SR, Van Putten T, Mintz J, et al. Low- and conventional-dose maintenance therapy with fluphenazine decanoate. Two-year outcome. *Arch Gen Psychiatry* 1987;44:518–521.

123. Glazer WM, Kane JM. Depot neuroleptic therapy: an underutilized treatment option. *J Clin Psychiatry* 1992;53:426–433.

124. Schooler NR, Keith SJ, Severe JB, et al. Relapse and rehospitalization during maintenance treatment of schizophrenia. The effects of dose reduction and family treatment. *Arch Gen Psychiatry* 1997;54:453–463.

125. Carpenter WRJ, Buchanan RW, Kirkpatrick B, et al. Diazepam treatment of early signs of exacerbation in schizophrenia. *Am J Psychiatry* 1999;156:299–303.

126. Essock SM, Hargreaves WA, Covell NH, et al. Clozapine's effectiveness for patients in state hospitals: results from a randomized trial. *Psychopharmacol Bull* 1996;32:683–697.

127. Steingard S, Allen M, Schooler NR. A study of the pharmacologic treatment of medication-compliant schizophrenics who relapse. *J Clin Psychiatry* 1994;55:470–472.

128. Conley RR, Love RC, Kelly DL, et al. Rehospitalization rates of patients recently discharged on a regimen of risperidone or clozapine. *Am J Psychiatry* 1999;156:863–868.

129. Dixon L, Adams C, Lucksted A. Update on family psychoeducation for schizophrenia. *Schizophr Bull* 2000;26:5–20.

130. Huxley NA, Rendall M, Sederer L. Psychosocial treatments in schizophrenia: a review of the past 20 years. *J Nerv Ment Dis* 2000;188:187–201.

131. McFarlane WR, Lukens E, Link B, et al. Multiple-family groups and psychoeducation in the treatment of schizophrenia. *Arch Gen Psychiatry* 1995;52:679–687.

132. Penn DL, Mueser KT. Research update on the psychosocial treatment of schizophrenia. *Am J Psychiatry* 1996;153:607–617.

133. Stein LI, Test MA. Alternative to mental hospital treatment. I. Conceptual model, treatment program, and clinical evaluation. *Arch Gen Psychiatry* 1980;37:392–397.

134. Heinssen RK, Liberman RP, Kopelowicz A. Psychosocial skills training for schizophrenia: lessons from the laboratory. *Schizophr Bull* 2000;26:21–46.

135. Herz MI, Lamberti JS, Mintz J, et al. A program for relapse prevention in schizophrenia: a controlled study. *Arch Gen Psychiatry* 2000;57:277–283.

136. Cramer JA, Rosenheck R. Compliance with medication regimens for mental and physical disorders. *Psychiatr Serv* 1998;49:196–201.

137. Hoge SK, Appelbaum PS, Lawlor T, et al. A prospective, multicenter study of patients' refusal of antipsychotic medication. *Arch Gen Psychiatry* 1990;47:949–956.

138. Pristach CA, Smith CM. Medication compliance and substance abuse among schizophrenic patients. *Hosp Community Psychiatry* 1990;41:1345–1348.

139. Van Putten T. Why do schizophrenic patients refuse to take their drugs? *Arch Gen Psychiatry* 1974;31:67–72.

140. Hogan TP, Awad AG, Eastwood R. A self-report scale predictive of drug compliance in schizophrenics: reliability and discriminative validity. *Psychol Med* 1983;13:177–183.

141. Adams SGJ, Howe JT. Predicting medication compliance in a psychotic population. *J Nerv Ment Dis* 1993;181:558–560.

142. Day JC, Kinderman P, Bentall R. A comparison of patients' and prescribers' beliefs about neuroleptic side-effects: prevalence, distress and causation. *Acta Psychiatr Scand* 1998;97:93–97.

143. Finn SE, Bailey JM, Schultz RT, et al. Subjective utility ratings of neuroleptics in treating schizophrenia. *Psychol Med* 1990;20:843–848.

144. Larsen EB, Gerlach J. Subjective experience of treatment, side-effects, mental state and quality of life in chronic schizophrenic out-patients treated with depot neuroleptics. *Acta Psychiatr Scand* 1996;93:381–388.

145. Perkins DO. Adherence to antipsychotic medications. *J Clin Psychiatry* 1999;60(Suppl 21):25–30.

146. Kemp R, David A, Haywood P. Compliance therapy: an intervention targeting insight and treatment adherence in psychotic patients. *Behav Cogn Psychother* 1996;24:331–350.

147. Kemp R, Kirov G, Everitt B, et al. Randomised controlled trial of compliance therapy. 18-month follow-up. *Br J Psychiatry* 1998;172:413–419.

148. Cramer JA, Rosenheck R. Enhancing medication compliance for people with serious mental illness. *J Nerv Ment Dis* 1999;187:53–55.

149. Meltzer HY. Defining treatment refractoriness in schizophrenia. *Schizophr Bull* 1990;16:563–565.

150. Hegarty JD, Baldessarini RJ, Tohen M, et al. One hundred years of schizophrenia: a metaanalysis of the outcome literature. *Am J Psychiatry* 1994;151:1409–1416.

151. Wolkin A, Barouche F, Wolf AP, et al. Dopamine blockade and clinical response: evidence for two biological subgroups of schizophrenia. *Am J Psychiatry* 1989;146:905–908.

152. Meltzer HY. Clozapine: pattern of efficacy in treatment-resistant schizophrenia. In: Meltzer HY, ed. *Novel antipsychotic drugs*. New York: Raven, 1992:33–46.

153. Carpenter WTJ, Conley RR, Buchanan RW, et al. Patient response and resource management: another view of clozapine treatment of schizophrenia. *Am J Psychiatry* 1995;152:827–832.

154. Meltzer HY. Clozapine: is another view valid? *Am J Psychiatry* 1995;152:821–825.

155. Marder SR, Meibach RC. Risperidone in the treatment of schizophrenia. *Am J Psychiatry* 1994;151:825–835.

156. Wirshing DA, Marshall BDJ, Green MF, et al. Risperidone in treatment-refractory schizophrenia. *Am J Psychiatry* 1999;156:1374–1379.

157. Bouchard RH, Merette C, Pourcher E, et al. Longitudinal comparative study of risperidone and conventional neuroleptics for treating patients with schizophrenia. The Quebec Schizophrenia Study Group. *J Clin Psychopharmacol* 2000;20:295–304.

158. Conley RR, Tamminga CA, Kelly DL, et al. Treatment-resistant schizophrenic patients respond to clozapine after olanzapine non-response. *Biol Psychiatry* 1999;46:73–77.

159. Henderson DC, Goff DC. Risperidone as an adjunct to clozapine therapy in chronic schizophrenics. *J Clin Psychiatry* 1996;57:395–397.

160. Still DJ, Dorson PG, Crismon ML, et al. Effects of switching inpatients with treatment-resistant schizophrenia from clozapine to risperidone. *Psychiatr Serv* 1996;47:1382–1384.

161. Klieser E, Lehmann E, Kinzler E, et al. Randomized, double-blind, controlled trial of risperidone versus clozapine in patients with chronic schizophrenia. *J Clin Psychopharmacol* 1995;15:45S–51S.

162. Miller DD. The clinical use of clozapine plasma concentrations in the management of treatment-refractory schizophrenia. *Ann Clin Psychiatry* 1996;8:99–109.

163. Conley R, Brecher M, Group ROS. Risperidone versus olanzapine in patients with schizophrenia or schizoaffective disorder. 11th ECNP Congress. Paris, France, 1998.

164. Wang PS, West JC, Tanielian T, et al. Recent patterns and predictors of antipsychotic medication regimens used to treat schizophrenia and other psychotic disorders. *Schizophr Bull* 2000;26:451–457.

165. Naber D, Holzbach R, Perro C, et al. Clinical management of clozapine patients in relation to efficacy and side-effects. *Br J Psychiatry Suppl* 1992;54–59.

166. Morera AL, Barreiro PJL. Risperidone and clozapine combination for the treatment of refractory schizophrenia. *Acta Psychiatr Scand* 1999;99:305–306.

167. Friedman J, Ault K, Powchik P. Pimozide augmentation for the treatment of schizophrenic patients who are partial responders to clozapine. *Biol Psychiatry* 1997;42:522–523.

168. Shiloh R, Zemishlany Z, Aizenberg D, et al. Sulpiride augmentation in people with schizophrenia partially responsive to clozapine. A double-blind, placebo-controlled study. *Br J Psychiatry* 1997;171:569–573.

169. Chong SA, Remington G. Clozapine augmentation: safety and efficacy. *Schizophr Bull* 2000;26:421–440.

170. Carman JS, Bigelow LB, Wyatt RJ. Lithium combined with neuroleptics in chronic schizophrenic and schizoaffective patients. *J Clin Psychiatry* 1981;42:124–128.

171. Growe GA, Crayton JW, Klass DB, et al. Lithium in chronic schizophrenia. *Am J Psychiatry* 1979;136:454–455.

172. Small JG, Kellams JJ, Milstein V, et al. A placebo-controlled study of lithium combined with neuroleptics in chronic schizophrenic patients. *Am J Psychiatry* 1975;132:1315–1317.

173. Schulz SC, Thompson PA, Jacobs M, et al. Lithium augmentation fails to reduce symptoms in poorly responsive schizophrenic outpatients. *J Clin Psychiatry* 1999;60:366–372.

174. Wilson WH. Addition of lithium to haloperidol in non-affective, antipsychotic non-responsive schizophrenia: a double blind, placebo controlled, parallel design clinical trial. *Psychopharmacology* 1993;111:359–366.

175. Christison GW, Kirch DG, Wyatt RJ. When symptoms persist: choosing among alternative somatic treatments for schizophrenia. *Schizophr Bull* 1991;17:217–245.

176. Hogarty GE, McEvoy JP, Ulrich RF, et al. Pharmacotherapy of impaired affect in recovering schizophrenic patients. *Arch Gen Psychiatry* 1995;52:29.

177. Luchins DJ. Carbamazepine in violent non-epileptic schizophrenics. *Psychopharmacol Bull* 1984;20:569–571.

178. Neppe VM. Carbamazepine as adjunctive treatment in nonepileptic chronic inpatients with EEG temporal lobe abnormalities. *J Clin Psychiatry* 1983;44:326–331.

179. Okuma T, Yamashita I, Takahashi R, et al. A double-blind study of adjunctive carbamazepine versus placebo on excited states of schizophrenic and schizoaffective disorders. *Acta Psychiatr Scand* 1989;80:250–259.

180. Goff D, Baldessarini R. Antipsychotics. In: Ciraulo D, Shader R, Greenblatt D, et al., eds. *Drug interactions in psychiatry*, first ed. Baltimore: Williams & Wilkins, 1995:129–174.

181. Arana GW, Goff DC, Friedman H, et al. Does carbamazepine-induced reduction of plasma haloperidol levels worsen psychotic symptoms? *Am J Psychiatry* 1986;143:650–651.

182. Wassef AA, Dott SG, Harris A, et al. Randomized, placebo-controlled pilot study of divalproex sodium in the treatment of acute exacerbations of chronic schizophrenia. *J Clin Psychopharmacol* 2000;20:357–361.

183. Ko GN, Korpi ER, Freed WJ, et al. Effect of valproic acid on behavior and plasma amino acid concentrations in chronic schizophrenic patients. *Biol Psychiatry* 1985;20:209–215.

184. Wassef AA, Dott SG, Harris A, et al. Critical review of GABA-ergic drugs in the treatment of schizophrenia. *J Clin Psychopharmacol* 1999;19:222–232.

185. Arana GW, Ornsteen ML, Kanter F, et al. The use of benzodiazepines for psychotic disorders: a literature review and preliminary clinical findings. *Psychopharmacol Bull* 1986;22:77–87.

186. Wolkowitz OM, Pickar D. Benzodiazepines in the treatment of schizophrenia: a review and reappraisal. *Am J Psychiatry* 1991;148:714–726.

187. Wolkowitz OM, Breier A, Doran A, et al. Alprazolam augmentation of the antipsychotic effects of fluphenazine in schizophrenic patients. Preliminary results. *Arch Gen Psychiatry* 1988;45:664–671.

188. Fink M, Sackeim HA. Convulsive therapy in schizophrenia? *Schizophr Bull* 1996;22:27–39.

189. Baker AA, Bird G, Lavin NI, et al. ECT in schizophrenia. *J Ment Sci* 1960;106:1506–1511.

190. Kino FF, Thorpe FT. Electrical convulsion therapy in 500 selected psychotics. *J Ment Sci* 1946;92:138–145.

191. Zeifert M. Results obtained from the administration of 12,000 doses of Metrazol to mental patients. *Psychiat Quart* 1941;15:772–778.

192. Abraham KR, Kulhara P. The efficacy of electroconvulsive therapy in the treatment of schizophrenia. A comparative study. *Br J Psychiatry* 1987;151:152–155.

193. Brandon S, Cowley P, McDonald C, et al. Leicester ECT trial: results in schizophrenia. *Br J Psychiatry* 1985;146:177–183.

194. Taylor P, Fleminger JJ. ECT for schizophrenia. *Lancet* 1980;1:1380–1382.

195. Dodwell D, Goldberg D. A study of factors associated with response to electroconvulsive therapy in patients with schizophrenic symptoms. *Br J Psychiatry* 1989;154:635–639.

196. Herzberg F. Prognostic variables for electro-shock therapy. *J Gen Psychol* 1954;50:79–86.

197. Kalinowsky LB, Worthing HJ. Results with electric convulsant treatment in 200 cases of schizophrenia. *Psychiat Quart* 1943;17:144–153.

198. Lowinger L, Huddleson JH. Outcome in dementia praecox under electro-shock therapy as related to mode of onset and to number of convulsions induced. *J Nerv Ment Dis* 1945;102:243–246.

199. Landy DA. Combined use of clozapine and electroconvulsive therapy. *Convulsive Ther* 1991;7:218–221.

200. Safferman AZ, Munne R. Combining clozapine with ECT. *Convulsive Ther* 1992;8:141–143.

201. George MS, Lisanby SH, Sackeim HA. Transcranial magnetic stimulation: applications in neuropsychiatry. *Arch Gen Psychiatry* 1999;56:300–311.

202. Klein E, Kreinin I, Chistyakov A, et al. Therapeutic efficacy of right prefrontal slow repetitive transcranial magnetic stimulation in major depression: a double-blind controlled study. *Arch Gen Psychiatry* 1999;56:315–320.

203. Pascual-Leone A, Rubio B, Pallardo F, et al. Rapid-rate transcranial magnetic stimulation of left dorsolateral prefrontal cortex in drug-resistant depression. *Lancet* 1996;348:233–237.

204. Hoffman RE, Boutros NN, Hu S, et al. Transcranial magnetic stimulation and auditory hallucinations in schizophrenia. *Lancet* 2000;355:1073–1075.

205. Garety PA, Fowler D, Kuipers E. Cognitive-behavioral therapy for medication-resistant symptoms. *Schizophr Bull* 2000;26:73–86.

206. Garety PA, Kuipers L, Fowler D, et al. Cognitive behavioural therapy for drug-resistant psychosis. *Br J Med Psychol* 1994;67:259–271.

207. Kuipers E, Garety P, Fowler D, et al. London-East Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis. I: effects of the treatment phase. *Br J Psychiatry* 1997;171:319–327.

208. Tarrier N, Beckett R, Harwood S, et al. A trial of two cognitive-behavioural methods of treating drug-resistant residual psychotic symptoms in schizophrenic patients: I. Outcome. *Br J Psychiatry* 1993;162:524–532.

209. Tarrier N, Yusupoff L, Kinney C, et al. Randomised controlled trial of intensive cognitive behaviour therapy for patients with chronic schizophrenia. *Br Med J* 1998;317:303–307.

210. Gould RA, Mueser KT, Bolten E, et al. Cognitive therapy for psychosis in schizophrenia: an effect size analysis. *Schizophr Res* 2001;48:335–342.

211. Marder SR, Davis JM, Chouinard G. The effects of risperidone on the five dimensions of schizophrenia derived by factor analysis: combined results of the North American trials. *J Clin Psychiatry* 1997;58:538–546.

212. Moller HJ. Neuroleptic treatment of negative symptoms in schizophrenic patients. Efficacy problems and methodological difficulties. *Eur Neuropsychopharmacol* 1993;3:1–11.

213. Tollefson GD, Sanger TM. Negative symptoms: a path analytic approach to a double-blind, placebo- and haloperidol-controlled clinical trial with olanzapine. *Am J Psychiatry* 1997;154:466–474.

214. Conley R, Gounaris C, Tamminga C. Clozapine response varies in deficit versus non-deficit schizophrenic subjects. *Biol Psychiatry* 1994;35:746–747.

215. Goff DC, Evins AE. Negative symptoms in schizophrenia: neurobiological models and treatment response. *Harv Rev Psychiatry* 1998;6:59–77.

216. Gelders YG. Thymosthenic agents, a novel approach in the treatment of schizophrenia. *Br J Psychiatry (Suppl)* 1989;33–36.

217. Svensson TH, Mathe JM, Andersson JL, et al. Mode of action of atypical neuroleptics in relation to the phencyclidine model of schizophrenia: role of 5-HT2 receptor and alpha 1-adrenoceptor antagonism. *J Clin Psychopharmacol* 1995;15:11S–18S.

218. Svensson TH, Tung CS, Grenhoff J. The 5-HT2 antagonist ritanserin blocks the effect of pre-frontal cortex inactivation on rat A10 dopamine neurons in vivo. *Acta Physiol Scand* 1989;136:497–498.

219. Goff DC, Brotman AW, Waites M, et al. Trial of fluoxetine added to neuroleptics for treatment-resistant schizophrenic patients. *Am J Psychiatry* 1990;147:492–494.

220. Evins A, Goff D. Adjunctive antidepressant drug therapies in the treatment of negative symptoms of schizophrenia. *CNS Drugs* 1996;6:130–147.

221. Goff DC, Midha KK, Sarid-Segal O, et al. A placebo-controlled trial of fluoxetine added to neuroleptic in patients with schizophrenia. *Psychopharmacology* 1995;117:417–423.

222. Lee MS, Kim YK, Lee SK, et al. A double-blind study of adjunctive sertraline in haloperidol-stabilized patients with chronic schizophrenia. *J Clin Psychopharmacol* 1998;18:399–403.

223. Buchanan RW, Kirkpatrick B, Bryant N, et al. Fluoxetine augmentation of clozapine treatment in patients with schizophrenia. *Am J Psychiatry* 1996;153:1625–1627.

224. McEvoy JP. The clinical use of anticholinergic drugs as treatment for extrapyramidal side effects of neuroleptic drugs. *J Clin Psychopharmacol* 1983;3:288–302.

225. Tandon R, Greden JF, Silk KR. Treatment of negative schizophrenic symptoms with trihexyphenidyl. *J Clin Psychopharmacol* 1988;8:212–215.

226. Gerlach J, Rasmussen PT, Hansen L, et al. Antiparkinsonian agents and long-term neuroleptic treatment. Effect of G 31.406, orphenadrine, and placebo on parkinsonism, schizophrenic symptoms, depression and anxiety. *Acta Psychiatr Scand* 1977;55:251–260.

227. Goff DC, Arana GW, Greenblatt DJ, et al. The effect of benztropine on haloperidol-induced dystonia, clinical efficacy and pharmacokinetics: a prospective, double-blind trial. *J Clin Psychopharmacol* 1991;11:106–112.

228. Johnstone EC, Crow TJ, Ferrier IN, et al. Adverse effects of anticholinergic medication on positive schizophrenic symptoms. *Psychol Med* 1983;13:513–527.

229. Tandon R, Greden JF. Cholinergic hyperactivity and negative schizophrenic symptoms. A model of cholinergic/dopaminergic interactions in schizophrenia. *Arch Gen Psychiatry* 1989;46:745–753.

230. Tandon R, Mann NA, Eisner WH, et al. Effect of anticholinergic medication on positive and negative symptoms in medication-free schizophrenic patients. *Psychiatr Res* 1990;31:235–241.

231. Goff DC, Amico E, Dreyfuss D, et al. A placebo-controlled trial of trihexyphenidyl in unmedicated patients with schizophrenia. *Am J Psychiatry* 1994;151:429–431.

232. Baker LA, Cheng LY, Amara IB. The withdrawal of benztropine mesylate in chronic schizophrenic patients. *Br J Psychiatry* 1983;143:584–590.

233. Strauss ME, Reynolds KS, Jayaram G, et al. Effects of anticholinergic medication on memory in schizophrenia. *Schizophr Res* 1990;3:127–129.

234. Mathew RJ, Wilson WH. Changes in cerebral blood flow and mental state after amphetamine challenge in schizophrenic patients. *Neuropsychobiology* 1989;21:117–123.

235. Sanfilipo M, Wolkin A, Angrist B, et al. Amphetamine and negative symptoms of schizophrenia. *Psychopharmacology* 1996;123:211–214.

236. Van Kammen DP, Boronow JJ. Dextro-amphetamine diminishes negative symptoms in schizophrenia. *Int Clin Psychopharmacol* 1988;3:111–121.

237. Wolkin A, Angrist B, Wolf A, et al. Effects of amphetamine on local cerebral metabolism in normal and schizophrenic subjects as determined by positron emission tomography. *Psychopharmacology* 1987;92:241–246.

238. Casey JF, Hollister LE, Klett CJ, et al. Combined drug therapy of chronic schizophrenics: controlled evaluation of placebo, dexto-amphetamine, imipramine, isocarboxazid and trifluoperazine added to maintenance doses of chlorpromazine. *Am J Psychiatry* 1961;117:997–1003.

239. Goff DC. Glutamate receptors in schizophrenia and antipsychotic drugs. In: Lidow MS, ed. *Neurotransmitter receptors in actions of antipsychotic medications*. New York: CRC Press, 2000:121–136.

240. Goff DC, Tsai G, Levitt J, et al. A placebo-controlled trial of D-cycloserine added to conventional neuroleptics in patients with schizophrenia. *Arch Gen Psychiatry* 1999;56:21–27.

241. Heresco-Levy U, Javitt DC, Ermilov M, et al. Efficacy of high-dose glycine in the treatment of enduring negative symptoms of schizophrenia. *Arch Gen Psychiatry* 1999;56:29–36.

242. Tsai GE, Yang P, Chung LC, et al. D-serine added to clozapine for the treatment of schizophrenia. *Am J Psychiatry* 1999;156:1822–1825.

243. Evins AE, Fitzgerald SM, Wine L, et al. Placebo-controlled trial of glycine added to clozapine in schizophrenia. *Am J Psychiatry* 2000;157:826–828.

244. Goff DC, Henderson DC, Evins AE, et al. A placebo-controlled crossover trial of D-cycloserine added to clozapine in patients with schizophrenia. *Biol Psychiatry* 1999;45:512–514.

245. Potkin SG, Jin Y, Bunney BG, et al. Effect of clozapine and adjunctive high-dose glycine in treatment-resistant schizophrenia. *Am J Psychiatry* 1999;156:145–147.

246. Kopelowicz A, Liberman RP, Mintz J, et al. Comparison of efficacy of social skills training for deficit and nondeficit negative symptoms in schizophrenia. *Am J Psychiatry* 1997;154:424–425.

247. Cook JA, Razzano L. Vocational rehabilitation for persons with schizophrenia: recent research and implications for practice. *Schizophr Bull* 2000;26:87–103.

248. Birchwood M, Mason R, MacMillan F, et al. Depression, demoralization and control over psychotic illness: a comparison of depressed and non-depressed patients with a chronic psychosis. *Psychol Med* 1993;23:387–395.

249. Roy A, Thompson R, Kennedy S. Depression in chronic schizophrenia. *Br J Psychiatry* 1983;142:465–470.

250. Tollefson GD, Andersen SW, Tran PV. The course of depressive symptoms in predicting relapse in schizophrenia: a double-blind, randomized comparison of olanzapine and risperidone. *Biol Psychiatry* 1999;46:365–373.

251. Wassink TH, Flaum M, Nopoulos P, et al. Prevalence of depressive symptoms early in the course of schizophrenia. *Am J Psychiatry* 1999;156:315–316.

252. Johnson DA. Studies of depressive symptoms in schizophrenia. *Br J Psychiatry* 1981;139:89–101.

253. Koreen AR, Siris SG, Chakos M, et al. Depression in first-episode schizophrenia. *Am J Psychiatry* 1993;150:1643–1648.

254. Siris SG. Diagnosis of secondary depression in schizophrenia: implications for DSM-IV. *Schizophr Bull* 1991;17:75–98.

255. Newcomer JW, Faustman WO, Yeh W, et al. Distinguishing depression and negative symptoms in unmedicated patients with schizophrenia. *Psychiatr Res* 1990;31:243–250.

256. Prosser ES, Csernansky JG, Kaplan J, et al. Depression, parkinsonian symptoms, and negative symptoms in schizophrenics treated with neuroleptics. *J Nerv Ment Dis* 1987;175:100–105.

257. Knights A, Hirsch SR. “Revealed” depression and drug treatment for schizophrenia. *Arch Gen Psychiatry* 1981;38:806–811.

258. Kramer MS, Vogel WH, DiJohnson C, et al. Antidepressants in ‘depressed' schizophrenic inpatients. A controlled trial. *Arch Gen Psychiatry* 1989;46:922–928.

259. Harrow M, Yonan CA, Sands JR, et al. Depression in schizophrenia: are neuroleptics, akinesia, or anhedonia involved? *Schizophr Bull* 1994;20:327–338.

260. Van Putten T, May RP. “Akinetic depression” in schizophrenia. *Arch Gen Psychiatry* 1978;35:1101–1107.

261. Kane J, Rifkin A, Quitkin F, et al. Extrapyramidal side effects with lithium treatment. *Am J Psychiatry* 1978;135:851–853.

262. Siris SG, Bermanzohn PC, Mason SE, et al. Maintenance imipramine therapy for secondary depression in schizophrenia. A controlled trial. *Arch Gen Psychiatry* 1994;51:109–115.

263. Siris SG, Morgan V, Fagerstrom R, et al. Adjunctive imipramine in the treatment of postpsychotic depression. A controlled trial. *Arch Gen Psychiatry* 1987;44:533–539.

264. Plasky P. Antidepressant usage in schizophrenia. *Schizophr Bull* 1991;17:649–657.

265. Siris SG, Bermanzohn PC, Gonzalez A, et al. The use of antidepressants for negative symptoms in a subset of schizophrenic patients. *Psychopharmacol Bull* 1991;27:331–335.

266. Mohamed S, Paulsen JS, O'Leary D, et al. Generalized cognitive deficits in schizophrenia: a study of first-episode patients. *Arch Gen Psychiatry* 1999;56:749–754.

267. Aleman A, Hijman R, de Haan EH, et al. Memory impairment in schizophrenia: a meta-analysis. *Am J Psychiatry* 1999;156:1358–1366.

268. Gold S, Arndt S, Nopoulos P, et al. Longitudinal study of cognitive function in first-episode and recent-onset schizophrenia. *Am J Psychiatry* 1999;156:1342–1348.

269. Harvey PD, Silverman JM, Mohs RC, et al. Cognitive decline in late-life schizophrenia: a longitudinal study of geriatric chronically hospitalized patients. *Biol Psychiatry* 1999;45:32–40.

270. Buchanan RW, Strauss ME, Kirkpatrick B, et al. Neuropsychological impairments in deficit vs nondeficit forms of schizophrenia. *Arch Gen Psychiatry* 1994;51:804–811.

271. Waddington JL, Youssef HA, Kinsella A. Cognitive dysfunction in schizophrenia followed up over 5 years, and its longitudinal relationship to the emergence of tardive dyskinesia. *Psychol Med* 1990;20:835–842.

272. King DJ. The effect of neuroleptics on cognitive and psychomotor function. *Br J Psychiatry* 1990;157:799–811.

273. Lidow MS, Elsworth JD, P.S. Down-regulation of the D1 and D5 dopamine receptors in the primate prefrontal cortex by chronic treatment with antipsychotic drugs. *J Pharmacol Exp Ther* 1997;281:597–603.

274. Castner SA, Williams GV, Goldman-Rakic PS. Reversal of antipsychotic-induced working memory deficits by short-term dopamine D1 receptor stimulation. *Science* 2000;287:2020–2022.

275. Saletu B, Grunberger J, Linzmayer L, et al. Comparative placebo-controlled pharmacodynamic studies with zotepine and clozapine utilizing pharmaco-EEG and psychometry. *Pharmacopsychiatry* 1987;20:12–27.

276. Goldberg TE, Greenberg RD, Griffin SJ, et al. The effect of clozapine on cognition and psychiatric symptoms in patients with schizophrenia. *Br J Psychiatry* 1993;162:43–48.

277. Meltzer HY, McGurk SR. The effects of clozapine, risperidone, and olanzapine on cognitive function in schizophrenia. *Schizophr Bull* 1999;25:233–255.

278. Purdon SE, Jones BD, Stip E, et al. Neuropsychological change in early phase schizophrenia during 12 months of treatment with olanzapine, risperidone, or haloperidol. The Canadian Collaborative Group for research in schizophrenia. *Arch Gen Psychiatry* 2000;57:249–258.

279. Goff DC, Bagnell AL, Perlis RH. Glutamatergic strategies for cognitive impairment in schizophrenia. *Psychiatr Ann* 1999;29:649–654.

280. Tsai G, Yang P, Chung LC, et al. D-Serine added to antipsychotics for the treatment of schizophrenia. *Biol Psychiatry* 1998;44:1081–1089.

281. Krystal JH, Karper LP, Seibyl JP, et al. Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch Gen Psychiatry* 1994;51:199–214.

282. Malhotra AK, Pinals DA, Adler CM, et al. Ketamine-induced exacerbation of psychotic symptoms and cognitive impairment in neuroleptic-free schizophrenics. *Neuropsychopharmacology* 1997;17:141–150.

283. Staubli U, Rogers G, Lynch G. Facilitation of glutamate receptors enhances memory. *Proc Natl Acad Sci USA* 1994;91:777–781.

284. Goff D, Berman I, Posever T, et al. A preliminary dose-escalation trial of CX 516 (ampakine) added to clozapine in schizophrenia. *Schizophr Res* 1999;36:280.

285. Cassidy JJ, Easton M, Capelli C, et al. Cognitive remediation of persons with severe and persistent mental illness. *Psychiat Quart* 1996;67:313–321.

286. Corrigan PW, Hirschbeck JN, Wolfe M. Memory and vigilance training to improve social perception in schizophrenia. *Schizophr Res* 1995;17:257–265.

287. Brenner H, Roder V, Hodel B, et al. *Integrated psychological therapy for schizophrenia patients*. Toronto: Hogrefe & Huber, 1994.

288. Spaulding WD, Reed D, Sullivan M, et al. Effects of cognitive treatment in psychiatric rehabilitation. *Schizophr Bull* 1999;25:657–676.

289. Hogarty GE, Flesher S. Practice principles of cognitive enhancement therapy for schizophrenia. *Schizophr Bull* 1999;25:693–708.

290. Hogarty GE, Flesher S. Developmental theory for a cognitive enhancement therapy of schizophrenia. *Schizophr Bull* 1999;25:677–692.

291. Van Tol HH, Bunzow JR, Guan HC, et al. Cloning of the gene for a human dopamine D4 receptor with high affinity for the antipsychotic clozapine. *Nature* 1991;350:610–614.

292. Seeman P, Guan HC, Van Tol HH. Dopamine D4 receptors elevated in schizophrenia. *Nature* 1993;365:441–445.

293. Danysz W. Sonepiprazole. *Curr Opin CPNS Invest Drugs* 2000;2:97–104.

294. Mansbach RS, Brooks EW, Sanner MA, et al. Selective dopamine D4 receptor antagonists reverse apomorphine-induced blockade of prepulse inhibition. *Psychopharmacology* 1998;135:194–200.

295. Feldpausch DL, Needham LM, Stone MP, et al. The role of dopamine D4 receptor in the induction of behavioral sensitization to amphetamine and accompanying biochemical and molecular adaptations. *J Pharmacol Exp Ther* 1998;286:497–508.

296. Kramer MS, Last B, Getson A, et al. The effects of a selective D4 dopamine receptor antagonist (L-745,870) in acutely psychotic inpatients with schizophrenia. D4 Dopamine Antagonist Group. *Arch Gen Psychiatry* 1997;54:567–572.

297. Bristow LJ, Kramer MS, Kulagowski J, et al. Schizophrenia and L-745,870, a novel dopamine D4 receptor antagonist. *Trends Pharmacol Sci* 1997;18:186–188.

298. Mansbach RS, Brooks EW, Sanner MA, et al. Selective dopamine D4 receptor antagonists reverse apomorphine-induced blockade of prepulse inhibition. *Psychopharmacology* 1998;135:194–200.

299. Coward D, Dixon K, Enz A, et al. Partial brain dopamine D_2 receptor agonists in the treatment of schizophrenia. *Psychopharmacol Bull* 1989;25:393–397.

300. Kinon BJ, Lieberman JA. Mechanisms of action of atypical antipsychotic drugs: a critical analysis. *Psychopharmacology* 1996;124:2–34.

301. Ozdemir V. Aripiprazole. *Curr Opin CPNS Invest Drugs* 2000;2:105–111.

302. Kikuchi T, Tottori K, Uwahodo Y, et al. 7-(4-[4-(2,3-Dichlorophenyl)-1-piperazinyl]butyloxy)-3,4-dihydro-2(1H)-quinolinone(OPC–14597), a new putative antipsychotic drug with both presynaptic dopamine autoreceptor agonistic activity and postsynaptic D2 receptor antagonistic activity. *J Pharmacol Exp Ther* 1995;274:329–336.

303. Semba J, Watanabe A, Kito S, et al. Behavioural and neurochemical effects of OPC-14597, a novel antipsychotic drug, on dopaminergic mechanisms in rat brain. *Neuropharmacology* 1995;34:785–791.

304. Lawler CP, Prioleau C, Lewis MM, et al. Interactions of the novel antipsychotic aripiprazole (OPC-14597) with dopamine and serotonin receptor subtypes. *Neuropsychopharmacology* 1999;20:612–627.

305. Momiyama T, Amano T, Todo N, et al. Inhibition by a putative antipsychotic quinolinone derivative (OPC-14597) of dopaminergic neurons in the ventral tegmental area. *Eur J Pharmacol* 1996;310:1–8.

306. Toru M, Miura S, Kudo Y. Clinical experiences of OPC-14597, a dopamine autoreceptor agonist in schizophrenic patients. *Neuropsychopharmacology* 1994;10:122S.

307. Petrie JL, Saha AR, McEvoy JP. Acute and long-term efficacy and safety of aripiprazole: a new atypical antipsychotic. *Schizophr Res* 1998;29:155.

308. Saha AR, Petrie JL, Ali MW. Safety and efficacy profile of aripiprazole, a novel antipsychotic. *Schizophr Res* 1999;36:295.

309. Wright JL, Caprathe BW, Downing DM, et al. The discovery and structure-activity relationships of 1,2,3,6-tetrahydro-4-phenyl-1-[(arylcyclohexenyl)alkyl]pyridines. Dopamine autoreceptor agonists and potential antipsychotic agents. *J Med Chem* 1994;37:3523–3533.

310. Pugsley TA, Davis MD, Akunne HC, et al. Cl-1007, a dopamine partial agonist and potential antipsychotic agent. I. neurochemical effects. *J Pharmacol Exp Ther* 1995;274:898–911.

311. Meltzer LT, Christoffersen CL, Corbin AE, et al. CI-1007, a dopamine partial agonist and potential antipsychotic agent. II. Neurophysiological and behavioral effects. *J Pharmacol Exp Ther* 1995;274:912–920.

312. Schmidt CJ, Sorensen SM, Kehne JH, et al. The role of 5-HT2A receptors in antipsychotic activity. *Life Sci* 1995;56:2209–2222.

313. Meltzer HY. Pre-clinical pharmacology of atypical antipsychotic drugs: a selective review. *Br J Psychiatry (Suppl)* 1996;29:23–31.

314. Ugedo L, Grenhoff J, Svensson TH. Ritanserin, a 5-HT2 receptor antagonist, activates midbrain dopamine neurons by blocking serotonergic inhibition. *Psychopharmacology* 1989;98:45–50.

315. Schmidt CJ, Fadayel GM, Sullivan CK, et al. 5-HT2 receptors exert a state-dependent regulation of dopaminergic function: studies with MDL 100,907 and the amphetamine analogue, 3,4-methylenedioxymethamphetamine. *Eur J Pharmacol* 1992;223:65–74.

316. O'Neill MF, Heron-Maxwell CL, Shaw G. 5-HT2 receptor antagonism reduces hyperactivity induced by amphetamine, cocaine, and MK-801 but not D1 agonist C-APB. *Pharmacol Biochem Behav* 1999;63:237–243.

317. Sorensen SM, Kehne JH, Fadayel GM, et al. Characterization of the 5-HT2 receptor antagonist MDL 100907 as a putative atypical antipsychotic: behavioral, electrophysiological and neurochemical studies. *J Pharmacol Exp Ther* 1993;266:684–691.

318. Arnt J. Differential effects of classical and newer antipsychotics on the hypermotility induced by two dose levels of D-amphetamine. *Eur J Pharmacol* 1995;283:55–62.

319. Maurel-Remy S, Bervoets K, Millan MJ. Blockade of phencyclidine-induced hyperlocomotion by clozapine and MDL 100,907 in rats reflects antagonism of 5-HT2A receptors. *Eur J Pharmacol* 1995;280:R9–11.

320. Gleason SD, Shannon HE. Blockade of phencyclidine-induced hyperlocomotion by olanzapine, clozapine and serotonin receptor subtype selective antagonists in mice. *Psychopharmacology* 1997;129:79–84.

321. Carlsson ML, Martin P, Nilsson M, et al. The 5-HT2A receptor antagonist M100907 is more effective in counteracting NMDA antagonist- than dopamine agonist-induced hyperactivity in mice. *J Neural Transm* 1999;106:123–129.

322. Schmidt CJ, Fadayel GM. The selective 5-HT2A receptor antagonist, MDL 100,907, increases dopamine efflux in the prefrontal cortex of the rat. *Eur J Pharmacol* 1995;273:273–279.

323. Schmidt CJ, Fadayel GM. Regional effects of MK-801 on dopamine release effects of competitive NMDA or 5-HT2A receptor blocade. *J Pharmacol Exp Ther* 1996;277:1541–1549.

324. Varty GB, Bakshi VP, Geyer MA. M100907, a serotonin 5-HT2A receptor antagonist and putative antipsychotic, blocks dizocilpine-induced prepulse inhibition deficits in Sprague-Dawley and Wistar rats. *Neuropsychopharmacology* 1999;20:311–321.

325. Wang RY, Liang X. M100907 and clozapine, but not haloperidol or raclopride, prevent phencyclidine-induced blockade of NMDA responses in pyramidal neurons of the rat medial prefrontal cortical slice. *Neuropsychopharmacology* 1998;19:74–85.

326. Carlsson A. Focusing on dopaminergic stabilizers and 5-HT2A receptor antagonists. *Curr Opin CPNS Invest Drugs* 2000;2:22–24.

327. Newman-Tancredi A, Chaput C, Verriele L, et al. Clozapine is a partial agonist at cloned, human serotonin 5-HT1A receptors. *Neuropharmacology* 1996;35:119–121.

328. Evenden JL. Effects of 8-hydroxy-2-(di-n-propylamino)tetralin (8-OH-DPAT) after repeated administration on a conditioned avoidance response (CAR) in the rat. *Psychopharmacology* 1992;109:134–144.

329. Lucas G, Bonhomme N, De Deurwaerdere P, et al. 8-OH-DPAT, a 5-HT1A agonist and ritanserin, a 5-HT2A/C antagonist, reverse haloperidol-induced catalepsy in rats independently of striatal dopamine release. *Psychopharmacology* 1997;131:57–63.

330. Burnet PW, Eastwood SL, Harrison PJ. 5-HT1A and 5-HT2A receptor mRNAs and binding site densities are differentially altered in schizophrenia. *Neuropsychopharmacology* 1996;15:442–455.

331. Simpson MD, Lubman DI, Slater P, et al. Autoradiography with [3H]8-OH-DPAT reveals increases in 5-HT(1A) receptors in ventral prefrontal cortex in schizophrenia. *Biol Psychiatry* 1996;39:919–928.

332. Millan MJ, Schreiber R, Dekeyne A, et al. S 16924 ((R)-2-[1-[2-(2,3-dihydro-benzo[1,4] dioxin-5-yloxy)-ethyl]-pyrrolidin-3yl]-1-(4-fluoro-phenyl)-ethanone), a novel, potential antipsychotic with marked serotonin (5-HT)1A agonist properties: II. Functional profile in comparison to clozapine and haloperidol. *J Pharmacol Exp Ther* 1998;286:1356–1373.

333. Millan MJ, Gobert A, Newman-Tancredi A, et al. S 16924 ((R)-2-[1-[2-(2,3-dihydro-benzo[1,4] dioxin-5-yloxy)-ethyl]-pyrrolidin-3yl]-1-(4-fluoro-phenyl)-ethanone), a novel, potential antipsychotic with marked serotonin (5-HT)1A agonist properties: I. Receptorial and neurochemical profile in comparison with clozapine and haloperidol. *J Pharmacol Exp Ther* 1998;286:1341–1355.

334. Bengtsson HJ, Kullberg A, Millan MJ, et al. The role of 5-HT1A autoreceptors and alpha1-adrenoceptors in the modulation of 5-HT release—III. Clozapine and the novel putative antipsychotic S 16924. *Neuropharmacology* 1998;37:349–356.

335. Cummings JL, Back C. The cholinergic hypothesis of neuropsychiatric symptoms in Alzheimer's disease. *Am J Geriatr Psychiatry* 1998;6:S64–S78.

336. Kaufer D, Cummings JL, Christine D. Differential neuropsychiatric symptom responses to tacrine in Alzheimer's disease: relationship to dementia severity. *J Neuropsychiatry Clin Neurosci* 1998;10:55–63.

337. Levy ML, Cummings JL, Kahn-Rose R. Neuropsychiatric symptoms and cholinergic therapy for Alzheimer's disease. *Gerontology* 1999;45(Suppl 1):15–22.

338. Tune LE, Sunderland T. New cholinergic therapies: treatment tools for the psychiatrist. *J Clin Psychiatry* 1998;59(Suppl 13):31–35.

339. Bodick NC, Offen WW, Levey AI, et al. Effects of xanomeline, a selective muscarinic receptor agonist, on cognitive function and behavioral symptoms in Alzheimer disease. *Arch Neurol* 1997;54:465–473.

340. Perry EK, Perry RH. Acetylcholine and hallucinations: disease-related compared to drug-induced alterations in human consciousness. *Brain Cogn* 1995;28:240–258.

341. Bymaster FP, Shannon HE, Rasmussen K, et al. Potential role of muscarinic receptors in schizophrenia. *Life Sci* 1999;64:527–534.

342. Sauerberg P, Jeppesen L, Olesen PH, et al. Muscarinic agonists with antipsychotic-like activity: structure-activity relationships of 1,2,5-thiadiazole analogues with functional dopamine antagonist activity. *J Med Chem* 1998;41:4378–4384.

343. Bymaster FP, Shannon HE, Rasmussen K, et al. Unexpected antipsychotic-like activity with the muscarinic receptor ligand (5R,6R)6-(3-propylthio-1,2,5-thiadiazol-4-yl)-1-azabicyclo[3.2.1]octane. *Eur J Pharmacol* 1998;356:109–119.

344. Chiodo LA, Bunney BS. Typical and atypical neuroleptics: differential effects of chronic administration on the activity of A9 and A10 midbrain dopaminergic neurons. *J Neurosci* 1983;3:1607–1619.

345. Skarsfeldt T. Differential effects of repeated administration of novel antipsychotic drugs on the activity of midbrain dopamine neurons in the rat. *Eur J Pharmacol* 1995;281:289–294.

346. Javitt DC, Zukin SR. Recent advances in the phencyclidine model of schizophrenia. *Am J Psychiatry* 1991;148:1301–1308.

347. Malhotra AK, Pinals DA, Weingartner H, et al. NMDA receptor function and human cognition. The effects of ketamine in healthy volunteers. *Neuropsychopharmacology* 1996;14:301–307.

348. Lahti AC, Holcomb HH, Medoff DR, et al. Ketamine activates psychosis and alters limbic blood flow in schizophrenia. *Neuroreport* 1995;6:869–872.

349. Lahti AC, Koffel B, LaPorte D, et al. Subanesthetic doses of ketamine stimulate psychosis in schizophrenia. *Neuropsychopharmacology* 1995;13:9–19.

350. Yenari MA, Bell TE, Kotake AN, et al. Dose escalation safety and tolerance study of the competitive NMDA antagonist selfotel (CGS 19755) in neurosurgery patients. *J Neuropharmacol* 1998;21:28–34.

351. Deutsch SI, Mastropaolo J, Schwartx BL, et al. A "glutamatergic hypothesis" of schizophrenia. Rationale for pharmacotherapy with glycine. *J Neuropharmacol* 1989;12:1–13.

352. Olney JW, Farber NB. Glutamate receptor dysfunction and schizophrenia. *Arch Gen Psychiatry* 1995;52:998–1007.

353. Coyle JT. The glutamatergic dysfunction hypothesis for schizophrenia. *Harv Rev Psychiatry* 1996;3:241–253.

354. Duncan GE, Moy SS, Knapp DJ, et al. Metabolic mapping of the rat brain after subanesthetic doses of ketamine: potential relevance to schizophrenia. *Brain Res* 1998;787:181–190.

355. Duncan GE, Miyamoto S, Leipzig JN, et al. Comparison of brain metabolic activity patterns induced by ketamine, MK-801 and amphetamine in rats: support for NMDA receptor involvement in responses to subanesthetic dose of ketamine. *Brain Res* 1999;843:171–183.

356. Miyamoto S, Leipzig JN, Lieberman JA, et al. Effects of ketamine, MK-801, and amphetamine on regional brain 2-deoxyglucose uptake in freely moving mice. *Neuropsychopharmacology* 2000;22:400–412.

357. Duncan GE, Leipzig JN, Mailman RB, et al. Differential effects of clozapine and haloperidol on ketamine-induced brain metabolic activation. *Brain Res* 1998;812:65–75.

358. Duncan GE, Miyamoto S, Leipzig JN, et al. Comparison of the effects of clozapine, risperidone, and olanzapine on ketamine-induced alterations in regional brain metabolism. *J Pharmacol Exp Ther* 2000;293:8–14.

359. Bakshi VP, Geyer MA. Antagonism of phencyclidine-induced deficits in prepulse inhibition by the putative atypical antipsychotic olanzapine. *Psychopharmacology* 1995;122:198–201.

360. Bakshi VP, Swerdlow NR, Geyer MA. Clozapine antagonizes phencyclidine-induced deficits in sensorimotor gating of the startle response. *J Pharmacol Exp Ther* 1994;271:787–794.

361. Corbett R, Camacho F, Woods AT, et al. Antipsychotic agents antagonize non-competitive N-methyl-D-aspartate antagonist-induced behaviors. *Psychopharmacology* 1995;120:67–74.

362. Duncan GE, Sheitman BB, Lieberman JA. An integrated view of pathophysiological models of schizophrenia. *Brain Res* Rev 1999;29:250–264.

363. Leeson PD, Iversen LL. The glycine site on the NMDA receptor: structure-activity relationships and therapeutic potential. *J Med Chem* 1994;37:4053–4067.

364. Javitt DC, Sershen H, Hashim A, et al. Reversal of phencyclidine-induced hyperactivity by glycine and the glycine uptake inhibitor glycyldodecylamide. *Neuropsychopharmacology* 1997;17:202–204.

365. Rosse RB, Theut SK, Banay-Schwartz M, et al. Glycine adjuvant therapy to conventional neuroleptic treatment in schizophrenia: an open-label, pilot study. *J Neuropharmacol* 1989;12:416–424.

366. Costa J, Khaled E, Sramek J, et al. An open trial of glycine as an adjunct to neuroleptics in chronic treatment-refractory schizophrenics. *J Clin Psychopharmacol* 1990;10:71–72.

367. Heresco-Levy U, Javitt DC, Ermilov M, et al. Double-blind, placebo-controlled, crossover trial of glycine adjuvant therapy for treatment-resistant schizophrenia. *Br J Psychiatry* 1996;169:610–617.

368. Javitt DC, Zylberman I, Zukin SR, et al. Amelioration of negative symptoms in schizophrenia by glycine. *Am J Psychiatry* 1994;151:1234–1236.

369. van Berckel BN, Hijman R, van der Linden JA, et al. Efficacy and tolerance of D-cycloserine in drug-free schizophrenic patients. *Biol Psychiatry* 1996;40:1298–1300.

370. Goff DC, Tsai G, Manoach DS, et al. Dose-finding trial of D-cycloserine added to neuroleptics for negative symptoms in schizophrenia. *Am J Psychiatry* 1995;152:1213–1215.

371. Goff DC, Tsai G, Manoach DS, et al. D-cycloserine added to clozapine for patients with schizophrenia. *Am J Psychiatry* 1996;153:1628–1630.

372. Hashimoto A, Oka T. Free D-aspartate and D-serine in the mammalian brain and periphery. *Prog Neurobiol* 1997;52:325–353.

373. Wolosker H, Sheth KN, Takahashi M, et al. Purification of serine racemase: biosynthesis of the neuromodulator D-serine. *Proc Natl Acad Sci USA* 1999;96:721–725.

374. Bergeron R, Meyer TM, Coyle JT, et al. Modulation of N-methyl-D-aspartate receptor function by glycine transport. *Proc Natl Acad Sci USA* 1998;95:15730–15734.

375. Berger AJ, Dieudonne S, Ascher P. Glycine uptake governs glycine site occupancy at NMDA receptors of excitatory synapses. *J Neurophysiol* 1998;80:3336–3340.

376. Javitt DC, Frusciante M. Glycyldodecylamide, a phencyclidine behavioral antagonist, blocks cortical glycine uptake: implications for schizophrenia and substance abuse. *Psychopharmacology* 1997;129:96–98.

377. Moghaddam B, Adams B, Verman A, et al. Activation of glutamatergic neurotransmission by ketamine. A novel step in the pathway from NMDA receptor blockade to dopaminergic and cognitive disruptions associated with the prefrontal cortex. *J Neurosci* 1997;17:2921–2927.

378. Grunze HC, Rainnie DG, Hasselmo ME, et al. NMDA-dependent modulation of CA1 local circuit inhibition. *J Neurosci* 1996;16:2034–2043.

379. Attwell PJ, Singh KN, Jane DE, et al. Anticonvulsant and glutamate release-inhibiting properties of the highly potent metabotropic glutamate receptor agonist (2S,2′R, 3′R)-2-(2′,3′-dicarboxycyclopropyl)glycine (DCG-IV). *Brain Res* 1998;805:138–143.

380. Battaglia G, Monn JA, Schoepp DD. In vivo inhibition of veratridine-evoked release of striatal excitatory amino acids by the group II metabotropic glutamate receptor agonist LY354740 in rats. *Neurosci Lett* 1997;229:161–164.

381. Anand A, Charney DS, Oren DA, et al. Attenuation of the neuropsychiatric effects of ketamine with lamotrigine: support for hyperglutamatergic effects of N-methyl-D-aspartate receptor antagonists. *Arch Gen Psychiatry* 2000;57:270–276.

382. Moghaddam B, Adams BW. Reversal of phencyclidine effects by a group II metabotropic glutamate receptor agonist in rats. *Science* 1998;281:1349–1352.

383. Bubser M, Keseberg U, Notz PK, et al. Differential behavioral and neurochemical effects of competitive and non-competitive NMDA receptor antagonists in rats. *Eur J Pharmacol* 1992;229:75–82.

384. Hauber W, Andersen R. The non-NMDA glutamate receptor antagonist GYKI 52466 counteracts locomotor stimulation and anticataleptic activity induced by the NMDA antagonist dizocilpine. *Naunyn Sch Arch Pharmacol* 1993;348:486–490.

385. Willins DL, Narayanan S, Wallace LJ, et al. The role of dopamine and AMPA/kainate receptors in the nucleus accumbens in the hypermotility response to MK801. *Pharmacol Biochem Behav* 1993;46:881–887.

386. Sharp JW, Petersen DL, Langford MT. DNQX inhibits phencyclidine (PCP) and ketamine induction of the hsp 70 heat shock gene in the rat cingulate and retrosplenial cortex. *Brain Res* 1995;687:114–124.

387. Hampson RE, Rogers G, Lynch G, et al. Facilitative effects of the ampakine CX516 on short-term memory in rats: correlations with hippocampal neuronal activity. *J Neurosci* 1998;18:2748–2763.

388. Hampson RE, Rogers G, Lynch G, et al. Facilitative effects of the ampakine CX516 on short-term memory in rats: enhancement of delayed-nonmatch-to-sample performance. *J Neurosci* 1998;18:2740–2747.

389. Johnson SA, Luu NT, Herbst TA, et al. Synergistic interactions between ampakines and antipsychotic drugs. *J Pharmacol Exp Ther* 1999;289:392–397.

390. Manji HK, Bebchuk JM, Moore GJ, et al. Modulation of CNS signal transduction pathways and gene expression by mood-stabilizing agents: therapeutic implications. *J Clin Psychiatry* 1999;60(Suppl 2):27–39.

391. Manji HK, Lenox RH. Ziskind-Somerfeld Research Award. Protein kinase C signaling in the brain: molecular transduction of mood stabilization in the treatment of manic-depressive illness. *Biol Psychiatry* 1999;46:1328–1351.

392. Bebchuk JM, Arfken CL, Dolan-Manji S, et al. A preliminary investigation of a protein kinase C inhibitor in the treatment of acute mania. *Arch Gen Psychiatry* 2000;57:95–97.

393. Seeman MV. Current outcome in schizophrenia: women vs men. *Acta Psychiatr Scand* 1986;73:609–617.

394. Harris MJ, Jeste DV. Late-onset schizophrenia: an overview. *Schizophr Bull* 1988;14:39–55.

395. Lindamer LA, Lohr JB, Harris MJ, et al. Gender, estrogen, and schizophrenia. *Psychopharmacol Bull* 1997;33:221–228.

396. Hafner H, Maurer K, Loffler W, et al. The influence of age and sex on the onset and early course of schizophrenia. *Br J Psychiatry* 1993;162:80–86.

397. Hallonquist JD, Seeman MV, Lang M, et al. Variation in symptom severity over the menstrual cycle of schizophrenics. *Biol Psychiatry* 1993;33:207–209.

398. Riecher-Rossler A, Hafner H, Dutsch-Strobel A, et al. Further evidence for a specific role of estradiol in schizophrenia? *Biol Psychiatry* 1994;36:492–494.

399. Hafner H, Behrens S, De Vry J, et al. An animal model for the effects of estradiol on dopamine-mediated behavior: implications for sex differences in schizophrenia. *Psychiatr Res* 1991;38:125–134.

400. Maurice T, Phan VL, Urani A, et al. Neuroactive neurosteroids as endogenous effectors for the sigma1 (sigma1) receptor: pharmacologic evidence and therapeutic opportunities. *Jpn J Pharmacol* 1999;81:125–155.

401. Roberts E, Bologa L, Flood JF, et al. Effects of dehydroepiandrosterone and its sulfate on brain tissue in culture and on memory in mice. *Brain Res* 1987;406:357–362.

402. Bologa L, Sharma J, Roberts E. Dehydroepiandrosterone and its sulfated derivative reduce neuronal death and enhance astrocytic differentiation in brain cell cultures. *J Neurosci Res* 1987;17:225–234.

403. Compagnone NA, Mellon SH. Dehydroepiandrosterone: a potential signalling molecule for neocortical organization during development. *Proc Natl Acad Sci USA* 1998;95:4678–4683.

404. Mao X, Barger SW. Neuroprotection by dehydroepiandrosterone-sulfate: role of an NFkappaB-like factor. *Neuroreport* 1998;9:759–763.

405. Bastianetto S, Ramassamy C, Poirier J, et al. Dehydroepiandrosterone (DHEA) protects hippocampal cells from oxidative stress-induced damage. *Mol Brain Res* 1999;66:35–41.

406. Bergeron R, de Montigny C, Debonnel G. Potentiation of neuronal NMDA response induced by dehydroepiandrosterone and its suppression by progesterone: effects mediated via sigma receptors. *J Neurosci* 1996;16:1193–1202.

407. Debonnel G, Bergeron R, de Montigny C. Potentiation by dehydroepiandrosterone of the neuronal response to N-methyl-D-aspartate in the CA3 region of the rat dorsal hippocampus: an effect mediated via sigma receptors. *J Endocrinol (Suppl)* 1996;150:S33–S42.

408. Flood JF, Smith GE, Roberts E. Dehydroepiandrosterone and its sulfate enhance memory retention in mice. *Brain Res* 1988;447:269–278.

409. Flood JF, Roberts E. Dehydroepiandrosterone sulfate improves memory in aging mice. *Brain Res* 1988;448:178–181.

410. Flood JF, Morley JE, Roberts E. Memory-enhancing effects in male mice of pregnenolone and steroids metabolically derived from it. *Proc Natl Acad Sci USA* 1992;89:1567–1571.

411. Reddy DS, Kulkarni SK. The effects of neurosteroids on acquisition and retention of a modified passive-avoidance learning task in mice. *Brain Res* 1998;791:108–116.

412. Maurice T, Junien JL, Privat A. Dehydroepiandrosterone sulfate attenuates dizocilpine-induced learning impairment in mice via sigma 1-receptors. Behav *Brain Res* 1997;83:159–164.

413. Tourney G, Erb JL. Temporal variations in androgens and stress hormones in control and schizophrenic subjects. *Biol Psychiatry* 1979;14:395–404.

414. Strauss EB, Sands DE, Robibson AM, et al. Use of dehydroisoandrosterone in psychiatric treatment: a preliminary survey. *Br Med J* 1952;64–66.

415. Sands DE. Further studies on endocrine treatment in adolescence and early adult life. *J Ment Sci* 1954;100:211–219.

416. Strauss EB, Stevenson WAH. Use of dehydroisoandrosterone in psychiatric practice. *J Neurol Neurosurg Psychiat* 1955;18:137–144.

417. Wolkowitz OM, Reus VI, Keebler A, et al. Double-blind treatment of major depression with dehydroepiandrosterone. *Am J Psychiatry* 1999;156:646–649.

418. Horrobin DF. Schizophrenia as a prostaglandin deficiency disease. *Lancet* 1977;1:936–937.

419. Fenton WS, Hibbeln J, Knable M. Essential fatty acids, lipid membrane abnormalities, and the diagnosis and treatment of schizophrenia. *Biol Psychiatry* 2000;47:8–21.

420. Horrobin DF, Ally AI, Karmali RA, et al. Prostaglandins and schizophrenia: further discussion of the evidence. *Psychol Med* 1978;8:43–48.

421. Horrobin DF. Schizophrenia as a membrane lipid disorder which is expressed throughout the body. *Prostaglandins Leukot Essent Fatty Acids* 1996;55:3–7.

422. Delion S, Chalon S, Guilloteau D, et al. Age-related changes in phospholipid fatty acid composition and monoaminergic neurotransmission in the hippocampus of rats fed a balanced or an n-3 polyunsaturated fatty acid-deficient diet. *J Lipid Res* 1997;38:680–689.

423. Delion S, Chalon S, Guilloteau D, et al. alpha-Linolenic acid dietary deficiency alters age-related changes of dopaminergic and serotoninergic neurotransmission in the rat frontal cortex. *J Neurochem* 1996;66:1582–1591.

424. Wainwright PE. Do essential fatty acids play a role in brain and behavioral development? *Neurosci Biobehav Rev* 1992;16:193–205.

425. Horrobin DF. The membrane phospholipid hypothesis as a biochemical basis for the neurodevelopmental concept of schizophrenia. *Schizophr Res* 1998;30:193–208.

426. Puri BK, Richardson AJ, Horrobin DF, et al. Eicosapentaenoic acid treatment in schizophrenia associated with symptom remission, normalisation of blood fatty acids, reduced neuronal membrane phospholipid turnover and structural brain changes. *Int J Clin Pract* 2000;54:57–63.

427. Fenton WS, Boronow J, Dickerson F, et al. Randomized trial of supplemental EPA for residual symptoms of schizophrenia. *Biol Psychiatry* 2000;47:159S–160S.

428. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 1987;44:660–669.

429. Thome J, Foley P, Riederer P. Neurotrophic factors and the maldevelopmental hypothesis of schizophrenic psychoses. *J Neural Transm* 1998;105:85–100.

430. Lieberman JA. Is schizophrenia a neurodegenerative disorder? A clinical and pathophysiological perspective. *Biol Psychiatry* 1999;46:729–739.

431. Stahl SM. When neurotrophic factors get on your nerves: therapy for neurodegenerative disorders. *J Clin Psychiatry* 1998;59:277–278.

432. Riva MA, Molteni R, Tascedda F, et al. Selective modulation of fibroblast growth factor-2 expression in the rat brain by the atypical antipsychotic clozapine. *Neuropharmacology* 1999;38:1075–1082.

433. Altar CA. Neurotrophins and depression. *Trends Pharmacol Sci* 1999;20:59–61.

434. Kongsamut S, Roehr JE, Cai J, et al. Iloperidone binding to human and rat dopamine and 5-HT receptors. *Eur J Pharmacol* 1996;317:417–423.

57

The Economics of the Treatment of Schizophrenia

Susan M. Essock

Linda K. Frisman

Nancy H. Covell

Susan M. Essock: Department of Psychiatry, Mount Sinai School of Medicine, and the Mental Illness Research, Education, and Clinical Center, VA Medical Center, Bronx, New York.

Linda K. Frisman and Nancy H. Covell: Connecticut Department of Mental Health and Addiction Services (DMHAS), Hartford, Connecticut, and the Department of Psychology, University of Connecticut, Storrs, Connecticut.

ECONOMICS OF MENTAL HEALTH AND SCHIZOPHRENIA

Part of "57 - The Economics of the Treatment of Schizophrenia "

Prior to the 1980s, economists paid scant attention to schizophrenia, or indeed to mental health in general (1). Early work in the field included efforts to consider the impact of organization and financing on system efficiency and to address the supply of personnel in caring for persons with mental illness. For example, McGuire (2) reviewed the market for psychotherapy and the insurability of mental health care, and Frank (3) examined the supply of psychiatrists. Researchers at RAND analyzed the impact of cost-sharing on demand for mental health care (4). The use of diagnosis-related groupings to pay for care under prospective payment was considered by Taube and his colleagues (5). Dickey and Goldman (6) reviewed the impact of various funding mechanisms in public mental health.

During the 1990s, work on insurance, regulation, and the organization of mental health services continued. Studies of insurance mandates for mental health care, such as Frank and colleagues' (7) simulation of mandates and related costs, provided valuable information to state legislators considering such laws. Observers of systems change considered major reorganizational efforts, such as those implemented through the Robert Wood Johnson Program on Chronic Mental Illness (8) and other types of organizational reforms (9 ,10 ,11 and 12). The 1990s also brought analysis of the increasing implementation of managed care with behavioral health carve-outs (13).

These examples of the contributions of mental health economists indicate the range of activities in the field of mental health economics, including the economics of schizophrenia. This chapter focuses on two aspects of the economics of schizophrenia: cost studies and cost-effectiveness studies. Studies of the cost of mental illnesses appeared before other types of work in the economics of mental health, and they have continued throughout the past two decades. Cost studies lay the foundation for cost-effectiveness and cost-benefit studies because they identify the range of resources that are consumed as a result of an illness. Cost-effectiveness analyses of mental health programs began to appear in the early 1980s with the hallmark study by Weisbrod and colleagues (14) on the cost-benefit of assertive community treatment teams. Collectively, studies of costs and cost-effectiveness are perhaps the most important foci of the economics of schizophrenia. First, the sizable cost to society captures the attention of policy makers and taxpayers, and convinces them of the huge fiscal impact of schizophrenia. Second, decisions by clinicians, managers, and policy makers that are informed by research on costs and cost-effectiveness lead to better distribution of the resources available for mental health care.

Costs Of Schizophrenia

Early studies of the costs of mental illness (15 ,16 and 17) did not distinguish between the costs of different diagnostic categories (18). More recent studies have estimated specific costs for schizophrenia and other illnesses. Rice has estimated the cost of schizophrenia in the United States at $32.5 billion in 1990 (19) and $44.9 billion in 1994 (20). Goeree and colleagues (21) have calculated costs in Canada in 1996 to be approximately $2.35 billion, whereas Knapp (22) has estimated the costs of schizophrenia in the United Kingdom to be 2.6 billion pounds. Some schizophrenia cost studies focus only on service costs, such as Rund and Ruud's (23) study of costs in Norway and Martin and Miller's (24) study

of Georgia Medicaid recipients. In contrast, Rice's and Knapp's studies include both direct costs (treatment and other service costs) and indirect costs, such as lost income. But no study of the cost of schizophrenia can claim to capture all costs. As noted by McGuire (18), even comprehensive studies of the cost of schizophrenia often underestimate two types of costs: the costs to families and the costs of publicly owned capital.

Cost Perspective

Cost-of-illness studies like Rice's and Knapp's typically employ the perspective of society; that is, they consider economic costs. Economic, or social, costs are the costs of resources consumed because of an illness. Cost-effectiveness and cost-benefit analysis should always state the perspective from which the study is undertaken. Although a societal perspective presumably provides the balanced view of the neutral scientist, it is also helpful to examine costs from perspectives of particular stakeholders. For example, in an analysis of the impact of Assertive Community Treatment in Connecticut, Essock and colleagues (25) present costs from the perspectives of society, the state, and the Department of Mental Health. Comparison of the results from multiple perspectives may identify areas of cost-shifting that results from certain programs and policies. For example, a treatment that reduces hospital days may shift costs from state-run inpatient facilities to private nonprofit outpatient settings.

Cost Components

Costs of Treatment and other Services

The examples provided by Rice and Knapp are instructive for those conducting cost-of-illness studies and cost-effectiveness studies in the area of schizophrenia. As they show, there are many ways in which the illness is associated with greater costs. First are the costs of treatment, including medication. Treatment may be offered by public, private, or voluntary sector settings, and many persons with schizophrenia receive care in multiple places. Besides treatment, services like case management, vocational rehabilitation, and psychosocial clubhouses generate significant costs. Medical and surgical costs also may be relevant in cost-effectiveness analyses, because utilization of these services may vary depending on the adequacy of mental health care (26 ,27).

Lost Productivity and Family Burden

Mental illnesses, like other disorders, cause people to lose workdays (28) and sometimes even to forfeit aspirations of having any career at all. Although lost productivity is usually addressed in comprehensive studies of the costs of schizophrenia, many studies of interventions for persons with schizophrenia have ignored productivity losses, because of the high rates of disability in this population. However, new strategies for improving the employment outcomes for persons with serious mental illness, such as Individual Placement and Support (29), have made employment a realistic goal of rehabilitation. These new successes suggest that loss of productivity should be included when calculating the cost of interventions for persons with schizophrenia. In addition to the productivity losses of the individual, cost studies should attend to the work losses of family members, and other contributions of time and in-kind services (30).

Capital Costs

Economic cost studies appropriately study the opportunity costs of all resources, that is, the value of those resources in their best alternative use. In a cost-effectiveness study of a new residential model for persons with serious mental illness, Cannon and her colleagues (31) carefully considered the value of capital costs of a public hospital, which would have been underestimated if valued through traditional methods of depreciation. Capital costs can be large enough to change the most basic findings of a cost study, as shown by Rosenheck and colleagues (32). Public administrators may not consider the value of buildings and property to be part of a cost equation because it is not always part of the operating costs, but the value of the property in alternative use may be considerable.

Other Components

Especially where an intervention is expected to have an impact on co-occurring substance use disorders, it is important to attend to criminal justice costs (33). Another neglected aspect of cost studies is the costs of administering transfer payments (such as social security). Although disability payments themselves do not represent the use of new resources, the cost of administering these payments is a cost that should be counted, especially if the intervention could change the rate of receipt of disability payments or other public benefits (34). For example, an intervention that returns people to work will not only increase their productivity (a benefit), but also decrease disability payments (decrease a cost).

The importance of including any type of cost in a cost-effectiveness study is related to the potential impact of the cost type on the study findings. The larger the cost per unit, or the more frequently it is used, the more carefully it should be assessed (35). But should cost-effectiveness analyses always be conducted? As indicated by the range of costs and cost perspectives that might be included, these studies can be expensive to implement. This expense is further increased because, to detect meaningful differences in a highly variable outcome such as cost, significantly more study participants

may be needed than for an effectiveness study alone. The usefulness depends in large part on the likelihood that the treatment or intervention under study will have an effect on costs—either positive or negative. Cost studies are critical in the analysis of novel antipsychotic agents because of the relatively high acquisition costs of the drugs, compared to conventional antipsychotic medications, and because of the potential of these drugs to reduce the number of days that people with schizophrenia are hospitalized (36).

COST-EFFECTIVENESS

Part of "57 - The Economics of the Treatment of Schizophrenia "

The success of interventions in schizophrenia, whether medications or psychosocial rehabilitation programs, is reflected in multiple domains. An antipsychotic may have an impact on cognition, on hallucinations, on affect, on disruptiveness, on sexual functioning, on extrapyramidal side effects, on weight, on employment—the list goes on and on. These are all measures of the effectiveness of the agent, some positive and some negative. Some, such as hallucinations and delusions, may be influenced much more directly by the medication than more distal outcomes such as housing or employment. Different individuals value changes in these domains differently (elimination of hearing voices that other people don't hear may be much more important to the person who is troubled by harassing voices making a steady stream of demeaning comments than to the person whose voices are good company or connect the patient to an imagined network of internationally renowned researchers). Similarly, some people are very troubled by changes in weight or sexual functioning, whereas such changes mean little to others. Hence, much as we would like a composite measure across all effectiveness domains, this reductionistic approach is fraught with untenable compromises. Just as there is variation among different patients and different providers, patients and payers ascribe different values to the same outcome (e.g., legislators who make funding allocations to public mental health systems may be more concerned about decreases in violence and patients may be more concerned with increases in quality of life). How much is a symptom-free day worth? It depends on who is asking and who is paying.

Cost-effectiveness analyses have evolved to deal with the multiple domains touched by a single treatment. Such analyses report the change in a given effectiveness measure associated with a particular cost investment in treatment. A medication may be cost-effective with respect to certain outcomes, cost-neutral for others, and costly for yet others. Lehman (37) reminds us that the current explosion of new knowledge about effective treatments and the advent of evidence-based quality standards for treating schizophrenia come at a time when cost containment is paramount in the health policy agenda. Policy makers need to know the impact of dollars invested in treatment—but not just in a single domain like reductions in hospital care. Those who make purchasing decisions for the public systems of care under which most of the country's treatment for schizophrenia is funded need information on multiple domains of effectiveness.

An alternative to cost-effectiveness analysis is cost-utility analysis, where a comprehensive outcome indicator is calculated as a preference-weighted sum of the outcome measures. An example of a cost-utility approach is the use of quality-adjusted life years (QALYs) (38 ,39). As noted above, different stakeholder groups value different outcomes differently; hence, approaches such as QALYs create an effectiveness metric representative of at best only one stakeholder group, and at worst the resulting metric is representative of no one. Although elegant in presentation, as with sausages, observing their creation can reduce enthusiasm for their use. One must find or create weights to apply to the various effectiveness measures and then decide on a combination scheme—deciding, for example, what weight gain is the equivalent of what change in extrapyramidal side effects (EPSs) and what change in psychotic symptoms. Typically, one does this either by interviewing individuals representative of the population under study (e.g., treatment refractory patients with schizophrenia) or by adopting someone else's measures as close enough. Because “close enough” is a very subjective call, it is important for researchers to disclose the sources of the weighting estimates so that readers can make their own call. For example, Rosenheck and colleagues' (40) report of changes in QALYs among mostly male veterans with schizophrenia used weights derived as part of a doctoral dissertation by Kleinman (Johns Hopkins University, 1995) of mainly African-American women, only about half of whom (55%) were diagnosed with schizophrenia (the rest were diagnosed with major depression, bipolar, and other affective disorders). We were unwilling to take the leap of faith needed to generalize from groups this disparate when presenting cost-effectiveness results from our own work (41). Nevertheless, Rosenheck and colleagues are to be commended for providing the information necessary to follow back their methods to see what was used. This is not always the case.

Another type of utility analysis is the measure of symptom-free days (42). Under such analyses, interventions are compared with respect to the number of symptom-free days they produce. Following the methodology of Lave and colleagues (43), Simon and colleagues (42) credited a study participant with having one depression-free day if the study participant had 2 days with a depression score of 0.5. Many people with depression, as well as many researchers, would take issue with saying that someone was symptom free for half a year if they reported having 50% of full symptoms for each day of that year. Symptom-free days may be a poor measure within schizophrenia studies simply because, unlike with depression, symptoms and functioning are poorly correlated,

and the likelihood of having a completely symptom-free day is rather small.

Disability-adjusted life years (DALYs), where one DALY equals one lost year of healthy life, can also be used to express years lost, both to premature death and to disabilities associated with living with schizophrenia (44). In contrast to QALYs and estimates of symptom-free days, DALYs are proxies for negative outcomes (45) and, as such, the calculation of cost-effectiveness centers around how many DALYs are saved by using a particular intervention. In a population where individuals may live for a long period of time with a relatively debilitating illness like schizophrenia, measures of mortality alone do not adequately capture the impact of the disability. DALYs are calculated by adding together the number of years between mortality and life expectancy (years of life lost, YLL) and the number of years lived with a disability (YLDs). Calculating YLDs requires making assumptions about the relative impact of illness onset, duration, and severity on healthy living (for example, making an assumption that a first psychotic episode at age 15 is worse than a first episode at age 25). As with QALYs, these metrics can be derived by surveying individuals with schizophrenia or their proxies, with the accompanying assumptions that how one weights hypothetical events is the same as the trade-offs one would make if one could trade fewer days of healthy life for more days of life with particular disabilities. Because such ratings are inherently untestable by rigorous methods, whether reliable or not, their validity remains suspect. Further, the calculation of DALYs “presupposes that life years of disabled people are worth less than life years of people without disabilities” (46), and may even rank some individuals' lives as worse than death (47). Schizophrenia brings with it an increased risk of suicide (48), which is consistent with DALYs ranking some lives as worse than death. However, assuming that person A and person B, in reality, would make the same choices as to what fates are worse than death presumes an ecologic validity to DALY ratings that may be unwarranted.

Cost-utility measures such as QALYs, DALYs, and measures like symptom-free days, have enormous appeal because of their ability to reduce multiple effectiveness domains to a single bullet measure. By deriving a single measure, one can compare any treatment approach to any other treatment approach. Where the measure is reduced to dollars (as in QALYs), one may even compare the values of interventions between different conditions (38), for example, if dollars expended on diabetes reap more benefits than dollars spent on schizophrenia. But the assumptions built into such bullet measures may have limited usefulness for informing decisions at the level of the individual patient, prescriber, or health care payer. These individuals weigh their particular circumstances, and may be unwilling to have others' preferences serve as proxies for their own. Instead, these stakeholders are asking more specific questions. For example, the mental health commissioner asks, “If I put an extra $3 million in the pharmacy budget for medication X, what can I expect this to buy me in terms of the other domains under my purview, and what is the downside risk? What will it buy me in terms of reductions in hospital use, improvements in vocational functioning, reductions in violent episodes, and reductions in side effects?” Similarly, patients and families paying for medications ask, “If I increase/decrease my spending by changing to medication X, what changes am I likely to see in the voices I hear, in my employability, in my sexual functioning, and in my body movements?”

An alternative to composite measures are measures that contrast costs invested to a variety of outcome domains, some of which will be more important to some stakeholders than to others. An analogy is a proposal for a city park to be funded from multiple sources. Depending on one's perspective (e.g., whether you would use the park, how the park would impact the value of your property, your safety, your recreational options, what you are called on to invest), the park may or may not be a good idea. And, depending on who is paying for what, and which outcome domains are most important to you, you may stand to get a lot or a little out of the dollars going into the park. The challenge is to present the data on costs and effects in such a way that the various payers (the city, private foundations, neighborhood organizations, individual contributors) can each look from their own perspective, see what the expected gains and losses are in the outcome domains they care about most (less street noise, more open space, more dogs, more people drawn to the neighborhood), perform their own idiosyncratic weighting of these factors, and decide if they are in favor of the park or not.

In contrast to cost-utility analysis, cost-effectiveness analysis does not reduce the impact of an intervention into one measure. Some outcomes may be clearly preferential or “dominant choices” (e.g., lower costs and higher effectiveness). Other outcomes are not as clearly dominant, and in these cases it may be useful to show the likely range of cost compared to multiple domains of effectiveness. One method of examining these ranges is to create sampling distributions for costs and effectiveness measures to show the precision of estimates as well as their mean. For example, bootstrap techniques use every study participant's data to create an empirical sampling distribution of the test statistic and plot these estimates as a cost-effectiveness plane. Bootstrapping techniques offer one means of describing confidence intervals for incremental cost-effectiveness ratios (ICERs) (49 ,50). Cost data are often highly positively skewed, and ICERs provide less biased estimates of confidence intervals in highly skewed cost data (43 ,51 ,52).

Figure 57.1 shows such an approach when considering the cost-effectiveness of clozapine compared to conventional agents among long-stay state hospital patients (41). The cluster of points displays the sampling distribution of the ICER. Most of the points fall in the lower-right quadrant, indicating that clozapine is most likely to be less costly and

more effective than conventional antipsychotic agents from the cost perspective (total societal cost) and for the effectiveness measure in question (reduction in EPS). Such displays of information give the reader/policy maker a sense of the tightness of the point estimate and the risk of falling in a quadrant other than the one indicating cost-effectiveness. One can use these sampling distributions to create cost-acceptability curves from the viewpoint of particular payers for particular outcomes (e.g., the likelihood that the intervention will be cost effective for the payer who is willing to risk $1, $10, or $100 to obtain an 80% likelihood of return in sexual functioning).

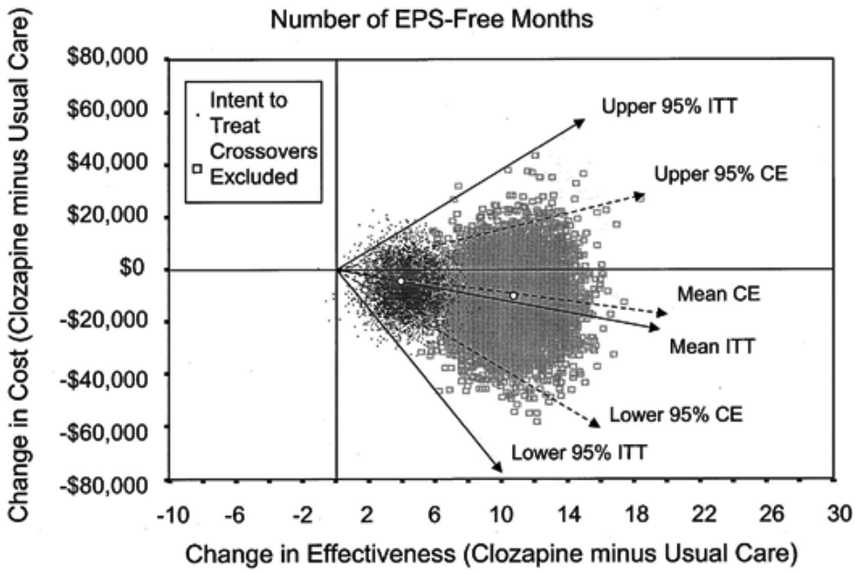


FIGURE 57.1. Ten thousand bootstrap replications plotted in the cost-effectiveness plane (intent-to-treat, *N* = 136 clozapine and *N* = 87 usual care; treatment crossovers excluded, *N* = 89 clozapine and *N* = 30 usual care). The x-axis and y-axis, respectively, show the difference between clozapine and usual-care groups in estimated number of extrapyramidal side effects (EPS)-free months and total cost during a 2-year period. The quadrant to the lower right of the origin (0,0) contains those estimates where clozapine was found to be less costly and more effective than the usual care (80% of the estimates for the intent-to-treat analyses and 81% of the estimates when treatment crossovers are excluded). (From Essock SM, Frisman LK, Covell NH, et al. Cost-effectiveness of clozapine compared with conventional antipsychotic medication for patients in state hospitals. *Arch Gen Psychiatry* 2000;57:992, with permission. Copyright 2000 American Medical Association.)

Saul Feldman (53) has held positions as the head of the National Institute of Mental Health (NIMH) Staff College and chief executive officer of one of the country's largest managed behavioral health care organizations. Thus, he has been in a position to make policy based on research, and to inform policy makers with research. He has posed the question, Is good research good if it does not inform policy and practice? It is incumbent on mental health services researchers to report their findings in ways that speak to funders and service system managers, which means providing estimates of the most likely outcome as well as the likelihood of alternative outcomes.

COST OF THE NEWER ANTIPSYCHOTIC MEDICATIONS

Part of "57 - The Economics of the Treatment of Schizophrenia "

In general, the acquisition cost of the newer antipsychotic medications is greater than that of conventional ones. These acquisition costs are reflected in formulary budgets. Once a relatively small component of treatment costs, formulary budgets in psychiatric settings have risen dramatically in the past decade, and the market share of the newer agents has risen as they have replaced the less costly conventional agents. Figure 57.2 shows the distribution of antipsychotic prescriptions paid for by Medicaid in 1998 (left) and the

dollars Medicaid paid for these prescriptions (right). These data show that the newer agents account for 58% of all antipsychotic prescriptions paid for by Medicaid but for $1.15 billion (90%) of the $1.28 billion in Medicaid costs for antipsychotic prescriptions. These charts dramatically display the disparity in medication costs associated with the newer versus the conventional agents.

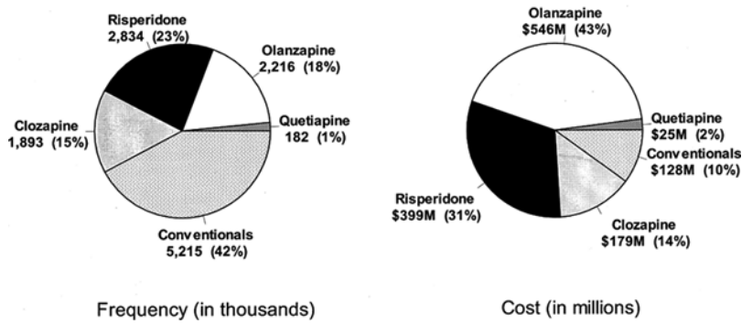


FIGURE 57.2. Distribution of (*left circle*) and total dollars paid (*right circle*) by Medicaid for antipsychotic medication prescriptions during 1998. Newer antipsychotic medications represented slightly over half of the total prescriptions, and they were responsible for 90% of the total cost.

This price difference between the older and the newer antipsychotic medications, which can be a 100-fold difference (e.g., when contrasting generic oral haloperidol with nongeneric Clozaril), prompted scores of studies asking the cost impact of using the newer medications when more than simply the cost of the medication was considered. For example, if using new and expensive medication X results in fewer days hospitalized than some alternative, then, all else being equal, using X will reduce overall costs as long as the cost savings associated with fewer days in the hospital is greater than the cost difference between medication X and the alternative.

Clozapine Cost Effectiveness Studies As Case Examples

The rub, of course, is that “all else” is rarely equal in effectiveness or cost-effectiveness studies, and the early cost projections concerning the impact of using clozapine often suffered from faulty assumptions about what was equivalent. Many of these studies were pre–post comparisons that examined changes in hospital use and lacked a comparison group (54 ,55 ,56 ,57 ,58 ,59 and 60). For example, the study by Meltzer and colleagues (59) of patients with schizophrenia who were taking clozapine collected retrospective cost data for 2 years before and after these 47 individuals began taking clozapine and concluded that clozapine was associated with a 23% drop in treatment costs. This study generated a series of letters criticizing the study's methodology (61 ,62 and 63) [see also the response by Meltzer and Cola (64)]. Critics focused on the problem of the regression toward the mean that can be expected whenever study participants are enrolled during a low point in their functioning (such as may have prompted the initiation of clozapine), and on the other potential temporal and case-mix confounds associated with mirror-image studies of individuals who were selected to begin a medication through other than a random-assignment process among all eligibles.

The results of the two randomized clinical trials of clozapine's cost-effectiveness each showed much more modest benefits associated with clozapine, both in the 2-year, open-label trial comparing clozapine to the usual care with a range of conventional antipsychotics among long-term patients in state hospitals (41 ,65 ,66), and in the 1-year masked trial comparing clozapine to haloperidol among veterans hospitalized for a year or less (67). Each trial showed clozapine to be somewhat more effective than the comparison agents, and this increase in effectiveness comes at no additional cost when costs are viewed from a societal perspective. Each trial also showed that clozapine is more effective than the usual care in minimizing days hospitalized, enough so that the reduction in hospital days more than covers the increased cost of the medication plus increased outpatient services. But, from more narrow perspectives (e.g., the hospital formulary budget, capitated outpatient service providers), clozapine would be viewed as increasing their costs. For cost-effectiveness studies to influence planning and policy making, the perspectives of these different payers need to be taken into account because it is these local incentives and disincentives that must be addressed to be sure that the fiscal incentives are lined up to promote good care. A hospital would have a great incentive to use clozapine for a heavy user of hospital services if it has a fixed budget (the case with most state hospitals), but a hospital paid a per diem would have no such incentive.

Lengthy randomized clinical trials in routine practice settings, such as the clozapine study in Connecticut state hospitals and the clozapine study in Veterans Administration (VA) hospitals, suffer from treatment crossovers. By the end of 6 months in the Connecticut study, only 11% of the usual care patients had begun a trial on clozapine, but by the end of 24 months in the study, 66% had. In the VA clozapine study, 72% of the patients assigned to masked haloperidol had ceased taking the masked medication by the end of the 1-year study period, with 49 of 157 (31%) of them switching to clozapine and the rest to conventional antipsychotics, including unmasked haloperidol (67). Because of the biases introduced by what is likely to be highly nonrandom discontinuation of the assigned treatment, the importance of intent-to-treat analyses and the unspecified biases of crossovers-excluded analyses are well documented (68). Regardless, when crossovers are common, analyses excluding crossovers offer a proxy for the best-case scenarios for each treatment condition by comparing only those who do well enough on treatment A to stay on it with only those who do well enough on treatment B to stay on it. Figure 57.1 illustrates this using data from the Connecticut clozapine study. The exclusion of treatment crossovers increases the apparent effectiveness of clozapine (the crossoversexcluded oval is shifted to the right of the intent-to-treat oval in Fig. 57.1) and decreased the estimate of the relative costliness of clozapine (the crossovers-excluded oval is shifted lower by about $5,500) (41). Clearly, individuals who leave their assigned treatment are different in terms of costs and outcomes from those who remain in their assigned treatment condition.

Another difficulty when trying to assess relative costs is the great variability in costs across patients. For example, in the VA study just cited, health care costs in the 6 months prior to randomization were approximately $27,000 with a standard deviation of about $17,000 (67). For the Connecticut clozapine study, the 95% confidence interval for patients assigned to clozapine was $96,847 to $114,308 for year 2 versus $103,665 to $121,144 for those assigned to the usual care. With such variability, cost differences are

very difficult to detect, even with the relatively large sample sizes of the VA and Connecticut trials (*N* = 423 and 227, respectively). Even for individuals who are heavy service users at study entry, mounting a trial powered to detect cost differences requires hundreds of individuals per treatment arm. If the trial were a study of outpatients who are infrequent users of expensive services like hospitals, it would require even larger samples to detect cost differences apart from medication.

From a public health perspective, an emphasis on point estimates of costs and effectiveness is misguided when the confidence intervals are so broad. Economists would call clozapine the dominant alternative in these randomized trials (because most of the range spanned by the cost confidence intervals includes the values where clozapine costs less than or the same as the usual care and the effectiveness measures favor clozapine or are neutral). The reduction of data to such a point estimate belies the broad distribution of possible outcomes that are likely to occur across patients. Planners and policy makers, as well as patients and their treating clinicians, need a sense of the range of possible outcomes and their relative likelihood to inform their decisions about what chances they want to take.

Costs Associated with Risperidone, Olanzapine, and Quetiapine

Figure 57.3 shows the frequency of prescribing by type of antipsychotic in three large states in different parts of the country among individuals whose medications are paid via

Medicaid. Because Medicaid formularies allow unrestricted access to any of these medications independent of location in the country and the same financial incentives apply, one would expect to see similar rates of prescribing these medications. Indeed, the distributions do appear quite similar to each other and to the national data (Fig. 57.2). That these distributions do not reflect what we know about the relative effectiveness of these agents suggests that other factors are strong influences on medication choice and that these influences combine to create similar patterns of antipsychotic prescribing under Medicaid nationwide. In addition to factors such as effectiveness, factors as disparate as patients' past histories of medication use, order of receiving Food and Drug Administration (FDA) approval, convenience of use, acquisition costs, relative marketing budgets, and side-effect profiles may also be at play. These figures serve as reminders that medications are started and discontinued for reasons other than effectiveness. (Data for these pie charts were extracted from the Health Care Financing Administration's (HCFA) Web site http://medicare.hcfa.gov/medicaid/drug5.htm; potential users take note that, as of March 2000, the Web site reported Medicaid expenditures in cents rather than dollars and does not label the cost units.)

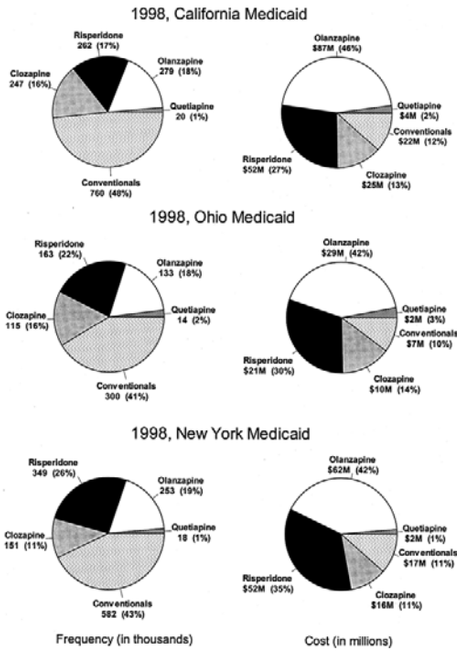


FIGURE 57.3. Distribution of (*left circles*) and total dollars paid (*right circles*) by Medicaid for antipsychotic medication prescriptions in California, Ohio, and New York during 1998.

Several studies of risperidone and olanzapine also suggest that the acquisition costs of these medications may be offset by reduction in use of more expensive health care services such as inpatient treatment. Many of these studies have methodologic shortcomings similar to those of the earlier cost studies of clozapine described above. Another concern is that industry sponsorship of many of these studies means that they do not meet the criteria for lack of an incentive for bias set forth by the *New England Journal of Medicine* (69), whose editors noted that the opportunities for introducing bias into economic studies are far greater than in studies of biological phenomena because of the unusually discretionary nature of model building and data selection in such analyses, and because drug costs in particular can be quite arbitrary as they are prices (not costs) set by the manufacturer. Hence, additional work is needed in this area.

In general, the studies by the medications' manufacturers show support for cost reductions favoring that manufacturer's medication [e.g., for risperidone (70) and for olanzapine (71 ,72)]. Although such studies form good starting points for further investigation, they need follow-up by independent investigators to assess how the agents' cost-effectiveness plays out in broader settings with representative patients, lest best-case examples be generalized to settings where they are not applicable and used to set policy there. An example of an important follow-up study is that of Conley and colleagues (73), who found that, among 84 treatment-refractory patients randomly assigned to a double-blind 8-week fixed-dose trial of either olanzapine or chlorpromazine, olanzapine appeared to have limited efficacy, showing only a 7% response. Hence, the reduction in treatment costs associated with olanzapine noted in the reviews of Palmer and colleagues (74) and Foster and Goa (75) would not be expected among treatment-refractory patients, even though these patients are heavy users of inpatient services. Under other scenarios, these patients are the very ones for whom new interventions are associated with cost savings because they have higher initial rates of utilization on which to show an impact (25 ,40). An independent study of risperidone compared to conventional antipsychotics among outpatients with schizophrenia using a matched comparison group found no difference in total treatment costs or effectiveness measures, although there was a trend for the risperidone-treated group to have higher costs, attributable to higher medication costs (76).

CONCLUSION AND ADDITIONAL RESOURCES

Part of "57 - The Economics of the Treatment of Schizophrenia "

This brief review of the economics of treating schizophrenia has emphasized some of the methodologic complexities that must be acknowledged and surmounted when addressing treatment costs. The emphasis has been on illustrating the importance of bias minimization and estimation when constructing studies and reporting results, and on the importance of reporting results from the perspective of different payers and giving estimates of the variability associated with any findings reported as point estimates. It is important to tell patients, prescribers, and payers not just the best estimate of costs and effectiveness, but the likelihood that their costs and outcomes will fall within their acceptable ranges for what they are willing to pay and/or risk to gain a given outcome.

Fortunately, this evolving literature has many active champions who publish widely and lead the way in documenting the ways in which costs and fiscal (dis)incentives have impact on access to treatment, the quality of care received, and patient outcomes. Although their work cannot be summarized here, useful source books include those by Drummond and colleagues (38), Frank and Manning (77), Gold and colleagues (39), and Hargreaves and colleagues (35). The journal *Health Affairs* continues to be a particularly valuable resource for reports on mental health economics and thoughtful analyses of the economic influences on the treatment of individuals with schizophrenia.

ACKNOWLEDGMENTS

Part of "57 - The Economics of the Treatment of Schizophrenia "

This research is the product of the collaboration of many individuals, both within and outside the Connecticut Department of Mental Health and Addiction Services (DMHAS). In particular, we would like to thank Carlos Jackson, Ph.D., of the University of Connecticut for his assistance with the data extraction and statistical analyses of the Medicaid prescription data. The research was funded

in part by U.S. Public Health Service (USPHS) grants R01 MH-48830 and R01 MH-52872 from the National Institute of Mental Health (NIMH) to Susan Essock, Ph.D., principal investigator, as well as by DMHAS. This publication does not express the views of the Department of Mental Health and Addiction Services or the State of Connecticut. The views and opinions expressed are those of the authors.

REFERENCES

1. McGuire TG. Growth of a field in policy research: the economics of mental health. *Administration and Policy in Mental Health* 1990;17:165–175.

2. McGuire TG. *Financing psychotherapy: costs, effects, and public policy*. Cambridge, MA: Ballinger, 1981.

3. Frank RG. Is there a shortage of psychiatrists? *Community Ment Health J* 1983;19:42–53.

4. Manning WG, Wells KB, Duan N, et al. Cost sharing and the use of ambulatory mental health services. *Am Psychol* 1984;39:1077–1089.

5. Taube CA, Lee ES, Forthofer RN. DRGs in psychiatry: an empirical evaluation. *Med Care* 1984;22:597–610.

6. Dickey B, Goldman HH. Public health care for the chronically mentally ill: financing operating costs. *Administration in Mental Health* 1986;14:63–77.

7. Frank RG, McGuire TG, Salkever DS. Benefit flexibility, cost shifting, and mandated mental health coverage. *J Ment Health Administration* 1991;18:264–271.

8. Goldman HH, Morrissey JP, Ridgely MS. Evaluating the Robert Wood Johnson Foundation program on chronic mental illness. *Millbank Q* 1994;72:37–47.

9. McGuire TG, Hodgkin D, Shumway D. Managing Medicaid mental health costs: the case of New Hampshire. *Administration and Policy in Mental Health* 1995;23:97–117.

10. Shepherd G, Muijen M, Hadley TR, et al. Effects of mental health services reform on clinical practice in the United Kingdom. *Psychiatr Serv* 1996;47:1351–1355.

11. Semke J, Fisher WH, Goldman HH, et al. The evolving role of the state hospital in the care and treatment of older adults: state trends, 1984 to 1993. *Psychiatr Serv* 1996;47:1082–1087.

12. Goldman HH, Frank RG, Gaynor MS. What level of government? Balancing the interests of the state and the local community. *Administration and Policy in Mental Health* 1995;23:127–135.

13. Mechanic D, ed. *Managed behavioral health care: Current realities and future potential*. New directions for mental health services, no. 78. San Francisco: Jossey-Bass, 1998.

14. Weisbrod BA, Test, MA, Stein LI. Alternative to hospital treatment: II. Economic benefit-cost analysis. *Arch Gen Psychiatry* 1980;37:400–405.

15. Cruze AM, Harwood HJ, Kristiansen PL, et al. *Economic costs to society of alcohol and drug abuse and mental illness—1977*. Research Triangle Park, NC: Research Triangle Institute, 1981.

16. Harwood HJ, Napolitano DM, Kristiansen PL, et al. *Economic costs to society of alcohol and drug abuse and mental illness—1980*. Research Triangle Park, NC: Research Triangle Institute, 1984.

17. Rice DP, Kelman S, Miller LS, et al. *The economic costs of alcohol and drug abuse and mental illness: 1985*. DHHS pub. no. (ADM) 90-1694. Rockville, MD: National Institute of Mental Health, 1990.

18. McGuire TG. Measuring the economic costs of schizophrenia. *Schizophr Bull* 1991;17:375–388.

19. Rice DP. The economic impact of schizophrenia. *J Clin Psychiatry* 1999;60(suppl 1):4–6.

20. Rice DP. Economic burden of mental disorders in the United States. *Economics Neurosci* 1999;1:40–44.

21. Goeree R, O'Brien BJ, Goering P, et al. The economic burden of schizophrenia in Canada. *Can J Psychiatry* 1999;44:464–472.

22. Knapp M. Costs of schizophrenia. *Br J Psychiatry* 1997;171:509–518.

23. Rund BR, Ruud T. Costs of services for schizophrenic patients in Norway. *Acta Psychiatr Scand* 1999;99:120–125.

24. Martin BC, Miller LS. Expenditures for treating schizophrenia: a population-based study of Georgia Medicaid recipients. *Schizophr Bull* 1998;24:479–488.

25. Essock SM, Frisman LK, Kontos NJ. Cost effectiveness of assertive community treatment teams. *Am J Orthopsychiatry* 1998;68:179–190.

26. Mumford E, Schlesinger HJ. Assessing consumer benefit: cost offset as an incidental effect of psychotherapy. *Gen Hosp Psychiatry* 1987;9:360–363.

27. Pallack MS, Cummings NA. Inpatient and outpatient psychiatric treatment: the effect of matching patients to appropriate level of treatment on psychiatric and medical-surgical hospital days. *Appl Prevent Psychol* 1992;1:83–87.

28. Kessler RC, Frank RG. The impact of psychiatric disorders on work loss days. *Psychol Med* 1997;27:861–873.

29. Drake RE, McHugo GJ, Becker DR, et al. The New Hampshire study of supported employment for people with severe mental illness. *J Consult Clin Psychol* 1996;64:391–399.

30. Clark RE. Family costs associated with severe mental illness and substance use. *Hosp Community Psychiatry* 1994;45:808–813.

31. Cannon NC, McGuire TG, Dickey B. Capital costs in economic program evaluation: the case of mental health services. In: Catterall J, ed. *Economic evaluation of public programs. New direction for program evaluation*, vol 26. San Francisco: Jossey-Bass, 1985:69–82.

32. Rosenheck RA, Frisman LN, Neale M. Estimating the capital component of mental health care costs in the public sector. *Administration and Policy in Mental Health* 1994;21:493–509.

33. Clark RE, Ricketts SK, McHugo GJ. Legal system involvement and costs for person in treatment for severe mental illness and substance use disorders. *Psychiatr Serv* 1999;50:641–647.

34. Frisman L, Rosenheck R. How transfer payments are treated in cost-effectiveness and cost-benefit analyses. *Administration and Policy in Mental Health* 1996;23:533–545.

35. Hargreaves WA, Shumway M, Hu T, et al. *Cost-outcome methods for mental health*. New York: Academic Press, 1998.

36. Byrom BD, Garratt CJ, Kilpatrick, AT. Influence of antipsychotic profile on cost of treatment in schizophrenia: a decision analysis approach. *Int J Psychiatry Clinical Pract* 1998;2:129–138.

37. Lehman AF. Quality of care in mental health: the case of schizophrenia. *Health Affairs* 1999;18:52–70.

38. Drummond MF, O'Brien B, Stoddart GL, et al. *Methods for the economic evaluation of health care programmes*, 2nd ed. New York: Oxford University Press, 1997.

39. Gold MR, Siegel JE, Russell LB, et al., eds. *Cost-effectiveness in health and mental health*. New York: Oxford University Press, 1996.

40. Rosenheck R, Cramer J, Allan E, et al. Cost-effectiveness of clozapine in patients with high and low levels of hospital use. *Arch Gen Psychiatry* 1999;56:565–572.

41. Essock SM, Frisman LK, Covell NH, et al. Cost-effectiveness of clozapine compared to conventional antipsychotic medication for patients in state hospitals. *Arch Gen Psychiatry* 2000;57:987–994.

42. Simon GE, VonKorff M, Rutter C, et al. Randomised trial of monitoring, feedback, and management of care by telephone to improve treatment of depression in primary care. *BMJ* 2000;320:550–554.

43. Lave JR, Frank RG, Schulberg HC, et al. Cost-effectiveness of treatments for major depression in primary care practice. *Arch Gen Psychiatry* 1998;55:645–651.

44. Shore, MF. Replacing cost with value. *Harvard Rev Psychiatry* 1999;6:334–336.

45. Murray CJL, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: global burden of disease study. *Lancet* 1997;349:1498–1504.

46. Arneson T, Nord E. The value of DALY life: problems with ethics and validity of disability adjusted life years. *BMJ* 1999;319:1423–1425.

47. Rock, M. Discounted lives? Weighing disability when measuring health and ruling on “compassionate” murder. *Soc Sci Med* 2000;51:407–417.

48. Radomsky ED, Haas GL, Mann JJ, et al. Suicidal behavior in patients with schizophrenia and other psychotic disorders. *Am J Psychiatry* 1999;156:1590–1595.

49. Briggs A, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Econ* 1998;7:723–740.

50. Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 1990;10:212–214.

51. Chaudhary MA, Stearns SC. Estimating confidence intervals for cost-effectiveness ratios: An example from a randomized trial. *Stat Med* 1996;15:1447–1458.

52. Pollack S, Bruce P, Borenstein M, et al. The resampling method of statistical analysis. *Psychopharmacol Bull* 1994;30:227–234.

53. Feldman S. Strangers in the night: research and managed mental health care. *Health Aff* 1999;18:48–51.

54. Aitchison KJ, Kerwin RW. Cost-effectiveness of clozapine. A UK clinic-based study. *Br J Psychiatry* 1997;171:125–130.

55. Blieden N, Flinders S, Hawkins K, et al. Health status and health care costs for publicly funded patients with schizophrenia started on clozapine. *Psychiatr Serv* 1998;49:1590–1593.

56. Ghaemi SN, Ziegler DM, Peachey TJ, et al. Cost-effectiveness of clozapine therapy for severe psychosis. *Psychiatr Serv* 1998;49:829–831.

57. Jonsson D, Walinder J. Cost-effectiveness of clozapine treatment in therapy-refractory schizophrenia. *Acta Psychiatr Scand* 1995;92:199–201.

58. Luchins DJ, Hanrahan P, Shinderman M, et al. Initiating clozapine treatment in the outpatient clinic: service utilization and cost trends. *Psychiatr Serv* 1998;49:1034–1038.

59. Meltzer HY, Cola P, Way L, et al. Cost effectiveness of clozapine in neuroleptic-resistant schizophrenia. *Am J Psychiatry* 1993;150:1630–1638.

60. Reid WH, Mason M, Toprac M. Savings in hospital bed-days related to treatment with clozapine. *Hosp Community Psychiatry* 1994;45:261–264.

61. Essock SM. Clozapine's cost effectiveness (letter). *Am J Psychiatry* 1995;152:152.

62. Rosenheck R, Charney DS, Frisman LK et al. Clozapine's cost effectiveness (letter). *Am J Psychiatry* 1995;152:152–153.

63. Schiller M, Hargreaves WA. Clozapine's cost effectiveness (letter). *Am J Psychiatry* 1995;152:151–152.

64. Meltzer HY, Cola P. Clozapine's cost effectiveness (reply). *Am J Psychiatry* 1995;152:153–154.

65. Essock SM, Hargreaves WA, Covell NH, et al. Clozapine's effectiveness for patients in state hospitals: results from a randomized trial. *Psychopharmacol Bull* 1996;32:683–697.

66. Essock SM, Hargreaves WA, Dohm FA, et al. Clozapine eligibility among state hospital patients. *Schizophr Bull* 1996;22:15–25.

67. Rosenheck RA, Cramer J, Xu W, et al. A comparison of clozapine and haloperidol in the treatment of hospitalized patients with refractory schizophrenia. *N Engl J Med* 1997;337:451–458.

68. Lavori, PW. Clinical trials in psychiatry: should protocol deviation censor patient data? *Neuropsychopharmacology* 1992;6:39–48.

69. Kassirer JP, Angell M. Journal's policy on cost-effectiveness analyses. *N Engl J Med* 1994;331(10):669–670.

70. Nightengale BS, Crumly JM, Liao J, et al. Current topics in clinical psychopharmacology. *Psychopharmacol Bull* 1998;34:373–382.

71. Hamilton SH, Revicki DA, Edgell ET, et al. Clinical and economic outcomes of olanzapine compared with haloperidol for schizophrenia. *Pharmacoeconomics* 1999;15:469–480.

72. Tunis SL, Bryan MJ, Gibson J, et al. Changes in perceived health and functioning as a cost-effectiveness measure for olanzapine versus haloperidol treatment of schizophrenia. *J Clin Psychiatry* 1999;60(suppl):38–45.

73. Conley R, Tamminga C, and Group, MS. Olanzapine compared with chlorpromazine in treatment-resistant schizophrenia. *Am J Psychiatry* 1998;155:914–920.

74. Palmer CS, Revicki DA, Genduso LA, et al. A cost-effectiveness clinical decision analysis model for schizophrenia. *Am J Manag Care* 1998;4:345–355.

75. Foster RH, Goa KL. Olanzapine: a pharmacoeconomic review of its use in schizophrenia. *Pharmacoeconomics*. 1999;15:611–640.

76. Schiller MJ, Shumway M, Hargreaves WA. Treatment costs and patient outcomes with use of risperidone in a public mental health setting. *Psychiatr Serv* 1999;50:228–230.

77. Frank RG, Manning WG, eds. *Economics and mental health*. Baltimore: Johns Hopkins University Press, 1992.

58

Mechanism of Action of Atypical Antipsychotic Drugs

Herbert Y. Meltzer

Herbert Y. Meltzer: Bixler Professor of Psychiatry and Pharmacology, Vanderbilt University School of Medicine, Nashville, Tennessee.

ATYPICAL ANTIPSYCHOTIC DRUGS: WHAT IS TO BE EXPLAINED

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

The designation of chlorpromazine, and subsequently haloperidol, thioridazine, loxapine, thiothixene, molindone, pimozide, and related compounds as *antipsychotic drugs* (1) reflects their predominant action in humans, namely the suppression of auditory hallucinations and delusions, in some, but not all, individuals with diagnoses of schizophrenia as well as other psychoses. These drugs are also called neuroleptics because they caused catalepsy in rodents and extrapyramidal side effects (EPSs) in humans (2). Their ability to diminish psychotic symptoms was convincingly shown to be initiated by blockade of dopamine (DA) D2 receptors in mesolimbic nuclei, especially the nucleus accumbens, stria terminalis, and the extended amygdala (2 ,3 ,4 and 5). Blockade of D2 receptors in terminal regions, e.g., the striatum, nucleus accumbens, and prefrontal cortex, by these agents initially causes compensatory increases in the activity of dopaminergic neurons in the substantia nigra and ventral tegmentum, respectively, followed by a gradual decrease in the activity of DA neurons, and, ultimately, complete inactivation of DA neuron firing in both regions (6). This so-called depolarization block was suggested to be the reason for the slow onset of antipsychotic action, which is observed in some, but not all, psychotic patients. The development of depolarization block following subchronic haloperidol treatment has been challenged by Melis et al. (7) on the basis of microdialysis studies of DA release in the nigrostriatal system, but those findings have been suggested by Moore et al. (8) to be an artifact.

These first-generation antipsychotic drugs are often referred to as *typical* antipsychotic drugs because they typically produce EPSs, e.g., acute dystonic reactions, subacute parkinsonism, and akathisia, and, after chronic use, tardive dyskinesia or dystonia (see Chapter 56) as a direct or indirect result of blockade of D2 receptors in the dorsal striatum, in vulnerable individuals. The immediate cause of acute and subacute EPSs is considered to be blockade of the dopaminergic inhibition of striatal cholinergic neurons, leading to increased cholinergic activity in the basal ganglia (2). Subsequently, clozapine was found to achieve an antipsychotic effect without causing EPSs, whereas loxapine, a clozapine congener, was equipotent in producing its antipsychotic action and EPS in humans and laboratory animals (10). This led Hippius and Angst to describe clozapine as an *atypical* antipsychotic drug (11). Preclinical scientists almost invariably refer to clozapine and other drugs that have antipsychotic properties and low EPSs as atypical antipsychotics but, as will be discussed, clinical investigators do not universally accept this designation.

The atypical profile of clozapine was initially attributed to its anticholinergic properties, which, along with other unknown features, caused selective depolarization of the A9 DA neurons in the substantia nigra, which project to the dorsal striatum, sparing those of the A10 ventral tegmentum, which project to the cortex and mesolimbic systems (12). The subsequent evidence that clozapine, compared to neuroleptic drugs such as haloperidol, had at least six advantages in addition to producing significantly less EPS and tardive dyskinesia, has attracted enormous interest to clozapine and other subsequently developed atypical antipsychotic drugs. These six effects of clozapine, not all of which are fully shared with other atypical antipsychotic drugs, are (a) absence of tardive dyskinesia; (b) lack of serum prolactin elevations in humans; (c) ability to eliminate positive symptoms without exacerbating motor symptoms in patients with Parkinson's disease who become psychotic due to exogenous dopaminergic agents such as levodopa (L-DOPA); (d) ability to decrease or totally eliminate psychotic symptoms in approximately 60% of the patients with schizophrenia who fail to respond to typical neuroleptic drug; (e) ability to improve primary and secondary negative symptoms;

and (f) ability to improve some domains of cognition in patients with schizophrenia, especially secondary memory and semantic memory (verbal fluency) (9 ,11 ,13). Some of the atypical antipsychotic drugs have also been shown to be more effective than the typical neuroleptic drugs in improving depression, stabilize mood, and decrease suicidality (9 ,10). These collective advantages of clozapine led to the search for the mechanism(s) involved in these effects and to find drugs that did not have the panoply of side effects of clozapine, especially agranulocytosis (9).

The other widely available antipsychotic drugs that are classified as atypical, by consensus, are, in order of their introduction, risperidone, olanzapine, sertindole, quetiapine, and ziprasidone. Melperone, a butyrophenone, introduced at about the same time as clozapine, has also been suggested to be an atypical antipsychotic drug because of its many clinical similarities with clozapine (14). Other agents with the essential clinical characteristics of an atypical antipsychotic drug that are currently at an advanced stage of clinical testing are iloperidone, ORG-5222, and aripiprazole (9 ,10). Zotepine and amisulpride, both of which are widely used antipsychotic drugs in Europe, are also sometimes grouped with the atypical antipsychotic drugs (9 ,10). With the exception of aripiprazole, a partial DA agonist (15), and amisulpride, a selective D2/D3 antagonist (16), all of the drugs listed above cause potent serotonin (5-hydroxytryptamine, 5-HT) receptor subtype 5-HT_{2A} relative to DA D2 receptor blockade (17 ,18 and 19).

There are a very large number of drugs in development as antipsychotics that have the property of being active in various animal models that predict antipsychotic action, e.g., blockade of amphetamine-induced locomotor activity or of the conditioned avoidance response, at doses 5- to 20-fold lower than that which produce catalepsy, a predictor of EPSs. All of these drugs are routinely referred to as putative *atypical* antipsychotic drugs, at least by preclinical scientists, because of their ability to produce an antipsychotic action at doses that do not cause significant EPSs in humans and a comparable dissociation in animal models of psychosis and EPSs, e.g., blockade of conditioned avoidance response and blockade of DA-induced locomotor activity, and the induction of catalepsy, respectively. These drugs differ greatly in chemical structure and, to some extent, pharmacologic profile, and thus cannot be referred to as a group by either chemical class or pharmacologic profile. However, some clinical investigators find the term *atypical antipsychotic drug* misleading because there are important clinical differences among the compounds with regard to the six clinical features of clozapine noted above, and they prefer the term *novel* or *new generation* over *atypical* to describe these agents. However, this temporal-based nomenclature is not routed in any meaningful or enduring characteristic of these agents. Others prefer to call them *multireceptor* antipsychotics, which is clearly preferable to *5-HT/DA* antagonists, another commonly used term. It is our view that these other designations have no specific advantages and some disadvantages compared to the classic term *atypical*. Thus, this chapter continues to use the term *atypical* to designate antipsychotic drugs that have a major advantage with regard to EPSs in patients with schizophrenia or Parkinson's disease, or both, to contrast with the typical antipsychotic drugs, and to update some of the key hypotheses for explaining some of the other highly valued advantages of these agents, as well as their unique side effects.

As can be expected, there has been an intensive effort to determine the basis for the differences between the typical and atypical antipsychotic drugs. This chapter reviews the major hypotheses, which are based on the pharmacologic profiles of the numerous classes of agents with atypical properties as well as current theories of the action of drugs effective in animal models of psychosis, but not yet adequately tested in humans, e.g., AMPA antagonists and metabotropic glutamate receptor antagonists (20 ,21).

The affinities of clozapine and some of its congeners for monoamine receptors are given elsewhere in this volume (see Chapter 56) (22). The affinities reported therein are for the D1, D2, D3, D4, 5-HT_{1A}, 5-HT_{1D}, 5-HT_{2A}, 5-HT_{2C}, 5-HT_{6}, 5-HT_{7}, α_1, α_2, H_1, and muscarinic M1 receptors. Of these, the greatest interest is in the role of D2, D4, 5-HT_{1A}, 5-HT_{2A}, 5-HT_{2C}, α_1, and α_2 receptors. Clozapine does have effects on glutamate and GABA neurons and interneurons, respectively, but space considerations preclude their discussion here.

ROLE OF D2 RECEPTORS

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

Most effective antipsychotics, typical as well as atypical, have affinities for the DA D2 receptor high enough to suggest that they produce effective blockade of these receptors *in vivo* (23 ,24). The model for atypical antipsychotic drug action proposed by Meltzer et al. (17) postulated that atypical antipsychotic drugs had to have some D2 receptor blockade *in vivo*, although weaker than 5-HT_{2A} receptor blockade, to achieve a low EPS profile and, possibly, some of the other advantages of clozapine. An exception to this may be amperozide, with is a potent 5-HT_{2A} antagonist and DA reuptake inhibitor with very low affinity for the D2 receptor (25). Recently, NRA0045, which has potent 5-HT_{2A}, D4, and α_1 but no D2 or D3 receptor blockade has been found to have atypical antipsychotic properties (26). Partial DA agonists, which may act as agonists at presynaptic DA receptors, and antagonists at postsynaptic DA receptors are a new class of antipsychotic drugs that has promise (15 ,27 ,28). However, clinical testing of these agents is just beginning, and current data support only the view that they are atypical in the classic sense, i.e., they are antipsychotic in preclinical or clinical testing at doses that produce weak or absent EPSs. Other evidence of the importance of α_1 antagonism for atypical antipsychotic drug activity will be discussed subsequently. The *in vitro* affinity of a drug at the DA D2 receptor

is a useful predictor of the dose that produces EPSs and control of positive symptoms for typical neuroleptic drugs (2 ,29), although it does not do so for some atypical antipsychotic drugs, e.g., ziprasidone. Furthermore, there is no agreement on how to determine the doses used in such correlations because of the differences in dosage requirements as a function of stage of illness, body mass index, and age. There is little agreement even on the best dose for haloperidol, the most widely used antipsychotic drug. A wide range of 2 to 15 mg/day has been suggested, far removed from the 20 to 40 mg/day thought to be most effective in 1966 (9 ,10). To date, no clinically proven antipsychotic with the possible exception of amperozide lacks significant D2 receptor antagonist properties. As will be discussed, the combination of D2 and 5-HT_{2A} receptor blockade, in the right ratio, produces some of the effects of clozapine and other atypical antipsychotic drugs in rodents, e.g., increases DA efflux in the cortex and striatum of rats (30) and blockade of the conditioned avoidance response, an indication of antipsychotic activity (31). There have been only limited tests of this hypothesis in humans, mainly using ritanserin, which is a mixed 5-$HT_{2A/2B/2C}$ antagonist (32). Nevertheless, various comprehensive reviews of the action of the atypical antipsychotic drugs have concluded that the combination of 5-HT_{2A}, D2, and α_1 receptor blockade is the probable basis of their antipsychotic action (10 ,18 ,19 ,22 ,33 ,34 ,35 ,36 and 37). The evidence for this hypothesis will be discussed subsequently.

Counter to the hypothesis of the importance of 5-HT_{2A} receptor antagonism to the action of clozapine and other atypical antipsychotic drugs is the proposal of Seeman and Tallerico (24) and Kapur and Seeman (38) that the basis of atypical antipsychotics may lie in their rapid dissociation from the DA D2 receptor and their relatively easily displacement by surges of endogenous DA. It has also been proposed that rapid and extensive displacement of clozapine and quetiapine from binding sites accounts for the reported low occupancy of striatal D2 receptors by these drugs (24). The authors also suggested that this might account for more rapid relapse following clozapine and quetiapine withdrawal (24). Although the evidence cited for clozapine-induced relatively rapid relapse is robust (39), the evidence with regard to quetiapine and rapid relapse has never been published and does not accord with general clinical experience. Seeman and Tallerico found that the affinity for and rate of dissociation of antipsychotics from the D2 receptor are highly correlated. Drugs with low affinity for the D2 receptor, e.g., clozapine and quetiapine, were found to have a higher dissociation rate constant than drugs with higher affinity, e.g., haloperidol. Rapid dissociation from the D2 receptor was reported to also permit easier displacement of clozapine and quetiapine by endogenous DA, thereby avoiding side effects related to DA receptor blockade such as EPSs and hyperprolactinemia (38). It was also reported that olanzapine, risperidone, and sertindole, all of which are well established as atypical antipsychotic drugs, are comparable to haloperidol in their rate of dissociation from the D2 receptor and are not displaced by raclopride or iodobenzamide, as are clozapine and quetiapine (24). Thus, this hypothesis could not explain the basis for their low EPSs. Moreover, for these agents to achieve their antipsychotic action, they would have to be less easily displaced from limbic and possibly cortical D2 receptors. There are no data to support this selectivity with regard to displacement as yet. Although there is evidence for higher occupancy of extrastriatal D2 receptors by clozapine and quetiapine in patients with schizophrenia (40 ,41), and for atypical antipsychotic drugs that show more potent 5-HT_{2A} receptor blockade in rodents (42), the same appears to be true for olanzapine, which does not show the higher off-rates of clozapine and quetiapine (R. Kessler and H. Meltzer, in preparation). Because clozapine produces a greater increase in DA release in the cortex than in the accumbens or striatum, at least in rodents and monkeys (43), clozapine might be expected to produce greater occupancy of D2 receptors in these regions than the cortex, but, as noted above, this is not the case. It is also not clear how this model could explain any of the advantages of clozapine with regard to efficacy in neuroleptic-resistant patients or for cognition.

Studies on the regulation of prefrontal cortical or limbic DA release also provide no evidence that blockade of D2 receptors alone, regardless of degree of occupancy, can mimic the effects of the multireceptor antagonists such as clozapine (31 ,44 ,45). Pharmacologic analysis of this important model for the action of atypical antipsychotic drugs on cognition and negative symptoms strongly supports the importance of combined blockade of 5-HT_{2A}, D2, and possibly α_1 receptors (36 ,37 ,44 ,45).

D3 AND D4 RECEPTOR MECHANISMS

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

Ever since the cloning and characterization of the distribution of the D3 and D4 receptors, which revealed a limbic and cortical distribution, there has been considerable speculation about the role of these receptors in schizophrenia and the mechanism of action of antipsychotic drugs (19 ,46 ,47 and 48). As specific antagonists have become available, it has been possible to test their efficacy in animal models of psychosis, cognition, and motor function, as well as to carry out some clinical trials in schizophrenia. Recently, Reavill et al. (49) reported that SB-277011-A, which has high affinity and selectivity for the D3 receptor and good brain bioavailability, has an atypical antipsychotic drug profile. This compound was active in preventing isolation-induced deficits in prepulse inhibition but was not effective in blocking either amphetamine- or phencyclidine (PCP)-induced locomotor activity or, by using microdialysis, to increase prefrontal cortical DA release in rats (49). However, subchronic administration of SB-277011-A selectively decreased the firing rate of A10, but not A9, DA neurons in the rat, indicating a clozapine-like profile (50). These are the most promising

data yet that a selective antagonist of D3 receptors might be useful in the treatment of psychosis.

On the basis of the finding that the affinity of clozapine for the cloned DA D4 receptor was two to three times greater than that for the D2 receptor, Van Tol et al. (51) suggested that blockade of the D4 receptor was the basis for the superiority of clozapine over the typical neuroleptic drugs in the treatment of schizophrenia. Other investigators have found less of a difference between the affinity of clozapine for D2 and D4 receptors (48). Several typical antipsychotic drugs, including haloperidol, have nearly equivalent affinity for D2 and D4 receptors, suggesting that D4 affinity per se does not convey any special advantages for an antipsychotic drug (48). The preclinical behavioral and electrophysiologic profile of highly selective D4 antagonists provides mixed evidence with regard to antipsychotic action associated with this receptor (52 ,53 and 54). A D4/α_1 antagonist, NRA0025, appears to have promise as an antipsychotic agent based on its preclinical profile (26). A clinical trial of a selective D4 antagonist showed no sign of activity (55). A compound with potent 5-HT_{2A} and D4 antagonist properties, finanserin, was also ineffective (56). Further clinical trials of compounds that have D4 or D3 antagonism, or both, together with high affinities for 5-HT_{1A} or α_1 receptors, and without D2 affinity seems indicated.

ROLE OF SEROTONIN IN ANTIPSYCHOTIC DRUG ACTION

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

Determining the biological basis for the advantages of clozapine and other atypical antipsychotic drugs cited above and described in detail by Miyamoto et al. in Chapter 56 has led to considerable interest in the role of 5-HT in antipsychotic drugs action. We now consider some of this evidence. Other reviews of this topic should also be consulted (18 ,19 ,22 ,34 ,35 ,36 and 37 ,57).

Serotonin Receptors Involved in Antipsychotic Drug Action

The hypothesis that a relatively high affinity for the 5-HT_{2A} receptor compared to an affinity for the D2 receptor was the basis for the difference between atypical and typical antipsychotic agents contributed to the development of the newer antipsychotic agents listed above, all of which support the previously mentioned hypothesis of high affinity for 5-HT_{2A} and low affinity for D2 receptors (17 ,58 ,59). However, other 5-HT receptors may be important to the action of clozapine and other recently introduced antipsychotic agents, or of potential value for developing more effective or better tolerated antipsychotic agents. These include the 5-HT_{1A}, 5-HT_{2A}, 5-HT_{2C}, 5-HT_3, 5-HT_6, and 5-HT_7 receptors (19 ,22 ,33). Although some of the atypical antipsychotic drugs developed on the basis of the 5-HT_{2A}/D2 hypothesis also have affinities for 5-HT_{2C}, 5-HT_3, 5-HT_6, or 5-HT_7 receptors that are in the same range as that for the 5-HT_{2A} receptor, this is not characteristic of all of these agents, and thus it is not likely that affinities for these receptors are primary factors contributing to the low EPS profile of the entire class of agents (19 ,22 ,60 ,61). However, this does not rule out that actions at various 5-HT receptors contribute to low EPSs of specific drugs, or other actions, e.g., cognitive improvement or improvement in negative symptoms. For example, 5-HT_{1A} receptor agonism has also been suggested to be able to contribute to an atypical antipsychotic drug profile (62), and some of the atypical antipsychotics are 5-HT_{1A} partial agonists as well as 5-HT_{2A}/5-HT_{2C} antagonists, e.g., clozapine, quetiapine, ziprasidone, and S16924 (63 ,64). The role of the 5-HT_4 receptor in cognition will be discussed subsequently. Furthermore, there is some evidence of interactions among the 5-HT_{1A}, 5-HT_{2A}, and 5-HT_{2C} receptors (19 ,35). Because of space limitations, this chapter focuses on these three 5-HT receptors and briefly considers the others.

ATYPICAL ANTIPSYCHOTICS AND THE 5-HT_{2A} RECEPTOR

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

5-HT_{2A} receptors have been implicated in the genesis of, as well as the treatment of, psychosis, negative symptoms, mood disturbance, and EPSs (17 ,19 ,33 ,34 ,35 and 36 ,45). The hallucinogenic effect of indole hallucinogens has been related to stimulation of 5-HT_{2A} rather than 5-HT_{2C} receptors (65). Numerous studies have examined the density of 5-HT_{2A} receptors in various cortical regions of patients with schizophrenia with decreased (66 ,67), increased (68), or normal levels reported. It is well established that some typical and atypical antipsychotic drugs can decrease the density of 5-HT_{2A} receptors (69 ,70), so the postmortem results noted above may be related to drug treatment. Positron emission tomography (PET) studies have not found decreased 5-HT_{2A} receptors in the cortex of never-medicated or unmedicated patients with schizophrenia (71). As mentioned above, the antipsychotic effect of clozapine has been attributed, in part, to its ability to block excessive 5-HT_{2A} receptor stimulation without excessive blockade of D2 receptors (17). This conclusion is consistent with the high occupancy of 5-HT_{2A} receptors produced by clozapine at clinically effective doses and its low occupancy of D2 receptors (in the 30% to 50% range as measured with [^{3}H]raclopride), the latter being significantly below the 80% to 100% occupancy usually produced by typical neuroleptic drugs (35 ,72 ,73 ,74 and 75). The occupancy of 5-HT_{2A} and D2 receptors has been studied with other novel antipsychotic drugs such as risperidone, olanzapine, sertindole, and quetiapine with results similar to those of clozapine; all are more potent 5-HT_{2A} and D2 antagonists at appropriate doses, but less so than clozapine (72 ,73 ,74 and 75).

Some of these agents (e.g., risperidone and olanzapine) produce high D2 occupancy at high doses (76 ,77).

The bell-shaped dose–response curve of risperidone, with higher doses being less effective than lower doses (78), is consistent with the hypothesis that excessive D2 receptor antagonism may diminish some of the beneficial effects of 5-HT_{2A} receptor blockade (17 ,19 ,33 ,35 ,36). The highly selective 5-HT_{2A}agonist M100907, formerly MDL 100907, has been found in a controlled study to have some efficacy for treating positive and negative symptoms in hospitalized schizophrenic patients (79). However, because it was less effective than haloperidol, no further testing in schizophrenia has been scheduled at present. Nevertheless, the concept that 5-HT_{2A} antagonism may be useful to treat some forms of psychosis, especially when combined with weak D2 receptor blockade, warrants further study. Other 5-HT_{2A} selective agents such as SR 46349B (80) are currently being tested. Additional clinical evidence supporting the role of 5-HT_{2A} receptor blockade in the action of clozapine and possibly other drugs with potent 5-HT_{2A} affinities is available from the several reports that the His452Tyr allele of the 5-HT_{2A} receptor, which is present in 10% to 12% of the population, is associated with a higher frequency of poor response to clozapine (81 ,82). Taken together, the evidence from clinical trial data suggests that 5-HT_{2A} receptor blockade may contribute to antipsychotic drug action.

There is additional basic research that is also consistent with the relevance of 5-HT_{2A} receptor blockade for antipsychotic drug action. Thus, M100907 or other selective 5-HT_{2A} receptor antagonists, either alone or in combination with selective antagonists of other receptors, have been found to be effective in various animal models of psychosis. These include (a) blockade of amphetamine-induced locomotor activity and the slowing of ventral tegmental area (VTA) (A10) dopaminergic neurons (34); (b) blockade of PCP- and dizocilpin (MK-801)–induced locomotor activity (83 ,84); (c) blockade of MK-801–induced prepulse inhibition (85); and (d) antipsychotic-like activity in the paw test (86) among others. Of particular interest is the report of Wadenberg et al. (31) that the combination of a median effective dose (ED_{50}) of raclopride, a D2 receptor antagonist, and M100907, but not M100907 alone, was effective in blocking the conditioned avoidance response. The authors concluded that 5-HT_{2A} antagonism alone could not achieve an antipsychotic action, but that minimal blockade of D2 receptors was required to achieve such an effect. This corresponds much more closely to the apparent clinical situation than do the models where M100907 alone was effective, e.g., blockade of PCP- or MK-801–induced locomotor activity (34 ,83 ,84 ,87 ,88 and 89). As mentioned previously, administration of even a single dose of atypical antipsychotic drugs, which are relatively more potent 5-HT_{2A} than D2 blockers, has been shown to down-regulate 5-HT_{2A} receptors in rat brain (69 ,70). This has now been shown to be due to internalization of the receptors (70). Recovery from this process requires the synthesis of new receptors.

An important effect of 5-HT_{2A} (and 5-HT_{2C}) receptors that may be relevant to their contribution to psychosis is their ability to influence dopaminergic activity in the mesolimbic and mesostriatal systems (19 ,33 ,90 ,91 ,92 ,93 and 94). Increased dopaminergic activity in the nucleus accumbens and other mesolimbic and possibly cortical regions may contribute to positive symptoms, including formal thought disorder (2 ,5). Increased dopaminergic activity in the striatum would be expected to diminish EPSs (2 ,5). The 5-$HT_{2A/2C}$ agonist DOI [1-(2,5-dimethoxy-4-iodophenyl)-2-aminopropane], which itself had no effect on basal DA release, potentiated amphetamine-induced DA release and attenuated the ability of apomorphine, a direct acting D1/D2/D3 agonist, to decrease DA release in the striatum (93). Increasing serotoninergic activity, e.g., by administration of selective serotonin reuptake inhibitors (SSRIs) alone or in combination with the 5-HT_{1A} receptor antagonist WAY 100635, has no effect on basal DA output in the striatum. However, the SSRIs can significantly enhance the increase in DA outflow induced by haloperidol. These findings indicate that in the striatum, endogenous 5-HT positively modulates DA outflow when nigrostriatal DA transmission is activated (94). There is now considerable evidence from both behavioral and neurochemical studies involving *N*-methyl-D-aspartate (NMDA) antagonists such as PCP and MK-801 that 5-HT_{2A} receptors modulate activated but not basal mesolimbic DA function (84 ,95). Thus, stimulated DA release, e.g., with stress, may be increased in the forebrain terminal regions secondary to enhanced stimulation of 5-HT_{2A} receptors. Agents that block the effect of excessive, but not basal, 5-HT_{2A} receptor stimulation may be the most useful clinically. M100907 has been found to diminish the increase in DA efflux in the nucleus accumbens produced by haloperidol (30) or S-sulpiride (92). Taken together, these data suggest that 5-HT_{2A} antagonism by itself may have antipsychotic action when dopaminergic activity is slightly to moderately increased. More studies are needed to define the ability of 5-HT_{2A} receptor antagonists to potentiate the action of low doses of D2 receptor blockers in animal models as well as in humans.

Jakab and Goldman-Rakic (96) have proposed that the 5-HT_{2A} receptors on cortical pyramidal neurons may play a crucial role in psychosis by virtue of their ability to modulate intracortical and cortical-subcortical glutamatergic neurotransmission. This could contribute to the ability of 5-HT_{2A} antagonists to attenuate some of the behavioral effects of PCP and ketamine. Aghajanian and Marek (97) have proposed a link between the glutamate hypothesis of schizophrenia and the hallucinogen hypothesis. Briefly, stimulation of 5-HT_{2A} receptors on layer V pyramidal cells increases the frequency of postsynaptic potentials (PSPs). These are mostly blocked by the AMPA/kainate glutamatergic receptor antagonist LY293558, indicating that

they are mainly excitatory PSPs (EPSPs), in contrast with the piriform cortex, where 5-HT produces inhibitory PSPs (IPSPs). The selective group II metabotropic agonist LY354740, which inhibits glutamate release by stimulating inhibitory presynaptic autoreceptors on glutamatergic nerve terminals, suppresses the 5-HT–induced increase in the frequency of EPSPs. However, Aghajanian and Marek suggest that the main way in which 5-HT_{2A} receptor agonists increase glutamate release is a retrograde action from stimulation of postsynaptic 5-HT_{2A} receptors. The type of glutamate release induced by 5-HT_{2A} receptor stimulation differs from ordinary depolarization-induced neurotransmitter release, which is called synchronous release. The type of release induced by 5-HT is delayed in onset, slow, and produces small excitatory postsynaptic currents (EPSCs) and is called asynchronous release. Aghajanian and Marek propose that hallucinogens such as LSD enhance asynchronous EPSCs. They suggest that stimulation of other types of 5-HT receptors may oppose this action and that the effect of 5-HT_{2A} receptor antagonists unmasks the effect of these other types of 5-HT receptors (97). A similar proposal has been made by Martin et al. (95). Although Aghajanian and Marek do not mention the 5-HT_{1A} receptor specifically, it would be a good candidate to counter the effect of 5-HT_{2A} receptor stimulation, as will be discussed. They propose that the thalamic filter hypothesis of Carlsson (98) might be related to the effect of 5-HT_{2A} receptor stimulation on thalamocortical afferents to affect cortical function or on corticostriatal or corticothalamic efferents to affect the thalamic filter.

SEROTONIN RECEPTORS AND COGNITIVE FUNCTION

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

Clozapine, risperidone, ziprasidone, quetiapine, and olanzapine have been shown to improve selected areas of cognitive function in patients with schizophrenia, with the available data suggesting differential effects on specific functions (13). The available data suggest that each of the atypical drugs has a different pattern of effects on cognitive dysfunction in schizophrenia, but more head-to-head studies are needed to confirm this impression. Whether relatively more potent 5-HT_{2A} receptor compared to D2 antagonism has a major or, indeed, any role in the cognitive effects of these agents is not known. However, this is the major characteristic that these drugs share in common. It may be that the effect of these agents on cognition is mainly dependent on their ability to increase the release of DA (99 ,100) and acetylcholine in prefrontal cortex (101), which may depend, in part, on their serotoninergic actions. The effect of the atypical agents to increase DA efflux in the medial prefrontal cortex (mPFC) of rats appears to be due mainly to actions at the terminal regions rather than cell bodies and not to be related to D2 receptor blockade because local administration of haloperidol is without effect whereas clozapine and olanzapine produce huge increases in DA efflux (102). Studies have suggested that the clozapine, olanzapine, risperidone, and ziprasidone, but not haloperidol, may enhance acetylcholine release in the rat prefrontal cortex (101 ,103). Interactions between the 5-HT and cholinergic systems have been previously reported (104). 5-HT_{1A}, 5-HT_{2C}, 5-HT_3, and 5-HT_4 receptors have also been reported to have significant effects on acetylcholine release in the rat prefrontal cortex (105 ,106 ,107 ,108 and 109).

Because cognitive enhancement is critical for functional improvement in schizophrenia, establishing the mechanism for the ability of the atypical antipsychotic drugs to increase DA efflux is of the greatest importance. The evidence concerning 5-HT receptors and cognition has been reviewed in detail elsewhere (110 ,111). 5-$HT_{2A/2C}$ antagonists have little adverse effect and no apparent beneficial effects on learning and memory (112). There is some evidence that 5-HT_4 agonists, e.g., RS 67333, can improve learning and memory in rodents (113). Impairment of working memory in humans following administration of the 5-HT_{1A}agonist flesinoxan has been reported (114). However, Sumiyoshi et al. (115) have found that tandospirone, a 5-HT_{1A} partial agonist, can improve other domains of cognition in patients with schizophrenia treated with typical neuroleptic drugs. Neurochemical differences in the patient populations, including possible abnormalities in the density of 5-HT_{1A} receptors in schizophrenia, concomitant administration of a neuroleptic to the patients with schizophrenia, and differences in the type of cognitive domain studied, may account for this discrepancy. Further study in patients with schizophrenia is clearly indicated.

5-HT_{2A} RECEPTOR BLOCKADE AND EXTRAPYRAMIDAL FUNCTION

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

There have been numerous suggestions to explain the low EPSs of clozapine, namely its anticholinergic properties, lack of ability to increase acetylcholine in the striatum, D1 or D4 receptor blockade, and its effects as an α_1- or α_2-adrenoceptor antagonist (2 ,18 ,33 ,116). In addition, several lines of evidence suggest that potent 5-HT_{2A} receptor blockade is relevant to the low EPS profile of clozapine, but that 5-HT_{2A} receptor blockade by itself cannot explain the low EPS liability of these agents (17 ,58 ,59). Meltzer et al. (17) studied a group of compounds that had antipsychotic activity in humans or in animal models that are thought to be predictive of antipsychotic activity, e.g., conditioned avoidance response or blockade of amphetamine-induced locomotor activity, and which produced weak EPSs in humans or weak catalepsy in animals relative to their antipsychotic efficacy. These compounds shared in common relatively weaker D2 compared to 5-HT_{2A} receptor affinities, whereas D1 receptor affinities did not contribute to this effect.

Among the drugs studied were melperone, a butyrophenone long used in Europe and Scandinavia as an antipsychotic and reported to produce low EPSs (14). Indeed, it has been found to be tolerable to patients with Parkinson's disease (117), even more so than risperidone and olanzapine. The ∩-shaped dose–response curve of risperidone as well as the increasing incidence of EPSs as the dose increases for olanzapine and ziprasidone, together with PET studies of DA receptor occupancy previously discussed, strongly suggests that lower D2 receptor occupancy, possibly in relation to high 5-HT_{2A} receptor is necessary to avoid EPSs with these compounds.

As previously discussed, numerous compounds of diverse chemical structure that share this pharmacologic profile have been identified or deliberately synthesized and tested for antipsychotic action and EPS liability. These include risperidone, olanzapine, sertindole, quetiapine, ziprasidone, and iloperidone. All of these compounds can produce fewer EPSs than haloperidol at comparable doses. Clozapine and quetiapine have been shown to produce the least EPSs in studies of patients with Parkinson's disease. Consistent with this concept, the 5-HT_{2A} antagonist mianserin has been reported to be effective in neuroleptic-induced akathisia (118). There are also a variety of preclinical data to support the importance of relatively high 5-HT_{2A} compared to D2 receptor affinity to preserve striatal function. For example, Ishikane et al. (119) reported that M100907 is able to block haloperidol-induced catalepsy only at low doses of haloperidol. Consistent with this, Spampinato et al. (120) reported that specific 5-HT_{2A} and 5-HT_{2C} antagonists were able to modulate the ability of haloperidol at 0.01 mg/kg but not at a higher dose (1.0 mg/kg) to increase striatal DA release in freely moving rats.

THE ROLE OF THE 5-HT_{2C} RECEPTOR IN ANTIPSYCHOTIC DRUG ACTION: 5-HT_{2A} AND 5-HT_{2C} INTERACTIONS

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

There has been some consideration given to the role of 5-HT_{2C} receptors in the action of atypical antipsychotic drugs. The 5-HT_{2C} receptor is found throughout the central nervous system (CNS), including the ventral tegmentum and the nucleus accumbens (121). With the availability of specific 5-HT_{2C} agonists and antagonists, evidence for a tonic inhibitory action of 5-HT_{2C} receptors on the burst firing of mesolimbic and mesocortical dopaminergic neurons has been obtained. Thus, the firing rate of VTA DA neurons is inhibited or increased by 5-HT_{2C} agonists or antagonists, respectively. This is consistent with microdialysis studies that show that 5-HT_{2C} antagonists increase extracellular concentrations of DA in the nucleus accumbens, striatum, and medial prefrontal cortex (91 ,122). Early studies found no significant differences between groups of novel antipsychotic drugs and typical neuroleptics with regard to the affinity for 5-HT_{2C} receptor or the difference between 5-HT_{2C} and D2 affinities (22 ,60). Of the approved novel antipsychotic drugs, some have equivalent affinities for the 5-HT_{2A} and 5-HT_{2C} receptors (clozapine, olanzapine, sertindole), whereas others are more selective for the 5-HT_{2A} receptor (risperidone, quetiapine, ziprasidone). This difference roughly corresponds with the potential to produce weight gain, in that clozapine and olanzapine cause the greatest weight gain and risperidone and ziprasidone the least (see Chapter 56). There is little available data for sertindole and quetiapine, but they appear to be intermediate. There is no apparent relationship between 5-HT_{2C} affinity relative to 5-HT_{2A} affinity with regard to EPSs because quetiapine and ziprasidone are comparable to olanzapine and sertindole in this regard. Similarly, there is no apparent relationship to efficacy in treatment-resistant schizophrenia. There could be a relationship to differences among the atypical antipsychotic drugs with regard to improvement in specific types of cognitive function in schizophrenia (13).

An interesting aspect of the 5-HT_{2C} receptor with regard to antipsychotic action is that 5-HT_{2C} antagonism may be functionally opposed to 5-HT_{2A} antagonism. Meltzer et al. (123) reported that atypical antipsychotic drugs were more likely to be weak 5-HT_{2C} and potent 5-HT_{2A} antagonists compared to typical neuroleptic drugs. Subsequently, neurochemical (120) and behavioral (96 ,124) data have been reported that support the notion of a functional antagonism of these two receptors that may coexist on the same neurons. Thus, Martin et al. (95) found that ritanserin, a mixed 5-$HT_{2A/2C}$ antagonist, blocked the ability of M100907 to antagonize the effect of MK-801 to increase locomotor activity in mice.

THE ROLE OF THE 5-HT_{1A} RECEPTOR IN ANTIPSYCHOTIC DRUG ACTION

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

The 5-HT_{1A} receptor is located pre- and postsynaptically. The presynaptic 5-HT_{1A} receptor is an autoreceptor located on cell bodies of raphe neurons; stimulation leads to inhibition of firing of 5-HT neurons. Stimulation of postsynaptic 5-HT_{1A} receptors generally leads to hyperpolarization of neurons, which is opposite of the effect of stimulation of 5-HT_{2A} receptors. There is extensive evidence that cannot be reviewed in detail here that indicates that 5-HT_{1A} receptor agonists and 5-HT_{2A} receptor antagonists produce similar neurochemical and behavioral effects on a variety of measures (125 ,126). For example, DOI injected bilaterally into the rat medial prefrontal cortex elicits a dose-dependent head twitch response. This effect is inhibited by M100907 and ketanserin, relatively selective for 5-HT_{2A} receptors at appropriate doses, but not the selective 5-HT_{2C} antagonist SDZ SER082. Pretreatment with the 5-HT_{1A} agonist 8-OH-DPAT also inhibited the head twitch response to DOI (127). Ahlenius (128) first suggested that stimulation

of 5-HT_{1A} receptors might produce an antipsychotic like action on the basis of behavioral studies in animals using the direct 5-HT_{1A} agonist 8-OH-DPAT. Subsequent studies demonstrated that 8-OH-DPAT enhanced the antipsychotic-like effect of the D2/D3 antagonist raclopride (129) and of haloperidol (130), and antagonized the catalepsy induced by the D1 agonist SCH23390 in rats (131). The ability of clozapine to reverse olanzapine-induced catalepsy is blocked by the selective 5-HT_{1A} antagonist WAY 100635, suggesting the effect of clozapine was mediated by stimulation of 5-HT_{1A} receptors. The beneficial effect of 5-HT_{1A} agonists appears to be mediated by inhibition of median raphe serotoninergic neurons (132).

5-HT_{1A} agonists have different regional effects on DA release in the rat brain. 5-HT_{1A} receptor stimulation appears to inhibit DA release in subcortical regions. Thus, Ichikawa et al. (133) demonstrated that the 5-HT_{1A} agonist 8-OH-DPAT inhibited the ability of amphetamine to increase extracellular DA levels in the nucleus accumbens and the striatum of conscious rats. The effect of 5-HT_{1A} receptor stimulation in the nucleus accumbens would be expected to enhance the antipsychotic effect of these agents by reducing dopaminergic activity. Several atypical antipsychotic drugs, including clozapine, ziprasidone, quetiapine, and tiospirone, are partial agonist at the 5-HT_{1A} receptor. Their affinities for the 5-HT_{1A} receptor are similar to their affinities for the human D2 receptor (22). Rollema et al. (134) demonstrated that the ability of clozapine to increase DA release in the rat prefrontal cortex was due, in part, to its 5-HT_{1A} agonist properties, as it could be blocked by WAY-100635, a 5-HT_{1A} antagonist. Ichikawa et al. (92) have extended these findings in a variety of ways; e.g., the ability of risperidone and the combination of M100907 and sulpiride to increase prefrontal cortical PFC DA efflux were both blocked by WAY100635. These findings suggest that the combination of D2 antagonism and 5-HT_{1A} agonism provides some of the key features of atypical antipsychotic agents. S16924, a 5-HT_{1A} partial agonist D2 antagonist, is an example of a putative atypical antipsychotic drug based on this model. It has atypical antipsychotic properties very similar to those of clozapine in a variety of relevant animal models (64). Whether this or similar compounds will have the same spectrum of efficacy and side effect advantages as the multireceptor antagonists that are relatively more potent 5-HT_{2A} than D2 antagonists remains to be determined. Significant differences should be expected. It is noteworthy that clozapine has both relatively more potent 5-HT_{2A} antagonism than D2 antagonism as well as 5-HT_{1A} partial agonism. This may be part of the mixture that accounts for its particular advantages over other atypical antipsychotic drugs. Wedzony et al. (135) found that WAY100135, a 5-HT_{1A} antagonist, attenuated the effect of MK-801, an NMDA antagonist on locomotor activity, prepulse inhibition, and the detrimental effect of MK-801, a noncompetitive NMDA antagonist on working memory and selective attention in rats. They cite other evidence that 5-HT_{1A} antagonists may improve learning and memory in animal models and suggest this may be due to blocking the inhibitory effects of 5-HT_{1A} receptor stimulation on the firing of hippocampal neurons. This suggests that a partial agonist, acting as an antagonist, may sometimes be of benefit with regard to effects relevant to schizophrenia.

THE ROLE OF SEROTONIN RELEASE IN ANTIPSYCHOTIC DRUG ACTION

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

The antagonism of multiple 5-HT receptors by clozapine would be expected to enhance the release of 5-HT by feedback mechanisms. Thus, it is surprising that Ferré and Artigas (136) reported that clozapine decreased 5-HT release in the nucleus accumbens. However, Ichikawa et al. (137) reported that clozapine (20 mg/kg) and risperidone (1 mg/kg) significantly increased extracellular 5-HT levels in the nucleus accumbens and medial prefrontal cortex, respectively, whereas amperozide (1 and 10 mg/kg) increased extracellular 5-HT levels in both regions. Hertel et al. (138) reported similar results with risperidone and suggested that this might be relevant to its ability to improve negative symptoms. If so, this is not the explanation for the effects of clozapine or olanzapine on negative symptoms because olanzapine, sulpiride, haloperidol, and M100907 had no effect on extracellular 5-HT levels in either region. The latter consideration also indicates that blockade of 5-HT_{2A} receptors is not the basis for the ability of clozapine, risperidone, or amperozide to increase 5-HT levels. The enhancement of 5-HT efflux in the prefrontal cortex may contribute to the ability of these agents to improve mood disorders and cognition.

α_2- AND α_1-ADRENERGIC MECHANISMS AND ATYPICAL ANTIPSYCHOTIC DRUGS

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

Most of the atypical antipsychotic drugs are potent antagonists of the α_1 or α_2 adrenoceptors, or both. Thus, risperidone 9-hydroxyrisperidone, clozapine, olanzapine, zotepine, quetiapine, ORG-5222, sertindole and ziprasidone are potent α_1 antagonists (22). Prazosin, an α_1 adrenoceptor antagonist, has, like clozapine and other atypical antipsychotic drugs, been shown to increase DA efflux in the shell but not the core of the nucleus accumben, signifying a limbic rather than a striatal effect of α_1 antagonism (91). These authors also suggested that α_1 antagonism may explain the atypical properties of sertindole, which has been reported to achieve as high an occupancy of D2 receptors as typical antipsychotic drugs (36). All of the atypical agents mentioned above are also potent α_2 antagonists, with the exception of zotepine and sertindole (22). Kalkman et al. (116) raised the possibility that the α_{2C} subtype may be particularly relevant to the anticataleptic as well as other actions of clozapine and iloperidone. However, McAllister and Rey (139) were unable to reverse the effects of loxapine or haloperidol

on catalepsy with α_2 antagonists and showed that the effect of clozapine to reverse loxapine-induced increase in catalepsy was due to its anticholinergic rather than its adrenoceptor blocking properties. Clozapine produces massive increases in plasma norepinephrine, which may indicate that it can cause effective stimulation of α-adrenoceptors receptors in brain (140). The addition of idazozan, an α_2 antagonist, to fluphenazine, a typical neuroleptic, was reported by Littman et al. (141) to have efficacy comparable to clozapine in a small group of neuroleptic-resistant patients with schizophrenia. These results need to be replicated. Idazoxan has also been shown to improve attentional and executive dysfunction in patients with dementia of the frontal type (142), suggesting that some of the cognitive enhancing effects of the atypical antipsychotic drugs might be related to their α_2 blocking properties. Another α_2 antagonist, atipamezole, has been reported to improve cognitive performance in aged rats (143). Polymorphisms of the α_1 and α_2 receptors have been reported not to predict response to clozapine (144).

In this regard, it is of interest that idazoxan has been shown to preferentially increase DA efflux in the rat mPFC by an action at the terminal area (145). This effect appears to be independent of dopaminergic activity (146). Westerink et al. (147) demonstrated that systemic administration of clozapine, risperidone, ziprasidone, and olanzapine, as well as haloperidol, produced a dose-dependent increase in noradrenaline in the mPFC of rats. This effect was closely coupled to the increase in DA efflux. Increased levels of norepinephrine might also be related to the cognitive and antidepressant effects of the atypical antipsychotic drugs (148 ,149).

CONCLUSION

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

Typical neuroleptic drugs such as haloperidol have been reliably shown to produce their antipsychotic action by blockade of D2 receptors in the mesolimbic system, suggesting that increased dopaminergic activity in these terminal areas of the ventral tegmental DA neurons are of importance to the etiology of schizophrenia. Striatal D2 receptor antagonism is the critical element in the EPSs produced by these drugs. Atypical antipsychotic drugs are those antipsychotics that achieve an antipsychotic action with quantitatively less EPSs in humans or a clear distinction between doses that affect mesolimbic and striatal dopaminergic function in rodents. Clozapine was the first atypical antipsychotic drug by the definition noted above, but more importantly it showed that antipsychotic drugs might also be effective in some patients with schizophrenia whose positive symptoms do not respond to neuroleptic-type agents and to improve negative symptoms, cognitive impairment, depression, and possibly suicidality of schizophrenia and other psychotic disorders as well (150 ,151 and 152). There is strong evidence of the role of 5-HT_{2A} receptors and suggestive evidence of the roles of the 5-HT_{1A}, 5-HT_{2C}, and α_1 receptors in various actions of clozapine, risperidone, olanzapine, quetiapine, ziprasidone, iloperidone, sertindole, and related atypical antipsychotic drugs. Atypical antipsychotic drugs that are potent 5-HT_{2A} antagonists relative to their D2 receptor blocking property appear to potentiate 5-HT_{1A}-mediated effects on dopaminergic neurons in the mesocortical, mesolimbic, and mesostriatal regions. The effects in the mesocortical regions appear to be mediated by modulation of glutamate release from pyramidal neurons. These agents have been found to preferentially increase DA efflux in the mPFC compared to limbic and striatal regions. They also increase acetylcholine release in the PFC. Effects on 5-HT_{2C}, 5-HT_3, 5-HT_4, 5-HT_6, and 5-HT_7 receptors may also be relevant to some of their actions, e.g., improvement of cognition, weight gain, etc. Other models of atypicality appear to be effective, including partial DA agonists such as aripiprazole. Selective D2/D3 antagonists such as amisulpride may also have atypical properties. At this time, multireceptor agents appear to be more promising as antipsychotic agents for the majority of psychiatric patients because of important interactions between neural circuits that employ multiple neurotransmitters.

ACKNOWLEDGMENTS

Part of "58 - Mechanism of Action of Atypical Antipsychotic Drugs "

This work is supported in part by grants from Mr. Donald Test, Mrs. Test, and the Warren Foundation. The assistance of Ms. Karen Espenant and Ms. Alice Hammond in the preparation of this manuscript is greatly appreciated.

REFERENCES

1. Jain AK, Kelwala S, Gershon S. Antipsychotic drugs in schizophrenia: current issues. *Int Clin Psychopharmacol* 1988;3:1–30.

2. Meltzer HY, Stahl SM. The dopamine hypothesis of schizophrenia: a review. *Schizophr Bull* 1976;2:19–76.

3. Carlsson A, Lindquist M. Effect of chlorpromazine or haloperidol on formation of 3-methoxytyramine and normetanephrine in mouse brain. *Acta Pharmacol Toxicol* 1963;20:140–144.

4. Creese I, Burt IR, Snyder SH. Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. *Science* 1976;192:481–483.

5. Davis KL, Kahn RS, Ko G, et al. Dopamine in schizophrenia: a review and reconceptualization. *Am J Psychiatry* 1991;148:1474–1486.

6. Chiodo LA, Bunney BS. Typical and atypical neuroleptics: differential effects of chronic administration on the activity of A9 and A10 midbrain dopaminergic neurons. *J Neurosci* 1983;3:1607–1619.

7. Melis M, Mereu G, Lilliu V, et al. Haloperidol does not produce dopamine cell depolarization-block in paralyzed, unanesthetized rats. *Brain Res* 1998;783:127–132.

8. Moore H, Todd CL, Grace AA. Striatal extracellular dopamine levels in rats with haloperidol-induced depolarization block of substantia nigra dopamine neurons. *J Neurosci* 1998;18:5068–5077.

9. Deleted in press.

10. Meltzer HY. Atypical antipsychotic drugs. In: *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1999:1277–1286.

11. Meltzer HY. The concept of atypical antipsychotics. In: *Advances in the neurobiology of schizophrenia*, vol 1. England: Wiley, 1995:265–273.

12. Bunney BS, Sesack SR, Silva NL. Midbrain dopaminergic systems; neurophysiology and electrophysiological pharmacology. In: *Psychopharmacology: the third generation of progress*. New York: Raven Press, 1987:113–126.

13. Meltzer HY, McGurk SR. The effect of clozapine, risperidone and olanzapine on cognitive function in schizophrenia. *Schizophr Bull* 1999;25:233–255.

14. Meltzer HY, Fang V, Young MA. Clozapine-like drugs. *Psychopharmacol Bull* 1980;16:32–34.

15. Lawler CP, Prioleau C, Lewis MM, et al. Interactions of the novel antipsychotic aripiprazole (OPC-14597) with dopamine and serotonin receptor subtypes. *Neuropsychopharmacology* 1999;20:612–627.

16. Di Giovanni G, Di Mascio M, Di Matteo V, et al. Effects of acute and repeated administration of amisulpride, a dopamine D2/D3 receptor antagonist, on the electrical activity of midbrain dopaminergic neurons. *J Pharmacol Exp Ther* 1998;28:51–57.

17. Meltzer HY, Matsubara S, Lee J-C. Classification of typical and atypical antipsychotic drugs on the basis of dopamine D-1, D-2 and serotonin-2 pKi values. *J Pharmacol Exp Ther* 1989;251:238–246.

18. Arndt J, Skarsfeldt T. Do novel antipsychotics have similar pharmacological characteristics? A review of the evidence. *Neuropsychopharmacology* 1998;18:63–101.

19. Meltzer HY, Fatemi S. The role of serotonin in schizophrenia and the mechanism of action of anti-psychotic drugs. In: Kane JM, Möller H-J, Awouters F, eds. *Serotonergic mechanisms in antipsychotic treatment*. New York: Marcel Dekker, 1996:77–107.

20. Moghaddam B, Adams BW. Reversal of phencyclidine effects by a group II metabotropic glutamate receptor agonist in rats. *Science* 1998;281:1349–1352.

21. Cartmell J, Monn JA, Schoepp DD. Attenuation of specific PCP-evoked behaviors by the potent mGlu2/3 receptor agonist, LY379268 and comparison with the atypical antipsychotic, clozapine. *Psychopharmacology* 2000;148:423–429.

22. Schotte A, Janssen PFM, Gommeren W, et al. Risperidone compared with new and reference antipsychotic drugs: in vitro and in vivo receptor binding. *Psychopharmacology* 1996;124:57–73.

23. Pickar D. Su TP, Weinberger DR, et al. Individual variation in D2 dopamine receptor occupancy in clozapine-treated patients. *Am J Psychiatry* 1996;153:1571–1578.

24. Seeman P, Tallerico T. Rapid release of antipsychotic drugs from dopamine D2 receptors: an explanation for low receptor occupancy and early clinical relapse upon withdrawal of clozapine or quetiapine. *Am J Psychiatry* 1999;156:876–884.

25. Svartengren J, Pettersson E, Bjork A. Interaction of the novel antipsychotic drug amperozide and its metabolite FG5620 with central nervous system receptors and monoamine uptake sites: relation to behavioral and clinical effects. *Biol Psychiatry* 1997;42:247–259.

26. Chaki S, Funakoshi T, Yoshikawa R, et al. In vivo receptor occupancy of NRA0045, a putative atypical antipsychotic, in rats. *Neuropharmacology* 1999;38:1185–1194.

27. Corbin AE, Meltzer LT, Ninteman FW, et al. PD 158771, a potential antipsychotic agent with D2/D3 partial agonist and 5-HT(1A) agonist actions. II. Preclinical behavioral effects. *Neuropharmacology* 2000;39:1211–1221.

28. Rivet J-M, Brocco M, Dekeyne A, et al. S33592, a benzopyrolle partial agonist at dopamine D2/D3 receptors and potential antipsychotic agent: II. Functional profile in comparison to aripiprazole, preclamol and raclopride *Neurosci Abst* 2000;26:273.

29. Seeman P, Lee T, Chau-Wong M, et al. Antipsychotic drug doses and neuroleptic/dopamine receptors. *Nature* 1976;261:717–719.

30. Liegeois J-F, Ichikawa J, Bonaccorso S, et al. M100907, a 5-HT2A antagonist, potentiates haloperidol-induced dopamine release in rat medial prefrontal cortex, but inhibits that in the nucleus accumbens. *Neurosci Abst* 2000;26:390.

31. Wadenberg MLPB, Richter JT, Young KA. Enhancement of antipsychotic-like properties of raclopride in rats using the selective serotonin-2A receptor antagonist MDL 100,907. *Biol Psychiatry* 1998;44:508–515.

32. Duinkerke SJ, Botter PA, Jansen AA, et al. Ritanserin, a selective 5-HT2/1C antagonist, and negative symptoms in schizophrenia. A placebo-controlled double-blind trial. *Br J Psychiatry* 1993;163:451–455.

33. Meltzer HY, Nash JF. Effects of antipsychotic drugs on serotonin receptors. *Pharmacol Rev* 1991;43:587–604.

34. Schmidt CJ, Sorensen, SM, Kehne JH, et al. The role of 5-HT_{2A} receptors in antipsychotic activity. *Life Sci* 1995;56:2209–2222.

35. Kapur S, Remington G. Serotonin-dopamine interaction and its relevance to schizophrenia. *Am J Psychiatry* 1996;153:466–476.

36. Meltzer HY. The role of serotonin in antipsychotic drug action. *Neuropsychopharmacology* 1999;21(2 suppl):106S–115S.

37. Svensson TH. Dysfunctional brain dopamine systems induced by psychotomimetic NMDA-receptor antagonists and the effects of antipsychotic drugs. *Brain Res Rev* 2000;31:320–329.

38. Kapur S, Seeman P. Antipsychotic agents differ in how fast they come off the dopamine D2 receptors. Implications for atypical antipsychotic action. *J Psychiatry Neurosci* 2000;25:161–166.

39. Meltzer HY, Lee MA, Ranjan R, et al. Relapse following clozapine withdrawal: effect of cyproheptadine plus neuroleptic. *Psychopharmacology* 1996;124:176–187.

40. Pilowsky LS, Mulligan RS, Acton PD, et al. Limbic selectivity of clozapine. *Lancet* 1997;350:490–491.

41. Pilowsky LS, O'Connell P, Davies N, et al. In vivo effects on striatal dopamine D2 receptor binding by the novel atypical antipsychotic drug sertindole—a 123I IBZM single photon emission tomography. SPET study. *Psychopharmacology* 1997;130:152–158.

42. Stockmeier CA, DiCarlo JJ, Zhang Y, et al. Characterization of typical and atypical antipsychotic drugs based on *in vivo* occupancy of serotonin and dopamine receptors. *J Pharmacol Exp Ther* 1993;266:1374–1384.

43. Youngren KD, Inglis FM, Pivirotto PJ, et al. Clozapine preferentially increases dopamine release in the rhesus monkey prefrontal cortex compared with the caudate nucleus. *Neuropsychopharmacol* 1999;20:403–412.

44. Carboni E, Rolando MTP, Silvagni A, et al. Increase of dialysate dopamine in the bed nucleus of stria terminalis by clozapine and related neuroleptics. *Neuropsychopharmacology* 1999;22:140–147.

45. Ichikawa, J, Ishii H, Fowler WL, et al. Functional 5-HT1A agonism, most likely produced by combined blockade of 5-HT2A and D2 receptors, may be a mechanism by which atypical antipsychotic drugs preferentially increase dopamine release in rat medial prefrontal cortex. *Neurosci Abst* 2000;26:389.

46. Bennett MR. Monoaminergic synapses and schizophrenia: 45 years of neuroleptics. *J Psychopharmacol* 1998;12:289–304.

47. Sokoloff P. Diaz J, Bordet R, et al. [Function and therapeutic potential of the dopamine D3 receptor (French).] *Comptes Rendus Seances Soc Biol Filiales* 1998;192:1111–1125.

48. Roth BL, Tandra S, Burgess LH, et al. D4 dopamine receptor binding affinity does not distinguish between typical and atypical antipsychotic drugs. *Psychopharmacology* 1995;120:365–368.

49. Reavill C, Taylor SG, Wood MD, et al. Pharmacological actions of a novel, high-affinity, and selective human dopamine D(3) receptor antagonist, SB-277011-A. *J Pharmacol Exp Ther* 2000;294:1154–1165.

50. Ashby CR Jr, Minabe Y, Stemp G, et al. Acute and chronic administration of the selective D(3) receptor antagonist SB-277011-A alters activity of midbrain dopamine neurons in rats: an in vivo electrophysiological study. *J Pharmacol Exp Ther* 2000;294:1166–1174.

51. Van Tol HHM, Bunzow JR, Guan HCV, et al. Cloning of the gene for a human dopamine D4 receptor with high affinity for the antipsychotic clozapine. *Nature (Lond)* 1991;350:610–614.

52. Belliotti TR, Wustrow DJ, Brink WA, et al. A series of 6- and 7-piperazinyl- and -piperidinylmethylbenzoxazinones with dopamine D4 antagonist activity: discovery of a potential atypical antipsychotic agent. *J Med Chem* 1999;42:5181–5187.

53. Millan MJ, Newman-Tancredi A, Brocco M, et al. S 18126 ([2-[4-(2,3-dihydrobenzo[1,4]dioxin-6-yl)piperazin-1-yl methyl]indan-2-yl]), a potent, selective and competitive antagonist at dopamine D4 receptors: an in vitro and in vivo comparison with L 745,870 (3-(4-[4-chlorophenyl]piperazin-1-yl)methyl-1H-pyrrolo[2, 3b]pyridine) and raclopride. *J Pharmacol Exp Ther* 1998;287:167–186.

54. Kawashima N, Okuyama S, Omura T, et al. Effects of selective dopamine D4 receptor blockers, NRA0160 and L-745,870, on A9 and A10 dopamine neurons in rats. *Life Sci* 1999;65:2561–2571.

55. Kramer MS, Last B, Getson A, et al., and the D4 Dopamine Antagonist Group. The effects of a selective D4 dopamine receptor antagonism (L-745,870) in acutely psychotic inpatients with schizophrenia. *Arch Gen Psychiatry* 1997;54:567–572.

56. Truffinet P, Tamminga CA, Fabre LF, et al. A placebo controlled study of the D4/5-HT_{2a} antagonist fananserin in the treatment of schizophrenia. *Am J Psychiatry* 1999;156(3):419–425.

57. Abi-Dargham A, Laruelle M, Aghajanian GK. The role of serotonin in the pathophysiology and treatment of schizophrenia. *J Neuropsychiatr Clin Neurosci* 1997;9:1–17.

58. Altar CA, Wasley AM, Neale RF, et al. Typical and atypical antipsychotic occupancy of D2 and S_2 receptors: an autoradiographic analysis in rat brain. *Brain Res Bull* 1986;16:517–525.

59. Rasmussen K, Aghajanian GK. Potency of antipsychotics in reversing the effects of a hallucinogenic drug on locus coeruleus neurons correlates with 5-HT_2 binding affinity. *Neuropsychopharmacology* 1988;1:101–107.

60. Roth BL, Ciaranello RD, Meltzer HY. Binding of typical and atypical antipsychotic agents with transiently expressed 5-HT_{1c} receptors. *J Pharmacol Exp Ther* 1992;260:1361–1365.

61. Roth B, Craigo SC, Choudhary MS, et al. Binding of typical and atypical antipsychotic agents to 5-hydroxytryptamine-6 and 5-hydroxytryptamine-7 receptors. *J Pharmacol Exp Ther* 1994;268:1401–1410.

62. Protais P, Chagraoui A, Arbaoui J. Dopamine receptor antagonist properties of S 14506, 8-OH-DPAT, raclopride and clozapine in rodent. *Eur J Pharmacol* 1994;271:167–173.

63. Newman-Tancredi A, Gavaudan S, Conte C, et al. Agonist and antagonist actions of antipsychotic agents at 5-HT_{1A} receptors: a [^{35}S]GTP_{γ}S. *Eur J Pharmacol* 1998;355:245–256.

64. Millan MJ, Gobert A, Newman-Tancredi A, et al. S16924 ((R)-2-{1-[2-(2,3-dihydro-benzo[1,4] dioxin-5-yloxy)-ethyl]-pyrrolidin-3yl}-1-(4-fluoro-phenyl)-ethanone), a novel potential antipsychotic with marked serotonin 5-HT_{1A} agonist properties: 1. Receptorial and neurochemical profile in comparison with clozapine and haloperidol. *J Pharmacol Exp Ther* 1998;286:1341–1355.

65. Fiorella D, Helsley S, Rabin RA, et al. The interactions of typical and atypical antipsychotics with the (-)2,5-dimethoxy-4-methamphetamine (DOM) discriminative stimulus. *Neuropharmacology* 1995;34:1297–1303.

66. Arora RC, Meltzer HY. Serotonin 2 (5-HT2) receptor binding in the frontal cortex of schizophrenic patients. *J Neural Transm* 1991;85:19–29.

67. Burnet PW, Eastwood SL, Harrison PJ. 5-HT1A and 5-HT2A receptor mRNAs and binding site densities are differentially altered in schizophrenia. *Neuropsychopharmacology* 1996;15:442–455.

68. Joyce JN, Shane A, Lexow N, et al. Serotonin uptake sites and serotonin receptors are altered in the limbic system of schizophrenics. *Neuropsychopharmacology* 1993;8:315–336.

69. Reynolds GP, Garrett NJ, Rupniak N, et al. Chronic clozapine treatment of rats down-regulates cortical 5-HT2 receptors. *Eur J Pharmacol* 1983;89:325–326.

70. Roth BL, Berry SA, Kroeze WK, et al. Serotonin 5-HT2A receptors: molecular biology and mechanisms of regulation. *Crit Rev Neurobiol* 1998;12:319–338.

71. Trichard C, Paillere-Martinot ML, Attar-Levy D, et al. No serotonin 5-HT2A receptor density abnormality in the cortex of schizophrenic patients studied with PET. *Schizophr Res* 1998;31:13–17.

72. Nyberg S, Farde L, Halldin C. PET study of 5-HT2 and D2 dopamine receptor occupancy induced by olanzapine in healthy subjects. *Neuropsychopharmacology* 1997;16:1–7.

73. Nyberg S, Nilsson U, Okubo Y, et al. Implications of brain imaging for the management of schizophrenia. *Int Clin Psychopharmacol* 1998;13(suppl 3):S15–20.

74. Kasper S, Tauscher J, Kufferle B, et al. Dopamine- and serotonin-receptors in schizophrenia: results of imaging-studies and implications for pharmacotherapy in schizophrenia. *Eur Arch Psychiatr Clin Neurosci* 1999;249(suppl 4):83–89.

75. Kapur S, Zipursky R, Jones C, et al. A positron emission tomography study of quetiapine in schizophrenia: a preliminary finding of an antipsychotic effect with only transiently high dopamine D2 receptor occupancy. *Arch Gen Psychiatry* 2000;57:553–559.

76. Nyberg S, Eriksson B, Oxenstierna G, et al. Suggested minimal effective dose of risperidone based on PET-measured D2 and 5-HT2A receptor occupancy in schizophrenic patients. *Am J Psychiatry* 1999;156:869–875.

77. Kapur S, Zipursky RB, Remington G, et al. 5-HT2 and D2 receptor occupancy of olanzapine in schizophrenia: a PET investigation. *Am J Psychiatry* 1998;155:921–928.

78. Marder SR, Meibach RC. Risperidone in the treatment of schizophrenia. *Am J Psychiatry* 1994;151:825–835.

79. Shipley J. M100907 phase IIB trial. Presented at the Hoechst Marion Roussel Conference on M100907, West Palm Beach Florida, April 1998.

80. Rinaldi-Carmona M, Congy C, Santucci V, et al. Biochemical and pharmacological properties of SR 46349B, a new potent and selective 5-hydroxytryptamine$_2$ receptor antagonists. *J Pharmacol Exp Ther* 1992;262:759–768.

81. Masellis FM, Macciardi FM, Meltzer HY, et al. Serotonin subtype 2 receptor genese and clinical response to clozapine in schizophrenic patients. *Neuropsychopharmacology* 1998;19:123–132.

82. Arranz MJ, Munro J, Birkett J, et al. Pharmacogenetic prediction of clozapine response. *Lancet* 2000;355:1615–1616.

83. Martin P, Water N, Carlsson A, et al. The apparent antipsychotic action of the 5-HT_{2a} receptor antagonist M100907 in a mouse model of schizophrenia is counteracted by ritanserin. *J Neural Transm* 1997;104:561–564.

84. Gleason SC, Shannon HE. Blockade of phencyclidine-induced hyperlocomotion by olanzapine, clozapine and serotonin receptor subtype selective antagonists in mice. *Psychopharmacology* 1997;129:79–84.

85. Varty GB, Higgins GA. Reversal of a dizocilpine-induced disruption of prepulse inhibition of an acoustic startle response by the 5-HT_2 receptor antagonist ketanserin. *Eur J Pharmacol* 1995;287:201–205.

86. Prinssen EPM, Ellenbroek BA, Cools AR. Combined antagonism of adrenoceptors and dopamine and 5-HT receptors underlies the atypical profile of clozapine. *Eur J Pharmacol* 1994;262:167–170.

87. Nabeshima T, Ishikawa K, Yamaguchi K, et al. Phencyclidine-induced head-twitch responses as 5-HT2 receptor-mediated behavior in rats. *Neurosci Lett* 1987;76:335–338.

88. Krebs-Thomson K, Lehmann-Masten V, Naiem S, et al. Modulation of phencyclidine-induced changes in locomotor activity and patterns in rats by serotonin. *Eur J Pharmacol* 1998;343:135–143.

89. Carlsson ML, Martin P, Nilsson M, et al. The 5-HT2A receptor antagonist M100907 is more effective in counteracting NMDA antagonist- than dopamine agonist-induced hyperactivity in mice. *J Neural Transm (Budapest)* 1999;106:123–129.

90. Marcus MM, Nomikos GG, Svensson TH. Differential actions of typical and atypical antipsychotic drugs on dopamine release in the core and shell of the nucleus accumbens. *Eur Neuropsychopharmacol* 1996;6:29–38.

91. De Deurwaerdere P, Spampinato U. Role of serotonin(2A) and serotonin (2B/2C) receptor subtypes in the control of accumbal and striatal dopamine release elicited in vivo by dorsal raphe nucleus electrical stimulation. *J Neurochem* 1999;73(3):1033–1042.

92. Ichikawa J, Ishii H, Bonaccorso S, et al. 5-HT2a and D2 receptor blockade increases cortical DA release via of 5-HT1A receptor activation: a possible mechanism of atypical antipsychotic induced cortical dopamine release. *J Neurochem* 2001;76:1521–1531.

93. Ichikawa J, Meltzer HY. DOI, a 5-$HT_{2a/2c}$ receptor agonist, potentiates amphetamine-induced dopamine release in rat striatum. *Brain Res* 1995;698:204–208.

94. Lucas G, De Deurwaerdere P, Porras G, et al. Endogenous serotonin enhances the release of dopamine in the striatum only when nigro-striatal dopaminergic transmission is activated. *Neuropharmacology* 2000;39:1984–1995.

95. Martin P, Waters N, Schmidt CJ, et al. Rodent data and general hypothesis: antipsychotic action exerted through 5-HT2A receptor antagonism is dependent on increased serotonergic tone. *J Neural Transm (Budapest)* 1998;105:365–396.

96. Jakab RL, Goldman-Rakic P. 5-Hydroxytryptamine 2a serotonin receptors in the primate cerebral cortex: possible site of action of hallucinogenic and antipsychotic drugs in pyramidal cell apical dendrites. *Proc Natl Acad Sci USA* 1998;95:735–740.

97. Aghajanian GK, Marek GJ. Serotonin model of schizophrenia: emerging role of glutamate mechanisms. *Brain Res Rev* 2000;31:302–312.

98. Carlsson A. The current status of the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology* 1988;1(3):179–186.

99. Moghaddam B, Bunney BS. Acute effects of typical and atypical antipsychotic drugs on the release of dopamine from prefrontal cortex, nucleus accumbens, and striatum of the rat: an in vivo microdialysis study. *J Neurochem* 1990;54:1755–1760.

100. Kuroki T, Meltzer HY, Ichikawa J. Effects of antipsychotic drugs on extracellular dopamine levels in rat medial prefrontal cortex and nucleus accumbens. *J Pharmacol Exp Ther* 1999;288:774–781.

101. Ichikawa J, O'Laughlin IA, Dai J, et al. Atypical, but not typical antipsychotic drugs selectively increase acetylcholine release in rat medial prefrontal cortex. *Neuropsychopharmacology, in press.*

102. Gessa GL, Devoto P, Diana M, et al. Disassociation of haloperidol, clozapine, and olanzapine effects on electrical activity of mesocortical dopamine neurons and dopamine release in the prefrontal cortex. *Neuropsychopharmacology* 1999;22:642–649.

103. Parada MA, Hernandez L, Puig de Parada M, et al. Selection action of acute systemic clozapine on acetylcholine release in the rat prefrontal cortex by reference to the nucleus accumbens and striatum. *J Pharmacol Exp Ther* 1997;281:582–588.

104. Altman HJ, Stone WS, Ögren S. Evidence for a possible functional interaction between serotonergic and cholinergic mechanism in memory retrieval. *Behav Neural Biol* 1987;48:49–62.

105. Consolo S, Arnaboldi S, Giorgi S, et al. 5-HT4 receptor stimulation facilitates acetylcholine release in rat frontal cortex. *Neuroreport* 1994;5:1230–1232.

106. Consolo S, Bertorelli R, Russi G, et al. Serotonergic facilitation of acetylcholine release in vivo from rat dorsal hippocampus via serotonin 5-HT3 receptors. *J Neurochem* 1994;62:2254–2261.

107. Consolo S, Ramponi S, Ladinsky H, et al. A critical role for D1 receptors in the 5-HT1A-mediated facilitation of in vivo acetylcholine release in rat frontal cortex. *Brain Res* 1996;707:320–323.

108. Zhelyazkova-Savova M, Giovannini MG, Pepeu G. Increase of cortical acetylcholine release after systemic administration of chlorophenylpiperazine in the rat: an in vivo microdialysis study. *Neurosci Lett* 1997;236:151–154.

109. Zhelyazkova-Savova M, Giovannini MG, Pepeu G. Systemic chlorophenylpiperazine increases acetylcholine release from rat hippocampus-implication of 5-HT2C receptors. *Pharmacol Res* 1999;40:165–170.

110. Buhot M-C. Serotonin receptors in cognitive behaviors. *Curr Opin Neurobiol* 1997;7:243–254.

111. Meneses A, Hong E. A pharmacological analysis of serotonergic receptors: effect of their activation of blockade in learning. *Prog Neuropsychopharmacol Biol Psychiatry* 1997;21:273–296.

112. Ruotsalainen S, Sirvio J, Jäkälä P, et al. Differential effect of three 5-HT receptor antagonists on the performance of rats in attentional and working memory tasks. *Eur Neuropsychopharmacol* 1997;7:99–108.

113. Marchetti E, Dumuis A, Bockaert J, et al. Differential modulation of the 5-HT4 receptor agonist on rat learning and memory. *Neuropharmacology* 2000;39:2017–2027.

114. Herremans AH, Hijzen TH, Olivier B, et al. Serotonergic drug effects on a delayed conditional task in the rat: involvement of the 5-HT_{1A} receptor in working memory. *J Psychopharmacol* 1995;9:242–250.

115. Sumiyoshi T, Matsui M, Yamashita I, et al. Effect of adjunctive treatment with serotonin$_{1A}$ agonist tandospirone on memory functions in schizophrenia. *J Clin Psychopharmacol* 2000;20:386–388.

116. Kalkman HO, Neumann V, Hoyer D, et al. The role of a_2-adrenoceptor antagonism in the anti-cataleptic properties of the atypical neuroleptic agent, clozapine, in the rat. *Br J Pharmacol* 1998;124:1550–1556.

117. Barbato L, Monge A, Stocchi F, et al. Melperone in the treatment of iatrogenic psychosis in Parkinson's disease. *Clin Neuropharmacol* 1996;11:201–207.

118. Poyurovsky M, Shardorodsky M, Fuchs C, et al. Treatment of neuroleptic-induced akathisia with the 5-HT2 antagonist mianserin. Double-blind, placebo-controlled study. *Br J Psychiatry* 1999;174:238–242.

119. Ishikane T, Kusumi I, Matsubara R, et al. Effects of serotonergic agents on the up-regulation of dopamine D2 receptors induced by haloperidol in rat striatum. *Eur J Pharmacol* 1997;321:163–169.

120. Spampinato U, De Deuwwaerdere P, Caccia S, et al. Opposite role of central 5-HT_{2a} and 5-HT_{2c} receptors subtypes in the control of haloperidol-induced release of dopamine in the rat striatum. *Neurosci Abst* 1998;24:48.4.

121. Pazos A, Probst A, Palacios JM. Serotonin receptors in the human brain—IV. Autoradiographic mapping of serotonin-2 receptors. *Neuroscience* 1987;21(1):123–139.

122. Millan MJ, Dekeyne A, Gobert A. Serotonin (5-HT)2C receptors tonically inhibit dopamine (DA) and noradrenaline (NA), but not 5-HT, release in the frontal cortex in vivo. *Neuropharmacology* 1998;37:953–955.

123. Meltzer HY, Roth B, Thompson P. Serotonin and dopamine receptor affinities predict atypical antipsychotic drug (AAD) activity. *Neurosci Abst* 1996;22:480.

124. Browning JL, Jones DH, Young KA, et al. 5-HT2C antagonists potentiate apomorphine-induced locomotor activity in rats by decreasing stereotypy. *Neurosci Abst* 2000;26:274.

125. Darmani NA, Martin BR, Pandy U, et al. Do functional relationships exist between 5-HT_{1A} and 5-HT_2 receptors? *PharmacoI Biochem Behav* 1990;26:901–906.

126. Meltzer HY, Maes M. Effects of pindolol on the L-5-HTP-induced increase in plasma prolactin and cortisol concentrations in man. *Psychopharmacology* 1994;114:635–643.

127. Willins DI, Meltzer HY. Direct injections of 5-HT_{2a} receptor agonists into the medial prefrontal cortex produces a head-twitch response in rats. *J PharmacoI Exp Ther* 1997;282:699–706.

128. Ahlenius S. Antipsychotic-like properties of the 5-HT_{1a} agonist 8-OH-DPAT in the rat. *PharmacoI Toxicol* 1988;64:3–5.

129. Wadenberg ML, Ahlenius S. Antipsychotic-like profile of combined treatment with raclopride and 8-OH-DPAT in the rat: enhancement of antipsychotic-like effects without catalepsy. *J Neural Transm* 1991;83:43–53.

130. Prinssen EPM, Kleven MS, Koek W. Effects of dopamine antagonists in a two-way active avoidance procedure in rats: interactions with 8-OH-DPAT, ritanserin, and prazosin. *Psychopharmacology* 1996;128:191–197.

131. Wadenberg ML. Antagonism by 8-OH-DPAT, but not ritanserin, of catalepsy induced by SCH 23390 in the rat. *J Neural Transm* 1992;89:49–59.

132. Wadenberg ML, Hillegaart V. Stimulation of median, but nor dorsal, raphe 5-HT_{1a} autoreceptors by the local application of 8-OH-DPAT reverses raclopride-induced catalepsy in the rat. *Neuropharmacology* 1995;34:495–499.

133. Ichikawa J, Kuroki T, Kitchen MT, et al. R(+)-8-OH-DPAT, a 5-HT_{1a} receptor agonist, inhibits amphetamine-induced serotonin and dopamine release in rat striatum and nucleus accumbens. *Eur J PharmacoI* 1995;287:179–184.

134. Rollema H, Lu Y, Schmidt AW. Clozapine increases dopamine release in prefrontal cortex by 5-HT_{1a} receptor activation. *Eur J PharmacoI* 1997;338:R3–R5.

135. Wedzony K, Mackowiak M, Zajaczkowski W, et al. WAY 100135, an antagonist of 5-HT1A serotonin receptors, attenuates psychotomimetic effects of MK-801. *Neuropsychopharmacology* 2000;23:547–559.

136. Ferré S, Artigas F. Clozapine decreased serotonin extracellular levels in the nucleus accumbens by a dopamine receptor-independent mechanism. *Neurosci Lett* 1995;187:61–64.

137. Ichikawa J, Kuroki T, Dai J, et al. Effect of antipsychotic drugs on extracellular serotonin levels in rat medial prefrontal cortex and nucleus accumbens. *Eur J PharmacoI* 1998;351:163–171.

138. Hertel P, Nomikos G, Schilstrom B, et al. Risperidone dose-dependently increases extracellular concentrations of serotonin in the rat frontal cortex: role of α_2-adrenoceptor antagonism. *Neuropsychopharmacology* 1997;17:44–55.

139. McAllister KH, Rey B. Clozapine reversal of the deficits in coordinated movement induced by D2 receptor blockade does not depend upon antagonism of α2 adrenoceptors. *Naunyn Schmiedebergs Arch PharmacoI* 1999;360(6):603–608.

140. Davidson M, Kahn RS, Stern RG, et al. Treatment with clozapine and its effect on plasma homovanillic acid and norepinephrine concentrations in schizophrenia. *Psychiatry Res* 1993;46:151–163.

141. Litman RE, Su TP, Potter WZ, et al. Idazoxan and response to typical neuroleptics in treatment-resistant schizophrenia. Comparison with the atypical neuroleptic, clozapine. *Br J Psychiatry* 1996;168:571–579.

142. Coull JT, Sahakian BJ, Hodges JR. The α(2) antagonist idazoxan remediates certain attentional and executive dysfunction in patients with dementia of frontal type. *Psychopharmacology* 1996;123:239–249.

143. Haapalinna A, Sirvio J, MacDonald E, et al. The effects of a specific α(2)-adrenoceptor antagonist, atipamezole, on cognitive performance and brain neurochemistry in aged Fisher 344 rats. *Eur J PharmacoI* 2000;387:141–150.

144. Bolonna AA, Arranz MJ, Munro J, et al. No influence of adrenergic receptor polymorphisms on schizophrenia and antipsychotic response. *Neurosci Lett* 2000;280:65–68.

145. Hertel P, Fagerquist MV, Svensson TH. Enhanced cortical dopamine output and antipsychotic-like effects of raclopride by α_2 adrenoceptor blockade. *Science* 1999;286:105–107.

146. Hertel P, Nomikos GG, Svensson TH. Idazoxan preferentially increases dopamine output in the rat medial prefrontal cortex at the nerve terminal level. *Eur J PharmacoI* 1999;371:153–158.

147. Westerink BHC, de Boer B, de Vries JB, et al. Antipsychotic drugs induce similar effects on the release of dopamine and noradrenaline in the medial prefrontal cortex of the rat brain. *Eur J PharmacoI* 1998;361:27–33.

148. Arnsten AF, Steere JC, Jentsch DJ, et al. Noradrenergic influences on prefrontal cortical cognitive function: opposing actions at postjunctional α1 versus α2-adrenergic receptors. *Adv PharmacoI* 1998;42:764–767.

149. Nutt DJ, Lalies MD, Lione LA, et al. Noradrenergic mechanisms in the prefrontal cortex. *J Psychopharmacol* 1997;11:163–168.

150. Kane J, Honigfeld G, Singer J, et al., and the Clozaril Collaborative Study Group. Clozapine for the treatment-resistant schizophrenic: a double-blind comparison with chlorpromazine. *Arch Gen Psychiatry* 1988;45:789–796.

151. Hagger C, Buckley P, Kenny JT, et al. Improvement in cognitive functions and psychiatric symptoms in treatment-refractory schizophrenic patients receiving clozapine. *Biol Psychiatry* 1993;34:702–712.

152. Meltzer HY, Okayli G. The reduction of suicidality during clozapine treatment in neuroleptic-resistant schizophrenia: impact on risk-benefit assessment. *Am J Psychiatry* 1995;152:183–190.

59

Neurochemical and Neuropharmacological Imaging in Schizophrenia

Daniel R. Weinberger

Marc Laruelle

Daniel R. Weinberger: Clinical Brain Disorders Branch, Intramural Research Program, National Institute of Mental Health, Bethesda, Maryland.

Marc Laruelle: Division of Functional Brain Mapping, New York State Psychiatric Institute; Departments of Psychiatry and Radiology, Columbia College of Physicians and Surgeons, New York, New York.

Over the last 15 years, the ability to measure specific molecules and proteins in the living human brain underwent enormous developments, opening direct windows into neurotransmitter functions and cellular processes associated with health and disease. These techniques are based either on the injection of radioactive moieties whose distribution is recorded with positron emission tomography (PET) or single photon emission computed tomography (SPECT), or on the direct detection of molecules based on their intrinsic magnetic properties with magnetic resonance spectroscopy (MRS).

Although lacking the level of resolution of postmortem studies and limited by the relatively small number of targets that can be currently studied *in vivo*, these imaging techniques provide unique opportunities for elucidation of pathophysiology associated with neuropsychiatric conditions. *In vivo* imaging techniques enable studying patients with well-documented psychopathology, relating biochemical observations to psychiatric symptoms and cognitive processes, and clarifying abnormalities associated with the disease as opposed to its treatment. Neurochemical imaging allows longitudinal studies to investigate mechanisms of actions and consequences of treatments, as well as to characterize neurochemical abnormalities in relation to treatment response and illness outcome. Neurochemical imaging further provides insight as to the pathophysiologic bases of alterations measured with flow and metabolism imaging studies, and can provide a direct link with animal models of the illness. Moreover, these techniques enable the study of patient's relatives, in order to clarify the endophenotypes associated with illness vulnerability. These potentialities are discussed in this chapter.

This chapter also critically summarizes results obtained using neurochemical imaging techniques in schizophrenia research, and the various insights on the pathophysiology and treatment of schizophrenia gained by these results. We first consider PET and SPECT investigations, and then MRS studies.

PET AND SPECT NEUROCHEMICAL IMAGING

Part of "59 - Neurochemical and Neuropharmacological Imaging in Schizophrenia "

The principles of PET and SPECT neurochemical imaging are reviewed elsewhere in this volume. Numerous PET and SPECT radiotracers are currently available to study key proteins in the living brain, such as receptors, transporters, and enzymes. Regarding schizophrenia, the majority of clinical investigations studied various aspects of dopaminergic transmission. Dopamine (DA) D2 receptors were the first neuroreceptors visualized in the living human brain (1). Since then, several DA-related radiotracers have been developed, allowing the study of many aspects of dopaminergic transmission (DA synthesis, DA release, D1 and D2 receptors, DA transporters). Given the availability of these tools and the important role that DA transmission is believed to play in schizophrenia, it is not surprising that most of the research effort focused on this system. Despite marked limitations, these studies provide a relatively consistent picture suggesting that schizophrenia, at least during periods of clinical exacerbation, is associated with dysregulation of DA transmission.

Imaging DA Transmission Parameters In Schizophrenia

The classic DA hypothesis of schizophrenia, formulated over 30 years ago, proposed that a hyperactivity of the dopaminergic transmission is associated with this illness (2 ,3).

This hypothesis was essentially based on the observation that all antipsychotic drugs provided at least some degree of D2 receptors blockade, a proposition that is still true today (4 ,5). As D2 receptor blockade is most effective against positive symptoms, the DA hyperactivity model appeared to be most relevant to the pathophysiology of positive symptoms. The fact that sustained exposure to DA agonists such as amphetamine can induce a psychotic state characterized by some salient features of positive symptoms of schizophrenia (emergence of paranoid delusions and hallucinations in the context of a clear sensorium) also contributed to the idea that positive symptoms might be due to sustained excess dopaminergic activity (6 ,7).

These pharmacologic effects indeed suggest, but do not establish, a dysregulation of DA systems in schizophrenia. For example, these observations are also compatible with the hypotheses that DA activity per se would be normal and increased only relatively to other systems that may be deficient, such as the glutamatergic or serotoninergic system (8 ,9). Under these conditions, D2 receptor blockade would reestablish a compromised balance between dopaminergic and glutamatergic or serotoninergic tone. Thus, these pharmacologic observations do not necessarily imply a specific disturbance of DA activity per se in the brain of patients with schizophrenia. Indeed, documentation of abnormalities of DA function in postmortem studies in schizophrenia has remained elusive (10 ,11 and 12). Because positive symptoms are mostly prominent in young patients and their intensity decreases with age, the ability to detect their biochemical correlates in postmortem studies (generally performed in older subjects) may be limited.

On the other hand, negative and cognitive symptoms are generally resistant to treatment by antipsychotic drugs. Functional brain imaging studies suggested that these symptoms are associated with prefrontal cortex (PFC) dysfunction (13). Studies in nonhuman primates demonstrated that deficits in DA transmission in PFC induce cognitive impairments reminiscent of those observed in patients with schizophrenia (14), suggesting that a deficit in DA transmission in the PFC might be implicated in cognitive impairments presented by these patients (15 ,16). In addition, a recent postmortem study described abnormalities of DA terminals in the PFC associated with schizophrenia (17). Thus, a current view on DA and schizophrenia is that subcortical mesolimbic DA projections might be hyperactive (resulting in positive symptoms) and that the mesocortical DA projections to the PFC are hypoactive (resulting in negative symptoms and cognitive impairment). Furthermore, these two abnormalities might be related, as the cortical DA system generally exerts an inhibitory action on subcortical DA systems (18 ,19).

The advent in the early 1980s of techniques based on PET and SPECT to measure indices of DA activity in the living human brain held considerable promise for investigating these questions.

Striatal D2 Receptor Density

Striatal D2 receptor density in schizophrenia has been extensively studied with PET and SPECT imaging. Studies comparing parameters of D2 receptor binding in patients with schizophrenia and healthy controls (*n* = 16 studies) are listed in Table 59.1 , and included a total of 228 patients (102 were neuroleptic-naive, and 126 were neuroleptic-free for variable periods of time). These patients were compared to 213 controls, matched for age and sex. Ten studies used PET and six studies used SPECT. Radiotracers included butyrophenones ([^{11}C]N-methyl-spiperone, [^{11}C]NMSP, *n* = 3, and [^{76}Br]bromospiperone, *n* = 3), benzamides ([^{11}C]raclopride, *n* = 3, and [^{123}I]IBZM, *n* = 5) or the ergot derivative [^{76}Br]lisuride, *n* = 2). A variety of methods and outcome measures were used to estimate D2 receptors density. Six studies used empirical ratio methods, i.e., the ratio of striatal to cerebellar activities at a given time after single bolus injection of the radiotracer. Five studies used model-based methods to measure the binding potential (BP), which is equal to the product of receptor density (B_{max}) and affinity ($1/K_d$). Five studies reported both B_{max} and K_d.

Class Radiotracer	Radiotracer	Study	Controls (*n*)	Patients (*n*) (DN/DF)	Method	Outcome	Controls (*n* Mean ± SD)[a]	Patients (*n* Mean ± SD)[a]	*p*	Effect Size[b]	Ratio SD
Butyrophenones	[^{11}C]NMSP	Wong et al. (20)	11	15 (10/5)	Kinetic	B_{max}	100 ± 50	253 ± 105	<.05	3.06	2.10
	[^{76}Br]SPI	Crawley et al. (21)	8	16 (12/4)	Ratio	S/C	100 ± 14	111 ± 12	<.05	0.79	0.86
	[^{76}Br]SPI	Blin et al. (201)	8	8 (0/8)	Ratio	S/C	100 ±1 4	104 ± 14	NS	0.28	1.00
	[^{76}Br]SPI	Martinot et al. (202)	12	12 (0/12)	Ratio	S/C	100 ± 11	101 ± 15	NS	0.14	1.41
	[^{11}C]NMSP	Tune et al. (203)	17	10 (8/2)	Kinetic	B_{max}	100 ± 80	173 ± 143	.08	0.91	1.79
	[^{11}C]NMSP	Nordstrom et al. (204)	7	7 (7/0)	Kinetic	B_{max}	100 ± 25	133 ± 63	NS	1.33	2.50
Benzamides	[^{11}C]Raclopride	Farde et al. (205)	20	18 (18/0)	Equilib.	B_{max}	100 ± 29	107 ± 18	NS	0.23	0.63
	[^{11}C]Raclopride	Hietala et al. (206)	10	13 (0/13)	Equilib.	B_{max}	100 ± 22	112 ± 43	NS	0.55	1.99
	[^{123}I]IBZM	Pilowsky et al. (207)	20	20 (17/3)	Ratio	S/FC	100 ± 8	99 ± 7	NS	−0.07	0.82
	[^{123}I]IBZM	Laruelle et al. (39)	15	15 (1/14)	Equilib.	BP	100 ± 26	115 ± 33	NS	0.56	1.25
	[^{123}I]IBZM	Knable et al. (208)	16	21 (1/20)	Equilib.	BP	100 ± 29	97 ± 38	NS	−0.12	1.31
	[^{11}C]Raclopride	Breier et al. (38)	12	11 (6/5)	Equilib.	BP	100 ± 18	100 ± 30	NS	0.02	1.69
	[^{123}I]IBZM	Abi-Dargham et al. (40)	15	15 (2/13)	Equilib.	BP	100 ± 20	102 ± 49	NS	0.09	2.50
	[^{123}I]IBZM	Abi-Dargham et al. (45)	18	18 (8/10)	Equilib.	BP	100 ± 13	104 ± 14	NS	0.33	1.11
Ergot alk.	[^{76}Br]Lisuride	Martinot et al. (209)	14	19 (10/9)	Ratio	S/C	100 ± 10	104 ± 12	NS	0.45	1.21
	[^{76}Br]Lisuride	Martinot et al. (210)	10	10 (2/8)	Ratio	S/C	100 ± 10	100 ± 13	NS	0.00	1.29

[a] Mean normalized to mean of control subjects.
[b] Effect size calculated as (mean patients − mean controls) / SD controls.
BP, binding potential; DF, drug free; DN, drug naive; IBZM, iodobenzamide; NMSP, *N*-methyl-spiperone; NS, not significant.

TABLE 59.1. IMAGING STUDIES OF D_2 RECEPTOR PARAMETERS IN DRUG NAIVE AND DRUG FREE PATIENTS WITH SCHIZOPHRENIA

Only two out of 13 studies detected a significant elevation of D2 receptor density parameters at a level *p* <.05. However, metaanalysis of the 16 studies reveals a small (on the order of 13%) but significant elevation of D2 receptors in patients with schizophrenia. If D2 receptor density did not differ between patients and controls (null hypothesis), one would expect approximately 50% of the studies to report lower D2 receptor levels in schizophrenics compared to controls. Instead, 12 out of 16 studies reported an increase (although not significant in 10 out of 12 cases), two reported no change, and only two studies reported a decrease in patients compared to controls. This distribution is unlikely (*p* <.05, sign test) under the null hypothesis. Moreover, under the null hypothesis, the effect sizes [mean value in schizophrenic group - mean value in control group/standard deviation (SD) in control group] should be distributed around 0. The average effect size of the 16 studies was 0.57 ± 0.78 (SD), and the probability to yield such effect size under the null hypothesis is again lower than .05. The aggregate magnitude of this elevation is thus 57% of the SD of controls. Given an average control SD of 23%, the effect is about 13%. To detect an effect size of 0.50 at the .05 significance level with a power of 80%, a sample of 64 patients and 64 controls would be needed. Clearly, none of the studies included enough patients to detect this small effect with appropriate power. Another observation is that the variability in the patient sample was larger than in the control sample in 13 out of 15 studies, which was also significant (*p* <.05, sign test). The average variance ratio (SD schizophrenics/SD controls) was 1.47 ± 0.58. The larger variance in patients compared to controls further increases the sample size needed to detect this small group difference with reasonable power.

No clinical correlates of increased D2 receptor binding parameters has been reliably identified. Thus, the simplest conclusions from these studies are that patients with schizophrenia show a modest elevation in D2 receptor density parameters of undetermined clinical significance, that all studies were underpowered, and that positive results occasionally reported (20 ,21) are due to a sampling effect. These conclusions are reached under the assumptions that all studies measured parameters from the "same" D2 receptors population.

However, there is another way to look at these data. Studies performed with butyrophenones (*n* = 6) have an effect size of 1.08 ± 1.06 (*n* = 6), whereas studies performed with other ligands (benzamides and lisuride, *n* = 10) have an effect size of 0.20 ± 0.26, a difference that is significant (*p* = .022). This observation suggests that schizophrenia might be associated with an increase in butyrophenone binding and no change in benzamide or lisuride binding.

Unfortunately, no studies have been reported in which the same subjects were scanned with both ligands. Such a study is warranted to directly test this view. Nevertheless, several hypotheses have been advanced to account for the existence of a differential increase in [^{11}C]NMSP binding *in vivo* in patient with schizophrenia in the face of normal benzamide binding. Because [^{11}C]raclopride and [^{123}I]IBZM binds to D2 and D3 receptors whereas [^{11}C]NMSP binds to D2, D3, and D4 receptors, this difference could reflect a selective elevation of D4 receptors in schizophrenia (22). This hypothesis has not been substantiated. The density of D4 receptors is negligible in the striatum, and, when measured with a specific ligand, not different in postmortem striatal samples from patients with schizophrenia and controls (23). Another hypothesis derives from the observation that D2 receptors, like several G-protein–coupled receptors, exist in monomers, dimers, and other oligomeric forms (24 ,25 ,26 and 27). Photoaffinity labeling experiments suggested that butyrophenones detect only monomers, whereas benzamides detect both monomers and dimers. Thus, increased butyrophenone binding and normal benzamide binding might reflect a higher monomer/dimer ratio in schizophrenia. This interesting hypothesis warrants further exploration. A third proposition evolved around the idea that the binding of these ligands would display different vulnerability to competition by endogenous DA (28 ,29). This proposition was based on two assumptions: (a) the concentration of DA in the proximity of D2 receptors might be higher in patients compared to controls, and (b) [^{11}C]NMSP might be less affected than [^{11}C]raclopride or [^{123}I]IBZM binding by endogenous DA competition. It follows that D2 receptor density measured *in vivo* with [^{11}C]raclopride and [^{123}I]IBZM would be "underestimated" to a greater extent in patients with schizophrenia than in control subjects. This hypothesis played an important role in bringing the endogenous competition concept to the attention of the imaging field.

Striatal DA Presynaptic Function

Imaging studies of presynaptic DA function in schizophrenia included measurements of dihydroxyphenylalanine (DOPA) decarboxylase activity, DA release at baseline and following pharmacologic challenges with amphetamine, and DA transporters (DAT). These studies are summarized in Table 59.2 .

Parameter	Study	Controls (*n*)	Patients (*n*) (DN/DF)	Radiotracer (/Challenge)	Method	Outcome	Controls (*n* Mean ± SD)[a]	Patients (*n* Mean ± SD)[a]	*p*	Effect Size[b]	Ratio SD
DOPA decarboxylase activity	Reith et al. (30)	13	5 (4/1)	[^{18}F]DOPA	Kinetic	k3	100 ± 23	120 ± 15	<.05	0.91	0.68
	Hietala et al. (31)	7	7 (7/0)	[^{18}F]DOPA	Graphical	K_i	100 ± 11	117 ± 20	<.05	1.54	1.82
	Dao-Castellana et al. (32)	7	6 (2/4)	[^{18}F]DOPA	Graphical	K_i	100 ± 11	103 ± 40	NS	0.30	3.80
	Lindstrom et al. (34)	10	12 (10/2)	[^{11}C]DOPA	Graphical	K_i	100 ± 17	113 ± 12	<.05	0.77	0.70
	Hietala et al. (33)	13	10 (10/0)	[^{18}F]DOPA	Graphical	K_i	100 ± 14	115 ± 28	<.05	1.09	1.25
Amphetamine -induced DA release	Laruelle et al. (39)	15	15 (2/13)	[^{123}I]IBZM/ amphetamine	Equilibrium	Delta BP	100 ± 113	271 ± 221	<.05	1.51	1.95
	Breier et al. (38)	18	18 (8/10)	[^{11}C]raclopride/ amphetamine	Equilibrium	Delta BP	100 ± 43	175 ± 82	<.05	1.73	1.90
	Abi-Dargham et al. (40)	16	21 (1/20)	[^{123}I]IBZM/ amphetamine	Equilibrium	Delta BP	100 ± 88	194 ± 145	<.05	1.07	1.64
Baseline DA concentration	Abi-Dargham et al. (45)	18	18 (8/10)	[^{123}I]IBZM/ α-MPT	Equilibrium	Delta BP	100 ± 78	211 ± 122	<.05	1.43	1.57
DAT density	Laakso et al. (211)	9	9 (9/0)	[^{18}F]CFT	Ratio	S/C	100 ± 12	101 ± 13	<.05	0.11	1.06
	Laruelle et al. (46)	22	22 (2/20)	[^{123}I]CIT	Equilibrium	BP	100 ± 17	93 ± 20	<.05	−0.43	1.21

[a]Mean normalized to mean of control subjects.
[b]Effect size calculated as (mean patients − mean controls) / SD controls.
DAT, dopamine transporter; DF, drug free; DN, drug naive; DOPA, dihydroxyphenylalanine; MPT, methyl-para-tyrosine.

TABLE 59.2. IMAGING STUDIES OF PRESYNAPTIC DA PARAMETERS IN DRUG NAIVE AND DRUG FREE PATIENTS WITH SCHIZOPHRENIA

DOPA Decarboxylase Activity

Five studies reported rates of DOPA decarboxylase in patients with schizophrenia, using [^{18}F]DOPA (30 ,31 ,32 and 33) or [^{11}C]DOPA (34) (Table 59.2). Four out of five studies reported increased accumulation of DOPA in the striatum of patients with schizophrenia, and the combined analysis yielded an effect size of 0.92 ± 0.45, which is significantly different from zero (*p* = .01). The variability of the DOPA accumulation was larger in the schizophrenic group compared to the control group. Several of these studies reported the observation of high DOPA accumulation in psychotic paranoid patients, and low accumulation in patients with negative or depressive symptoms and catatonia. Although the relationship between DOPA decarboxylase and DA synthesis rate is unclear (DOPA decarboxylase is not the rate-limiting step of DA synthesis), these observations are compatible with higher DA synthesis activity of DA neurons in schizophrenia, at least in subjects experiencing psychotic symptoms.

Amphetamine-Induced DA Release

As discussed above, endogenous DA competition is a source of errors for *in vivo* measurement of D2 receptors. On the other hand, the recognition of this phenomenon implies that D2 receptor imaging, combined with pharmacologic manipulation of DA release, could provide a functional evaluation of DA presynaptic activity. Indeed, over the last decade, numerous groups demonstrated that acute increase in synaptic DA concentration is associated with decreased *in vivo* binding of [^{11}C]raclopride and [^{123}I]IBZM. These interactions have been demonstrated in rodents, nonhuman primates, and humans, using a variety of methods to increase synaptic DA [amphetamine, DAT blockers, levodopa (L-DOPA), nicotine agonists, serotonin receptor subtype 2A (5-HT_{2A}) antagonists, direct electrical stimulation of DA neurons] (see ref. 35 for review of this abundant literature). It has also been consistently observed that the *in vivo* binding of spiperone and other butyrophenones is not as affected as the binding of benzamides by acute fluctuations in endogenous DA levels (35).

The decrease in [^{11}C]raclopride and [^{123}I]IBZM *in vivo*

binding following acute amphetamine challenge has been well validated as a measure of the change in D2 receptor stimulation by DA due to amphetamine-induced DA release. Manipulations that are known to inhibit amphetamine-induced DA release, such as pretreatment with the DA synthesis inhibitor α-methyl-para-tyrosine (α-MPT) or with the DAT blocker GR12909 also inhibit the amphetamine-induced decrease in [^{123}I]IBZM or [^{11}C]raclopride binding (36 ,37). These experiments support the assumption that the amphetamine effect on [^{11}C]raclopride and [^{123}I]IBZM binding is mediated by DA release. Combined microdialysis and imaging experiments in primates demonstrated that the magnitude of the decrease in ligand binding was correlated with the magnitude of the increase in extracellular DA induced by the challenge (37 ,38), suggesting that this noninvasive technique provides an appropriate measure of the changes in synaptic DA levels.

Three out of three studies demonstrated that amphetamine-induced decrease in [^{11}C]raclopride or [^{123}I]IBZM binding was elevated in untreated patients with schizophrenia compared to well-matched controls (38 ,39 and 40). A significant relationship was observed between magnitude of DA release and transient induction or worsening of positive symptoms. The increased amphetamine-induced DA release was observed in both male and female patients, and in both first-episode/drug-naive patients and patients previously treated by antipsychotic drugs (41). Combined analysis of the results of two studies revealed that patients who were experiencing an episode of illness exacerbation (or a first episode of illness) at the time of the scan showed elevated amphetamine-induced DA release, whereas patients in remission showed DA release values not different from those of controls (41). This exaggerated response of the DA system to amphetamine exposure did no appear to be a nonspecific effect of stress, as higher self-reports of anxiety before the experiments were not associated with larger effect of amphetamine. Furthermore, nonpsychotic subjects with unipolar depression, who reported levels of anxiety similar to, if not higher than, the schizophrenic patients at the time of the scan, showed normal amphetamine-induced displacement of [^{123}I]IBZM (42).

These findings were generally interpreted as reflecting a larger DA release following amphetamine in the schizophrenic group. Another interpretation of these observations would be that schizophrenia is associated with increased affinity of D2 receptors for DA. Development of D2 receptors imaging with radiolabeled agonists is needed to settle this issue (42). Another limitation of this paradigm is that it measures changes in synaptic DA transmission following a nonphysiologic challenge (i.e., amphetamine) and do not provide any information about synaptic DA levels at baseline, i.e., in the unchallenged state.

Baseline DA Release

Several laboratories reported that, in rodents, acute depletion of synaptic DA is associated with an acute increase in the *in vivo* binding of [^{11}C]raclopride or [^{123}I]IBZM to D2 receptors (see ref. 35 for review). The increased binding was observed *in vivo* but not *in vitro,* indicating that it was not due to receptor up-regulation (43), but to removal of endogenous DA and unmasking of D2 receptors previously occupied by DA. The acute DA depletion technique was developed in humans using α-MPT to assess the degree of occupancy of D2 receptors by DA (43 ,44). Using this technique, higher occupancy of D2 receptor by DA was recently reported in patients with schizophrenia experiencing an episode of illness exacerbation, compared to healthy controls (45). Again assuming normal affinity of D2 receptors for DA, the data are consistent with higher DA synaptic levels in patients with schizophrenia. This observation was present in both first-episode/drug-naive and previously treated patients.

Following DA depletion, higher D2 receptor availability was observed in patients with schizophrenia compared to controls. This observation supported the proposition that, in schizophrenia, elevated D2 receptor density might be masked by DA occupancy when imaging studies are performed with ligands vulnerable to endogenous competition (28). However, the increase in D2 receptors measured with [^{123}I]IBZM in DA-depleted patients was moderate (12%), suggesting that other factors than vulnerability to endogenous DA competition are involved in the butyrophenone-benzamides binding differences discussed above.

Interestingly, higher occupancy of D2 receptors by DA in patients with schizophrenia was not associated with the intensity of positive symptoms, but was predictive of good therapeutic response of these symptoms following 6 weeks of treatment with atypical antipsychotic medications. The fact that high levels of synaptic DA at baseline predicted better or faster response to atypical antipsychotic drugs suggested that the D2 receptor blockade induced by these drugs remains a key component of their initial mode of action.

DA Transporters

The data reviewed above are consistent with higher DA output in the striatum of patients with schizophrenia, which could be explained by increased density of DA terminals. Because striatal DATs are exclusively localized on DA terminals, this question was investigated by measuring binding of [^{123}I]-CIT (46) or [18F]CFT (48) in patients with schizophrenia. Both studies reported no differences in DAT binding between patients and controls. In addition, Laruelle et al. (46) reported no association between amphetamine-induced DA release and DAT density. Thus, the increased presynaptic output suggested by the studies reviewed above

does not appear to be due to higher terminal density, an observation consistent with postmortem studies that failed to identify alteration in striatal DAT binding in schizophrenia (47 ,48 ,49 ,50 ,51 and 52).

Subcortical DA Dysregulation as a Failure of Inhibitory Pathways

Although the studies reviewed above generally confirmed the classic DA hypothesis of schizophrenia, it is important to examine these results in light of the more recent views of schizophrenia as a neurodevelopmental illness, involving dysconnectivity of multiple cortico-subcortical and intracortical networks. Although it cannot be definitively ruled out that the DA dysregulation revealed by these studies would stem from a primary abnormality of DA neurons, it seems more likely that these abnormalities are a consequence of cortico-subcortical dysconnectivity. Moreover, given the weight of evidence implicating PFC connectivity as a central deficient node in the schizophrenic brain, it is tempting to speculate that dysregulation of DA neurons' firing activity might stem from a failure of the PFC to regulate this process. In fact, it has long been hypothesized that dysregulation of subcortical DA function in schizophrenia may be secondary to a failure of the PFC to adequately control subcortical dopaminergic function (53 ,54).

Activity of midbrain DA neurons is under dual influence of PFC via an activating pathway (the “accelerator”) and an inhibitory pathway (“the brake”), allowing fine-tuning of dopaminergic activity by the PFC (55). The activating pathway is provided by direct glutamatergic projections onto the dopaminergic cells. The inhibitory pathway is provided by glutamatergic projections to midbrain γ-aminobutyric acid (GABA)ergic interneurons or striatomesencephalic GABA neurons. The inhibition of dopaminergic cell firing following amphetamine is an important feedback mechanism by which the brain reduces the effect of amphetamine on DA release. The inhibition of dopaminergic cell firing induced by amphetamine is mediated both by stimulation of presynaptic D2 autoreceptors and by stimulation of this inhibitory pathway (56). Following administration of amphetamine (i.e., under conditions in which the inhibitory pathway should be activated), *N*-methyl-D-aspartate (NMDA) receptor blockade results in a failure of activation of the inhibitory pathway, resulting in exaggerated amphetamine-induced DA release (57). Kegeles et al. (58) recently confirmed this mechanism in humans: ketamine pretreatment significantly enhanced amphetamine-induced decrease in [^{123}I]IBZM BP, from -5.5 ± 3.5% under control conditions to -12.8 ± 8.8% under ketamine pretreatment (p = .023). The increase in amphetamine-induced DA release induced by ketamine (greater than twofold) was comparable in magnitude to the exaggerated response seen in patients with schizophrenia. These data are consistent with the hypothesis that the alteration of DA release revealed by the amphetamine challenge in schizophrenia results from a disruption of glutamatergic neuronal systems regulating dopaminergic cell activity and are consistent with the hypothesis that schizophrenia might be associated with NMDA receptor hypofunction (59 ,60 and 61).

The failure of glutamatergic control of DA release might stem from mechanisms other than NMDA hypofunction. For example, glutamatergic projections from the PFC to the ventral tegmental area (VTA) are under tonic inhibition by prefrontal GABA and DA activity (see ref. 62 and references therein). It follows that deficits in GABAergic or dopaminergic function in the PFC (both deficits also implicated in schizophrenia) are expected to have similar consequences to an NMDA deficiency on the subcortical DA response to amphetamine. Thus, in patients with schizophrenia, various or multiple mechanisms (NMDA receptor hypofunction, GABAergic or dopaminergic deficits in the PFC) may lead to the dysregulation of subcortical DA revealed by the amphetamine challenge.

As reviewed below, direct evidence has been provided that disinhibition of subcortical DA activity is associated with prefrontal pathology in schizophrenia. In patients with schizophrenia, low *N*-acetylaspartate (NAA) concentration in the dorsolateral prefrontal cortex (DLPFC), a marker of DLPFC pathology, is associated with increased amphetamine-induced DA release (63). Studies in primates have documented the consequences of neurodevelopmental alteration in PFC connectivity on subcortical DA release (64 ,65). Adult rhesus monkeys with neonatal ablation of the amygdala-hippocampal formation within 3 weeks of birth exhibit lower NAA concentration in the PFC and abnormal relationships between prefrontal and subcortical DA functions; whereas local perfusion of amphetamine into the PFC induced a decrease in striatal DA in control monkeys and in monkeys with adult lesions, PFC amphetamine perfusion increased striatal DA release in monkeys with neonatal lesions. This study documents that dysregulation of subcortical DA function might be a delayed and enduring consequence of neurodevelopmental abnormalities of PFC connectivity.

Positive Symptoms and Neuroplasticity

The fluctuating nature of the DA abnormalities documented by the imaging studies reviewed above implies that some level of neuroplasticity is involved in the emergence of this increased subcortical DA activity during episodes of illness exacerbation (66). Neurochemical sensitization of mesolimbic DA systems has been proposed by several authors as one mechanism that might underlie the progression of a “silent” vulnerability into an overt symptomatology (67 ,68 ,69 ,70 and 71). Sensitization of DA systems is a positive feedback loop, in which increased DA activity leads to more DA

activity (70 ,72). During late adolescence, the failure of cortical development associated with schizophrenia liability might limit the brain capacity to modulate stress-related increased activity of mesolimbic DA neurons. This failure of normal homeostatic and buffering mechanisms would result in an increased vulnerability of DA neurons to develop a process of endogenous sensitization, a response not observed in humans under normal circumstances. The endogenous sensitization process drives the prodromal and initial phases of the illness, characterized by increased DA activity and culminating in the expression of positive symptoms. Sustained D2 receptor blockade interrupts this positive feedback loop. Upon neuroleptic discontinuation, the brain becomes again vulnerable to the stress-induced reemergence of this endogenous sensitization process and clinical relapse.

It should be emphasized that the relationship between stimulation of D2 receptors and psychotic symptoms is complex and presumably also involves neuroplasticity. NMDA antagonists such as ketamine or 5-HT_{2A} agonists such as lysergic acid diethylamide (LSD) induce psychotic symptoms in healthy subjects immediately upon drug exposure. In contrast, sustained administration of DA agonists is required to induce psychotic symptoms in healthy subjects (7 ,73). This observation suggests that sustained overstimulation of D2 receptors leads to remodeling of prefrontal-ventrostriatal-thalamic-prefrontal loops and their modulation by hippocampal afferents projections, neuronal ensembles that are believed to underlie the psychotic experience (74 ,75). In the amphetamine studies, DA-mediated stimulation of D2 receptors explained only about 30% of the variance in the positive symptom change in untreated patients with schizophrenia (41), indicating that factors downstream from the DA synapse play a role in the exacerbation of these symptoms following amphetamine. In the α-MPT study, global severity of positive symptoms did not correlate with occupancy of striatal D2 receptors by DA (45), suggesting that, in some patients, the experience of positive symptoms is no longer (or has never been) dependent on DA overstimulation. Patients with psychotic symptoms in the presence of apparently normal DA function failed to show significant improvement in these symptoms following 6 weeks of D2 receptor blockade (45). Thus, although these imaging studies have generally confirmed the time-honored dopamine hypothesis of schizophrenia, they also contributed to pointing out the limitations of an oversimplified model linking psychosis and excess DA activity.

Prefrontal DA D1 Receptor Density

As discussed above, several lines of evidence from preclinical, clinical, and postmortem studies converge to suggest that a deficiency in DA transmission in the prefrontal cortex is involved in the pathophysiology of negative symptoms and cognitive impairment in schizophrenia (14 ,16). Furthermore, a deficient prefrontal DA function is a potential mechanism to account for the subcortical DA disinhibition discussed above, as cortical DA function has an inhibitory impact on subcortical DA function (18 ,19). *In vivo* measurement of prefrontal DA function would provide the tools to directly test these hypothesis.

The majority of DA receptors in the PFC are of the D1 subtype (76 ,77). At the ultrastructural level, they are mostly located on pyramidal spines, and are mostly abundant on the distal dendrites (78 ,79 and 80). In postmortem studies, no evidence was found of an alteration in D1 receptors in the DLPFC of patients with schizophrenia (81 ,82), and the expression of the D1 receptor gene is unaltered (83). In contrast, a PET study with [^{11}C]SCH 23390 reported decreased density of D1 receptors in younger patients with schizophrenia (84). No significant differences were found in the other regions examined (anterior cingulate, temporal, occipital, and striatum). In addition, low-PFC D1 density was associated with the severity of negative symptoms and poor performance on the Wisconsin Card Sort Test (WCST). This finding is important, because it represents the first direct evidence of an association between negative symptoms, working memory deficits, and selective alteration in prefrontal DA function. However, the camera used in this study had a limited resolution, and the low specific to nonspecific ratio of [^{11}C]SCH3390 makes the measurement of D1 receptor in PFC with this ligand quite vulnerable to noise (85). Several groups are currently attempting to replicate this finding, using better cameras and a superior D1 receptor radiotracer, [^{11}C]NNC 112 (86).

In addition, measurement of receptor availability reveals only one aspect of neurotransmission. As D1 receptors are the most abundant DA receptors in the PFC, the availability of a D1 receptor radiotracer vulnerable to competition by endogenous DA (i.e., a D1 receptor "[^{11}C]raclopride") would be invaluable to assess presynaptic DA function in the PFC. Unfortunately, such a ligand is currently lacking (87 ,88).

Studies of Nondopaminergic Receptors in Schizophrenia

Receptors related to the GABA and 5-HT systems have been studied *in vivo* in schizophrenia. Postmortem studies reported abnormalities of both systems in schizophrenia. A robust body of findings suggests deficiency of GABAergic function in the PFC in schizophrenia (see refs. 89 and 90 for reviews). *In vivo* evaluation of GABAergic systems in schizophrenia has so far been limited to evaluation of benzodiazepine receptor densities with SPECT and [^{123}I]iomazenil, and three out of three studies comparing patients with schizophrenia and controls reported no significant regional differences (91 ,92 and 93). Although some significant correlations with symptoms clusters and regional benzodiazepine densities have been observed (91 ,92 ,94 ,95), these relationships

have not been replicated by other studies. Thus, together, these studies are consistent with an absence of marked abnormalities of benzodiazepine receptor concentration in the cortex and patients with schizophrenia. Alterations of GABAergic systems in schizophrenia might not involve benzodiazepine receptors (96), or be restricted to certain cortical layers or classes of GABAergic cells that are beyond the resolution of current radionuclides based imaging techniques. Recent developments in GABA imaging with MRS (described below) are a promising new avenue to study *in vivo* GABAergic function in schizophrenia.

Abnormalities of 5-HT transporters (SERT), 5-HT_{2A} receptors and, more consistently, 5-HT_{1A} receptors have been described in postmortem studies in schizophrenia (see references in ref. 97). Given the relatively recent development of radiotracers to study 5-HT receptors, only a limited number of imaging studies have been published. The concentration of SERT in the midbrain measured by [^{123}I]β-CIT is unaltered in patients with schizophrenia (46). Studies with more specific ligands are warranted to assess the distribution of SERT in other brain areas, such as the PFC, where their density has been reported to be reduced in three out of four postmortem studies (97). Decrease in 5-HT_{2A} receptors has been reported in the PFC in four out of eight postmortem studies (97 ,98). Three PET studies in drug-naive or drug-free patients with schizophrenia reported normal cortical 5-HT_{2A} receptor binding (98 ,99 and 100), whereas one study reported a significant decrease in PFC 5-HT_{2A} binding in a small group (*n* = 6) of drug-naive schizophrenic patients (101). The most consistent abnormality of 5-HT parameters reported in postmortem studies in schizophrenia is an increase in the density of 5-HT_{1A} receptors in the PFC, reported in seven out of eight studies (97). Several groups are currently evaluating the binding of this receptor *in vivo* with PET and [^{11}C]WAY100907.

Receptor Occupancy By Antipsychotic Drugs

Maybe the most widespread use of neuroreceptor imaging in schizophrenia over the last decade has been the assessment of neuroreceptor occupancy achieved by typical and atypical antipsychotic drugs, a topic that has been the subject of recent reviews (102 ,103). Neuroreceptor studied included essentially D2 receptors, but also 5-HT_{2A} and D1 receptors. The main conclusions from this line of research are as follows: (a) Studies repeatedly confirmed the existence of a threshold of occupancy of striatal D2 receptors (about 80%) above which extrapyramidal side effects are likely to occur (104). (b) In general, studies failed to observe a relationship between degree of D2 receptor occupancy and clinical response (105 ,106). Yet, most studies were performed at doses achieving more than 50% occupancy, and the minimal level of occupancy required for therapeutic response remains undefined. Two studies performed with low doses of relatively selective D2 receptor antagonists (haloperidol and raclopride) suggested that 50% to 60% occupancy was required to observe a rapid clinical response (107 ,108). (c) Clozapine, at clinically therapeutic doses, achieved only 40% to 60% D2 receptor occupancy (104 ,106 ,109), which, in conjunction with its anticholinergic properties, accounts for its low liability for extrapyramidal symptoms (EPSs). (d) Occupancy of 5-HT_{2A} receptors by "5-HT_{2A}/D2 balanced antagonists" such as risperidone does not confer protection against EPS, because the threshold of D2 receptor occupancy associated with EPS is not markedly different between these drugs and drugs devoid of 5-HT_{2A} antagonism (110 ,111 ,112 and 113). (e) Studies with quetiapine suggested that, at least with this agent, transient high occupancy of D2 receptors might be sufficient to elicit clinical response (114 ,115).

An interesting question relates to putative differences in degree of occupancy achieved by atypical antipsychotic drugs in striatal and extrastriatal areas. Pilowsky et al. (116) reported lower occupancy of striatal D2 receptors compared to temporal cortex D2 receptors in seven patients treated with clozapine, using the high-affinity SPECT ligand [^{123}I]epidipride. In contrast, typical antipsychotics were reported to achieve similar occupancy in striatal and extrastriatal areas, as measured with [^{11}C]FLB 457 (117) or [^{123}I]epidipride (118). It should be noted, however, that these very high affinity ligands do not allow accurate determination of D2 receptor availability in the striatum. In contrast, [^{18}F]fallypride enables accurate determination of D2 receptor availability in both striatal and extrastriatal areas (119), and preliminary PET experiments in primates with [^{18}F]fallypride indicate that clozapine and risperidone achieve similar D2 receptor occupancy in striatal and extrastriatal regions (120). Finally, it is important to point out that the most robust evidence relative to the site of therapeutic effect of antipsychotic drugs in rodents points toward the nucleus accumbens (121 ,122), whereas the imaging studies reviewed above contrasted striatal versus mesotemporal D2 receptor binding. Improved resolution of PET cameras currently allows dissociating signals from ventral and dorsal striatum (123 ,124), and it is now feasible to specifically study the clinical correlates of D2 receptor occupancy in ventral striatum in humans.

Another unresolved question is the discrepant values of D2 receptor occupancy obtained with [^{11}C]raclopride versus [^{11}C]NMSP. The haloperidol plasma concentration associated with 50% inhibition of [^{11}C]NMSP binding (3 to 5 mg/mL) (125) is ten times higher than that associated with 50% inhibition of [^{11}C]raclopride binding (0.32 ng/mL) (126). Quetiapine, at a dose of 750 mg, decreased [^{11}C]raclopride-specific binding by 51%, but failed to affect [^{11}C]NMSP-specific binding (127). These observations contribute to the debate regarding differences between benzamides and butyrophenones binding to D2 receptors.

Future Developments

Despite the remarkable achievements of the last decade, imaging neurosignaling processes with PET and SPECT in schizophrenia are still limited by the relative low number of probes available. For example, despite major research efforts, direct measurement of parameters of glutamate transmission are still not available. Radiotracers enabling evaluation of second messengers and intracellular pathways are only beginning to emerge (128). A growing collaboration between academic centers and industry currently holds the promise of increasing access to molecules for evaluation as candidate radiotracers.

Studies of the DA systems in schizophrenia illustrates how dynamic measurement of neurotransmission can be more informative than simple measurement of receptor density. With the exception of the cholinergic system (129 ,130), the paradigm used with [^{11}C]raclopride and [^{123}I]IBZM has been difficult to extend to other neuroreceptor systems, maybe because the fundamental mechanisms underlying acute change in *in vivo* binding following transmitter fluctuations are still not perfectly understood (66). Additional research is warranted to better characterize the factors that confer vulnerability of radiotracers *in vivo* binding to functional status of neurotransmission.

Finally, a general limitation of these radionuclide-based techniques is the technical sophistication required as well as the high cost of these investigations. However, the growing success of PET in oncology results in a larger availability of PET cameras for neuropsychiatric clinical and basic research. This growing availability should be associated with a vigorous research effort toward the development of more F-18 based probes, since the relatively longer half-life of F-18 compared to C-11 does not require that these ligands be radiolabeled locally. For SPECT, the development of technetium-based neuroreceptor ligands (131) will further enhance the availability of these techniques to the nuclear medicine community.

MAGNETIC RESONANCE SPECTROSCOPY

Part of "59 - Neurochemical and Neuropharmacological Imaging in Schizophrenia "

Magnetic resonance spectroscopy (MRS) is a chemical assay technique. It is the only clinically available method for the direct measurement of chemical moieties in the living brain. MRS is based on the same physical principles as magnetic resonance imaging (MRI), which involves characterizing atoms and molecules based on how they interact with a magnetic field. This interaction occurs because the nuclei of atoms with an odd number of nucleons (i.e., protons and neutrons), such as ^{1}H, ^{19}F, ^{13}C, ^{7}Li, ^{23}Na, have angular momentum, so-called spin, which generates a small magnetic field around the nucleus. The spin properties of specific atoms and of the specific molecules that they compose are unique and are exploited in an MRS experiment.

When a strong external magnetic field is applied to a tissue (e.g., the main magnet of an MR scanner), nuclei with spin align themselves with the external field and assume an equilibrium state of net magnetization, which is proportional to the strength of the field and the spin properties of the nucleus. An MRS experiment involves four steps, analogous to an MRI procedure. First, specific nuclei are excited with a brief “pulse” of a radiofrequency (RF) magnetic field supplied by an RF transmitter coil. This excitation causes magnetized spins to transiently assume a higher energy state, from which they “relax” to a lower energy state of equilibrium magnetization. Because the energy states are quantitized, only specific RFs will excite the nucleus to another state and only these frequencies will be emitted during relaxation (“resonance”). These resonant frequencies are unique for each atom, and vary in proportion to the strength of the external field. The motion of spins in the process of returning to equilibrium (“relaxation”) induces a current in a receiver coil, which represents the MR signal.

The second step involves spatially encoding the signal so that its origin can be mapped to a particular locale in the field of view. This is accomplished with the application of linear magnetic gradients that add localizing characteristics to the signal. The third step involves translating the signal acquired over time into a representation of its component frequencies and amplitudes, the so-called Fourier transformation. The fourth step involves the mathematical and statistical analyses of the data.

In MRI, the signal used to reconstruct images is from hydrogen atoms (^{1}H), which are found in many molecules, but most abundantly in water and lipids. Although the signal from hydrogen contains frequencies corresponding to many different molecules, the water and lipids signals dominate and the signals from hydrogen in other molecules are ignored. MRS, however, is based on resolving these other molecular signals. These molecules are identifiable because of the phenomenon called “chemical shift.” The electron cloud that surrounds the nucleus of an atom partially shields the nucleus from the external magnetic field. This shielding effect will cause the nucleus to experience a slightly different external field, and thus to resonate at a slightly shifted frequency. Because the degree of shielding varies from one molecule to another, depending on the electron sharing of the chemical bonds in the molecule, the exact shift in the resonant frequency of a target nucleus (e.g., ^{1}H, ^{31}P) will reflect its chemical environment. In addition to electron shielding, complex interactions between neighboring spins (“j-coupling”) also may affect the resonant frequency of a nucleus in certain molecules. A Fourier transformation of the emitted signal resolves the various components of the signal into a spectrum having peaks of specific frequency and amplitude, each peak corresponding to a different chemical environment of the target nucleus. The area under the peak is directly proportional to the concentration of spins having that frequency, and thus of that specific chemical

moiety. It is customary to label peaks in the spectrum based on their relative displacement (“chemical shift”) in parts per million from a standard peak (e.g., in the proton spectra, water is assigned a value of 4.75), with higher frequencies to the left. The use of relative shift units allows for peaks to be identified regardless of the strength of the magnet. This also makes it possible to compare *in vivo* and *in vitro* data to establish the chemical identity of each peak.

In schizophrenia research, MRS has been used to approach four general topics: (a) assays of brain chemicals, (b) evidence of tissue pathology, (c) functional analysis of specific neuronal populations, and (d) drug effects. Before considering the results of these efforts, it is important to consider some general methodologic issues that are germane to the MRS literature.

MRS Methodology

MRS is a methodologically complex technique that requires many quality control steps, both at the front end in data collection and at the back end in data analysis. It is not a turnkey technique, and careful scrutiny of the raw data is essential in every experiment. There are no absolute standards for the quality of spectral data, e.g., for the sharpness of a peak, or the degree of noise in the baseline. Nevertheless, spectra derived from a particular atom (e.g., ^{1}H or ^{19}F) have a characteristic pattern, with sharp peaks at defined relative frequencies, appropriate separation of peaks, acceptable signal to noise, and a stable baseline. Because the goal of MRS is to assay the concentrations of chemicals that are in the range of three to five orders of magnitude less abundant than water, it is necessary to make compromises to maximize the signal-to-noise ratio (SNR) in the data. One of the first decisions to be made in an MRS experiment is how to sample from the brain. Most MRS studies in the schizophrenia literature have sampled from relatively large (5 to 80 mL) single volume elements (“voxels”), though so-called MRS imaging (MRSI) approaches have also been utilized. The single-voxel approach benefits from enhanced signal-to-noise (STN) and reduced scanner time (5 to 20 minutes per voxel), but it suffers from imprecise and limited anatomic localization, contamination from signal outside of the voxel, and, most importantly, partial volume averaging. Large voxels contain multiple tissue compartments, including cerebrospinal fluid (CSF), white matter, gray matter, and vascular tissue, which have different concentrations of chemical constituents. Thus, the signals from a voxel are the average of signals from all of these compartments. Although there are statistical procedures aimed at “segmenting” these compartments within a voxel, such procedures are rough approximations (132). One of the fundamental difficulties in comparing data from single-voxel studies between subject groups and across studies in the literature is uncertainty about the comparability of voxel characteristics. ^{1}H-MRSI studies generally employ smaller voxels (<1.5 mL) and have the advantage of being more anatomically precise. This means that regions based on known anatomy, rather than arbitrary voxel placement, can be compared. There also are less partial volume effects. However, because of the smaller voxels, imaging studies have reduced SNR and require longer scanner times (20 to 60 minutes for a multislice study).

Because the amplitudes and frequencies of peaks in a spectrum are critically dependent on the magnetic field surrounding a target nucleus, magnetic field uniformity within a voxel and between voxels is essential. Field homogeneity is routinely honed with “shimming” procedures. It is also possible to make additional corrections based on field mapping, which corrects for frequency shifts between voxels. Most studies in the literature have used shim procedures supplied by the manufacturer of the scanner, and it is doubtful that these procedures have been optimal and comparable across studies (133). Coil design is another important parameter in data acquisition and STN. Small surface coils have been used in many of the phosphorus spectroscopy studies, because they improve SNR. However, surface coils introduce potential variance in studies that compare signals from specific locales across individuals. This is because as transmitters, surface coils excite spins with a gradient of intensity from the surface, and again as a receiver, the sensitivity drops off across the volume of activated tissue. Clearly, differences in coil placement and in head geometry may contribute to variations in the data.

There has been much discussion in the MRS literature about absolute quantification of chemical concentrations. In practice, this is almost impossible to do. References in the schizophrenia MRS literature to studies that have employed absolute quantification are inaccurate, as all studies have employed some normalization procedure, often based on measures acquired in normal subjects. To accurately measure absolute concentration of a chemical peak, it is necessary to control for all the variables that will affect the amplitude of the peak from a given voxel. These include factors that affect field homogeneity, relaxation times of the molecule(s) responsible for the peak, transmission efficiency and reception sensitivity of the RF coil (“coil loading”), and partial volume effects. These characteristics vary across the brain within an individual and between individuals even in the same locale in the brain. Moreover, any attempt to control for these factors assumes that the voxel remains in place, i.e., that the subject does not move during the scan, which is difficult for many individuals. Thus, even with the help of absolute standards for calibration, it is very difficult to correct for differences in these parameters (134). All studies to date have employed some normalization routine, whether to noise in the baseline, to the total observed signal, to another metabolite in the spectra, or to a signal from a molecule in the spectra that is thought to be metabolically neutral (e.g., water in the ^{1}H spectra). Although each of these normalization procedures involves assumptions and trade-offs, they attempt to control for unavoidable variations

in the local field, in coil loading, in relaxation times, and, to a lesser extent, for partial volume averaging effects. A further discussion of technical issues involved in collecting and analyzing MRS data is available in several reviews of the topic (135 ,136 and 137).

^{31}P Spectroscopy: Assays Of Phospholipids And High Energy Phosphates

The first *in vivo* MRS study of schizophrenia was of ^{31}P spectra. At clinical MR field strength (1.5 to 4.0 tesla), the ^{31}P spectrum contains several peaks across a wide chemical shift range [approximately 30 parts per million (ppm)]. The major metabolite peaks represent resonances for (a) phosphomonoester (PME) compounds (at 6.5 ppm), which includes phosphocreatine, phosphorylethanolamine; (b) phosphodiesters (PDE, 2.6 ppm), which include glycerolphosphocholine, glycerolphosphoethanolamine, and various membrane phospholipids; (c) phosphate residues on adenosine triphosphate (ATP) (-16.3 ppm for beta, -7.8 for alpha, and -2.7 for gamma); (d) inorganic phosphate at 4.9 ppm; and (e) phosphocreatine, which, as the chemical shift standard, has a resonance set at 0 ppm. PME concentrations, which are thought to reflect in part membrane precursors, increase with tissue growth, including in early brain development and gliomas, but also with tissue destruction, e.g., in Alzheimer's disease and HIV. In white matter, however, PMEs are metabolites of both synthesis and breakdown of sphingomyelin. PDEs are thought to reflect in part products of membrane breakdown and have been shown to be decreased in certain brain tumors and to increase shortly after birth. However, both PME and PDE peaks include other phosphorylated proteins associated with cell organelles and with membrane phospholipids that are not clearly related to membrane turnover. This fact is underscored by studies of patients with Huntington's disease (138) and of at least some patients with Alzheimer's disease (139), which have observed no abnormalities in these metabolites. Thus, changes in PME and PDE peaks are not easily interpreted. The phosphocreatine (PCr), ATP, and Pi peaks confer information about the state of high-energy phosphate metabolism and pH in the tissue. It also should be noted that because ^{31}P spectra are acquired with large voxels (i.e., 15 to 80 cm^3), the relative contributions of various tissue components (i.e., gray matter, white matter, neurons, glia, endothelia) to the signal is uncertain.

O'Callaghan and colleagues (140) reported the first ^{31}P spectroscopy study in schizophrenia, on a sample of 18 patients and 10 control subjects. They used a surface coil to acquire data from 87-cm^3 voxels arbitrarily placed in both temporal lobes. There were no differences found in any of the peaks. Pettegrew et al. (141) reported a highly cited study of 11 first-episode patients, using a surface coil placed over the front of the head. They reported that their voxel of 20 cm^3 sampled the dorsolateral prefrontal cortex, though its exact location is unclear, as subsequent reports stated that both right and left prefrontal cortices were sampled (142). With only a single 20-cm^3 voxel, this would indicate a more midline localization. The investigators reported that patients had decreased PMEs and Pi, and increased PDEs and ATP. They interpreted their findings to reflect greater membrane turnover and less energy utilization, results that they speculated were consistent with hypotheses about excessive synaptic pruning and decreased frontal lobe metabolism.

Because of these initial reports, over fifteen ^{31}P studies of patients with schizophrenia have appeared in the literature. Although several studies have reported partial replication of the findings of Pettegrew et al. (141 ,142), particularly with respect to a reduced PME peak, the majority of the reports have failed to do so. The data with respect to high-energy phosphate peaks are especially inconsistent and generally negative. Studies that have looked outside the frontal lobe, at parietal and temporal lobes and at the basal ganglia, also have yielded most often negative or inconsistent results. Several recent reviews provide tabulated summaries of this literature (143 ,144). The reasons for the inconsistencies are unclear. The usual explanations—differences in patient sample (e.g., medicated or unmedicated, acute or chronic), differences in acquisition parameters (e.g., surface or volume coil, single voxel, or multivoxel)—seem inadequate, as both positive and negative reports have appeared regardless of the characteristics of the samples and in the context of a variety of acquisition parameters. A more likely explanation for the inconsistencies involves the low sensitivity and reproducibility of ^{31}P spectroscopy and problems in achieving a standardized placement of a region of interest (ROI) or of an acquisition plane. There are very limited data about the reliability of ^{31}P measurements in patients, with one study reporting a coefficient of variation of 30% across two scans (145). Because it is impossible to accurately register single voxels or single planes to an anatomic reference, there is unavoidable error in the localization process, which is further compounded by subject motion during the scan. Large voxel sizes also introduce error in terms of partial volume effects. The small differences that have been reported in most of the positive studies, on the order of 10% to 25%, would seem especially sensitive to such methodologic difficulties.

Two recent studies are particularly noteworthy in their efforts to control for localization variance and to improve the sensitivity of the spectral data. Volz et al. (146) acquired a plane of ^{31}P spectra, composed of 30 voxels of 19 cm^3 each, and drew ROIs on an anatomic reference scan that was acquired in a coplanar orientation, on 11 medication-free patients, including seven acute patients who were medication naive. Though the planes are not strictly registered in space, because subject motion may have tilted their relative orientations during the scans, the authors' approach to ROI

placement is more reliable than that in earlier studies. With the exception of one mesial prefrontal voxel having a decreased PDE peak and one voxel localized to the basal ganglia having a decreased PME peak, no other cerebral differences in phospholipids were found. Four prefrontal voxels having decreases in ATP and PCr also were found. The authors argued that their data suggested decreased membrane catabolism, but, given the number of voxels analyzed, chance results cannot be excluded. In an earlier study of chronic, medicated patients, they also found a decreased prefrontal PDE peak, but the high-energy phosphate data were in the opposite direction (147). Finally, Potwarka et al. (148), using signal enhancement techniques that reduce the effects of coupling between phosphorus and proton spins, were able to separate structural membrane phospholipids from other constituents of the PME and PDE peaks, with 50-cc voxels acquired as part of a plane. Although the authors found no decrease in PMEs as reported by Pettegrew et al., they did find increases in the frontal PDE peak, but not related to membrane breakdown products such as glycerophosphocholine (GPC) and glycerolphosphoethanolamine (GPE), as suggested by Pettegrew et al. and others. Rather, the difference was reflected in the membrane phospholipid components of the peak. These differences were not found in motor or occipital cortices, providing some evidence of internal validity. This study also found no differences between patients and controls in total PMEs, again in contrast to the Pettegrew et al. study, but a component of the PME peak, phosphocholine, was reduced. Potwarka et al. (148) proposed that their data implicated membrane abnormalities selectively in DLPFC, perhaps involving presynaptic vesicular phospholipids. However, to confuse the story even further, Bluml et al. (149), using a similar proton decoupled ^{31}P MRS approach, reported increases in GPC and GPE acquired with a large (97-cc) voxel in the middle of the cerebrum.

Several studies have attempted to link ^{31}P data to clinical characteristics of patients, but these also have been inconsistent. For example, Deicken et al. (150) reported a correlation between prefrontal PME signals and performance on the WCST, suggesting that prefrontal membrane abnormalities were reflected in prefrontal function. However, in the same patient sample, Deicken et al. did not find an abnormality of PME signals. Volz et al. (147) could not find a correlation between any ^{31}P signals and WCST performance. Potwarka et al. (148) found a correlation between an ATP peak in right prefrontal cortex and negative symptom ratings, but similar relationships were not found in other studies (151).

It is difficult to arrive at a synthetic analysis of the ^{31}P data in schizophrenia. Technical error is probably the critical factor in the variable results that have been reported. It is doubtful that the small differences between patients and controls could escape corruption by the many methodologic limitations of the current techniques. Future studies using higher field magnets, with better sensitivity and resolution, combined with signal enhancement and peak separation procedures may lead to more reliable methods.

Proton Spectroscopy: Assays of Cells and Cellular Metabolism

Proton spectroscopy has been a more widely applied technique in schizophrenia research and the results are much more consistent. A variety of chemicals in the proton spectrum can be assayed with clinical magnets, including several amino acids, membrane and myelin metabolites, and several high-energy substrates. Although the sensitivity of proton spectroscopy is approximately 20 times that of phosphorus, allowing for much better resolution, the metabolites of interest need to be resolved in a smaller chemical shift range (less than 10 ppm), in the presence of large concentrations of brain water (approximately 10^4 times greater concentration than the other metabolites) and mobile lipids from the skull and scalp. Until recently, most ^{1}H MRS techniques used special procedures to suppress the signal from water and lipids, procedures that can affect other signals. With the availability of analogue to digital converters having greater dynamic range, it is now possible to acquire the water signal and still resolve the other metabolites (152). In addition to preservation of neighboring signals, this approach also has the advantage of making the water signal available as an internal standard for normalization and potentially for tissue segmentation.

The proton spectrum is characterized by several relatively large and distinct peaks and several complexes of smaller overlapping peaks. The metabolite signals acquired with ^{1}H MRS vary depending on the echo time of the pulse sequence used for the acquisition. Many of the resolvable elements have short T2 (e.g., myoinositol, glutamate, glutamine, and GABA) and emit no observable signal with longer echo times. On the other hand, long echo time acquisitions produce signals from several compounds that are very distinctly resolved. The long echo time metabolite spectrum is dominated by a peak at approximately 2 ppm corresponding to the methyl group (CH_3) of several *N*-acetyl containing compounds, principally *N*-acetylaspartate (NAA) and to a small degree, *N*-acetyl aspartate glutamate (NAAG) and possibly *N*-acetylneuraminic acid (NANA) (134). NAA is an intracellular neuronal marker, found almost exclusively in mature neurons and their processes (153), with the highest concentrations in pyramidal glutamate neurons (154). NAA is the second most abundant amino acid in the brain (155). Its concentration is higher in gray matter than in white matter (156), and NAA signals increase during childhood, remaining relatively stable throughout adult life (156 ,157 and 158). The exact implications of changes in NAA signals is uncertain, as its cellular function is still unclear. It is synthesized in mitochondria from glutamate and either pyruvate or 3-hydroxybutyrate via L-aspartate-*N*-amino

transferase and also is a by-product of NAAladase catabolism of NAAG, which occurs within glia (159). Whether NAA signals are absolutely specific to neurons is unclear. Mature astroglia do not contain NAA, though small concentrations have been reported in oligodendroglial cultures (160). NAA is a nonspecific though highly sensitive marker of neuronal pathology. Virtually all neurologic conditions involving neuronal pathology that have been studied, including multiple sclerosis, motor neuron disease, Alzheimer's disease, Huntington's disease, cerebellar degenerations, multiple sclerosis, epilepsy, and various encephalopathies, show changes in NAA signals in the regions of brain pathology. Moreover, NAA changes are sensitive measures of dynamic neuropathologic processes (161 ,162 and 163), for example, correlating over time with cognitive change in Alzheimer's disease (164) and with the number of trinucleotide repeats in Huntington's disease (165). Although early studies interpreted NAA findings as indicative of cell loss, recent data have established that NAA reductions can reverse following various forms of brain damage and can change with clinical improvement and treatment (161 ,162 ,166 ,167 and 168). This has led to speculation that NAA reductions occur as a manifestation of changes in neuronal volume or in NAA concentrations within a neuron, perhaps reflecting reduced mitochondrial energy metabolism (159) or a change in the abundance and patterns of neuronal connectivity (169 ,170). It is interesting to note that in various conditions associated with tissue volume loss and reduced NAA signals (e.g., epilepsy, Alzheimer's disease, schizophrenia), these two parameters are not tightly correlated. The only condition in which NAA concentrations are increased is Canavan's disease, which involves a mutation in a gene on chromosome 17 controlling the synthesis of *N*-acetyl-L-aspartate amidohydrolase (aspartoacylase), the enzyme that breaks down NAA.

Two other prominent peaks are seen in the long echo time proton spectrum. At 3.2 ppm is a peak corresponding to the trimethylamine group of various choline (CHO)-containing compounds, mostly membrane phospholipids. This signal reflects the concentrations of several phospholipid moieties, including glycerophosphocholine, phosphocholine, and phosphatidylcholine. In pathologic conditions associated with membrane turnover or gliosis (e.g., Alzheimer's disease, gliomas, epilepsy), the CHO peak tends to be elevated. At 3.0 ppm, a peak corresponding to creatine and phosphocreatine (CRE) appears. These metabolites participate as energy buffers in many energy-consuming processes in the brain, but consistent changes in the CRE signal are generally found only in the presence of tissue loss.

At short echo time, several other peaks are observed, in addition to the peaks that persist into the long echo time spectra. The most studied of these is the myoinositol peak (3.6 ppm) and a peak complex (called tGlx) at around 2.2 to 2.4 and 3.75 ppm corresponding to overlapping signals from glutamate, glutamine, and GABA. Myoinositol is a hexol present in high concentrations in human brain, and accounts for most of the myoinositol peak, though other complex inositol phosphates also contribute to the signal. Some of these inositol phosphates may represent second-messenger signaling molecules that may vary with the state of cellular activity. The myoinositol peak, however, tends to change in conditions associated with active membrane turnover and gliosis, and is consistently increased in Alzheimer's disease. The glutamate and GABA peaks include soluble forms of these amino acids involved both in neurotransmission and in peptide synthesis. Glutamine is an intermediary in glial-based recycling of the carbon skeletons of these amino acids, and has been proposed as a more sensitive marker of turnover of the glutamate amino acid pool.

Evidence from ^{1}H MRS of Neuronal Pathology

Nasrallah et al. (171) reported the first ^{1}H MRS study of both mesial temporal lobes in schizophrenia, an investigation of 11 chronic patients and 11 controls, using a 12-cm^{3} voxel. They found decreases in NAA peaks in both voxels, with the difference on the right significant at the .05 level, and on the left at the .06 level. Since the report of Nasrallah et al., over 20 ^{1}H MRS studies of patients with schizophrenia have appeared. Most of them have addressed metabolite changes, primarily NAA, in the frontal and temporal lobes. Table 59.3 summarizes studies of frontal and temporal lobe NAA signals. Several recent reviews describe these reports in greater detail (143 ,144 ,172). Although most of these studies have involved single voxel data, with the inherent methodologic issues described above, several recent relatively high resolution (approximately 1-mL voxel) multivoxel and multislice imaging studies (MRSI) have also been reported. Reliability data for NAA are much superior to those of other metabolites in either the proton or the phosphorus spectra; for example, the coefficient of variation for repeat studies in patients using an MRSI technique is on the order of 10% (173). It is interesting in this regard that all of the MRSI studies to date that have examined cortical regions in the frontal and temporal lobes have reported reduced NAA signals in these regions (174 ,175 ,176 ,177 ,178 and 179). These studies are exemplified by a series of reports from Bertolino and colleagues (174 ,175 ,176 and 177) using a technique that acquires over 700 1.4-mL voxels in approximately 25 minutes. The technique allows for registration of spectroscopic and anatomic images, for reliable and anatomically correct ROI definition, and for relatively diminished partial volume effects. The studies of Bertolino et al. included chronic medicated patients, unmedicated and several neuroleptic-naive patients, and childhood-onset patients. In each of these samples, NAA signals were reduced selectively and bilaterally in dorsolateral prefrontal and perihippocampal cortices. Using a similar technical approach in studies of chronic patients, Deicken et al.

replicated these results (178 ,179), and also found reduced NAA signals in cingulated cortex (180) and in thalamus (181), which were not found in the Bertolino et al. studies. The only imaging study that reported no decrease in gray matter NAA did not regionally parcellate the cortex, and reported only an average measure for the entire cortex of the top half of the brain (182). Indeed, this result is consistent with the findings of Bertolino et al., in which NAA levels throughout most regions of brain were not different between patients and controls. Although most of the single voxel studies also found reductions in NAA signals in frontal and temporal lobes, the few negative studies (183 ,184 ,185 ,186 and 187) may reflect methodologic variance related to small sample sizes or to voxel placement, etc. The single voxel studies even of NAA signals also tend to have much less reliable data [e.g., a coefficient of variation (cv) in a negative study by Kegeles et al. (187) of greater than 30%]. It is clear from this literature that reduced NAA peaks are found in the frontal and temporal lobes of many patients with schizophrenia (see table for a summary).

Study	Method	Sample Size Patients/Controls	Temporal Lobe	Frontal Lobe	Other
Nasrallah et al. (171)	12 cc voxel	11/11	↓	—	
Choe et al. (212)	8 cc voxel	23/10	—	↓	↑ tGLx
Renshaw et al. (213)	8 cc voxel	13/15	↓	—	First episode patients
Yurgelun-Todd et al. (214)	8 cc voxel	16/14	↓	—	Mainly extrahippocampal voxel
Buckley et al. (215)	11 cc voxel	28/20	NC	↓	
Maier et al. (216)	4–9 cc voxel	25/32	↓	—	
Maier and Ron (217)	5–7 cc voxel	26/38	↓	—	No differential change w/aging
Fukuzako et al. (218)	27 cc voxel	30/30	↓	NC	No abnormalities in drug naive patients; white matter voxel in frontal lobe
Stanley et al. (184)	8 cc voxel	29/24	—	NC	↑ glutamine in chronic patients
Bartha et al. (183)	4.5 cc voxel	10/10	NC	—	First episode, mesial frontal voxel; increased glutamine
Cecil et al. (219)	4.5–8 cc voxel	10/24	↓	↓	Drug naive patients
Brooks et al. (220)	8 cc voxel	16/12	↓	↓	Childhood schizophrenia
Thomas et al. (221)	8 cc voxel	13/12	↓	↓	Childhood schizophrenia
Bertolino et al. (174)	MRSI 1.4 cc/voxel	10/10	↓	↓	Chronic patients
Bertolino et al. (175)	MRSI 1.4 cc/voxel	12/12	↓	↓	Medication-free patients
Bertolino et al. (176)	MRSI 1.4 cc/voxel	14/14	↓	↓	Childhood onset schizophrenia
Callicott et al. (177)	MRSI 1.4 cc/voxel	103/71	↓	↓	Hippocampal NAA decreased—prefrontal NAA unchanged in healthy siblings
Deicken et al. (179)	MRSI 1.3 cc/voxel	23/18	↓	—	No change in hippocampal volume
Deicken et al. (178)	MRSI 1.3 cc/voxel	24/15	—	↓	
Heimberg et al. (186)	8 cc^3 voxel	13/142	NC	NC	White matter frontal voxel; temporal voxel excluded hippocampus
Kegeles et al. (143)	1.4 cc multivoxel	10/10	NC	—	Poor reliability [cv >30%]
Fukuzako et al. (222)	8 cc^3 voxel	64/51	↓	—	
Block et al. (198)	30 cc^3 voxel	25/19	—	↓	No abnormality in healthy relatives
Bartha et al. (185)	6 cc^3 voxel	11/11	NC	—	No glutamine differences

↓, reduced; ↑, increased; —, not assayed; MRSI, magnetic resonance spectroscopy imaging; NAA, *N*-acetylaspartate; NC, no change.

TABLE 59.3. ^{1}H MRS STUDIES OF FRONTAL AND TEMPORAL LOBE NAA SIGNALS IN SCHIZOPHRENIA

There have been few studies of other peaks in the proton spectra, and the results have been inconclusive. Because of interest in glutamate and cortical function in schizophrenia, several groups have attempted to measure the tGlx peak, using quantitation methods based on a priori knowledge of the relative contributions of the metabolite components, but the reliability of the method is limited [e.g., cv >50% (183)]. Stanley et al. (184) found increased glutamine signals in chronically treated patients, but no differences in acute patients. Bartha et al. (183) reported increased glutamine in cingulated cortex of a group of ten patients. Rakow et al. (188) reported decreased tGlx in dorsolateral prefrontal cortex and Kegeles et al. (187) reported no changes in mesial temporal lobe. The data with choline and creatine peaks have generally been negative, and the occasional positive result has been inconsistent across studies.

There has been considerable interest in trying to understand the meaning of NAA decreases in patients with schizophrenia. The NAA changes have consistently been shown not to correlate with stage of illness, with medication status, and with length of illness. Moreover, hippocampal and prefrontal volume do not appear to correlate either. In the Bertolino et al. (175) study of unmedicated patients, several of the subjects had been chronically psychotic and untreated for over 10 years, and no relationship between NAA signals and length of untreated psychosis was found. These data suggest that NAA reductions are not the result of a linearly progressive pathologic process. They also provide evidence against the notion that chronic untreated psychosis is "neurotoxic." Because postmortem studies do not provide convincing evidence of neuronal loss or of neuronal degeneration in the areas where NAA signals are consistently reduced (i.e., hippocampal formation and prefrontal cortices) (189), the NAA changes probably reflect more subtle aspects of neuronal biology (see below). As such, they are unique *in vivo* evidence of cellular changes probably restricted to neurons in the schizophrenic brain. Although NAA reductions are found in both medicated and unmedicated patients, the assumption that NAA changes are independent of medical treatment may be incorrect. Bertolino et al. (167) recently reported that NAA signals in dorsolateral prefrontal cortex increase slightly but significantly after only several weeks of neuroleptic treatment. This slight increase is further evidence that NAA levels may reflect dynamic neuronal events, and is consistent with evidence that other pharmacologic treatments can alter NAA signals [e.g., lithium (190)]. Although NAA changes clearly occur independent of neuroleptic treatment, the neuroleptic effect illustrates that physiologic factors may contribute to the variations in NAA signals.

NAA Signals and Abnormalities of Specific Functional Systems

Because NAA signals originate perhaps exclusively in neurons, principally glutamate neurons, it is possible to probe the functional connectivity of cortical glutamate neurons with NAA as a surrogate marker of the integrity or activity of such connectivity. This has been done in animals and in human studies. In animals, optic nerve injury reduces NAA concentrations in the lateral geniculate (191), and prefrontal cortical injury reduces NAA signals in basal ganglia, suggestive of loss of afferent terminals or transynaptic cellular changes (169). Bertolino et al. (170) measured NAA signals in DLPFC of adult monkeys that had undergone mesial temporal lobe removals as neonates and in monkeys that had undergone temporal lobe removals as adults. Reduced NAA signals were found only in the animals with neonatal removals, suggesting that NAA could reflect more complex plastic neuronal modifications, perhaps at the level of local circuit architecture, than simply loss of afferent terminals. Similar results have also been observed in rodents (192).

This evidence of relationships between NAA signals and connections between neurons led to a series of studies aimed at testing hypotheses about the centrality of prefrontal connectivity in the pathophysiology of schizophrenia (193). Bertolino et al. (63 ,194 ,195) used NAA signals as a marker of glutamate projection neuronal function/integrity in predicting the activity of distributed neuronal systems implicated in positive and negative symptoms. NAA signals selectively in DLPFC were found to predict the availability of DA receptors in the striatum in patients with schizophrenia, assayed with radionuclide imaging (194). Specifically, lower DLPFC NAA predicted greater availability of DA receptors in an unstimulated, resting state, speculated to reflect less DA release and cell firing. The same measure, i.e., low NAA, also predicted the exaggerated response to amphetamine found in striatum with radioreceptor imaging (63). These data suggested that NAA in DLPFC predicted the steady-state and stimulus-induced responses of dopamine neurons in the ventral brainstem. Similar relationships were not found for any other cortical regions or in normal controls, implicating the neuronal pathology associated with the illness as instrumental in constraining these relationships. Although DA release was inferred indirectly from radioligand binding availability, the assumptions were confirmed in studies of monkeys undergoing *in vivo* microdialysis, in which NAA concentrations in DLPFC directly predicted DA release in the striatum (194).

NAA signals in DLPFC also were found in patients with schizophrenia to selectively predict the activation of the distributed working memory cortical network, studied both with PET (195) and with functional MRI (fMRI) (196). These data suggest that prefrontal glutamate neurons, by virtue of their intracortical connectivities, modulate the capacity and efficiency of distributed cortical networks involved in working memory. Working memory deficits have been consistently linked to other aspects of the negative symptoms of schizophrenia, and in fact NAA signals in DLPFC also specifically predict negative symptoms ratings in patients (197).

These various clinical studies of phenomena predicted by NAA signals in DLPFC converge on a tantalizing interpretation of these various findings—that NAA reductions reflect subtle cellular pathology of intrinsic DLPFC neurons and their local circuitry that, by virtue of intracortical and corticofugal projections, modify the activity of distributed cortical networks and of DA neurons, implicated in the negative and positive symptoms of schizophrenia, respectively. Thus, DLPFC neurons appear to be an effector neuronal population that is associated with the manifest biology of the illness. It is interesting to note that NAA signals in hippocampal formation, the other region consistently implicated in studies of patients with schizophrenia, do not show these predictable relationships. A study by Callicott et al. (177)

may shed some light on this apparent inconsistency. Callicott et al. studied 60 healthy siblings of patients with schizophrenia and found that although NAA signals were also reduced in the hippocampal formation of the healthy siblings, changes in DLPFC were not found. These data suggest that consistent with other evidence of hippocampal functional abnormalities in relatives of patients with schizophrenia (e.g., memory deficits, P-50–evoked potentials), NAA changes in the hippocampal formation may reflect biology of genetic risk. In contrast, consistent with the predictable relationships of NAA in DLPFC with other functional systems implicated in schizophrenia, NAA changes in DLPFC reflect biology of manifest illness. The lack of a difference in frontal lobe NAA in healthy relatives of patients with schizophrenia has recently been reported by another group as well (198). Therefore, these results have added hippocampal NAA measures to the list of potential phenotypic markers of genetic risk for further exploration in genetic studies of mental illness.

Future Developments

MRS is a rapidly evolving technology and its future applications in schizophrenia research should lead to important discoveries. New developments in the near future will likely emerge from methodologic advances leading to improved sensitivity and resolution, measurement of novel chemical moieties, indexing neuronal metabolism, and characterizing drug effects. The availability of high-field human magnets (3 to 7 tesla) will substantially improve the STN of MRS and allow for improved reliability and resolution. At the National Institute of Mental Health (NIMH), proton spectral images are currently acquired at 3 T with 0.7-mL voxels, and the cv of repeat NAA measurements in the hippocampus is about 3%. Various hardware upgrades have also made it possible to shim individual slices rather than slabs of tissue, and to acquire both early and late echo spectra within the same acquisition, without suppressing the water signal. Use of techniques to control for motion effects (e.g., navigator signals) may further improve reliability and make it possible to avoid lipid suppression approaches, as well. The improved STN and signal acquisition of new methods will result in more sensitive and reliable acquisition of other spectral peaks. This will improve the potential reliability of phosphorus moieties in the proton spectrum and will make calculations of the components of the tGlx peak more reliable. Preliminary results using spectral editing approaches to the GABA peak suggest that clinically meaningful data about GABA metabolism can be derived from this peak (199). Greater sensitivity and SNR also will permit spectral analyses of externally administered molecules. C13 glucose can be given as a glucose load and the fate of the C13 carbon skeleton tracked over time as changing concentrations of moieties in the C13 spectra (200). This may provide near–real-time information about glucose metabolism, and also potentially about the turnover of several neurotransmitters. Preliminary studies of fluorinated compounds, such as fluoxetine and fluphenazine, have demonstrated the feasibility of measuring the concentrations of such compounds in brain. The potential applications of such measurements will become clearer in the context of improved methodology. Finally, the demonstration that NAA changes with antipsychotic drug treatment represents the first *in vivo* evidence of an intracellular effect of these drugs. Future studies will aim to understand the basis for this change, as well as other factors that affect NAA signals.

ACKNOWLEDGMENTS

Part of "59 - Neurochemical and Neuropharmacological Imaging in Schizophrenia "

This work is supported by the National Institute of Mental Health (M.L., K02 MH01603-01). Dr. Weinberger has served as a consultant for Janssen and Eli Lilly & Company.

REFERENCES

1. Wagner H Jr, Burns HD, Dannals RF, et al. Imaging dopamine receptors in the human brain by positron tomography. *Science* 1983;221:1264–1266.

2. Rossum V. The significance of dopamine receptor blockade for the mechanism of action of neuroleptic drugs. *Arch Int Pharmacodyn Ther* 1966;160:492–494.

3. Carlsson A, Lindqvist M. Effect of chlorpromazine or haloperidol on formation of 3-methoxytyramine and normetanephrine in mouse brain. *Acta Pharmacol Toxicol* 1963;20:140–144.

4. Seeman P, Lee T. Antipsychotic drugs: direct correlation between clinical potency and presynaptic action on dopamine neurons. *Science* 1975;188:1217–1219.

5. Creese I, Burt DR, Snyder SH. Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. *Science* 1976;19:481–483.

6. Connell PH. *Amphetamine psychosis*. London: Chapman and Hill, 1958.

7. Angrist BM, Gershon S. The phenomenology of experimentally induced amphetamine psychosis-Preliminary observation. *Biol Psychiatry* 1970;2:95–107.

8. Meltzer H. Clinical studies on the mechanism of action of clozapine: the dopamine-serotonin hypothesis of schizophrenia. *Psychopharmacology* 1989;99:S18–S27.

9. Carlsson A. The current status of the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology* 1988;1:179–186.

10. Laruelle M. Imaging dopamine transmission in schizophrenia. A review and meta-analysis. *Q J Nucl Med* 1998;42:211–221.

11. Davis KL, Kahn RS, Ko G, et al. Dopamine in schizophrenia: a review and reconceptualization. *Am J Psychiatry* 1991;148:1474–1486.

12. Harrison PJ. The neuropathology of schizophrenia. A critical review of the data and their interpretation. *Brain* 1999;122:593–624.

13. Weinberger DR, Berman KF. Prefrontal function in schizophrenia: confounds and controversies. *Philos Trans R Soc Lond [B]* 1996;351:1495–1503.

14. Goldman-Rakic PS, Selemon LD. Functional and anatomical aspects of prefrontal pathology in schizophrenia. *Schizophr Bull* 1997;23:437–458.

15. Weinberger DR. Implications of the normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* 1987;44:660–669.

16. Knable MB, Weinberger DR. Dopamine, the prefrontal cortex and schizophrenia. *J Psychopharmacol* 1997;11:123–131.

17. Akil M, Pierri JN, Whitehead RE, et al. Lamina-specific alterations in the dopamine innervation of the prefrontal cortex in schizophrenic subjects. *Am J Psychiatry* 1999;156:1580–1589.

18. Deutch AY. Prefrontal cortical dopamine systems and the elaboration of functional corticostriatal circuits: implications for schizophrenia and Parkinson's disease. *J Neural Transm Gen Sect* 1993;91:197–221.

19. Wilkinson LS. The nature of interactions involving prefrontal and striatal dopamine systems. *J Psychopharmacol* 1997;11:143–50.

20. Wong DF, Wagner HN, Tune LE, et al. Positron emission tomography reveals elevated D2 dopamine receptors in drug-naive schizophrenics. *Science* 1986;234:1558–1563.

21. Crawley JC, Owens DG, Crow TJ, et al. Dopamine D2 receptors in schizophrenia studied *in vivo*. *Lancet* 1986;2:224–225.

22. Seeman P, Guan HC, Van Tol HHM. Dopamine D4 receptors elevated in schizophrenia. *Nature* 1993;365:411–445.

23. Lahti RA, Roberts RC, Cochrane EV, et al. Direct determination of dopamine D-4 receptors in normal and schizophrenic postmortem brain tissue: a [H-3]NGD-94-1 study. *Mol Psychiatr* 1998;3:528–533.

24. Ng GY, O'Dowd BF, Caron M, et al. Phosphorylation and palmitoylation of the human D2L dopamine receptor in Sf9 cells. *J Neurochem* 1994;63:1589–1595.

25. Ng GY, O'Dowd BF, Lee SP, et al. Dopamine D2 receptor dimers and receptor-blocking peptides. *Biochem Biophys Res Commun* 1996;227:200–204.

26. Zawarynski P, Tallerico T, Seeman P, et al. Dopamine D2 receptor dimers in human and rat brain. *FEBS Lett* 1998;441:383–386.

27. Lee SP, O'Dowd BF, Ng GY, et al. Inhibition of cell surface expression by mutant receptors demonstrates that D2 dopamine receptors exist as oligomers in the cell [In Process Citation]. *Mol Pharmacol* 2000;58:120–128.

28. Seeman P, Guan H-C, Niznik HB. Endogenous dopamine lowers the dopamine D2 receptor density as measured by [3H]raclopride: implications for positron emission tomography of the human brain. *Synapse* 1989;3:96–97.

29. Seeman P. Brain dopamine receptors in schizophrenia: PET problems. *Arch Gen Psychiatry* 1988;45:598–560.

30. Reith J, Benkelfat C, Sherwin A, et al. Elevated dopa decarboxylase activity in living brain of patients with psychosis. *Proc Natl Acad Sci USA* 1994;91:11651–11654.

31. Hietala J, Syvalahti E, Vuorio K, et al. Presynaptic dopamine function in striatum of neuroleptic-naive schizophrenic patients. *Lancet* 1995;346:1130–1131.

32. Dao-Castellana MH, Paillere-Martinot ML, Hantraye P, et al. Presynaptic dopaminergic function in the striatum of schizophrenic patients. *Schizophr Res* 1997;23:167–174.

33. Hietala J, Syvalahti E, Vilkman H, et al. Depressive symptoms and presynaptic dopamine function in neuroleptic-naive schizophrenia. *Schizophr Res* 1999;35:41–50.

34. Lindstrom LH, Gefvert O, Hagberg G, et al. Increased dopamine synthesis rate in medial prefrontal cortex and striatum in schizophrenia indicated by L-(beta-^{11}C) DOPA and PET [In Process Citation]. *Biol Psychiatry* 1999;46:681–688.

35. Laruelle M. Imaging synaptic neurotransmission with *in vivo* binding competition techniques: a critical review. *J Cereb Blood Flow Metab* 2000;20:423–451.

36. Villemagne VL, Wong DF, Yokoi F, et al. GBR12909 attenuates amphetamine-induced striatal dopamine release as measured by [(11)C]raclopride continuous infusion PET scans. *Synapse* 1999;33:268–273.

37. Laruelle M, Iyer RN, Al-Tikriti MS, et al. Microdialysis and SPECT measurements of amphetamine-induced dopamine release in nonhuman primates. *Synapse* 1997;25:1–14.

38. Breier A, Su TP, Saunders R, et al. Schizophrenia is associated with elevated amphetamine-induced synaptic dopamine concentrations: evidence from a novel positron emission tomography method. *Proc Natl Acad Sci USA* 1997;94:2569–2574.

39. Laruelle M, Abi-Dargham A, van Dyck CH, et al. Single photon emission computerized tomography imaging of amphetamine-induced dopamine release in drug free schizophrenic subjects. *Proc Natl Acad Sci USA* 1996;93:9235–9240.

40. Abi-Dargham A, Gil R, Krystal J, et al. Increased striatal dopamine transmission in schizophrenia: confirmation in a second cohort. *Am J Psychiatry* 1998;155:761–767.

41. Laruelle M, Abi-Dargham A, Gil R, et al. Increased dopamine transmission in schizophrenia: relationship to illness phases. *Biol Psychiatry* 1999;46:56–72.

42. Hwang DR, Kegeles LS, Laruelle M. [^{11}C]N-propyl-apomorphine. A positron-labeled dopamine agonist for PET imaging of D2 receptors. *Nucl Med Biol* 2000;27:533–539.

43. Laruelle M, D'souza CD, Baldwin RM, et al. Imaging D-2 receptor occupancy by endogenous dopamine in humans. *Neuropsychopharmacology* 1997;17:162–174.

44. Fujita M, Verhoeff NP, Varrone A, et al. Imaging extrastriatal dopamine D(2) receptor occupancy by endogenous dopamine in healthy humans. *Eur J Pharmacol* 2000;387:179–188.

45. Abi-Dargham A, Rodenhiser J, Printz D, et al. Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. *Proc Natl Acad Sci USA* 2000;97:8104–8109.

46. Laruelle M, Abi-Dargham A, van Dyck C, et al. Dopamine and serotonin transporters in patients with schizophrenia: an imaging study with [(123)I]beta-CIT. *Biol Psychiatry* 2000;47:371–379.

47. Hirai M, Kitamura N, Hashimoto T, et al. [3H]GBR-12935 binding sites in human striatal membranes: binding characteristics and changes in parkinsonians and schizophrenics. *Jpn J Pharmacol* 1988;47:237–243.

48. Czudek C, Reynolds GP. [3H] GBR 12935 binding to the dopamine uptake site in post-mortem brain tissue in schizophrenia. *J Neural Transm* 1989;77:227–230.

49. Pearce RK, Seeman P, Jellinger K, et al. Dopamine uptake sites and dopamine receptors in Parkinson's disease and schizophrenia. *Eur Neurol* 1990;30(suppl 1):9–14.

50. Joyce JN, Lexow N, Bird E, et al. Organization of dopamine D1 and D2 receptors in human striatum:receptor autoradiographic studies in Huntington's disease and schizophrenia. *Synapse* 1988;2:546–557.

51. Chinaglia G, Alvarez FJ, Probst A, et al. Mesostriatal and mesolimbic dopamine uptake binding sites are reduced in Parkinson's disease and progressive supranuclear palsy: a quantitative autoradiographic study using [3H]mazindol. *Neuroscience* 1992;49:317–327.

52. Knable MB, Hyde TM, Herman MM, et al. Quantitative autoradiography of dopamine-D1 receptors, D2 receptors, and dopamine uptake sites in postmortem striatal specimens from schizophrenic patients. *Biol Psychiatry* 1994;36:827–835.

53. Weinberger DR, Berman KF, Zec RF. Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia: I. Regional cerebral blood flow evidence. *Arch Gen Psychiatry* 1986;43:114–124.

54. Grace AA. Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* 1991;41:1–24.

55. Carlsson A, Waters N, Carlsson ML. Neurotransmitter interactions in schizophrenia–therapeutic implications. *Biol Psychiatry* 1999;46:1388–1395.

56. Bunney BS, Aghajanian GK. d-Amphetamine-induced depression of central dopamine neurons: evidence for mediation by both autoreceptors and a striato-nigral feedback pathway. *Naunyn Schmiedebergs Arch Pharmacol* 1978;304:255–261.

57. Miller DW, Abercrombie ED. Effects of MK-801 on spontaneous and amphetamine-stimulated dopamine release in striatum measured with *in vivo* microdialysis in awake rats. *Brain Res Bull* 1996;40:57–62.

58. Kegeles L, Abi-Dargham A, Zea-Ponce Y, et al. Modulation of amphetamine-induced striatal dopamine release by ketamine in humans: implications for schizophrenia. *Biol Psychiatry* 2000;48:627–640.

59. Olney JW, Farber NB. Glutamate receptor dysfunction and schizophrenia. *Arch Gen Psychiatry* 1995;52:998–1007.

60. Jentsch JD, Roth RH. The neuropsychopharmacology of phencyclidine: from NMDA receptor hypofunction to the dopamine hypothesis of schizophrenia. *Neuropsychopharmacology* 1999;20:201–225.

61. Javitt DC, Zukin SR. Recent advances in the phencyclidine model of schizophrenia. *Am J Psychiatry* 1991;148:1301–1308.

62. Karreman M, Moghaddam B. The prefrontal cortex regulates the basal release of dopamine in the limbic striatum: an effect mediated by ventral tegmental area. *J Neurochem* 1996;66:589–598.

63. Bertolino A, Breier A, Callicott JH, et al. The relationship between dorsolateral prefrontal neuronal N-acetylaspartate and evoked release of striatal dopamine in schizophrenia. *Neuropsychopharmacology* 2000;22:125–132.

64. Saunders RC, Kolachana BS, Bachevalier J, et al. Neonatal lesions of the medial temporal lobe disrupt prefrontal cortical regulation of striatal dopamine. *Nature* 1998;393:169–171.

65. Bertolino A, Saunders RC, Mattay VS, et al. Altered development of prefrontal neurons in rhesus monkeys with neonatal mesial temporo-limbic lesions: a proton magnetic resonance spectroscopic imaging study. *Cereb Cortex* 1997;7:740–748.

66. Laruelle M. The role of endogenous sensitization in the pathophysiology of schizophrenia: implications from recent brain imaging studies. *Brain Res Rev* 2000;31:371–384.

67. Lieberman JA, Kinon BL, Loebel AD. Dopaminergic mechanisms in idiopathic and drug-induced psychoses. *Schizophr Bull* 1990;16:97–110.

68. Glenthoj BY, Hemmingsen R. Dopaminergic sensitization: implications for the pathogenesis of schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 1997;21:23–46.

69. Lieberman JA, Sheitman BB, Kinon BJ. Neurochemical sensitization in the pathophysiology of schizophrenia: deficits and dysfunction in neuronal regulation and plasticity. *Neuropsychopharmacology* 1997;17:205–229.

70. Robinson TE, Becker JB. Enduring changes in brain and behavior produced by chronic amphetamine administration: a review and evaluation of animal models of amphetamine psychosis. *Brain Res Rev* 1986;11:157–198.

71. Pierce RC, Kalivas PW. A circuitry model of the expression of behavioral sensitization to amphetamine-like psychostimulants. *Brain Res Rev* 1997;25:192–216.

72. Kalivas PW, Sorg BA, Hooks MS. The pharmacology and neural circuitry of sensitization to psychostimulants. *Behav Pharmacol* 1993;4:315–334.

73. Griffith JD, Cavanaugh J, Held J, et al. Dextroamphetamine:evaluation of psychomimetic properties in man. *Arch Gen Psychiatry* 1972;26:97–100.

74. O'Donnell P, Grace AA. Dysfunctions in multiple interrelated systems as the neurobiological bases of schizophrenic symptom clusters. *Schizophr Bull* 1998;24:267–283.

75. Grace AA, Moore H, O'Donnell P. The modulation of corticoaccumbens transmission by limbic afferents and dopamine: a model for the pathophysiology of schizophrenia. *Adv Pharmacol* 1998;42:721–724.

76. De Keyser J, Ebinger G, Vauquelin G. Evidence for a widespread dopaminergic innervation of the human cerebral neocortex. *Neurosci Lett* 1989;104:281–285.

77. Hall H, Sedvall G, Magnusson O, et al. Distribution of D1- and D2-dopamine receptors, and dopamine and its metabolites in the human brain. *Neuropsychopharmacology* 1994;11:245–256.

78. Bergson C, Mrzljak L, Smiley JF, et al. Regional, cellular, and subcellular variations in the distribution of D1 and D5 dopamine receptors in primate brain. *J Neurosci* 1995;15:7821–7836.

79. Smiley JF, Levey AI, Ciliax BJ, et al. D1 dopamine receptor immunoreactivity in human and monkey cerebral cortex: predominant and extrasynaptic localization in dendritic spines. *Proc Natl Acad Sci USA* 1994;91:5720–5724.

80. Krimer LS, Jakab RL, Goldman-Rakic PS. Quantitative three-dimensional analysis of the catecholaminergic innervation of identified neurons in the macaque prefrontal cortex. *J Neurosci* 1997;17:7450–7461.

81. Laruelle M, Casanova M, Weinberger D, et al. Postmortem study of the dopaminergic D1 receptors in the dorsolateral prefrontal cortex of schizophrenics and controls. *Schizophr Res* 1990;3:30–31.

82. Knable MB, Hyde TM, Murray AM, et al. A postmortem study of frontal cortical dopamine D1 receptors in schizophrenics, psychiatric controls, and normal controls. *Biol Psychiatry* 1996;40:1191–1199.

83. Meador-Woodruff JH, Haroutunian V, Powchik P, et al. Dopamine receptor transcript expression in striatum and prefrontal and occipital cortex. Focal abnormalities in orbitofrontal cortex in schizophrenia. *Arch Gen Psychiatry* 1997;54:1089–1095.

84. Okubo Y, Suhara T, Suzuki K, et al. Decreased prefrontal dopamine D1 receptors in schizophrenia revealed by PET. *Nature* 1997;385:634–636.

85. Karlsson P, Farde L, Halldin C, et al. D1-dopamine receptors in schizophrenia examined by PET. *Schizophr Res* 1997;24:179.

86. Abi-Dargham A, Martinez D, Mawlawi O, et al. Measurement of striatal and extrastriatal dopamine D1 receptor binding potential with [^{11}C]NNC 112 in humans: validation and reproducibility. *J Cereb Blood Flow Metab* 2000;20:225–243.

87. Abi-Dargham A, Simpson N, Kegeles L, et al. PET studies of binding competition between endogenous dopamine and the D1 radiotracer [^{11}C]NNC 756. *Synapse* 1999;32:93–109.

88. Chou YH, Karlsson P, Halldin C, et al. A PET study of D1-like dopamine receptor ligand binding during altered endogenous dopamine levels in the primate brain. *Psychopharmacology* 1999;146:220–227.

89. Lewis DA. GABAergic local circuit neurons and prefrontal cortical dysfunction in schizophrenia. *Brain Res Rev* 2000;31:270–6.

90. Benes FM. Emerging principles of altered neural circuitry in schizophrenia. *Brain Res Rev* 2000;31:251–269.

91. Busatto GF, Pilowsky LS, Costa DC, et al. Correlation between reduced *in vivo* benzodiazepine receptor binding and severity of psychotic symptoms in schizophrenia. *Am J Psychiatry* 1997;154:56–63.

92. Verhoeff NP, Soares JC, D'souza CD, et al. [^{123}I]Iomazenil SPECT benzodiazepine receptor imaging in schizophrenia. *Psychiatry Res* 1999;91:163–173.

93. Abi-Dargham A, Laruelle M, Krystal J, et al. No evidence of altered *in vivo* benzodiazepine receptor binding in schizophrenia. *Neuropsychopharmacology* 1999;20:650–661.

94. Schröder J, Bubeck B, Demisch S, et al. Benzodiazepine receptor distribution and diazepam binding in schizophrenia: an exploratory study. *Psychiatr Res Neuroimaging* 1997;68:125–131.

95. Ball S, Busatto GF, David AS, et al. Cognitive functioning and GABAA/benzodiazepine receptor binding in schizophrenia: a ^{123}I-iomazenil SPET study. *Biol Psychiatry* 1998;43:107–117.

96. Benes FM, Wickramasinghe R, Vincent SL, et al. Uncoupling of GABA(A) and benzodiazepine receptor binding activity in the hippocampal formation of schizophrenic brain. *Brain Res* 1997;755:121–129.

97. Abi-Dargham A, Krystal J. Serotonin receptors as target of antipsychotic medications. In: Lidow MS, eds. *Neurotransmitter receptors in actions of antipsychotic medications*. Boca Raton, FL: CRC Press LLC, 2000:79–107.

98. Lewis R, Kapur S, Jones C, et al. Serotonin 5-HT2 receptors in schizophrenia: a PET study using [18F]setoperone in neuroleptic-naive patients and normal subjects. *Am J Psychiatry* 1999;156:72–78.

99. Trichard C, Paillere-Martinot ML, Attar-Levy D, et al. No serotonin 5-HT2A receptor density abnormality in the cortex of schizophrenic patients studied with PET. *Schizophr Res* 1998;31:13–17.

100. Okubo Y, Suhara T, Suzuki K, et al. Serotonin 5-HT2 receptors in schizophrenic patients studied by positron emission tomography. *Life Sci* 2000;66:2455–2464.

101. Ngan ET, Yatham LN, Ruth TJ, et al. Decreased serotonin 2A receptor densities in neuroleptic-naive patients with schizophrenia: a PET study using [(18)F]setoperone. *Am J Psychiatry* 2000;157:1016–1018.

102. Kapur S, Zipursky RB, Remington G. Clinical and theoretical implications of 5-HT2 and D2 receptor occupancy of clozapine, risperidone, and olanzapine in schizophrenia. *Am J Psychiatry* 1999;156:286–293.

103. Nyberg S, Nilsson U, Okubo Y, et al. Implications of brain imaging for the management of schizophrenia. *Int Clin Psychopharmacol* 1998;13(suppl 3):S15–20.

104. Farde L, Nordström AL, Wiesel FA, et al. Positron emission tomography analysis of central D1 and D2 dopamine receptor occupancy in patients treated with classical neuroleptics and clozapine. *Arch Gen Psychiatry* 1992;49:538–544.

105. Wolkin A, Barouche F, Wolf AP, et al. Dopamine blockade and clinical response: evidence for two biological subgroups of schizophrenia. *Am J Psychiatry* 1989;146:905–908.

106. Pilowsky LS, Costa DC, Ell PJ, et al. Clozapine, single photon emission tomography, and the D2 dopamine receptor blockade hypothesis of schizophrenia. *Lancet* 1992;340:199–202.

107. Nordstrom AL, Farde L, Wiesel FA, et al. Central D2-dopamine receptor occupancy in relation to antipsychotic drug effects: a double-blind PET study of schizophrenic patients. *Biol Psychiatry* 1993;33:227–235.

108. Kapur S, Zipursky R, Jones C, et al. Relationship between dopamine D(2) occupancy, clinical response, and side effects: a double-blind PET study of first-episode schizophrenia. *Am J Psychiatry* 2000;157:514–520.

109. Nordstrom AL, Farde L, Nyberg S, et al. D1, D2, and 5-HT2 receptor occupancy in relation to clozapine serum concentration: a PET study of schizophrenic patients [see comments]. *Am J Psychiatry* 1995;152:1444–1449.

110. Nyberg S, Farde L, Eriksson L, et al. 5-HT2 and D2 dopamine receptor occupancy in the living human brain. A PET study with risperidone. *Psychopharmacology* 1993;110:265–672.

111. Kapur S, Remington G, Zipursky RB, et al. The D2 dopamine receptor occupancy of risperidone and its relationship to extrapyramidal symptoms: a PET study. *Life Sci* 1995;57:L103–107.

112. Knable MB, Heinz A, Raedler T, et al. Extrapyramidal side effects with risperidone and haloperidol at comparable D2 receptor occupancy levels. *Psychiatr Res Neuroimag* 1997;75:91–101.

113. Kapur S, Zipursky RB, Remington G, et al. 5-HT2 and D2 receptor occupancy of olanzapine in schizophrenia: a PET investigation. *Am J Psychiatry* 1998;155:921–8.

114. Gefvert O, Bergstrom M, Langstrom B, et al. Time course of central nervous dopamine-D2 and 5-HT2 receptor blockade and plasma drug concentrations after discontinuation of quetiapine (Seroquel) in patients with schizophrenia. *Psychopharmacology (Berl)* 1998;135:119–126.

115. Kapur S, Zipursky R, Jones C, et al. A positron emission tomography study of quetiapine in schizophrenia: a preliminary finding of an antipsychotic effect with only transiently high dopamine D2 receptor occupancy. *Arch Gen Psychiatry* 2000;57:553–559.

116. Pilowsky LS, Mulligan RS, Acton PD, et al. Limbic selectivity of clozapine. *Lancet* 1997;350:490–491.

117. Farde L, Suhara T, Nyberg S, et al. A PET study of [C-11]FLB 457 binding to extrastriatal D-2-dopamine receptors in healthy subjects and antipsychotic drug-treated patients. *Psychopharmacology* 1997;133:396–404.

118. Bigliani V, Mulligan RS, Acton PD, et al. *In vivo* occupancy of striatal and temporal cortical D2/D3 dopamine receptors by typical antipsychotic drugs. [^{123}I]epidepride single photon emission tomography (SPET) study. *Br J Psychiatry* 1999;175:231–238.

119. Abi-Dargham A, Hwang DR, Huang Y, et al. Reliable quantification of both striatal and extrastriatal D2 receptors in humans with [18F]fallypride. *J Nucl Med* 2000;41:139P.

120. Mukherjee J, Christian BT, Narayanan TK, et al. Measurement of striatal and extrastriatal D-2 receptor occupancy by clozapine and risperidone using [18F]fallypride and PET. *Neuroimage* 2000;11:S53.

121. Robertson GS, Matsumura H, Fibiger HC. Induction patterns of Fos-like immunoreactivity in the forebrain as predictors of atypical antipsychotic activity. *J Pharmacol Exp Ther* 1994;271:1058–1066.

122. Deutch AY, Lee MC, Iadarola MJ. Regionally specific effects of atypical antipsychotic drugs on striatal fos expression: the nucleus accumbens shell as a locus of antipsychotic action. *Mol Cell Neurosci* 1992;3:332–341.

123. Martinez D, Hwang DR, Broft A, et al. PET imaging of amphetamine-induced endogenous dopamine release in mesolimbic and nigro-striatal dopamine systems in humans. *J Nucl Med* 2000;41:105P.

124. Drevets WC, Price JC, Kupfer DJ, et al. PET measures of amphetamine-induced dopamine release in ventral versus dorsal striatum. *Neuropsychopharmacology* 1999;21:694–709.

125. Wolkin A, Brodie JD, Barouche F, et al. Dopamine receptor occupancy and plasma haloperidol levels. *Arch Gen Psychiatry* 1989;46:482–484.

126. Fitzgerald PB, Kapur S, Remington G, et al. Predicting haloperidol occupancy of central dopamine D2 receptors from plasma levels. *Psychopharmacology (Berl)* 2000;149:1–5.

127. Hagberg G, Gefvert O, Bergstrom M, et al. N-[^{11}C]methylspiperone PET, in contrast to [^{11}C]raclopride, fails to detect D2 receptor occupancy by an atypical neuroleptic. *Psychiatry Res* 1998;82:147–160.

128. Lourenco CM, DaSilva JN, Warsh JJ, et al. Imaging of cAMP-specific phosphodiesterase-IV:comparison of [^{11}C]rolipram and [^{11}C]Ro 20-1724 in rats. *Synapse* 1999;31:41–50.

129. Suhara T, Inoue O, Kobayashi K, et al. An acute effect of triazolam on muscarinic cholinergic receptor binding in the human brain measured by positron emission tomography. *Psychopharmacology (Berl)* 1994;113:311–317.

130. Ding YS, Logan J, Bermel R, et al. Dopamine receptor-mediated regulation of striatal cholinergic activity: positron emission tomography studies with norchloro[18F]fluoroepibatidine. *J Neurochem* 2000;74:1514–1521.

131. Kung MP, Stevenson DA, Plossl K, et al. [99mTc]TRODAT-1: a novel technetium-99m complex as a dopamine transporter imaging agent. *Eur J Nucl Med* 1997;24:372–380.

132. Lundbom N, Barnett A, Bonavita S, et al. MR image segmentation and tissue metabolite contrast in 1H spectroscopic imaging of normal and aging brain. *Magn Reson Med* 1999;41:841–845.

133. Spielman DM, Adalsteinsson E, Lim KO. Quantitative assessment of improved homogeneity using higher-order shims for spectroscopic imaging of the brain. *Magn Reson Med* 1998;40:376–382.

134. Varho T, Komu M, Sonninen P, et al. A new metabolite contributing to N-acetyl signal in 1H MRS of the brain in Salla disease [published erratum appears in *Neurology* 1999;53(5):1162]. *Neurology* 1999;52:1668–1672.

135. Salibu NaMAB. *Clinical MR spectroscopy*. New York: Wiley-Liss, 1998.

136. Gunther H. *NMR spectroscopy: basic principle, concepts and applications in chemistry*, second ed. New York: Wiley, 1995.

137. Gadian DG. *NMR and its application to living systems*, second ed. New York: Oxford University Press, 1995.

138. Bottomley PA. Human *in vivo* NMR spectroscopy in diagnostic medicine: clinical tool or research probe? *Radiology* 1989;170:1–15.

139. Murphy DG, Bottomley PA, Salerno JA, et al. An *in vivo* study of phosphorus and glucose metabolism in Alzheimer's disease using magnetic resonance spectroscopy and PET. *Arch Gen Psychiatry* 1993;50:341–349.

140. O'Callaghan E, Redmond O, Ennis R, et al. Initial investigation of the left temporoparietal region in schizophrenia by 31P magnetic resonance spectroscopy. *Biol Psychiatry* 1991;29:1149–1152.

141. Pettegrew JW, Keshavan MS, Panchalingam K, et al. Alterations in brain high-energy phosphate and membrane phospholipid metabolism in first-episode, drug-naive schizophrenics. A pilot study of the dorsal prefrontal cortex by *in vivo* phosphorus 31 nuclear magnetic resonance spectroscopy [see comments]. *Arch Gen Psychiatry* 1991;48:563–568.

142. Pettegrew JW, Keshavan MS, Minshew NJ. 31P nuclear magnetic resonance spectroscopy: neurodevelopment and schizophrenia. *Schizophr Bull* 1993;19:35–53.

143. Kegeles LS, Shungu DC, Anjilvel S, et al. Hippocampal pathology in schizophrenia: magnetic resonance imaging and spectroscopy studies. *Psychiatry Res* 2000;98:163–175.

144. Keshevan MS, Stanley JA, Pettegrew J. Magnetic resonance spectroscopy in schizophrenia: methodological issues and findings-Part II. *Biol Psychiatry* 2000;48:369–380.

145. Kato Tea. Phosphorus-31 magnetic resonance spectroscopy and ventricular enlargement in bipolar disorder. *Psychiatry Research:Neuroimaging* 1994;55:41–50.

146. Volz HR, Riehemann S, Maurer I, et al. Reduced phosphodiesters and high-energy phosphates in the frontal lobe of schizophrenic patients: a (31)P chemical shift spectroscopic-imaging study. *Biol Psychiatry* 2000;47:954–961.

147. Volz HP, Hubner G, Rzanny R, et al. High-energy phosphates in the frontal lobe correlate with Wisconsin Card Sort Test performance in controls, not in schizophrenics: a 31-phosphorus magnetic resonance spectroscopic and neuropsychological investigation. *Schizophr Res* 1998;31:37–47.

148. Potwarka JJ, Drost DJ, Williamson PC, et al. A 1H-decoupled 31P chemical shift imaging study of medicated schizophrenic patients and healthy controls. *Biol Psychiatry* 1999;45:687–693.

149. Bluml S, Tan J, Harris K, et al. Quantitative proton-decoupled 31P MRS of the schizophrenic brain *in vivo*. *J Comput Assist Tomogr* 1999;23:272–275.

150. Deicken RF, Merrin EL, Floyd TC, et al. Correlation between left frontal phospholipids and Wisconsin Card Sort Test performance in schizophrenia. *Schizophr Res* 1995;14:177–181.

151. Stanley JA, Williamson PC, Drost DJ, et al. An *in vivo* study of the prefrontal cortex of schizophrenic patients at different stages of illness via phosphorus magnetic resonance spectroscopy. *Arch Gen Psychiatry* 1995;52:399–406.

152. van der Veen JW, et al. Proton MR spectroscopic imaging without water suppression. *Radiology* 2000;217:296–300.

153. Urenjak J, Williams SR, Gadian DG, et al. Proton nuclear magnetic resonance spectroscopy unambiguously identifies different neural cell types. *J Neurosci* 1993;13:981–989.

154. Moffett JR, Namboodiri MA. Differential distribution of N-acetylaspartylglutamate and N-acetylaspartate immunoreactivities in rat forebrain. *J Neurocytol* 1995;24:409–433.

155. Tallan HH. Studies on the distribution of N-acetyl-aspartic acid in brain. *J Biol Chem* 1957;224:41–45.

156. Kreis DG. Absolute quantitation of water and metabolites in the human brain. II. Metabolite concentrations. *J Magn Reson* 1993;102:9–19.

157. Huppi PS, Posse S, Lazeyras F, et al. Magnetic resonance in preterm and term newborns: 1H-spectroscopy in developing human brain. *Pediatr Res* 1991;30:574–578.

158. Hashimoto T, Tayama M, Miyazaki M, et al. Developmental brain changes investigated with proton magnetic resonance spectroscopy. *Dev Med Child Neurol* 1995;37:398–405.

159. Clark JB. N-acetyl aspartate: a marker for neuronal loss or mitochondrial dysfunction. *Dev Neurosci* 1998;20:271–276.

160. Bhakoo KK, Pearce D. *In vitro* expression of N-acetyl aspartate by oligodendrocytes: implications for proton magnetic resonance spectroscopy signal *in vivo*. *J Neurochem* 2000;74:254–262.

161. De Stefano N, et al. Chemical pathology of acute demyelinating lesions and its correlations with disability. *Ann Neurol* 1995;38:901–909.

162. Hugg JW, Kuzniecky RI, Gilliam FG, et al. Normalization of contralateral metabolic function following temporal lobectomy demonstrated by 1H magnetic resonance spectroscopic imaging. *Ann Neurol* 1996;40:236–239.

163. Jenkins BG, Klivenyi P, Kustermann E, et al. Nonlinear decrease over time in N-acetyl aspartate levels in the absence of neuronal loss and increases in glutamine and glucose in transgenic Huntington's disease mice. *J Neurochem* 2000;74:2108–2119.

164. Doraiswamy PM, Charles HC, Krishnan KR. Prediction of cognitive decline in early Alzheimer's disease [letter]. *Lancet* 1998;352:1678.

165. Jenkins BG, Rosas HD, Chen YC, et al. 1H NMR spectroscopy studies of Huntington's disease: correlations with CAG repeat numbers. *Neurology* 1998;50:1357–1365.

166. Vion-Dury J, Nicoli F, Salvan AM, et al. Reversal of brain metabolic alterations with zidovudine detected by proton localised magnetic resonance spectroscopy [letter]. *Lancet* 1995;345:60–61.

167. Bertolino A, et al. The effect of treatment with antipsychotic drugs on brain N-acetylaspartate measures in patients with schizophrenia. *Biol Psychiatry* 2001;49:39–46.

168. Roffman JL, Lipska BK, Bertolino A, et al. Local and downstream effects of excitotoxic lesions in the rat medial prefrontal cortex on *In vivo* 1H-MRS signals. *Neuropsychopharmacology* 2000;22:430–439.

169. Bertolino A, Saunders RC, Mattay VS, et al. Altered development of prefrontal neurons in rhesus monkeys with neonatal mesial temporo-limbic lesions: a proton magnetic resonance spectroscopic imaging study. *Cerebral Cortex* 1997;7:740–748.

170. Bertolino A, Roffman JL, Lipska BK, et al. Postpubertal emergence of N-acetylaspartate decreases in prefrontal cortex of rats with neonatal hippocampal damage. *Submitted*.

171. Nasrallah HA, Skinner TE, Schmalbrock P, et al. Proton magnetic resonance spectroscopy (1H MRS) of the hippocampal formation in schizophrenia: a pilot study. *Br J Psychiatry* 1994;165:481–485.

172. Bertolino A, Weinberger DR. Proton magnetic resonance spectroscopy in schizophrenia. *Eur J Radiol* 1999;30:132–141.

173. Bertolino A, Callicott JH, Nawroz S, et al. Reproducibility of proton magnetic resonance spectroscopic imaging in patients with schizophrenia. *Neuropsychopharmacology* 1998;18:1–9.

174. Bertolino A, Nawroz S, Mattay VS, et al. Regionally specific pattern of neurochemical pathology in schizophrenia as assessed by multislice proton magnetic resonance spectroscopic imaging. *Am J Psychiatry* 1996;153:1554–1563.

175. Bertolino A, Callicott JH, Elman I, et al. Regionally specific neuronal pathology in untreated patients with schizophrenia: a proton magnetic resonance spectroscopic imaging study. *Biol Psychiatry* 1998;43:641–648.

176. Bertolino A, Kumra S, Callicott JH, et al. Common pattern of cortical pathology in childhood-onset and adult-onset schizophrenia as identified by proton magnetic resonance spectroscopic imaging. *Am J Psychiatry* 1998;155:1376–1383.

177. Callicott JH, Egan MF, Bertolino A, et al. Hippocampal N-acetyl aspartate in unaffected siblings of patients with schizophrenia: a possible intermediate neurobiological phenotype [published erratum appears in *Biol Psychiatry* 1999;45(2):following 244]. *Biol Psychiatry* 1998;44:941–950.

178. Deicken RF, Zhou L, Corwin F, et al. Decreased left frontal lobe N-acetylaspartate in schizophrenia. *Am J Psychiatry* 1997;154:688–690.

179. Deicken RF, Pegues M, Amend D. Reduced hippocampal N-acetylaspartate without volume loss in schizophrenia. *Schizophr Res* 1999;37:217–223.

180. Deicken RF, Zhou L, Schuff N, et al. Proton magnetic resonance spectroscopy of the anterior cingulate region in schizophrenia. *Schizophr Res* 1997;27:65–71.

181. Deicken RF, Johnson C, Eliaz Y, et al. Reduced concentrations of thalamic N-acetylaspartate in male patients with schizophrenia. *Am J Psychiatry* 2000;157:644–647.

182. Lim KO, Adalsteinsson E, Spielman D, et al. Proton magnetic resonance spectroscopic imaging of cortical gray and white matter in schizophrenia. *Arch Gen Psychiatry* 1998;55:346–352.

183. Bartha R, Williamson PC, Drost DJ, et al. Measurement of glutamate and glutamine in the medial prefrontal cortex of never-treated schizophrenic patients and healthy controls by proton magnetic resonance spectroscopy. *Arch Gen Psychiatry* 1997;54:959–965.

184. Stanley JA, Williamson PC, Drost DJ, et al. An *in vivo* proton magnetic resonance spectroscopy study of schizophrenia patients. *Schizophr Bull* 1996;22:597–609.

185. Bartha R, al-Semaan YM, Williamson PC, et al. A short echo proton magnetic resonance spectroscopy study of the left mesial-temporal lobe in first-onset schizophrenic patients. *Biol Psychiatry* 1999;45:1403–1411.

186. Heimberg C, Komoroski RA, Lawson WB, et al. Regional proton magnetic resonance spectroscopy in schizophrenia and exploration of drug effect. *Psychiatry Res* 1998;83:105–115.

187. Kegeles LS, Humaran TJ, Mann JJ. *In vivo* neurochemistry of the brain in schizophrenia as revealed by magnetic resonance spectroscopy. *Biol Psychiatry* 1998;44:382–398.

188. Rakow RL, et al. N-acetylaspartate reduction in dorsolateral prefrontal cortex of patients with schizophrenia as revealed by short-echo time 1H-MRS. *Biol Psychiatry* 1999;45:121S.

189. Weinberger DR. Cell biology of the hippocampal formation in schizophrenia. *Biol Psychiatry* 1999;45:395–402.

190. Moore GJ, Bebchuk JM, Hasanat K, et al. Lithium increases N-acetyl-aspartate in the human brain: *in vivo* evidence in support of bcl-2's neurotrophic effects? [In Process Citation]. *Biol Psychiatry* 2000;48:1–8.

191. Rango M, Spagnoli D, Tomei G, et al. Central nervous system trans-synaptic effects of acute axonal injury: a 1H magnetic resonance spectroscopy study. *Magn Reson Med* 1995;33:595–600.

192. Bertolino A, Roffman JL, Lipska BK, et al. Postpubertal emergence of prefrontal neuronal deficits and altered dopaminergic behaviors in rats with neonatal hippocampal lesions. *Soc Neurosci* 1999.

193. Weinberger DR, Lipska BK. Cortical maldevelopment, anti-psychotic drugs, and schizophrenia: a search for common ground. *Schizophr Res* 1995;16:87–110.

194. Bertolino A, Knable MB, Saunders RC, et al. The relationship between dorsolateral prefrontal N-acetylaspartate measures and striatal dopamine activity in schizophrenia [see comments]. *Biol Psychiatry* 1999;45:660–667.

195. Bertolino A, Esposito G, Callicott JH, et al. Specific relationship between prefrontal neuronal N-acetylaspartate and activation of the working memory cortical network in schizophrenia [see comments]. *Am J Psychiatry* 2000;157:26–33.

196. Callicott JH, Bertolino A, Matley VS, et al. Physiological dysfunction of the dorsolateral prefrontal cortex in schizophrenia revisited. *Cerebral Cortex* 2000;10:1078–1092.

197. Callicott JH, Bertolino A, Egan MF, et al. A selective relationship between prefrontal N-acetylaspartate measures and negative symptoms in schizophrenia. *Am J Psychiatry* 2000;157:1646–1651.

198. Block W, Bayer TA, Tepest R, et al. Decreased frontal lobe ratio of N-acetyl aspartate to choline in familial schizophrenia: a proton magnetic resonance spectroscopy study [In Process Citation]. *Neurosci Lett* 2000;289:147–151.

199. Mattson RH, Petroff O, Rothman D, et al. Vigabatrin: effects on human brain GABA levels by nuclear magnetic resonance spectroscopy. *Epilepsia* 1994;35:S29–32.

200. Sibson NR, Shen J, Mason GF, et al. Functional energy metabolism: *in vivo* 13C-NMR spectroscopy evidence for coupling of cerebral glucose consumption and glutamatergic neuronal activity. *Dev Neurosci* 1998;20:321–330.

201. Blin J, Baron JC, Cambon H, et al. Striatal dopamine D2 receptors in tardive dyskinesia: PET study. *J Neurol Neurosurg Psychiatry* 1989;52:1248–1252.

202. Martinot J-L, Peron-Magnan P, Huret J-D, et al. Striatal D2 dopaminergic receptors assessed with positron emission tomography and 76-Br-bromospiperone in untreated patients. *Am J Psychiatry* 1990;147:346–350.

203. Tune LE, Wong DF, Pearlson G, et al. Dopamine D2 receptor density estimates in schizophrenia: a positron emission tomography study with ^{11}C-N-methylspiperone. *Psychiatry Research* 1993;49:219–237.

204. Nordstrom AL, Farde L, Eriksson L, et al. No elevated D2 dopamine receptors in neuroleptic-naive schizophrenic patients revealed by positron emission tomography and [^{11}C]N-methylspiperone [see comments]. *Psychiatry Res* 1995;61:67–83.

205. Farde L, Wiesel F, Stone-Elander S, et al. D2 dopamine receptors in neuroleptic-naive schizophrenic patients. A positron emission tomography study with [^{11}C]raclopride. *Arch Gen Psychiatry* 1990;47:213–219.

206. Hietala J, Syv_lahti E, Vuorio K, et al. Striatal D2 receptor characteristics in neuroleptic-naive schizophrenic patients studied with Positron Emission Tomography. *Arch Gen Psychiatry* 1994;51:116–123.

207. Pilowsky LS, Costa DC, Ell PJ, et al. D2 dopamine receptor binding in the basal ganglia of antipsychotic-free schizophrenic patients. An I-123-IBZM single photon emission computerized tomography study. *Br J Psychiatry* 1994;164:16–26.

208. Knable MB, Egan MF, Heinz A, et al. Altered dopaminergic function and negative symptoms in drug-free patients with schizophrenia. [^{123}I]-iodobenzamide SPECT study. *Br J Psychiatry* 1997;171:574–577.

209. Martinot Jl, Paill_re-Martinot ML, Loc'h C, et al. The estimated density of D2 striatal receptors in schizophrenia. A study with positron Emission tomography and 76Br-bromolisuride. *Br J Psychiatry* 1991;158:346–350.

210. Martinot JL, Paillère-Martinot ML, Loch'H C, et al. Central D2 receptors and negative symptoms of schizophrenia. *Br J Pharmacol* 1994;164:27–34.

211. Laakso A, Vilkman H, Alakare B, et al. Striatal dopamine transporter binding in neuroleptic-naive patients with schizophrenia studied with positron emission tomography. *Am J Psychiatry* 2000;157:269–271.

212. Choe BY, Kim KT, Suh TS, et al. 1H magnetic resonance spectroscopy characterization of neuronal dysfunction in drug-naive, chronic schizophrenia. *Acad Radiol* 1994;1:211–216.

213. Renshaw PF, Yurgelun-Todd DA, Tohen M, et al. Temporal lobe proton magnetic resonance spectroscopy of patients with first-episode psychosis. *Am J Psychiatry* 1995;152:444–446.

214. Yurgelun-Todd DA, Renshaw PF, Gruber SA, et al. Proton magnetic resonance spectroscopy of the temporal lobes in schizophrenics and normal controls. *Schizophr Res* 1996;19:55–59.

215. Buckley PF, Moore C, Long H, et al. 1H-magnetic resonance spectroscopy of the left temporal and frontal lobes in schizophrenia: clinical, neurodevelopmental, and cognitive correlates. *Biol Psychiatry* 1994;36:792–800.

216. Maier M, Ron MA, Barker GJ, et al. Proton magnetic resonance spectroscopy: an *in vivo* method of estimating hippocampal neuronal depletion in schizophrenia [published erratum appears in *Psychol Med* 1996;26(4):877]. *Psychol Med* 1995;25:1201–1209.

217. Maier M, Ron MA. Hippocampal age-related changes in schizophrenia: a proton magnetic resonance spectroscopy study. *Schizophr Res* 1996;22:5–17.

218. Fukuzako H, Takeuchi K, Hokazono Y, et al. Proton magnetic resonance spectroscopy of the left medial temporal and frontal lobes in chronic schizophrenia: preliminary report. *Psychiatry Res* 1995;61:193–200.

219. Cecil KM, Lenkinski RE, Gur RE, et al. Proton magnetic resonance spectroscopy in the frontal and temporal lobes of neuroleptic naive patients with schizophrenia. *Neuropsychopharmacology* 1999;20:131–140.

220. Brooks WM, Hodde-Vargas J, Vargas LA, et al. Frontal lobe of children with schizophrenia spectrum disorders: a proton magnetic resonance spectroscopic study. *Biol Psychiatry* 1998;43:263–269.

221. Thomas MA, Ke Y, Levitt J, et al. Preliminary study of frontal lobe 1H MR spectroscopy in childhood-onset schizophrenia. *J Magn Reson Imaging* 1998;8:841–846.

222. Fukuzako H, Kodama S, Fukuzako T, et al. Subtype-associated metabolite differences in the temporal lobe in schizophrenia detected by proton magnetic resonance spectroscopy. *Psychiatry Res* 1999;92:45–56.

Section VII
Anxiety and stress disorder

Dennis Charney

Dennis Charney: Mood and Anxiety Disorder Research Program, National Institute of Mental Health, Bethesda, Maryland.

Anxiety and stress disorder - Introduction

The chapters in this section reflect the advances that have been made in our understanding of the molecular neurobiology and neural circuits underlying fear and anxiety states and the potential for badly needed advances in the diagnosis, pathophysiology, and treatment of anxiety disorders.

Many readers will be surprised by the data cited in the Kessler and Greenberg chapter on the economic burden of anxiety and stress disorders. Recent surveys demonstrate that anxiety and stress disorders are the most commonly occurring of all mental disorders. Anxiety disorders are temporally the primary disorders in many people with a lifetime history of any mental disorder. Indeed the combined occurrence of high lifetime prevalence with an early age at onset and high chronicity makes anxiety disorders unique.

The chapter by Stein and Lang reviews in detail the course of anxiety and stress disorders over the lifetime. One point they make that is especially important to emphasize is the critical need for further research in childhood anxiety disorders. The presence of an anxiety disorder in childhood or adolescence is a predictor of the persistence into adulthood of not only anxiety disorders but other psychiatric disorders as well, particularly depression. It is not known if early effective treatment of childhood and adolescent anxiety disorders will prevent the development of psychiatric disorders later in life.

Important insights into the pathogenesis of anxiety disorders come from preclinical investigations reviewed by Davis in his chapter. The identification of the brain structures and neural circuits involved in the generation of fear and anxiety-related behaviors are most noteworthy. The delineation of specific neural pathways mediating conditioned and unconditioned fear can logically guide the design and conduct of clinical studies investigating the neurobiological mechanisms of anxiety disorders such as generalized anxiety disorder (unconditional fear) and phobic disorders (conditioned fear). Increased knowledge of the neural mechanisms and neurotransmitters involved in extinction may offer novel therapeutic approaches. If we can discover pharmacologic methods to enhance extinction, more effective drug treatments for conditioned fear or anxiety may be found. There is evidence of extinction being mediated by γ-aminobutyric acid (GABA) release. This suggests the possibility of augmenting extinction-based psychotherapy with GABA agonist drug treatment.

Determination of the genetic contribution to anxiety disorders is critically important to progress in understanding etiology and improving treatment.

Merikangas and Pine review the evidence that the major subtypes of the anxiety disorders aggregate in families. However, the magnitude of heritability is relatively moderate, indicating a strong environmental contribution to etiology. Unfortunately, thus far linkage and association studies in anxiety disorders have not been fruitful. Future studies should determine if components of anxiety syndromes are controlled by specific genes. The revolution occurring in genomics research as a consequence of sequencing the human genome and the identification of over 1½ million single nucleotide polymorphisms (SNPs) should make discovery of anxiety-related disease genes a reality. Strategies that are complementary to linkage analyses and utilize data from linkage studies are indicated. Such approaches couple the genotyping of candidate functional SNPs with linkage and equilibrium mapping in chromosomal regions implicated in linkage studies. In parallel with the genetic studies, enhanced efforts to identify endophenotypic biological vulnerability markers are indicated. The studies cited in the chapter on temperament, anxiety sensitivity, autonomic reactivity, psychophysiologic function, ventilatory function, neurochemical, and neuroendocrine factors are good examples of this approach.

Bakshi and Kalin point out in their chapter the advantages

of using putative animal endophenotypes of stress and anxiety to identify genetic abnormalities associated with anxiety disorders. Rodent and nonhuman primate studies of mother-infant interactions are particularly compelling, given the important clinical implications if these interactions are found to be a critical factor in future fearful disposition. Targeted mutations leading to anxiety-like endophenotypes in transgenic mice have suggested roles for serotonin receptor subtype 1A ($5\text{-}HT_{1A}$), corticotropin-releasing hormone (CRH), GABA, neuropeptide Y, cholecystokinin, and substance P neural systems in the generation on anxiety-fear behaviors. These studies provide clues for the discovery of new medications and pharmacogenomic approaches to treatment. Accelerated drug development efforts focusing on corticotropin-releasing factor 1 (CRF-1) receptor antagonists and benzodiazepine agonists with an anxioselective subunit profile are indicated. The feasibility of pharmacogenomic investigations designed to evaluate the relationships among functional polymorphisms of the $5\text{-}HT_{1A}$ receptor, benzodiazepine receptor, and GABA synthesis enzymes and therapeutic responses to specific drugs should be explored.

It is imperative that these findings from the preclinical studies of anxiety and fear states be translated into increased knowledge of the neural circuits and associated neural mechanisms that can account for the signs and symptoms in patients with anxiety disorders. Rauch and Shin reviewed the neuroimaging findings relevant to anxiety and stress disorders. Their chapter emphasizes the areas of congruence between animal studies and clinical neuroimaging investigations. For example, imaging studies in healthy subjects support a role for the amygdala in fear conditioning and the frontal cortex in extinction. In the imaging studies of patients with anxiety disorders, progress has been made in identifying specific neural circuits. Functional relationships among the amygdala, hippocampus, and medial prefrontal cortex have been reported in patients with posttraumatic stress disorder (PTSD). The imaging findings in patients with obsessive-compulsive disorder (OCD) are consistent with “pathology” in cortico-striatal-thalamo-cortical circuitry. Unfortunately, a critical gap in knowledge exists regarding the relevant neural circuits involved in panic disorder, social anxiety disorder, and generalized anxiety disorder. The neurochemical systems associated with anxiety and fear circuits are reviewed in the chapter by Charney and Drevets. As predicted from the preclinical studies, abnormalities in norepinephrine, benzodiazepine, glucocorticoid, and CRH systems have been identified in patients with anxiety disorders. However, most of the findings reviewed should be deemed preliminary, and they require replication. None of the reported neurobiological distinctions between patients and controls is robust enough to be of diagnostic relevance.

Tallman, Cassella, and Kehne review the mechanism of action of anxiolytic drugs and the status of new and novel therapeutic agents. They highlight the therapeutic potential and current status of CRH antagonist drug development. They also note the potential of developing targets for the CRF-2 receptor and other peptides, such as vasoactive intestinal peptide (VIP), involved in the regulation of stress. In regard to benzodiazepine drug development, they note an ideal drug might have limited effects on the α_1 subtype with increased responsiveness at α_2 and α_3 subunits. Glutamate receptor agonists and modulators are proposed as viable targets for anxiety disorders. For example, point mutations in the glycine-binding site of the NR1 subunit result in mice that have reduced glycine affinity and an anxiolytic profile. Group II metabotrophic glutamate agonists are in early clinical development for the treatment of anxiety disorders. Other novel drug targets include antagonists of AMPA receptors and antagonists of strychnine-sensitive glycine site, both of which show anxiolytic profiles in animal models.

Ultimately, the goal of research on the neurobiological underpinnings of anxiety disorders is to lead to more effective, more rapidly acting treatments, to achieve a more complete response, and to be able to predict responses to specific treatments. In their review, Gorman, Kent, and Coplan highlight the extremely broad spectrum of action of norepinephrine and serotonin transport inhibitors in anxiety disorders. These drugs are limited by a delayed onset and in complete response in many patients. This suggests that norepinephrine and serotonin have broad modulatory effects on other neuronal systems, which are more primarily related to the pathogenesis of anxiety disorders.

In summary, a reading of these chapters reveals that there have not been fundamental advances in our ability to diagnose anxiety disorders based on known etiology. The mainstay of medication treatment continues to be classes of medications that have existed for decades. There are major gaps in our knowledge of anxiety disorders in children. The mechanisms responsible for the occurrence of anxiety disorders in childhood and adolescence leading to increased risk for other psychiatric disorders in adulthood are unknown. However, the potential for progress is great. A multidisciplinary team approach utilizing the findings from preclinical investigations on neural circuitry, neurochemistry, and genetics to inform clinical investigations of genetic vulnerability, environmental risk factors, neuroimaging, pharmacogenomics, and novel drug design and testing will be a pathway to discovery.

60

Anxiety and Stress Disorders: Course Over the Lifetime

Murray B. Stein

Ariel J. Lang

Murray B. Stein and Ariel J. Lang: Department of Psychiatry, University of California–San Diego, San Diego, California.

This chapter traces the course and trajectory of anxiety disorders across the life span. The chapter discusses these disorders in three age groups: (a) childhood, (b) young adulthood to middle age, and (c) older adulthood.

ANXIETY AND STRESS DISORDERS IN CHILDHOOD

Part of "60 - Anxiety and Stress Disorders: Course Over the Lifetime "

Temperamental and Environmental Precursors

There are a number of factors that relate to the development of anxiety in general. Although this remains a controversial area, current evidence suggests that anxiety does not appear to be specifically heritable; what clusters in families is a more general predisposition to mood or anxiety disorders. Evidence of this comes from twin studies (1) as well as from studies of the incidence of disorders in the families of anxious patients (2). Specific phobia, however, may be an exception to this general heritability; family members with specific phobias tend to be associated with increased risk only for specific phobias (3).

Any biological predisposition is likely moderated by environmental factors. For example, Beidel and Turner (2) found that parental psychopathology was a risk factor for the development of disorder only among the lower socioeconomic status (STS) portion of their sample. It has been suggested that environmental factors play a significant role in the manifestation of specific psychopathology (1). Anxiety in particular is believed to be related to a combination of negative affect, a sense of lack of control over situations or environments, and attentional self-focus. Early experiences of being unable to influence or control situations, therefore, may lead to the development of anxiety (4).

Specific patterns within the family may lead to increased risk of development of childhood anxiety. Expressed emotion (EE), which is an interactional style composed of emotional overinvolvement and high levels of criticism, has been associated with an increased likelihood of childhood anxiety disorders (5) and with long-term functioning in those with anxiety disorders (6). Parents may also influence a child's anxiety by providing positive (e.g., the child gets attention) or negative (e.g., the child is allowed to avoid anxiety-provoking situations) reinforcement for expressions of anxiety, by providing inadequate affection and by excessive control/overprotection (4).

Can it be said that some people are born anxious? The answer is a qualified yes. The last decade or so of research by Jerome Kagan and colleagues has led to the identification of a temperament, “behavioral inhibition” (BI), that appears to be related to the subsequent onset of anxiety disorders. BI involves reacting to unfamiliar or novel situations with behavioral restraint and physiologic arousal (5). When confronted with an unfamiliar person or object, a BI child will withdraw, cling, be reluctant to interact, show emotional distress, and stop other activities. BI has also been associated with physiologic differences, such as high and stable heart rate, increased salivary cortisol and urinary catecholamines, pupillary dilation, and laryngeal muscle tension (7). These findings have led to the hypothesis that BI is related to a low threshold for arousal in the amygdala and hypothalamus (8). This characteristic, which appears to be present in 10% to 15% of children, has been identified in children as young as 14 months, has been shown to persist throughout childhood (9), and is more commonly found in offspring of anxious parents (8). The inhibited temperament has been associated with risk of developing an anxiety disorder, most commonly social phobia (10).

Some have suggested that childhood anxiety and depression are so closely related that they are best considered as part of the same construct, often referred to as internalizing disorder. This lack of differentiation appears to be characteristic of younger children, with increased specificity developing over time (11 ,12). At least by middle childhood, there is support for the set of anxiety-related diagnoses in the

Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV). Spence (12) conducted a confirmatory factor analysis with data from children of 8 to 12 years. The best model included six correlated factors—panic-agoraphobia, social phobia, separation anxiety, obsessive-compulsive problems, generalized anxiety, and fear of physical injury (including dogs, dentists, heights, doctors)—and a single higher-order factor reflecting overall anxiety.

There is a clear need for further research in childhood anxiety. Estimates of prevalence and recovery vary widely because of a lack of standardization of criteria, assessment instruments, and methodology (e.g., clinician rating vs. self-report, child vs. parent or teacher report). A few general conclusions can be drawn about childhood internalizing disorders. Internalizing symptoms appear to remain fairly stable over time (13 ,14). Among boys, internalizing symptoms are not only predictive of later internalizing symptoms but also of subsequent externalizing problems (15). Although there may be high rates of recovery associated with a particular anxiety disorder, children who recover are at increased risk of developing other psychiatric diagnoses, most commonly other anxiety disorders or depressive disorders (16). Similarly, the presence of an anxiety or depressive disorder [overanxious disorder (OAD)/generalized anxiety disorder (GAD), panic, major depression] in adolescence is nonspecifically predictive of anxious or depressive disorders in adulthood (17).

The remainder of this section discusses disorder-specific information, including clinical manifestations, prevalence, and course.

Generalized Anxiety Disorder (GAD)

GAD is characterized by excessive anxiety or worry, which is difficult to control and is accompanied by symptoms of tension and physiologic arousal (18). The GAD diagnosis is currently used in place of overanxious disorder (OAD), which was eliminated with the publication of the DSM-IV. Children and adolescents with GAD most frequently worry about future events, personal safety, and social evaluation, and often present with multiple somatic complaints, such as headaches and stomachaches (19).

There is little information about the prevalence of GAD in children and adolescents because the diagnosis of OAD was used previously. Prevalence estimates of OAD tended to be quite variable, 2% to 19% (19), and often very high, partially because functional impairment was not necessary for the diagnosis (20). Recent estimates of the prevalence of GAD are in the range of 2.1% to 10.8% (21). It appears to be more prevalent in older children and in girls (19). Onset typically falls between 10.8 and 13.4 years of age (4).

Information about course is not yet available for the GAD diagnosis, but some extrapolation from OAD is possible. Cohen et al. (22) looked at three different statistical measures of persistence. They found that OAD was very stable in a subset of children and younger adolescents (ages 11 to 16), but that there was a substantial amount of new onset among older adolescents (ages 17 to 19). Their conclusion was that the disorder may be trait-like for those who exhibit symptoms early and that the development of the disorder in others may be triggered in adolescence. Cantwell and Baker (23) also found considerable stability; 25% of children who had been diagnosed with OAD approximately 4 years earlier had recovered (although the group size, eight, limits the usefulness of this estimate). In contrast, Last et al. (16) reported that 80% of their OAD sample had recovered 3 to 4 years later; however, 25% had developed another anxiety disorder and 25% had developed a depressive disorder.

GAD/OAD is a frequently co-occurring disorder. Of those with a primary diagnosis of OAD, there is often an additional diagnosis of separation anxiety disorder (37% to 44%), social phobia (4% to 57%), simple phobia (9% to 43%) or a depressive disorder (1% to 69%) (24). There are a number of potential reasons for the high rates of comorbidity, including true covariation of distinct disorders, the presence of a single underlying construct, overlap of diagnostic categories, and measurement error (12 ,25).

Obsessive-Compulsive Disorder (OCD)

OCD is recognizable by the presence of obsessions and compulsions, which are distressing or cause marked interference in one's life (18). Among young children, the disorder often presents with excessive and rigid ritualized behavior. For children in general, compulsions alone are more common than obsessions alone (26). The most common compulsions include washing/cleaning, repeating/redoing, and checking, and the most common obsessions include germs/contaminants and fear of harm to the self or to another (26). OCD symptoms change over time in 90% of children (4).

OCD is relatively uncommon, with estimated prevalences of approximately 0.3% for children and 0.35% to 1.9% for adolescents (27 ,28). Mean age of onset is approximately 10 years. Information about the course of OCD is variable and may be best described as chronic but fluctuating (4). Leonard et al. (6) followed children and adolescents who had been part of National Institute of Mental Health (NIMH) treatment studies. During the 2- to 7-year follow-up period, the patients on average received two different modalities of treatment (medication, behavioral therapy, other individual therapy, and family therapy), with 96% having had additional psychopharmacologic treatment and 46% some form of behavioral therapy. In spite of ongoing treatment, at follow-up 43% met the criteria for OCD, 18% had subclinical OCD, and 28% had OC features. Of the 11% who were symptom-free, only three (of 54) patients had no symptoms and were not on current medications. However, treatment was associated with decreased functional impairment. Last et al. (16), on the other hand, found

a 75% recovery rate in their sample. Other estimates of continued OCD at follow-up (1.5 to 7 years) range from 31% to 68% (16).

Comorbidity is a frequent problem with OCD. The most common co-occurring conditions include other anxiety disorders (38%), tic disorders (24% to 30%), mood disorders (26% to 29%), and specific developmental disabilities (24%) (26). Leonard et al. (6) found that both a lifetime history of tic disorder and current affective disorder at baseline were associated with poorer outcome.

Panic Disorder (with and without Agoraphobia) (PD)

PD involves recurrent and unexpected panic attacks, which cause significant distress for the affected individual. This frequently leads to avoidance of various situations for fear of developing symptoms and being unable to escape (18). Young children tend to articulate their panic-related fears in a different way than do adolescents or adults by virtue of their developmental level; they are more likely to express concerns about sudden somatic symptoms and less likely to describe fears of dying, losing control, or going crazy (4).

PD is uncommonly reported in children, to the point that there has been some debate as to whether it exists before puberty. Evidence of the existence of PD in children comes from both retrospective reports of adults with PD and from case reports (29). There is no epidemiologic study to date, and it is likely that many cases go undetected because of the predominance of somatic symptoms in presentation. Prevalence of PD in adolescents is low, 0.7% girls and 0.4% boys (0.6 overall) according to one large high school sample (28). The rate of panic attacks is expectedly higher, 11.6% for at least one full attack and an additional 3.2% for at least one limited symptom attack (30). PD appears to be two to three times more common in females (29). Hayward et al. (31) examined the relationship between the occurrence of panic attacks and sexual development in girls and found a positive relationship. There were no reported panic attacks among the least sexually mature girls, and a rate of 8% among those who were most developed. The authors proposed a number of theories for this association, including hormonal changes, psychosocial factors, and emergence of the ability to think abstractly, but further work is necessary to draw any definitive conclusions.

In the one study of the course of early PD, 30% continued to have PD and 30% had another psychiatric disorder 3 to 4 years later, but the generalizablity of this result is questionable because of the small size of the study population (ten) (16). Retrospective reports suggest that earlier onset is associated with poorer outcomes, including greater functional impairment and increased incidence of alcohol abuse and suicidality (32). Available information suggests that there is a high rate of comorbidity in adolescents with PD, particularly with affective disorders; again, this should be interpreted with caution because it is based on a single, small (*N* = 28) study of adolescents (33).

Posttraumatic Stress Disorder (PTSD)

To meet the criteria for a diagnosis of PTSD, a person must have been exposed to a traumatic event and as a result is exhibiting symptoms of reexperiencing, numbing/avoidance, and arousal (18). A recent confirmatory factor analytic study supported the presence of these three basic clusters of symptoms in children and adolescents, although it found that arousal is often manifested as general somatic complaints (34). PTSD presentations that are specific to children include reenactment of the trauma in play, physical attempts (e.g., covering eyes or ears) to avoid memories of the trauma, reduced interest in activities, behavioral regression (e.g., thumb-sucking, enuresis), and sleep disturbance (35).

Diagnosis of PTSD depends on exposure to a traumatic stressor. Each year, 6% to 7% of Americans are exposed to traumatic events (36), but the incidence is much greater in certain subpopulations. For example, studies of urban youth report exposure rates of up to 75% (37). Not everyone who is exposed to trauma goes on to develop PTSD. Estimates vary tremendously depending on the type of trauma and the elapsed time between the event and assessment. In Saigh et al.'s (37) review of the literature, they found PTSD prevalence among exposed youth to be 0% to 70.8% for crime-related events, 8.3% to 75% for war, and 0% to 95% for disasters. Overall, it appears that exposed children may be somewhat more likely to develop PTSD than are exposed adults (36). PTSD is more common in those who have been exposed to more severe trauma (34).

The course of PTSD in children over the short- or long-term has not been well studied, but it appears that prognostically important factors are whether the trauma involves a single occurrence or is repeated, and whether it involves abuse. Although the evidence is not entirely consistent, it appears that a single exposure is less likely to lead to long-term symptomatology (36).

Comorbidity is generally high, particularly with other anxiety disorders (7.7% to 41.6%) and affective disorders (16.7% to 85%), but there is also substantial co-occurrence of attention-deficit/hyperactivity disorder (5.8% to 34.6%), conduct disorder (0% to 26.9%), and oppositional defiant disorder (25%) (37). Additional disorders may be integrally related to the trauma, such as fears about safety of the self or loved ones or grief about loss (35). Other psychopathology may also be a function of other factors, such as a disrupted or disorganized childhood or engagement in risky behaviors, which increase the risk for both psychopathology and traumatic exposure (38).

Separation Anxiety Disorder (SAD)

SAD is the only current anxiety disorder that is uniquely diagnosed in children and adolescents. The hallmark feature

of this disorder is excessive concern about separation from attachment figures. This is frequently manifested as distress at separation and excessive worry that harm will befall the attachment figure or that some negative event will lead to separation (18). These children frequently avoid going to school, fear being left alone or sleeping alone, and exhibit a panic-like physiologic response to separation (32).

Prevalence estimates for SAD are 2.0% to 5.4% for children and 1.3% to 4.6% for adolescents, with some evidence of higher rates in girls, those of lower SES, and those with less educated parents (21 ,32). The onset of SAD is usually early and associated with a major stressor (4). Of nine children with SAD followed by Cantwell and Baker (23), only one was still diagnosable 4 to 5 years later; this was the highest rate of recovery of any of the disorders that they followed. Similarly, Last et al. (16) found an approximately 96% recovery rate among their 24 children and adolescents with SAD, although 25% had developed another disorder, most commonly depressive, 3 to 4 years later. SAD frequently co-occurs with other disorders, most often other anxiety disorders (OAD, 23% to 33%; specific phobias, 12.5% to 27%; social phobia, 8%) (24) or a depressive disorder (approximately one-third) (32).

SAD has been suggested to be a childhood manifestation of PD. Evidence that has been cited in support of this idea includes the symptomatic similarity between a panic attack and the response to separation in a child with SAD; the frequency of a history of SAD in panic patients; the clustering of SAD, PD, and depressive disorders in families; and the similarities in effective pharmacologic treatments for the two conditions (39 ,40 and 41). Documented cases of panic episodes unrelated to separation, however, argue against this hypothesis (29). Nonetheless, SAD appears to be a risk factor for the later development of PD, at least among females (4).

Specific Phobia

A specific phobia is diagnosed if a child consistently displays significant and excessive fear in response to a specific object or situation (18). The most common fears among children are heights, small animals, doctors/dentists, dark, loud noises, and thunder/lightening (19). The prevalence of this disorder is in the range of 0.3% to 9.1%, with somewhat higher rates in girls and younger children (19 ,21 ,42).

Unlike the other anxiety disorders reviewed here, children with specific phobias remain a fairly distinct group. Last et al. (16) found that those with specific phobias were least likely to recover within 3 to 4 years (69.2%) but also were least likely to show onset of a different disorder within the follow-up period. Similarly, Pine et al. (17) found that simple phobias in adolescence were related only to adult simple phobias.

Social Phobia

Social phobia involves “marked and persistent fear of one or more performance or social situations in which the person is exposed to unfamiliar people or to possible scrutiny by others” (18). The anxious response in such situations is associated with cognitions involving concerns about being humiliated or embarrassed. Childhood social phobia is associated with significant impairment and distress, and frequently leads to extensive phobic avoidance and deficient social skill development (43).

Although there are few good epidemiologic studies of social phobia in childhood, data from community studies in adolescents suggest that it is quite common (1% to 2%), with a noticeable jump in prevalence rates sometime between ages 12 to 13 and ages 14 to 17 (44). One longitudinal study suggested that many cases of social phobia in childhood remit within 3 to 4 years (86.4%) (16). However, when social phobia is present in adolescence, it is a strong predictor of social phobia in adulthood (17). These data, taken together, suggest that social phobia in childhood may be a more transitory phenomenon than social phobia in adolescence. If these findings are confirmed in future studies, they may suggest critical developmental time frames during which preventative efforts may be applied.

ANXIETY AND STRESS DISORDERS IN YOUNG TO MIDDLE-AGE ADULTS

Part of "60 - Anxiety and Stress Disorders: Course Over the Lifetime "

Epidemiology

Several epidemiologic studies have documented the high rate of anxiety disorders among adults in the general population. In reports from the Epidemiologic Catchment Area (ECA) Study, anxiety disorders were found to occur as a lifetime diagnosis in 14.6% of the adult U.S. population aged 18 years or older (45). More recently, the National Comorbidity Survey (NCS) found that 24.9% of adults in the age group 15 to 54 years had a lifetime anxiety disorder diagnosis (46). The two studies used somewhat different sampling methods, and different diagnostic interviews, probably therein explaining at least some of the variance in rates between studies (47).

Comorbidity

Comorbidity among the anxiety, mood, and substance use disorders is extensive (47). For example, two-thirds of persons in the community with generalized anxiety disorder in a 12-month period also had major depressive disorder during that time frame (48). In a clinical sample of 85 patients with major depression, 29% met criteria for a current anxiety disorder and 34% had at least one anxiety disorder during their lives (49).

What is particularly noteworthy about this relationship

is the temporal sequencing of disorders. Certain anxiety disorders, social phobia in particular (which has a median onset of between 13 and 15 years of age), almost inevitably begin *prior* to the onset of the mood or substance use disorder (50 ,51). In one study of depressed patients, social phobia was the most common lifetime anxiety disorder (occurring in 15% of cases) followed closely by panic disorder with agoraphobia (in 12%) (49). Social phobia occurred on average 2 years prior to the onset of major depressive disorder in these patients. Similar findings have emerged from community studies (52), suggesting that particular anxiety disorders such as social phobia may be considered risk factors for the subsequent development of major depression. It remains to be established what the mechanisms might be for this observed relationship. Does being socially anxious lead to increased isolation or decreased self-worth, thereby leading to an increase in subsequent major depression? Is social phobia merely the earliest manifestation of an anxiety-mood disorder diathesis? These questions will only be answered with future research that focuses broadly on psychosocial and biological vulnerabilities for anxiety and mood disorders.

Another interesting aspect of the anxiety-depression link lies in the relationship between major depressive disorder (MDD) and PTSD. Extensive comorbidity between PTSD and MDD is the norm in studies of various traumatized groups, including persons exposed to combat (53 ,54), disasters (55), and intimate partner violence (56). Community studies also demonstrate strong ties between these two disorders, with approximately 35% to 50% of cases of PTSD in the general population being comorbid with MDD (57).

Studies that have examined the temporal association between major depression and PTSD have posited several causal pathways. It has been observed that preexisting major depressive disorder increases an individual's risk for PTSD following exposure to traumatic events (58 ,59). The converse has also been observed, namely that preexisting PTSD is a risk factor for the later development of MDD (58 ,60). The mechanisms for this apparent reciprocal risk have yet to be explained, but might involve a general vulnerability to stress that can result in major depression (61) or PTSD in susceptible individuals.

Functioning, Quality of Life, and Cost to Society

Data have been collected in the past several years that highlight the disability and reduced quality of life associated with anxiety disorders in young and middle-aged adults. Studies in clinical samples of patients with PD, PTSD, OCD, and social phobia all depict these as serious conditions that rob individuals of enjoyment and pleasure and impair functioning in multiple domains (62). Epidemiologic studies, where the range of severity is expected to be wider and where many milder cases are expected to be seen, also provide persuasive evidence of the seriousness of anxiety disorders (52 ,63 ,64). The annual cost of anxiety disorders in the United States was estimated at $42.3 billion in 1990, or $1,542 per sufferer (65). Other than simple phobia, all anxiety disorders analyzed were associated with impairment in workplace performance. These observations, gleaned from a variety of clinical and nonclinical perspectives, portray anxiety disorders in adults as serious mental disorders worthy (and in need) of greater societal willingness to develop and apply better interventions to prevent or mitigate their impact on the lives of individuals.

ANXIETY AND STRESS DISORDERS IN OLDER ADULTS

Part of "60 - Anxiety and Stress Disorders: Course Over the Lifetime "

Although anxiety is among the most prevalent of psychiatric disorders in the elderly, research in this area has lagged far behind that of depression and dementia (66). But in the past few years, several important studies have been conducted that provide novel information about the prevalence, features, and course of anxiety disorders in older adults.

Epidemiology

Whereas it had previously been believed that anxiety disorders decline in prevalence with age, several possible explanations for this finding have been put forward. It has been suggested that this might be an artifact of measurement error, owing to differences in the way older individuals report anxiety (67). Previous epidemiologic studies may also have underestimated the prevalence of anxiety disorders in the elderly by limiting participation to community-dwelling older adults, who may have lower rates of anxiety disorder than those living in institutions (68).

Fortunately, data have recently become available from a new community survey that provides a more accurate and detailed perspective on anxiety disorders in older adults (69). The Longitudinal Aging Study Amsterdam (LASA) is based on a random sample of 3,107 older adults (ages 55 to 85), stratified for age and sex. The overall prevalence of anxiety disorders in the community was estimated at 10.2%. GAD was most common in a 6-month time period (7.3%), followed by social phobia (3.1%), PD (1.0%), and OCD (0.6%). For comparison purposes, it is noteworthy that the 6-month prevalence of major depression in the same study was 2.0%. Thus, anxiety disorders were far more common than depressive disorders in the elderly, underscoring the point made earlier that it is surprising that the elderly have received so little attention in the clinical and research literature to date.

This study also examined vulnerability factors for anxiety disorders in older adults (69). Many of the vulnerability factors for anxiety disorders in younger adults are common to older adults (e.g., female sex, lower levels of education),

but several unique risk factors were also encountered (e.g., having suffered extreme experiences during World War II). These investigators were also able to show that current stresses commonly experienced by older people (e.g., recent losses in the family and chronic physical illness) also played a part in the onset or exacerbation of anxiety disorders. Current life stresses, then, should be evaluated as possible contributors not only to depression, but also to anxiety in the elderly.

Comorbidity

Comorbidity patterns of older adults with anxiety disorders are remarkably similar to those of younger adults. In the LASA, 48% of those with MDD also met criteria for anxiety disorders, whereas 26% of those with anxiety disorders also met criteria for MDD (70).

The entity known as “anxious depression” warrants special mention in this context. Although definitions vary, anxious depression usually refers to MDD with accompanying anxiety. Anxious depression is a particularly common presentation in the elderly (66). Although anxious depression is frequently severe and impairing, its outcome is no worse than nonanxious depression when treated appropriately (71).

Special Features of Anxiety Disorders in the Elderly

The fact that medical illness becomes more common with increasing age can put a special twist on the presentation and origins of certain anxiety disorders in the elderly. First and foremost, it must be recognized that many medical illnesses (e.g., thyroid disease, chronic obstructive pulmonary disease, and stroke, to name just a few) may be associated with *de novo* anxiety symptoms or with the exacerbation of a preexisting anxiety disorder (66). Most “new” anxiety disorders in older life are either GAD or agoraphobia, whereas most other anxiety syndromes seen in the elderly (e.g., PD and OCD) reflect recurrence or worsening of an anxiety disorder that had its onset earlier in life (72).

Agoraphobia in older adults is usually a different phenomenon, with different etiology, from agoraphobia in younger adults. In younger adults, agoraphobia is almost always a complication of PD (73)—the individual comes to avoid situations that are associated with the possible occurrence of panic or difficulty escaping should a panic attack occur. In the elderly, the new onset of agoraphobia is rarely associated with spontaneous panic attacks, but instead is a maladaptive reaction to some form of medical illness experience that renders the individual fearful of being unable to function safely away from home (66). An example is an elderly woman who breaks her hip, and even after it has satisfactorily healed, is afraid to maneuver without help and therefore avoids leaving the house alone.

SUMMARY

Part of "60 - Anxiety and Stress Disorders: Course Over the Lifetime "

Anxiety disorders span the full range of human existence from childhood to old age, though symptoms may vary considerably owing to developmental differences and related factors. Anxiety disorders in children are often transitory phenomena, with the majority showing remission by adolescence or early adulthood. Yet, in a minority, extremely shy and fearful temperament in childhood can merge almost imperceptibly into social phobic and panic disorders in adolescence. Anxiety disorders in youth appear to be a risk factor for the subsequent development of major depression (and, although less certain, possibly also substance use disorders) in late adolescence and young adulthood. By adulthood, comorbidity is the rule, with most individuals experiencing multiple anxiety disorders, or concurrent mood and anxiety disorders. For the most part, anxiety disorders are chronic, and these persist from young adulthood into old age. But even in later life, new onset of anxiety disorders can occur, often in the context of medical illness or other sources of life stress.

ACKNOWLEDGMENTS

Part of "60 - Anxiety and Stress Disorders: Course Over the Lifetime "

Dr. Stein has received research support from the following companies: Bristol-Myers Squibb; Eli Lilly and Company; Forrest Laboratories; Hoffman-LaRoche Pharmaceuticals; Novartis; Parke-Davis; Pfizer; SmithKline Beecham; and Solvay Pharmaceuticals. He is currently or has been in the past a consultant for Forrest Laboratories; Hoffmann-La Roche Pharmaceuticals; Janssen Research Foundation; SmithKline Beecham and Solvay Pharmaceuticals. Finally, he receives or has received speaking honoraria from Bristol-Myers Squibb; Eli Lilly and Company; Hoffmann-La Roche Pharmaceuticals; Pfizer; Pharmacia & Upjohn; SmithKline Beecham and Solvay Pharmaceuticals.

REFERENCES

1. Kendler KS, Neale MC, Kessler RC, et al. Major depression and generalized anxiety disorder: same genes, (partly) different environments? *Arch Gen Psychiatry* 1992;49:716–722.

2. Beidel DC, Turner SM. At risk for anxiety: I. Psychopathology in the offspring of anxious parents. *J Am Acad Child Adolesc Psychiatry* 1997;36:918–924.

3. Fyer AJ, Mannuzza S, Gallops MS, et al. Familial transmission of simple phobias and fears: a preliminary report. *Arch Gen Psychiatry* 1990;47:252–256.

4. Albano AM, Chorpita BF, Barlow DH. Childhood anxiety disorders. In: Mash EJ, Barkley RA, eds. *Child psychopathology*. Guildford Pressα1996:196–241.

5. Hirshfeld DR, Biederman J, Brody L, et al. Expressed emotion toward children with behavioral inhibitions: Associations with maternal anxiety disorder. *J Am Acad Child Adolesc Psychiatry* 1997;36:910–917.

6. Leonard HL, Swedo SE, Lenane MC, et al. A 2- to 7-year follow-up study of 54 obsessive-compulsive children and adolescents. *Arch Gen Psychiatry* 1993;50:429–439.

7. Rosenbaum JF, Biederman J, Gersten M, et al. Behavioral inhibition in children of parents with panic disorder and agoraphobia: A controlled study. *Arch Gen Psychiatry* 1988;45:463–470.

8. Biederman J, Rosenbaum JF, Chaloff J, et al. Behavioral inhibition as a risk factor for anxiety disorders. In: March JS, ed. *Anxiety disorders in children and adolescents*. New York: Guilford Press, 1995:61–81.

9. Kagan J, Reznick JS, Snidman N. Biological bases of childhood shyness. *Science* 1988;240:167–171.

10. Schwartz CE, Snidman N, Kagan J. Adolescent social anxiety as an outcome of inhibited temperament in childhood. *J Am Acad Child Adolesc Psychiatry* 1999;38:1008–1016.

11. Cole DA, Truglio R, Peeke L. Relation between symptoms of anxiety and depression in children: a multitrait-multimethod-multigroup assessment. *J Consult Clin Psychol* 1997;65:110–119.

12. Spence SH. Structure of anxiety symptoms among children: A confirmatory factor analytic study. *J Abnorm Psychol* 1997;106:280–297.

13. Feehan M, McGee R, Williams SM. Mental health disorders from age 15 to age 18 years. *J Am Acad Child Adolesc Psychiatry* 1993;32:1118–1126.

14. Ferdinand RF and Verhulst FC. Psychopathology from adolescence into young adulthood: An 8-year follow-up study. *Am J Psychiatry* 1995;152:1586–1594.

15. McGee R, Feehan M, Williams S, et al. DSM-III disorders from age 11 to age 15 years. *J Am Acad Child Adolesc Psychiatry* 1992;31:50–59.

16. Last CG, Perrin S, Hersen M, et al. A prospective study of childhood anxiety disorders. *J Am Acad Child Adolesc Psychiatry* 1996;35:1502–1510.

17. Pine DS, Cohen P, Gurley D, et al. The risk of early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. *Arch Gen Psychiatry* 1998;55:56–64.

18. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, fourth ed. Washington, DC: APA, 1994.

19. Silverman WK, Ginsburg GS. Specific phobia and generalized anxiety disorder. *Anxiety Disord Child Adolesc* 995;151–180.

20. Beidel DC. Social phobia and overanxious disorder in school-age children. *J Am Acad Child Adolesc Psychiatry* 991;30:545–552.

21. Costello EJ, Angold A. Epidemiology. In: March JS, ed. *Anxiety disorders in children and adolescents*. New York: Guilford Press, 1995:109–124.

22. Cohen P, Cohen J, Brook J. An epidemiological study of disorders in late childhood and adolescence: II. Persistence of disorders. *J Child Psychol Psychiatry* 1993;34:869–877.

23. Cantwell DP, Baker L. Stability and natural history of DSM-III childhood diagnoses. *J Am Acad Child Adolesc Psychiatry* 1989;28:691–700.

24. Curry JF, Murphy LB. Comorbidity of anxiety disorders. In: March JS, ed. *Anxiety disorders in children and adolescents*. New York: Guilford Press, 1995:301–317.

25. Ollendick TH, Yule W. Depression in British and American children and its relation to anxiety and fear. *J Consult Clin Psychol* 1990;58:126–129.

26. March JS, Leonard HL, Swedo SE. Obsessive-compulsive disorder. In: March JS, ed. *Anxiety disorders in children and adolescents*. New York: Guilford Press, 1995:251–275.

27. Flament MF, Whitaker A, Rapoport JL, et al. Obsessive compulsive disorder in adolescence: an epidemiological study. *J Am Acad Child Adolesc Psychiatry* 1988;27:764–771.

28. Whitaker A, Johnson J, Shaffer D, et al. Uncommon troubles in young people: prevalence estimates of selected psychiatric disorders in a nonreferred adolescent population. *Arch Gen Psychiatry* 1990;47:487–496.

29. Moreau D, Weissman MM. Panic disorder in children and adolescents: a review. *Am J Psychiatry* 1992;149:1306–1314.

30. Hayward C, Killen JD, Taylor CB. Panic attacks in young adolescents. *J Am Acad Child Adolesc Psychiatry* 1998;37:1308–1316.

31. Hayward C, Killen JD, Hammer LD, et al. Pubertal stage and panic attack history in sixth-and seventh-grade girls. *Am J Psychiatry* 1992;149:1239–1243.

32. Black B. Separation anxiety disorder and panic disorder. In: March JS, ed. *Anxiety disorders in children and adolescents*. New York: Guilford Press, 1995:212–234.

33. Bradley SJ, Hood J. Psychiatrically referred adolescents with panic attacks: presenting symptoms, stressors, and comorbidity. *J Am Acad Child Adolesc Psychiatry* 1993;32:826–829.

34. Anthony JL, Lonigan CJ, Hecht SA. Dimensionality of posttraumatic stress disorder symptoms in children exposed to disaster: results from confirmatory factor analyses. *J Abnorm Psychol* 1999;108:326–336.

35. Amaya-Jackson L, March JS. Posttraumatic stress disorder. In: March JS, ed. *Anxiety disorders in children and adolescents*. New York: Guilford Press, 1995:276–300.

36. Fletcher KE. Childhood posttraumatic stress disorder. In: Mash EJ, Barkley RA, eds. *Child psychopathology*. Guilford Press; 1996:242–276.

37. Saigh PA, Yasik AE, Sack WH, et al. Child-adolescent posttraumatic stress disorder: prevalence, risk factors, and comorbidity. In: Saigh PA, Bremner JD, eds. *Posttraumatic stress disorder: a comprehensive text*. 1999:18–43.

38. Widom CS. Posttraumatic stress disorder in abused and neglected children gown up. *Am J Psychiatry* 1999;156:1223–1229.

39. Abelson JL, Alessi NE. Discussion of "Child Panic Revisited." *J Am Acad Child Adolesc Psychiatry* 1992;31:114–116.

40. Black B, Robbins DR. Panic disorder in children and adolescents. *J Am Acad Child Adolesc Psychiatry* 1990;29:36–44.

41. Klein RG. Is panic disorder associated with childhood separation anxiety disorder? *Clin Neuropharmacol* 1995;18(suppl 2):S7–S14.

42. Muris P, Schmidt H, Merckelback H. The structure of specific phobia symptoms among children and adolescents. *Behav Res Ther* 1999;37:863–868.

43. Beidel DC, Turner SM, Morris TL. Psychopathology of childhood social phobia. *J Am Acad Child Adolesc Psychiatry* 1999;38:643–650.

44. Essau CA, Conradt J, Petermann F. Frequency and comorbidity of social phobia and social fears in adolescents. *Behav Res Ther* 1999;37:831–843.

45. Robins LN, Helzer JE, Weissman MM, et al. Lifetime prevalence of specific psychiatric disorders in three sites. *Arch Gen Psychiatry* 1984;41:949–958.

46. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of psychiatric disorders in the United States: results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994;51:8–19.

47. Regier DA, Rae DS, Narrow WE, et al. Prevalence of anxiety disorders and their comorbidity with mood and addictive disorders. *Br J Psychiatry*1998;173(suppl 34):24–28.

48. Kessler RC, DuPont RL, Berglund PA, et al. Impairment in pure and comorbid generalized anxiety disorder and major depression at 12 months in two national surveys. *Am J Psychiatry* 1999;156:1915–1923.

49. Schatzberg AF, Samson JA, Rothschild AJ, et al. McLean Hospital Depression Research Facility: early-onset phobia disorders and adult-onset major depression. *Br J Psychiatry* 1998;173(suppl 34):29–34.

50. Stein MB, Chavira DA. Subtypes of social phobia and comorbidity with depression and other anxiety disorders. *J Affect Disord* 1998;50:S11–S16.

51. Kessler RC, Stang P, Wittchen HU, et al. Lifetime comorbidities between social phobia and mood disorders in the U.S. National Comorbidity Survey. *Psychol Med* 1999;29:555–567.

52. Wittchen HU, Stein MB, Kessler RC. Social fears and social phobia in a community sample of adolescents and young adults: prevalence, risk factors, and co-morbidity. *Psychol Med* 1999;29:309–323.

53. Shalev AY, Freedman S, Peri T, et al. Prospective study of posttraumatic stress disorder and depression following trauma. *Am J Psychiatry* 1998;155:630–637.

54. Southwick SM, Yehuda R, Giller ELJ. Characterization of depression in war-related posttraumatic stress disorder. *Am J Psychiatry* 1991;148:179–183.

55. Green BL, Lindy JD. Post-traumatic stress disorder in victims of disasters. *Psychiatr Clin North Am* 1994;17:301–309.

56. Stein MB, Kennedy C. Major depressive and posttraumatic stress disorder comorbidity in female victims of intimate partner violence. *J Affect Disord* (*in press*).

57. Kessler RC, Sonnega A, Bromet E, et al. Posttraumatic stress disorder in the National Comorbidity Survey. *Arch Gen Psychiatry* 1995;52:1048–1060.

58. Breslau N, Davis GC, Peterson EL, et al. Psychiatric sequelae of posttraumatic stress disorder in women. *Arch Gen Psychiatry* 1997;54:81–87.

59. Bromet E, Sonnega A, Kessler RC. Risk factors for DSM-III-R posttraumatic stress disorder: findings from the National Comorbidity Survey. *Am J Epidemiol* 1998;147:353–361.

60. Mellman TA, Randolph CA, Brawman-Mintzer O, et al. Phenomenology and course of psychiatric disorders associated with combat-related posttraumatic stress disorder. *Am J Psychiatry* 1992;149:1568–1574.

61. Kendler KS, Neale MC, Kessler RC, et al. Causal relationship between stressful life events and the onset of major depression. *Am J Psychiatry* 1999;156:837–841.

62. Mendlowicz MV, Stein MB. Quality of life in individuals with anxiety disorders. *Am J Psychiatry* 2000;157:669–682.

63. Markowitz JS, Weissman MM, Ouellette R. Quality of life in panic disorder. *Arch Gen Psychiatry* 1989;46:984–992.

64. Stein MB, Kean Y. Disability and quality of life in social phobia. *Am J Psychiatry* 2000;157:1606–1613.

65. Greenberg PE, Sisitsky T, Kessler RC, et al. The economic burden of anxiety disorders in the 1990s. *J Clin Psychiatry* 1999;60:427–435.

66. Flint AJ. Anxiety disorders in late life. *Can Fam Physician* 1999;45:2672–2679.

67. Stanley MA, Beck JG, Zebb BJ. Psychometric properties of four anxiety measures in older adults. *Behav Res Ther* 1996;34:827–838.

68. Kogan JN, Edelstein BA and McKee DR. Assessment of anxiety in older adult: current status. *J Anxiety Disord* 2000;14:109–132.

69. Beekman ATF, Bremmer MA, Deeg DJH, et al. Anxiety disorders in later life: report from Longitudinal Aging Study Amsterdam. *Int J Geriatr Psychiatry* 2000;13:717–726.

70. Beekman ATF, de Beurs E, van Balkom AJ, et al. Anxiety and depression in later life: co-occurrence and communality of risk factors. *Am J Psychiatry* 2000;157:89–95.

71. Flint AJ, Rifat SL. Two-year outcome of elderly patients with anxious depression. *Psychiatr Res* 1997;66:23–31.

72. Flint AJ. Epidemiology and comorbidity of anxiety disorders in the elderly. *Am J Psychiatry* 1994;151:640–649.

73. Horwath E, Lish JD, Johnson J, et al. Agoraphobia without panic: clinical reappraisal of an epidemiologic finding. *Am J Psychiatry* 1993;150:1496–1501.

61

Genetic and other Vulnerability Factors for Anxiety and Stress Disorders

Kathleen R. Merikangas

Daniel Pine

Kathleen R. Merikangas and Daniel Pine: National Institute of Mental Health, Bethesda, Md.

Despite dramatic advances in our understanding of genetics and neurobiology, the etiology of the anxiety disorders is still relatively unknown. To date, there remain no pathognomonic markers with which a presumptive diagnosis of an anxiety disorder may be made. This highlights the importance of the empirical epidemiologic approach to investigating the definitions and risk factors for the expression of anxiety across the life course. Anxiety disorders are developmental conditions that often emerge during childhood and follow varied developmental trajectories (1 ,2). Research on early-life vulnerability factors that predict the trajectory of anxiety symptoms across development holds promise for elucidating mechanistic pathways in anxiety.

In evaluating the risk factors for the development of anxiety disorders, there are several issues requiring consideration. First, there is substantial overlap between the anxiety disorders and other psychiatric disorders both concomitantly and longitudinally. Second, manifestations of anxiety change substantially across the life course, particularly during childhood and adolescence. Therefore, a developmental perspective is essential in evaluating links between risk factors and anxiety disorders. Third, the assessment of anxiety requires evaluation of the context in which the individual experiences anxiety as well as the subjective response to anxiety-inducing situations. As such, anxiety becomes a disorder when there is a mismatch between inherent threat posed by a particular stimulus or situation and the cognitive or somatic response.

Research on vulnerability factors has undergone a relatively marked transformation in recent years, due to conceptual changes in causal theories of mental disorders. Such conceptual changes are reflected in three major themes that organize current research on vulnerability factors in anxiety. First, although studies through the early 1990s often emphasized the role of one or another particular risk factor, more recent studies emphasize the manner in which multiple risk factors might interact to cause mental syndromes, including anxiety, as part of a mechanistic pathway. For example, although dysregulation in fear conditioning has been linked to anxiety for more than two decades (3), such dysregulation is now viewed as part of a larger chain of intrinsic and extrinsic events that may ultimately culminate in an anxiety disorder (4). Second, as a corollary to this view, vulnerability markers are now conceptualized as tied to families of anxiety disorders, as opposed to specific conditions. This change in perspective follows the observation that validators of individual mental syndromes related to differential course, familial aggregation, or psychophysiology relate more closely to families of disorders than to particular disorders. Third, marked advances over the past 20 years in neuroscience have stimulated a closer integration of basic and clinical work on vulnerability markers in anxiety disorders. Progress in elucidating neural circuits related to anxiety has facilitated research on vulnerability markers for anxiety disorders that integrates data from basic and clinical science.

This chapter examines the major risk factors for the development of anxiety disorders across the life span. Particular attention is paid to the specificity of vulnerability factors and to developmental differences in expression of the disorders themselves.

MAGNITUDE AND DEMOGRAPHIC RISK FACTORS

Part of "61 - Genetic and other Vulnerability Factors for Anxiety and Stress Disorders "

Magnitude of Anxiety in General Population

The anxiety disorders are the most common psychiatric disorders both in the United States and elsewhere (2 ,5 ,6). The

results of the two large-scale community-based surveys of psychiatric disorders of adults in the United States, the Epidemiological Catchment Area study (ECA) (2) and the National Comorbidity Study (NCS) (5), reveal that the total prevalence rates of anxiety disorders are greater than those of the affective disorders, behavior disorders, and substance use disorders. Phobias tend to be the most common anxiety disorder, whereas panic disorder is fairly rare in the general population. There is substantial overlap both cross-sectionally and longitudinally between the anxiety disorders and other disorders, as well as between the subtypes of anxiety disorders themselves (6). On average, there is a threefold increased risk of having a second disorder compared to that of manifesting an anxiety disorder alone across the lifetime.

Comorbidity between anxiety disorders and other psychiatric disorders has been demonstrated in both clinical and community samples. Anxiety disorders are most strongly associated with affective disorders and with substance use disorders (6), though they are generally associated with all other major classes of disorders including depression, disruptive behaviors, eating disorders, and substance use. Comorbidity between anxiety disorders and other disorders in the *Diagnostic and Statistical Manual of Mental Disorders*, third edition revised or fourth edition (DSM-IIIR or -IV) may be even more common in adolescents than in adults (7). A review of comorbidity of anxiety and depression by Brady and Kendall (8) suggests that anxiety and depression may be part of a developmental sequence in which anxiety is expressed earlier in life than depression. Although the association between anxiety and depression is quite consistent, the evidence of links between anxiety disorders and behavior problems is inconclusive.

Sex Differences in Anxiety Disorders

Similar to the affective disorders, females tend to exhibit greater rates of anxiety disorders, though there is some variability according to specific subtypes. Table 61.1 presents the sex-specific lifetime rates of the major subtypes of anxiety disorders assessed in the ECA and NCS (5 ,6). Although the magnitude of the rates of anxiety disorders varies substantially between the two studies, the sex ratio is strikingly similar: women have an approximately twofold elevation in lifetime rates of panic, generalized anxiety disorder, agoraphobia, and simple phobia than men in both studies. In contrast, there is a nearly equal sex ratio for the lifetime prevalence of social phobia.

	Epidemiologic Catchment Area Study (5)			National Comorbidity Survey (6)		
	Males	Females	Sex Ratio (F:M)	Males	Females	Sex Ratio (F:M)
Anxiety disorders						
Anxiety disorders, total	1.8	10.3	5.7	19.2	30.5	1.6
Generalized anxiety disorder	4.3	6.8	1.6	3.6	6.6	1.8
Panic disorder	0.9	2.0	2.2	2.0	5.0	2.5
Phobic disorders						
Agoraphobia without panic	3.2	7.9	2.5	3.5	7.0	2.0
Simple phobia	7.8	14.5	1.9	6.7	15.7	2.3
Social phobia	2.5	2.9	1.2	11.1	15.5	1.4

TABLE 61.1. SEX-SPECIFIC LIFETIME PREVALENCE RATES OF ANXIETY DISORDERS IN COMMUNITY SURVEYS IN THE UNITED STATES

There is increasing evidence from community-based studies that anxiety symptoms and disorders are also the most common problems in childhood and adolescence as well. The rates of anxiety disorders in community or school-based surveys of children and adolescents as defined by contemporary diagnostic criteria range from 0.1% to 13.3% in males and 0.4% to 28.6% in females (9). Similar to the sex ratio for adults, girls tend to have more of all subtypes of anxiety disorders, irrespective of the age composition of the sample. For example, in a recent epidemiologic study, females compared to males had greater rates of current anxiety disorders (i.e., 12.2% vs. 8.5%), past anxiety disorders (5.2% vs. 2.7%), as well as anxiety symptom scores on a dimensional rating (mean = 1.9 vs. 0.9) (10). Nevertheless, Lewinsohn et al. (10) reported that despite the greater rates of anxiety in girls across all ages, there was no difference between boys and girls in the average age at onset of anxiety (mean for girls = 8.0 ± 3.9; mean for boys = 8.5 ± 3.8).

Age-Specific Patterns of Expression of Anxiety Disorders

Retrospective reports of adults with anxiety disorders suggest that the onset of anxiety disorders generally occurs in childhood or adolescence. Although there is substantial

variation across studies, the results of prospective community-based research reveal differential peak periods of onset of specific subtypes of anxiety: separation anxiety and specific phobias in middle childhood (i.e., ages 7 to 9); overanxious disorder in late childhood (i.e., 10 to 13); social phobia in middle adolescence (i.e., 15 to 16); and panic attacks, sometimes progressing to panic disorder, in late adolescence (i.e., 17 to 18) (1 ,11 ,12 ,13 and 14). Anxiety disorders, particularly the phobias, tend to persist across the life course. However, there are major differences among the anxiety subtypes in terms of specificity and chronicity. Whereas the phobic states tend to be fairly stable and nonprogressive, generalized anxiety and panic tend to be less specific and less stable over time (1 ,15 ,16).

Several follow-up studies of children and adolescents have shown that anxiety symptoms and disorders in general tend to exhibit some stability, but with substantial switching across categories of anxiety disorders over time (17 ,18). A recent 8-year follow-up study of a community sample of youth ages 9 to 18 at study entry provides compelling evidence of the stability of the subtypes of anxiety disorders (17). The stability of both social phobia and simple phobia was highly specific over time, whereas overanxious disorder was associated with major depression, social phobia, and generalized anxiety in early adulthood.

Epidemiologic surveys of adults reveal that the female preponderance of anxiety disorders is present across all stages of life but is most pronounced throughout early and mid-adulthood. The rates of anxiety disorders in males are also rather constant throughout adult life, whereas the rates in females peak in the fourth and fifth decades of life and decrease thereafter. The increased rates in females are present across all ages and do not diminish as the rates of anxiety decrease in late life. The importance of pure anxiety disorders in late life was described by Beekman et al. (19), who found different risk factors for anxiety disorders than for either depression or comorbid anxiety and depression in a community sample of adults over age 55.

Social Class and Ethnicity

Rates of anxiety disorders in general are greater among those at lower levels of socioeconomic status (20). Several community studies have yielded greater rates of anxiety disorders, particularly phobic disorders, among African-Americans (5). With respect to children, Compton et al. (14) found that Caucasian children were more likely to report symptoms of social phobia, whereas African-American children had more separation anxiety symptoms. Pine et al. (1) reported that phobias were greater among those at lower levels of social class. The reasons for ethnic and social class differences have not yet been evaluated systematically; however, both methodologic factors as well as differences in exposure to stressors have been advanced as possible explanations.

FAMILIAL AND GENETIC FACTORS

Part of "61 - Genetic and other Vulnerability Factors for Anxiety and Stress Disorders "

The familial aggregation of all of the major subtypes of anxiety disorders has been well established (21). As reviewed below, the results of more than a dozen controlled family studies of probands with specific subtypes of anxiety disorders converge in demonstrating a 3- to 5-fold increased risk of anxiety disorders among first-degree relatives of affected probands compared to controls. The importance of the role of genetic factors in the familial clustering of anxiety has been demonstrated by numerous twin studies of anxiety symptoms and disorders (22 ,23). However, the relatively moderate magnitude of heritability also strongly implicates environmental etiologic factors. Table 61.2 summarizes the results of family and twin studies of anxiety disorders.

Type of Study	Comparison	Number of Studies	Average Relative Risk	Range
Family	Rel of probands vs. rel of controls	13 panic	5.4	(4.2–17.8)
		4 social phobia	3.1	(2.5–9.7)
		3 general anxiety	4.3	(2.7–5.6)
		3 OCD	3.5	(1.0–5.1)
Twin	MZ vs. DZ	3 panic	2.4	(2.2–2.5)
		4 phobias	2.6	(1.4–9.5)
		1 OCD	4.9	—

DZ, dizygotic; MZ, monozygotic; OCD, obsessive-compulsive disorder.

TABLE 61.2. SUMMARY OF FAMILY AND TWIN STUDIES OF ANXIETY DISORDERS

Review of Family and Twin Studies of Anxiety Disorders in Adults

Panic Disorder

Of the subtypes of anxiety, panic disorder is the anxiety syndrome that has been shown to have the strongest degree of familial aggregation. A recent review of family studies of

panic disorder by Gorwood et al. (24) cited 13 studies that included 3,700 relatives of 780 probands with panic disorder compared to 3,400 relatives of 720 controls. The lifetime prevalence of panic was 10.7% among relatives of panic probands compared to 1.4% among relatives of controls, yielding a relative risk of 6.8. In addition, early-onset panic, panic associated with childhood separation anxiety, or panic associated with respiratory symptoms has each been shown to have a higher familial loading than other varieties of panic disorder (25).

Although there has been some inconsistency reported by twin studies of panic disorder (26), two studies applying modern diagnostic criteria demonstrated considerably higher rates for monozygotic compared to dizygotic twins (27 ,28). Furthermore, current estimates derived from the Virginia Twin Registry show panic disorder to have the highest heritability of all anxiety disorders at 44% (29).

Phobic Disorders

Though there are far fewer controlled family and twin studies of the other anxiety subtypes, all of the phobic states (i.e., specific phobia, agoraphobia) have also been shown to be familial (30 ,31 ,32 and 33 ; see refs 34 ,35 and 36 for reviews). The average relative risk of phobic disorders in the relatives of phobics is 3.1. Stein et al. (33) found that the familial aggregation of social phobia could be attributed to the generalized subtype of social phobia. Data from the Virginia Twin Study report the estimated total heritability for phobias to be 35% (29 ,37).

Generalized Anxiety Disorder

There is also evidence of both the familial aggregation and heritability of generalized anxiety disorder in a limited number of studies. The average familial odds ratio is approximately 5 (32 ,38), and the heritability was 0.32 among female twin pairs (37).

Obsessive-Compulsive Disorder

Likewise, there are also very few controlled family studies of obsessive-compulsive disorder. Two of the three studies (39 ,40) reported familial relative risks of 3 to 4, whereas Black et al. (41) found no evidence for familial aggregation. Nestadt et al. (40) found that both the age of onset and obsessions were associated with greater familiality. Twin studies have yielded weak evidence for heritability of obsessive compulsive disorder (42 ,43 and 44).

Linkage and Association Studies

Based on indirect evidence implicating the adrenergic system in panic disorder (45), several linkage studies have investigated the role of mutations in adrenergic receptor loci on chromosomes 4, 5, or 10 (46), but without success. Other work has similarly excluded linkage with γ-aminobutyric acid receptor A ($GABA_A$) genes (47). Reports from a genomic survey of panic disorder using 600 markers have not yielded evidence of linkage (48). Similarly, linkage studies have excluded the possibility that panic disorder was due to mutations in adrenergic receptor loci on chromosomes 4, 5, or 10 (46), and other work has similarly excluded linkage with $GABA_A$ receptor genes (47). Recent reports from a genomic survey of panic disorder using 600 markers have not yielded evidence of linkage (48).

Family Studies and Phenotypic Definitions

The lack of success in identifying specific genes for anxiety disorders is not surprising given their complexity. Similar to several other psychiatric disorders, the anxiety disorders are complicated by etiologic and phenotypic heterogeneity, a lack of valid diagnostic thresholds, unclear boundaries between discrete anxiety subtypes, and comorbidity with other forms of psychopathology. Impediments to estimating genetic influences in youth are demonstrated by dramatic differences in heritability according to the informant regarding child psychopathology. For example, Eaves et al. (49) found that the heritability of both separation anxiety and overanxious disorder was far greater for parent-reported rather than child-reported disorder.

The family study approach, particularly when employed with systematic community-based samples, is one of the most powerful strategies to minimize heterogeneity because etiologic factors for the development of a particular disorder can be assumed to be relatively homotypic within families. There is a dearth of studies that have employed within-family designs to examine either phenotypic expression or some of the putative biological factors underlying the major anxiety disorders. For example, both Perna et al. (50 ,51) and Coryell (52) have shown that healthy relatives of probands with panic disorder have increased sensitivity to CO_2 challenge, suggesting that CO_2 sensitivity may be a promising trait marker for the development of panic, as described below. Smoller and Tsuang (36) discuss the value of family and twin studies in identifying phenotypes for genetic studies.

Both family and twin studies have been used to examine sources of overlap within the anxiety disorders, and between the anxiety disorders and other syndromes including depression, eating disorders, and substance abuse. Fyer et al. (31 ,53) have demonstrated the independence of familial aggregation of panic and phobias. With respect to comorbidity, whereas panic disorder, generalized anxiety, and depression have been shown to share common familial and genetic liability (23 ,54 ,55), there is substantial evidence for the independent etiology of anxiety disorders and substance use disorders (36 ,55 ,56). Similar results have emerged from

studies of symptoms of anxiety and depression in youth in which both anxiety and depression was found to result from common genetic diathesis (57 ,58).

In a comprehensive consideration of what may be inherited, Marks (59) reviews the components of anxiety that have been investigated in both human and animal studies. Evidence from twin studies has indicated that somatic manifestations of anxiety may lie under some degree of genetic control. These studies demonstrate that physiologic responses, such as pulse, respiration rate, and galvanic skin response, are more alike in monozygotic than in dizygotic twin pairs. Furthermore, twin studies of personality factors have shown high heritability of anxiety reaction. Finally, the results of animal studies have suggested that anxiety or emotionality is under genetic control. Selective breeding experiments with mammals have demonstrated that emotional activity analogous to anxiety is controlled by multiple genes (59). These findings suggest that anxiety and fear states are highly heterogeneous and that future studies need to investigate the extent to which the components of anxiety result from common versus unique genetic factors and the role of environmental factors, either biologic or social, in either potentiating or suppressing their expression.

High-Risk Studies of Anxiety Disorders

Given the early age of onset for anxiety disorders, studies of children of parents with anxiety have become an increasingly important source of information on the premorbid risk factors and early forms of expression of anxiety. Increased rates of anxiety symptoms and disorders among offspring of parents with anxiety disorders have been demonstrated by Turner et al. (60), Biederman et al. (61), Sylvester et al. (62), Last et al. (63), Warner et al. (64), Beidel and Turner (65), Beidel (66), Capps et al. (67), Merikangas et al. (68), Unnewehr et al. (69), and Warner et al. (70). Table 61.3 presents the risk of anxiety disorders among offspring of parents with anxiety disorders compared to controls averages 3.5 (range 1.3 to 13.3), suggesting specificity of parent-child concordance within broad subtypes of anxiety disorders.

Study	Sample						
	Proband				Offspring		
Author (year)	Anxiety	Other	Other	Spouse	*N*	Age	Relative Risk
Sylvester et al., 1987 (73)	Panic	MDD	—	No Dx	91	7–17	13.3
Turner et al., 1987 (60)	Agoraphobia/OCD	Dysthymia	—	Not evaluated	43	7–12	4.8
Capps et al., 1996 (67)	Agoraphobia	Agoraphobia/Panic	—	Not evaluated	43	8–24	—
Warner et al., 1995 (64)	Panic/MDD	Panic	Early-onset MDD	Dx	145	6–29	1.3
Beidel et al., 1997 (65)	Anxiety + depression	Anxiety	MDD	No Dx	129	7–12	4.0
Merikangas et al., 1998 (68)	Panic/social phobia	Alcohol or drugs	Substance + anxiety	Dx	192	7–17	2.0
Unnewehr et al., 1998 (69)	Panic	Simple	—	Not evaluated	87	5–15	9.2

Dx, diagnosis; MDD, major depressive disorder.

TABLE 61.3. CONTROLLED HIGH-RISK STUDIES OF ANXIETY

However, similar to studies of adults that show common familial and genetic risk factors for anxiety and depression (27 ,71 ,72), studies in children have also revealed a lack of specificity with respect to depression (60 ,64 ,65 ,73). Studies that employed a comparison group of parent probands with depressive disorders have shown that rates of anxiety disorders are also increased among the offspring of these parents (60 ,62 ,65 ,70); conversely, offspring of parents with anxiety disorders and depression have elevated rates of depression when compared to those of controls (62) or to offspring of anxiety-disordered parents without depression (61). Similar findings emerged from the family study by Last et al. (63), who found an increase in rates of major depression among the adult relatives of children with anxiety. Weissman et al. (74) have even suggested that childhood anxiety represents one of the earliest manifestations of familial risk for depression. These findings are usually interpreted as providing evidence for age-specific expression of common risk factors for anxiety in childhood and depression with or without comorbid anxiety in adulthood.

The high rates of anxiety disorders among offspring of parents with anxiety suggest that there may be underlying psychological or biological vulnerability factors for anxiety disorders in general, which may already manifest in children

prior to puberty. Previous research has shown that children at risk for anxiety disorders throughout life are characterized by behavioral inhibition (75), autonomic reactivity (66 ,76), somatic symptoms (60 ,77), social fears (60 ,62), enhanced startle reflex (76), and respiratory sensitivity (160). Empirical research on each of these domains of risk is reviewed in the next section.

VULNERABILITY MARKERS

Part of "61 - Genetic and other Vulnerability Factors for Anxiety and Stress Disorders "

The current section reviews recent studies on vulnerability markers in anxiety disorders. This includes data on temperamental factors and biological profiles. The first section reviews evidence regarding individual-level vulnerability factors, whereas the subsequent section examines data linking exogenous or environmental factors with risk for anxiety. As noted above, both sets of vulnerability markers operate within complex causal chains involving multiple interacting risk factors. Moreover, in such complex chains, the boundary between intrinsic and exogenous risk factors can become blurred. For example, the effects of exogenous factors, including life events; social rearing experiences, such as trauma or parental nurturance; and factors such as the use of illicit substances may operate through effects on intrinsic factors, such as the regulation of neural systems that monitor danger.

Intrinsic, individual-oriented vulnerability markers for anxiety disorders can be conceived across a range of perspectives, focusing on increasingly more specified biological systems. At the most complex or global level, specific temperamental or personality characteristics, such as neuroticism, harm avoidance, and behavioral inhibition have been linked to risk for anxiety. At a more specified level, vulnerability can be modeled through the assessment of cognitive function, in the form of attention and memory, or peripheral physiologic function, as reflected in autonomic reactivity profiles, changes in the startle reflex, or changes in ventilatory control. These cognitive and physiologic functions, in turn, reflect functional aspects of neurochemical or neuroanatomic systems that are presumably homologous with systems linked to fear and anxiety across a range of mammalian species. Data from humans at each of these levels is reviewed within the context of research on fear and anxiety in other species.

Temperament/Personality

Behavioral Inhibition

One of the earliest indicators of vulnerability to the development of anxiety is behavioral inhibition, characterized by increased physiologic reactivity or behavioral withdrawal in the face of novel stimuli or challenging situations (79). Behavioral inhibition may be a manifestation of a biological predisposition characterized by both overt behavioral (e.g., cessation of play, latency to interact in the presence of unfamiliar objects and people) and physiologic indicators (e.g., low heart rate variability, accelerated heart rate, increased salivary cortisol level, pupillary dilation, increased cortisol level). There is an increased frequency of behavioral inhibition among children of parents with anxiety disorders compared to those of normal controls (61 ,62 ,65 ,75 ,80 ,81 and 82).

Few studies have evaluated the differences in manifest inhibition and approach/avoidance in both clinical and nonclinical samples, leaving gaps in the conceptualization of the construct of inhibition. Some studies have shown that there is more stability of behavioral inhibition across early childhood among girls than among boys (83). The expression of behavioral inhibition studied prospectively may reveal patterns of anxiety symptomatology similar to those endorsed in adult populations. In a prospective study of a large community cohort of subjects from age 3 months to 13 years, Prior et al. (84) found that maternal ratings of persistent shyness and shyness in late childhood were associated with the development of anxiety disorders in adolescence.

Anxiety Sensitivity

Anxiety sensitivity is another potential sensitive and specific trait marker for the development of anxiety disorders (85). Anxiety sensitivity is characterized by beliefs that anxiety sensations are indicative of harmful physiologic, psychological, or social consequences (e.g., fainting or an impending heart attack). The misinterpretation of bodily cues that characterizes anxiety sensitivity may lead to a self-perpetuating “fear of fear” cycle. Thus, the fear of benign arousal sensations produces anxiety, which in turn increases the frequency and intensity of physiologic sensations, and subsequently fuels apprehension regarding the significance of these sensations. This process may ultimately result in a full-blown panic attack.

Anxiety sensitivity is thought to represent a stable trait-like factor that is qualitatively different from general fear and anxiety (86). It has been proposed that anxiety sensitivity may interact with environmental experiences (e.g., hearing misinformation about the negative outcome of certain bodily sensations) to shape beliefs about the dangers of anxiety sensations. Thus, anxiety sensitivity may be involved in the development of certain anxiety disorders, particularly panic disorder (87 ,88). Of particular interest is the finding of the specificity of anxiety sensitivity with respect to development of anxiety disorders but not depression in a nonclinical sample (88). Likewise, Pollock et al. (89) reported that anxiety sensitivity appears to be specific to anxiety, as it did not contribute unique variance above self-rated anxiety symptoms in the prediction of depressive symptoms.

Anxiety sensitivity has been shown to be under genetic (90) and familial influence; anxiety sensitivity was found to

constitute a potential premorbid marker for the development of anxiety disorders in high-risk but not low-risk youth (89). Prospective studies of youth have also demonstrated the prognostic significance of anxiety sensitivity in predicting the development of anxiety disorders. Based on the results of a 5-year prospective study of adolescents, Hayward et al. (91) concluded that anxiety sensitivity appeared to be a specific risk factor for the development of panic attacks in adolescents. These findings from prospective research, particularly the specificity with respect to anxiety, together with the importance of genetic and familial liability suggest that anxiety sensitivity is an important vulnerability factor that should be examined in future studies.

Comorbid Disorders

Psychiatric

The magnitude of comorbidity in adults and adolescents with anxiety suggests that investigation of the role of other disorders in enhancing the risk of the initial development and persistence of anxiety disorders over time may be fruitful. The difficulty in dating onset of specific disorders, particularly from retrospective data, diminishes our ability to determine the temporal relations between disorders. Nevertheless, some prospective studies have examined the links between anxiety disorders and earlier expression of other forms of psychopathology. For example, whereas some studies suggest that childhood depression may presage the onset of panic attacks, the results of a fairly large prospective study suggest a bilateral temporal association between panic attacks and depression (91).

Other disorders that may enhance the risk of development of anxiety disorders include eating disorders (92), depression, and substance use and abuse. With respect to substance use disorders, Rao et al. (93) found that anxiety disorders may comprise a mediator of the link between depression and the subsequent development of substance use disorders in a clinical sample. The potential mechanisms through which anxiety may be associated with smoking in adolescents were examined by Patton et al. (94), who found that both anxiety and depression were associated with smoking initiation through increased susceptibility to peer influences. Conversely, some research suggests that substance use may trigger anxiety disorders in susceptible youth. For example, a prospective study of a community sample revealed that posttraumatic stress disorder (PTSD) may be triggered by substance abuse in about 50% of the cases (95). Similarly, Johnson et al. (96) found that adolescent smoking predicted adult onset of panic attacks, panic disorder, and agoraphobia (96). Thus, although comorbidity between anxiety and both depression and substance problems is quite common in children and adolescents, further research on the mechanisms for links between specific disorders both across and within genders is necessary.

Medical Symptoms/Disorders

Several studies have also suggested that there is an association between childhood medical conditions and the subsequent development of anxiety. Kagan et al. (97) reported an association between allergic symptoms, particularly hay fever, and inhibited temperament in young children. In a retrospective review of pre- and perinatal and early childhood risk factors for different forms of psychiatric disorders in adolescence and early adulthood, Allen et al. (98) found that anxiety disorders in adolescents were associated specifically with illness during the first year of life, particularly high fever. Likewise, Allen and Matthews (102) reported that adolescents and young adults with anxiety disorders were more likely to have suffered from infections during early childhood than others. The prevalence of high fevers in childhood along with other diseases associated with immune system were also elevated among offspring of parents with anxiety disorders in the Yale High Risk Study (76). Kagan (101) proposed that the high levels of cortisol associated with anxiety may lead to immunologic sensitivity to environmental stimuli. Taylor et al. (99) reported that immunologic diseases and infections were specifically associated with emotional disorders because children with developmental or behavioral disorders had no elevation in infections or allergic diseases. On the other hand, Cohen et al. (100) suggest that such medical problems show stronger associations with depressive as opposed to anxiety disorders during adolescence. These findings suggest that it may be fruitful to examine links between immunologic function and the development of anxiety disorders.

Prospective studies have revealed that the anxiety disorders may comprise risk factors for the development of some cardiovascular and neurologic diseases. Haines et al. (103) reported that phobic anxiety was associated with ischemic heart disease, particularly fatal ischemic events. Bovasso and Eaton (104) employed cardiac and respiratory symptoms and illness to subtype panic attacks and their association with depression in a large community-based sample. They found that “respiratory panic attacks were associated with the subsequent risk of myocardial infarction.” Likewise, phobic disorder is strongly associated with migraine, with the onset of phobias predating that of migraine (105 ,106). The results of both family studies and prospective cohort studies suggest that there may be a subtype of migraine with shared liability for anxiety and depression (105).

Autonomic Reactivity

Reactions to threatening stimuli among various organisms, including primates and lower mammals, involve changes in the autonomic nervous system. These changes can be detected through an analysis of time series for heart rate, heart period variability, blood pressure, and catecholamine levels.

There is a long history of research in this area, and much of the initial work concerned the assessment of physiologic changes associated with acute anxiety states. Hence, acute episodes of anxiety, both in the laboratory and in natural settings, are typically characterized by acute changes in heart rate, blood pressure, and heart period variability (107). These changes result from coordinated changes in the parasympathetic and sympathetic innervation of the cardiovascular system.

More recent work on physiologic changes during acute anxiety states has attempted to identify specific physiologic patterns associated with one or another emotion. The identification of such emotion-specific patterns may provide insights on emotion-specific patterns of brain activity. For example, some forms of anxiety, such as acute panic, may be characterized by marked parasympathetic withdrawal in the face of sympathetic enhancement. Other emotions, such as anger, may be characterized by a distinct physiologic “finger print,” reflecting the involvement of distinct brain systems across emotions (108 ,109 and 110). In general, consistent associations are found across development between acute anxiety states and changes in peripheral autonomic indices, including heart rate, blood pressure, or heart period variability. As a result, some suggest that perturbations in autonomic regulation may index an underlying vulnerability to develop anxiety disorders. This underlying vulnerability is thought to relate to the functioning of particular neural circuits within the brain that exert effects on both subjective internal states and physiologic activity. Potentially relevant neural circuits have been identified through basic science studies on the neural basis of fear and anxiety.

Despite consistent evidence of an association between acute anxiety states and changes in autonomic physiology, the degree to which such changes index vulnerability for anxiety, as opposed to the acute state of anxiety, remains unclear. If such changes in autonomic physiology primarily reflect downstream manifestations of relatively high degrees of acute fear, they would provide limited advantages as vulnerability markers. On the other hand, at least some of the underlying autonomic abnormalities in panic disorder persist after remission and may be independent of the current state. This suggests that changes in autonomic physiology may mirror subtle person-specific differences in brain processes related to the processing of risk or to the experience of fear. As such, autonomic indices might index vulnerability in a fashion that is more sensitive than indices derived through self-report measures. A series of recent studies provide preliminary evidence consistent with this possibility.

Autonomic physiologic profiles have been studied among individuals who face high risk for anxiety disorders. Physiologic profiles have been tied to at least three indicators of risk: temperamental factors, family history, and traumatic events. In terms of temperamental factors, Kagan (111) noted the relationship between behavioral inhibition, which predicts later anxiety, and a distinct autonomic physiology profile. Children with behavioral inhibition exhibit an autonomic physiology characteristic of the profile found during acute anxiety. Specifically, behaviorally inhibited children exhibit under conditions of novelty a shift from parasympathetic to sympathetic control of the cardiovascular system, manifest as an increase in heart rate and a reduction in high-frequency components of the heart period variability power spectrum. Such abnormalities in autonomic physiology are viewed as downstream reflections of perturbations based within the limbic system. In terms of family history, Bellodi et al. (112) found similar temperamental and physiologic abnormalities among children of parents with panic disorder. Such data are consistent with other studies finding high rates of behavioral inhibition among offspring of patients with anxiety disorders. Finally, in terms of traumatic events, physiologic reactions to an acute stress may index underlying vulnerability to develop anxiety states. Consistent with this possibility, Shalev et al. (113) found that enhanced cardiovascular activity in the emergency room immediately following a motor vehicle accident predicted the development of PTSD.

Taken together, available data clearly delineate associations between acute anxiety and autonomic physiology profiles, but the implications of this work for the study of risk remain unclear. Moreover, the underlying assumption in this work posits an effect of perturbations in brain systems on both autonomic physiology and anxiety symptoms. As such, more work is also needed relating brain function to autonomic physiology.

Psychophysiologic Function

Research on fear conditioning has facilitated an integration of basic and clinical work on vulnerability for anxiety. Fear conditioning develops following the pairing of a neutral “conditioned” stimulus (CS+), such as a tone or a light, and an aversive “unconditioned” stimulus (UCS), such as a shock, a loud noise, or an air puff. Across a range of mammalian species, including humans, fear conditioning results from changes in a relatively simple neural circuit that involves distinct amygdala nuclei, including the basolateral and central nucleus. Basic science research on the role of this circuit in learned fears has also called attention to the role played by related but relatively distinct neural circuits in the responses to other forms of danger. For example, reactions to intrinsically dangerous contexts, such as a brightly lit room for a rodent, involve a relatively extended period of vigilance. These reactions may more intimately involve the basolateral nucleus and the bed nucleus of the stria terminalis than the central nucleus of the amygdala. Such reactions in animals may model worry in humans as characteristically found in many anxiety disorders (114).

Similarly, acute reactions to intrinsically dangerous stimuli often involve rapid changes in behavior designed to facilitate escape or defense. Such reactions in animals may involve the hypothalamus and lower brainstem structures; such reactions in animals may be model acute panic in humans.

Work in neuroscience delineating circuits involved in mammal's response to danger has stimulated a series of studies on risk for anxiety in humans. Much of this work quantifies physiologic reactions to innate and learned fears with the goal of comparing physiology across high- and low-anxiety groups. Based on skin conductance data, Eysenck and Eysenck (115) suggested that abnormal habituation of conditioned fear responses confers risk for anxiety. Similarly, Raine et al. (116) suggest that deficiencies in learned fear, as modeled by skin conductance, relates to low anxiety and high risk for chronic behavior problems. However, due to methodologic advantages, more recent studies rely on startle as a physiologic index of activity in brain circuits tied to fear. Most importantly, it has been possible to map circuits that regulate startle in more precise detail, relative to circuits that regulate skin conductance or other indicators of autonomic response, such as heart rate. Cross-species parallels in startle regulation facilitate integration of basic and clinical work (76 ,114 ,117 ,118 ,119 ,120 ,121 ,122 ,123 and 124). For example, molecular genetic studies on fear conditioning in mice generate specific hypotheses on the genetics of risk human disorders (125 ,126 ,127 ,128 ,129 ,130 and 131). Fear-relevant stimuli in animals may potentiate startle through effects on genes in limbic structures, such as the amygdala, that are involved in fear conditioning. In adult humans, distinct stimuli effect startle across emotional disorders, but abnormal startle in some form is seen in many disorders, including phobias (132), PTSD (122 ,123 and 124), depression (133 ,134), and panic disorder (135). Moreover, there is some evidence that startle specifically indexes risk for anxiety. Three studies found startle abnormalities in children born to adults with an array of anxiety disorders (76 ,119), and a fourth study found startle abnormalities in inhibited children, who face high risk for anxiety disorders (111).

In a high-risk study of offspring of parents with anxiety disorders compared to psychiatric and normal controls, the startle reflex and its potentiation by aversive states was used as a possible vulnerability marker to anxiety disorders in adolescent offspring of parents with anxiety disorders (122). Startle was found to discriminate between children at high- and low-risk for anxiety disorders, as well as to discriminate between children at risk for anxiety compared to those at risk for alcoholism. However, different abnormalities in startle amplitude for high-risk males and females were observed. Startle levels were elevated among high-risk females, whereas high-risk males exhibited greater magnitude of startle potentiation during aversive anticipation. Two possible explanations for the gender differences in the high-risk groups were suggested by the authors: (a) differential sensitivity among males and females to explicit threat versus the broader contextual stimuli that are mediated by different neurobiologic pathways, and (b) different developmental levels in males and females in which the vulnerability to anxiety may be physiologically expressed earlier in females.

Nevertheless, more work in this area is needed, given inconsistencies across genders and across conditions under which startle is most discriminatory (76). These data are also consistent with the findings of Watson et al. (136). Overall, the data suggest that startle indices may provide an important window for assessing dysfunction in limbic circuits broadly related to mood and anxiety regulation.

Ventilatory Function

As in the area of autonomic physiology, a wealth of research delineates associations between respiratory perturbation and acute anxiety. This association has been most convincingly demonstrated in panic disorder, where various forms of respiratory stimulation, including lactate infusion (137) and CO_2 inhalation, consistently produce high degrees of anxiety and more pronounced perturbations in respiratory physiologic parameters. Of note, these associations extend beyond the specific diagnosis of panic disorder, because enhanced sensitivity to respiratory perturbation is also found in conditions that exhibit strong familial or phenomenologic associations with panic disorder, including limited symptom panic attacks; certain forms of situational phobias; childhood anxiety disorders, particularly separation anxiety disorder; and high ratings on anxiety sensitivity scales.

Compared to the work on autonomic physiology, a larger body of research implicates abnormalities in respiration in risk or vulnerability for anxiety. At least four sets of findings suggest that respiratory indices index risk for anxiety, independent of any association between current state and respiratory function. First, asymptomatic adult relatives of patients with panic disorder consistently exhibit enhanced subjective sensitivity to respiratory stimulation, in the form of exogenously inhaled CO_2 (51 ,138 ,139). Second, among patients with panic disorder, stronger family loading is found in panic patients with evidence of respiratory dysregulation, as opposed to those with no sign of respiratory dysregulation (51 ,140). Third, respiratory indices linked to panic disorder are strongly heritable, raising questions on the potential shared genetic vulnerability for panic attacks and respiratory dysregulation. Fourth, Pine et al. (78) reported increased carbon dioxide sensitivity in children with anxiety disorders. Such data are also consistent with work on respiratory disease (141) and smoking (96 ,142), which suggest that abnormalities in respiration predispose to later anxiety. Based on this work, abnormalities in respiration appear to provide some information on the vulnerability for anxiety states that are related to acute panic.

Despite the consistency of findings in this area, a number of questions remain. The most consistent data emerge for

subjective indices of respiratory sensitivity, manifest as a tendency to report dyspnea during stress or during respiratory stimulation. The mechanisms that contribute to such enhanced sensitivity remain poorly specified. At a cognitive level, such hypersensitivity may result from an overall sensitivity to somatic sensations, consistent with data linking high degrees of anxiety sensitivity to future panic attacks (143). On the other hand, enhanced sensitivity to respiratory sensations appears more closely tied to panic attacks than sensitivity to other somatic factors; the tie between anxiety sensitivity and respiratory sensitivity also appears relatively weak in some studies. At the physiologic level, such hypersensitivity may result from perturbations in brain systems involved in respiratory regulation or primary as opposed to learned fear states. Unfortunately, the precise role of fear systems in both respiratory regulation and human anxiety states also remains poorly specified.

Neurochemical and Neurohormonal Factors

As reviewed in other sections of this book, extensive data document associations between alterations in various neurochemical factors and ongoing anxiety disorders. This includes data on the serotoninergic, noradrenergic, and GABAergic systems. Moreover, there is some evidence to implicate neurochemical alterations in the causal chain contributing to anxiety disorders. In animal models, genetic manipulations of serotoninergic receptors, the serotonin reputake transporter gene, and components of the GABA complex each produce behavioral and physiologic effects reminiscent of clinical anxiety states. Similarly, clinical studies find that acute pharmacologic manipulations in these neurochemical systems produce concomitant change in acute anxiety. For example, the inverse GABA agonist flumazenil precipitates anxiety in patients with panic disorder, whereas GABA agonists are potent treatments for various forms of anxiety. These findings are consistent with evidence of a deficiency in GABAergic modulation among adults with anxiety disorders. Similarly, manipulations of the serotoninergic system, either through tryptophan depletion or treatment with medications, also produce both acute and more chronic changes in anxiety. Finally, manipulations of the noradrenergic system produce similar changes in both children and adults. Sallee et al. (144) found that the α_2-agonist yohimbine elevated selectively anxiety symptoms and was associated with blunting of growth hormone in children with anxiety disorders. Interestingly, as with the response of children to CO_2 inhalation (160), the response to yohimbine appeared particularly abnormal in children with separation anxiety disorder. However, evidence of perturbed noradrenergic function in children with depression or facing high familial risk for depression (145) suggest that these findings may not be specific to anxiety but rather may relate to broad risk for mood and anxiety disorders.

Despite the consistency of these findings relating neurochemical factors to anxiety, relatively few studies have examined the manner in which individual differences in neurochemical function predict vulnerability to anxiety. There is evidence from studies in adult patients that some of these neurochemical abnormalities persist after remission. For example, much like symptomatic patients, remitted patients with panic disorder exhibit abnormal secretory profiles in terms of the growth hormone and the hypothalamic-pituitary-adrenal (HPA) axis. These neurohormonal abnormalities are thought to reflect trait-related abnormalities in neurochemical systems involved in neurohormonal regulation. Finally, there have been numerous studies of patients and at-risk relatives using lactate challenge to induce anxiety (146 ,147 and 148). The limited information provided on neural pathways by this provocation test limits its value in informing the pathophysiology of anxiety disorders.

Although these studies raise the possibility that risk for anxiety may result at least partially from underlying neurochemical abnormalities, other studies are needed to confirm this possibility. For example, there are almost no studies of neurochemical function in high-risk youth, a key source of information regarding the underlying role of biological parameters in the development of anxiety disorders. One exception is the study of Reichler et al. (77), who assessed several biological factors in their high-risk study of panic disorder including lactate metabolism, mitral valve prolapse, urinary catecholamines, and monoamine oxidase. Although none of these parameters discriminated high-risk from low-risk youth, the lack of differences may have been attributable in part to low statistical power.

Likewise, very few studies have compared neurochemical function in asymptomatic relatives of patients with and without anxiety disorders. Similarly, no studies have examined family loading for anxiety disorders in patients stratified in terms of their neurochemical functioning.

Beyond this work examining monamine systems' influence on neurohormonal regulation and vulnerability for anxiety, a relatively extensive body of work examines the precise relationship between anxiety and HPA axis regulation. Corticotropin-releasing factor (CRF) represents a key neuropeptide in the regulation of this system. CRF infusions in animals produce behavioral and physiologic effects in animals that bear similarities to human anxiety states. Similarly, genetic manipulations that alter CRF produce similar effects. As such, this work suggests that an underlying dysregulation in the HPA axis, possibly centrally involving CRF, may contribute to vulnerability for anxiety. Consistent with basic science studies, clinical research notes a relationship between acute anxiety states and alterations in HPA axis function. For example, a variety of acute stressors induce consistent elevations of cortisol; patients with PTSD exhibit multiple signs of HPA axis dysregulation; multiple

other anxiety disorders exhibit other signs of HPA axis dysregulation.

Vigilance/Attention

Studies of the association between attention regulation and anxiety have revealed that adults with anxiety disorders exhibit enhanced vigilance for threat cues, as indexed by effects of fear-related words or pictures on reaction times. These effects have been attributed to amygdala influences on attention allocation (149 ,150 ,151 ,152 and 153). Enhanced attentional bias in acute anxiety represents a particularly robust finding, noted in more than 20 studies using various paradigms across virtually all anxiety disorders. These effects appear particularly robust in two paradigms, the emotional Stroop and the dot-probe tests. From a theoretical perspective, this enhanced bias is considered a vulnerability marker that antedates the developmental of anxiety disorders among adults. Consistent with this possibility, an enhanced bias for threat cues is found early in the course of anxiety disorders, particularly among children with anxiety disorders. On the other hand, this enhanced bias is generally not found in remitted patients (153), and studies have yet to document enhanced bias for threat cues in at-risk but asymptomatic individuals.

ENVIRONMENTAL EXPOSURES

Part of "61 - Genetic and other Vulnerability Factors for Anxiety and Stress Disorders "

Perinatal Exposures

There is virtually no evidence that either prenatal factors or delivery complications comprise risk factors for the development of anxiety disorders. The results of three studies that retrospectively assessed perinatal events converged in linking such exposures to behavioral outcomes, but not to subsequent anxiety. For example, Allen et al. (98) found that children who suffered from a variety of exposures ranging from prenatal substance use to postnatal injuries were more likely to develop behavior disorders, particularly attention deficit disorder and conduct problems, but not anxiety disorders. Likewise, the results of the Yale High-Risk Study yielded no association between pre- and perinatal risk factors and the subsequent development of anxiety disorders (76).

Life Events/Stressors

The role of life experiences in the etiology of anxiety states, particularly phobias and panic disorder, has been widely studied (154 ,155 ,156 and 157). Life events have often been designated a causal role in the onset of phobias, which are linked inherently to particular events or objects. More broadly, life experiences that to some extent threaten one's notion of safety and security in the world are often at least retrospectively perceived to trigger or precipitate the onset of anxiety disorders. In evaluating the evidence on the causal role of life experiences, it is critical to consider separately the subtypes of anxiety disorders. Although it is likely that life stress may exacerbate phobic and generalized anxiety states, Marks (59) concludes that phobic states resulting from exposure are far more rare than those that emerge with no apparent exposure. In contrast, posttraumatic stress disorder (PTSD) is defined as a sequela of a catastrophic life event.

The major impediment to evaluation of the causal role of life events in anxiety (or depression) is the retrospective nature of most research addressing this issue. For example, Lteif and Mavissakalian (158) found that patients with panic or agoraphobia exhibited an increased tendency to report life events in general; this suggests that studies that limit assessment of life events to those preceding onset of a disorder may be misleading because they fail to provide comparison for the time period of onset. Moreover, stressful life events may interact with other risk factors such as family history of depression in precipitating episodes of panic (159). In one of the few prospective studies, Pine et al. (160) did demonstrate a predictive relationship between life events during adolescence and both depressive as well as generalized anxiety disorder symptoms. Interestingly, the association with anxiety was limited to females, consistent with differential vulnerability to stress across genders.

In terms of specific environmental risk factors, there has been abundant literature on the role of parenting in enhancing vulnerability to anxiety disorders. Based on Bowlby's (161) theory that anxiety is a response to disruption in the mother-child relationship, it has been postulated that maternal overprotection is related to anxiety, particularly separation anxiety. Using the Parental Bonding Instrument of Parker et al. (162), several studies of clinical samples have found that adult patients with anxiety disorders recall their parents as less caring and more overprotective than did controls (163). These findings have been supported in nonclinical samples as well (164 ,165). However, all of these studies caution that a causal link cannot be established because of the lack of independent assessment of parent behaviors and offspring anxiety.

Another parental behavior that may enhance risk of anxiety in offspring is parental sensitization of anxiety through enhancing cognitive awareness of the child to specific events and situations such as bodily functions, social disapproval, the importance of routines, and necessity for personal safety (164). Bennet and Stirling (164) found that subjects with anxiety disorders and those with high trait anxiety reported greater maternal and paternal overprotection and increased maternal sensitization to anxiety stimuli than controls.

Another feature of the parental relationship that has received widespread attention in recent research has been exposure to severe childhood trauma through either separation or abuse (161 ,166). There is increasing animal research on the impact of early adverse experiences on brain systems and subsequent development (167 ,168). Pynoos et al. (169)

present a comprehensive developmental life-trajectory model for evaluating the effects of childhood traumatic stress and anxiety disorders. They propose different avenues by which dangerous circumstances, childhood traumatic experiences, and PTSD can intersect with other anxiety disorders across the life span. The developmental perspective is critical in light of different levels of neural response to experience at different stages of development (170).

SUMMARY AND FUTURE DIRECTIONS FOR RESEARCH ON ANXIETY VULNERABILITY

Part of "61 - Genetic and other Vulnerability Factors for Anxiety and Stress Disorders "

Despite the rich array of constructs associated with anxiety (Table 61.4), our ability to predict those who will suffer from anxiety disorders in adulthood is severely limited. The principles of multifinality (i.e., many outcomes of the same risk factor) and equifinality (i.e., diverse risk profiles leading to the same endpoint) apply to many of the risk pathways investigated herein (171). Although we distinguish between intrinsic and extrinsic risk factors for the development of anxiety disorders, there is increasing evidence that there is a bidirectional association between the factors subsumed under these two domains. Only a small proportion of those with known vulnerability factors truly develop anxiety disorders in adulthood, despite the vast majority of those with adulthood anxiety reporting onset in childhood and early adolescence.

Individual
Genetic factors
Temperament
Behavioral inhibition
Anxiety sensitivity
Preexisting psychiatric/Medical disorder
Autonomic reactivity
Respiratory sensitivity
Neurobiological factors
Neuroendocrine factors
Exogenous
Exposure to stress
Drug use
Parenting
Modeling
Sensitization
Life events

TABLE 61.4. VULNERABILITY FACTORS FOR ANXIETY DISORDERS

The major impediments to identifying specific risk factors for anxiety are exclusive reliance on retrospective data, blurred boundaries between normal and pathologic anxiety, difficulty distinguishing between risk factors and early manifestations of anxiety, limited interdisciplinary conceptualization of models of risk and pathogenesis, lack of evidence of the specificity of risk factors with respect to anxiety disorders or subtypes thereof, and limited tools for direct measurement of brain function. Moreover, many of the risk factors have been shown to operate differently according to gender and age, as well as the specific subtype of anxiety. Elucidation of the different risk profiles will provide valuable information on classification, etiology, treatment, and prevention.

Future research should do the following:

- Establish more accurate and developmentally sensitive methods of assessment of anxiety, with a focus on developing objective measures of the components of anxiety.
- Apply within-family design to minimize etiologic heterogeneity and to refine diagnostic boundaries and thresholds.
- Investigate specificity of putative markers with respect to other psychiatric disorders and the longitudinal stability of specific subtypes of anxiety disorders.
- Develop research on hormonally mediated neurobiological function in order to understand gender differences predisposing women to experience decreased resiliency to fear-provoking stimuli.
- Examine mechanisms for associations between panic attacks with extrinsic exposures (i.e., substance use), developmental periods (i.e. pubertal development), and cessation in later life.

ACKNOWLEDGMENTS

Part of "61 - Genetic and other Vulnerability Factors for Anxiety and Stress Disorders "

This work was supported primarily by grant DA05348 and in part by grants AA07080, AA09978, DA09055, MH36197, Research Scientist Development Awards K02 DA00293 (to Dr. Merikangas), from Alcohol, Drug Abuse, and the Mental Health Administration of the United States Public Health Service.

REFERENCES

1. Pine DS, Cohen P, Gurley D, et al. The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. *Arch Gen Psychiatry* 1998;55:56–64.

2. Wittchen HU, Stien MB, Kessler RC. Social fears and social phobia in a community sample. *Psychol Med* 1999;29:309–323.

3. Marks I. The development of normal fear: a review. *J Child Psychol Psychiatry* 1987;28:667–697.

4. Gorman JM, Kent JM, Sullivan GM, et al. Neuroanatomical hypothesis of panic disorder, revised. *Am J Psychiatry* 2000;157:493–505.

5. Eaton WW, Dryman A, Weissman MM. Panic and phobia. In: Robins LN, Regier DA, eds. *Psychiatric disorders in America: the epidemiological catchment area study*. New York: Free Press, 1991:155–179.

6. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994;51:8–19.

7. Rhode P, Lewinsohn PM, Seeley JR. Comparability of telephone and face-to-face interviews in assessing axis I and axis II disorders. *Am J Psychiatry* 1997;154:1593–1598.

8. Brady EU, Kendall PC. Comorbidity of anxiety and depression in children and adolescents. *Psychol Bull* 1992;111:244–255.

9. Wittchen H-U, Lachner G, Wunderllich U, et al. Test re-test reliability of the computerized DSM-IV version of the Munich-Composite International Diagnostic Interview (M-CIDI). *Soc Psychiatry Psychiatric Epidemiol* 1998;33:568–578.

10. Lewinsohn PM, Lewinsohn M, Gotlib IH, et al. Gender differences in anxiety disorders and anxiety symptoms in adolescents. *J Abnorm Psychol* 1998;107:109–117.

11. Last CG, Perrin S, Hersen M, et al. DSM-III-R anxiety disorders in children: sociodemographic and clinical characteristics. *J Am Acad Child Adolesc Psychiatry* 1992;31.

12. Cohen P, Cohen J, Kasen S, et al. An epidemiological study of disorders in late childhood and adolescence—I. Age and gender-specific prevalence. *J Child Psychol Psychiatry Allied Disc* 1993;34:851–867.

13. Costello EJ. Developments in child psychiatric epidemiology. *J Am Acad Child Adolesc Psychiatry* 1989;28:836–841.

14. Compton SN, Nelson AH, March JS. Social phobia and separation anxiety symptoms in community and clinical samples of children and adolescents. *J Am Acad Child Adolesc Psychiatry* 2000;39:1040–1046.

15. Angst J, Vollrath M. The natural history of anxiety disorders. *Acta Psychiatr Scand* 1991;84:446–452.

16. Last CG, Perrin S, Hersen M, et al. A prospective study of childhood anxiety disorders. *J Am Acad Child Adolesc Psychiatry* 1996;35:1502–1510.

17. Cantwell D, Baker L. Stability and natural history of DSM-III childhood diagnoses. *J Am Acad Child Adolesc Psychiatry* 1989;28:691–700.

18. Ialongo N, Edelsohn G, Werthamer-Larsson L, et al. The significance of self-reported anxious symptoms in first grade children: prediction to anxious symptoms and adaptive functioning in fifth grade. *J Child Psychol Psychiatry Allied Disc* 1995;36:427–437.

19. Beekman ATF, de Beurs E, van Balkom AJLM, et al. Anxiety and depression in later life: co-occurrence and communality of risk factors. *Am J Psychiatry* 2000;157:89–95.

20. Horwath E, Weissman MM. Epidemiology of depression and anxiety disorder. In: Tsuang MT, Tohen M, Zahner GEP, eds. *Textbook in psychiatric epidemiology*. New York: Wiley-Liss, 1995.

21. Merikangas K, Swendsen J. Contributions of epidemiology to the neurobiology of mental illness. In Charney D, Nestler E, Bunney S, eds. *Neurobiology of mental illness*. New York: Oxford, 1999:100–107.

22. Kendler KS, Neale MC, Heath AC, et al. A twin-family study of alcoholism in women. *Am J Psychiatry* 1994;151:707–715.

23. Kendler KS, Eaves LJ, Walters EE, et al. The identification and validation of distinct depressive syndromes in a population-based sample of female twins. *Arch Gen Psychiatry* 1996;53:391–399.

24. Gorwood P, Feingold J, Ades J. Epidemiologie genetique et psychiatrie (I): portees et limites des etudes de concentration familiale exemple du trouble panique. *L'Encephale* 1999;10:21–29.

25. Goldstein RB, Wickramaratne PJ, Horwath E, et al. Familial aggregation and phenomenology of “early”-onset (at or before age 20 years) panic disorder. *Arch Gen Psychiatry* 1997;54:271–278.

26. McGuffin P, Asherson P, Owen M, et al. The strength of the genetic effect. Is there room for an environmental influence in the aetiology of schizophrenia? *Br J Psychiatry* 1994;164:593–599.

27. Kendler KS, Neale MC, Kessler RC, et al. Major depression and phobias: the genetic and environmental sources of comorbidity. *Psychol Med* 1993;23:361–371.

28. Skre I, Onstad S, Torgersen S, et al. A twin study of DSM-III-R anxiety disorders. *Acta Psychiatr Scand* 1993;88:85–92.

29. Kendler KS, Walters EE, Neale MC, et al. The structure of the genetic and environmental risk factors for six major psychiatric disorders in women: Phobia, generalized anxiety disorder, panic disorder, bulimia, major depression, and alcoholism. *Arch Gen Psychiatry* 1995;52:374–383.

30. Fyer AJ, Mannuzza S, Chapman TF, et al. A direct interview family study of social phobia. *Arch Gen Psychiatry* 1993;50:286–293.

31. Fyer AJ, Mannuzza S, Chapman TF, et al. Specificity in familial aggregation of phobic disorders. *Arch Gen Psychiatry* 1995;52:564–573.

32. Noyes R Jr, Clarkson C, Crowe RR. A family study of generalized anxiety disorder. *Am J Psychol* 1987;144:1019–1024.

33. Stein MB, Chartier MJ, Hazen AL, et al. A direct-interview family study of generalized social phobia. *Am J Psychiatry* 1998;155:90–97.

34. Merikangas KR, Angst J. Comorbidity and social phobia: Evidence from clinical, epidemiologic and genetic studies. *Eur Arch Psychiatry Clin Neurosci* 1995;244:297–303.

35. Woodman C, Crowe R. The genetics of the anxiety disorders. *Baillieres Clin Psychiatry* 1996;2:47–57.

36. Smoller JW, Tsuang MT. Panic and phobic anxiety: Defining phenotypes for genetic studies. *Am J Psychiatry* 1998;155:1152–1162.

37. Kendler KS, Neale MC, Kessler RC, et al. The genetic epidemiology of phobias in women. *Arch Gen Psychiatry* 1992;49:273–281.

38. Mendlewicz J, Papdimitriou G, Wilmotte J. Family study of panic disorder: comparison with generalized anxiety disorder, major depression and normal subjects. *Psychiatr Genet* 1993;3:73–78.

39. Pauls DL, Alsobrook JP, Goodman W, et al. A family study of obsessive-compulsive disorder. *Am J Psychiatry* 1995;152:76–84.

40. Nestadt G, Samuels J, Riddle M, et al. A family study of obsessive-compulsive disorder. *Arch Gen Psychiatry* 2000;57:358–363.

41. Black DW, Noyes RJ, Goldstein RB, et al. A family study of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:362–368.

42. Carey G, Gottesman II. Twin and family studies of anxiety, phobic, and obsessive disorders. In: Klein DF, Rabkin JG, eds. *Anxiety: new research and changing concepts*. New York: Raven Press, 1981:117–136.

43. Lenane MC, Swedo SE, Leonard H, et al. Psychiatric disorders in first-degree relatives of children and adolescents with obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry* 1990;29:766–772.

44. Bellodi L, Sciuto G, Diaferia G, et al. Psychiatric disorders in the families of patients with obsessive-compulsive disorder. *Psychiatry Res* 1992;42:111–120.

45. Goddard AW, Woods SC, Charney DS. A critical review of the role of norepinephrine in panic disorder: focus on its interaction with serotonin. In: Westenberg HGM, Den Boer JA, Murphy DL, eds. *Advances in the neurobiology of anxiety disorders*. New York: Wiley, 1996:107–137.

46. Wang ZW, Crowe RR, Noyes RJ. Adrenergic receptor genes as candidate genes for panic disorder: a linkage study. *Am J Psychiatry* 1992;149:470–474.

47. Schmidt SM, Zoega T, Crowe RR. Excluding linkage between panic disorder and the gamma-aminobutyric acid beta I reactor locus in five Icelandic pedigrees. *Acta Psychiatr Scand* 1993;88:225–228.

48. Nurnberger J, Byerley WF. Molecular genetics of anxiety disorders. *Psychiatr Genet* 1995;5:5–7.

49. Eaves LJ, Silberg JL, Meyer JM, et al. Genetics and development psychopathology: 2. The main effects of genes and environment on behavioral problems in the Virginia Twin Study of Adolescent Behavioral Development. *J Child Psychol Psychiatry Allied Disc* 1997;38:965–980.

50. Perna G, Bertani A, Caldirola L. Family history of panic disorder and hypersensitivity to CO2 in patients with panic disorder. *Am J Psychol* 1996;153:1060–1064.

51. Perna G, Cocchi S, Bertani A, et al. Sensitivity to 35% CO2 in health first-degree relatives of patients with panic disorder. *Am J Psychiatry* 1995;152:623–625.

52. Coryell W. Hypersensitivity to carbon dioxide as a disease-specific trait marker. *Biol Psychiatry* 1997;41:259–263.

53. Fyer AJ, Mannuzza S, Chapman TF, et al. Panic disorder and social phobia: effects of comorbidity on familial transmission. *Anxiety* 1996;2:173–178.

54. Maier W, Minges J, Lichtermann D. The familial relationship between panic disorder and unipolar depression. *J Psychiatr Res* 1995;29:375–388.

55. Merikangas KR, Stevens DE, Fenton B, et al. Comorbidity and familial aggregation of alcoholism and anxiety disorders. *Psychol Med* 1998;28:773–788.

56. Kushner MG, Abrams K, Borchardt C. The relationship between anxiety disorders and alcohol use disorders: a review of major perspectives and findings. *Clin Psychol Rev* 2000;20:149–171.

57. Thapar A, McGuffin P. Anxiety and depressive symptoms in childhood—a genetic study of comorbidity. *J Child Psychol Psychiatry* 1997;38:651–656.

58. Eley TC, Stevenson J. Exploring the covariation between anxiety and depression symptoms: a genetic analysis of the effects of age and sex. *J Child Psychol Psychiatry* 1999;40:1273–1282.

59. Marks IM. Genetics of fear and anxiety disorders. *Br J Psychiatry* 1986;149:406–418.

60. Turner SM, Beidel DC, Costello A. Psychopathology in the offspring of anxiety disorders patients. *J Consult Clin Psychology* 1987;55:229–235.

61. Biederman J, Rosenbaum JF, Bolduc EA, et al. A high risk study of young children of parents with panic disorders and agoraphobia with and without comorbid major depression. *Psychiatry Res* 1991;37:333–348.

62. Sylvester CE, Hyde TS, Reichler RJ. Clinical psychopathology among children of adults with panic disorder. In: Dunner DL, Gershon ES, Barrett JE, eds. *Relatives at risk for mental disorder*. New York: Raven Press, 1988:87–102.

63. Last C, Hersen M, Kazdin A, et al. Anxiety disorders in children and their families. *Arch Gen Psychiatry* 1991;48:928–934.

64. Warner V, Mufson L, Weissman M. Offspring at high risk for depression and anxiety: mechanisms of psychiatric disorder. *J Am Acad Child Adolesc Psychiatry* 1995;34:786–797.

65. Beidel DC, Turner SM. At risk for anxiety: I. Psychopathology in the offspring of anxious parents. *J Am Acad Child Adolesc Psychiatry* 1997;36:918–924.

66. Beidel DC. Psychophysiological assessment of anxious emotional states in children. *J Abnorm Psychol* 1988;97:80–82.

67. Capps L, Sigman M, Sena R, et al. Fear, anxiety and perceived control in children of agoraphobic parents. *J Child Psychol Psychiatry Allied Disc* 1996;37:445–452.

68. Merikangas KR, Dierker LC, Szatmari P. Psychopathology among offspring of parents with substance abuse and/or anxiety: A high risk study. *J Child Adolesc Psychiatry* 1998;39:711–720.

69. Unnewehr S, Schneider S, Florin I, et al. Psychopathology in children of patients with panic disorder or animal phobia. *Psychopathology* 1998;31:69–84.

70. Warner LA, Kessler RC, Hughes M, et al. Prevalence and correlates of drug use and dependence in the United States. *Arch Gen Psychiatry* 1995;52:219–229.

71. Stavrakaki C, Vargo B. The relationship of anxiety and depression: a review of literature. *Br J Psychiatry* 1986;149:7–16.

72. Merikangas KR. Comorbidity for anxiety and depression: Review of family and genetic studies. In: Maser JD, Cloninger CR, eds. *Comorbidity of mood and anxiety disorders*. Washington, DC: American Psychiatric Press, 1990:331–348.

73. Sylvester CE, Hyde TS, Reichler RJ. Clinical psychopathology among children of adults with panic disorder. *J Am Acad Child Psychiatry* 1987;26:668–675.

74. Weissman MM, Warner V, Wickramaratne P, et al. Offspring of depressed parents: ten years later. *Arch Gen Psychiatry* 1997;54:932–940.

75. Rosenbaum JF, Biederman J, Gersten M. Behavioral inhibition in children of parents with panic disorder and agoraphobia: a controlled study. *Arch Gen Psychiatry* 1988;45:463–470.

76. Merikangas KR, Avenevoli S, Dierker L, et al. Vulnerability factors among children at risk for anxiety disorders. *Biol Psychiatry* 1999;46:1523–1535.

77. Reichler RJ, Sylvester CE, Hyde TS. Biological studies on offspring of panic disorder probands. In: Dunner DL, Gershon ES, Barrett JE, eds. *Relatives at risk for mental disorders*. New York: Raven Press, 1988.

78. Pine DS, Cohen E, Cohen P, et al. Social phobia and the persistence of conduct problems. *J Child Psychol Psychiatry Allied Disc* 2000;41:657–665.

79. Kagan J, Reznick SJ. Shyness and temperament. In: Jones WH, Cheek JM, Briggs SR, eds. *Shyness: perspectives on research and treatment*. New York: Plenum Press, 1986.

80. Turner SM, Beidel DC, Dancu CV, et al. Psychopathology of social phobia and comparison to avoidant personality disorder. *J Abnorm Psychol* 1986;95:389–394.

81. Biederman J, Rosenbaum J, Bolduc-Murphy E, et al. A three-year follow-up of children with and without behavioral inhibition. *J Am Acad Child Adolesc Psychiatry* 1993;32:814–821.

82. Rosenbaum JF, Biederman J, Hirshfeld DR, et al. Behavioral inhibition in children: a possible precursor to panic disorder or social phobia. *J Clin Psychiatry* 1991;52:5–9.

83. Hirshfeld DR, Rosenbaum JF, Biederman J, et al. Stable behavioral inhibition and its association with anxiety disorder. *J Am Acad Child Adolesc Psychiatry* 1992;31:103–111.

84. Prior M, Smart D, Sanson A, et al. Does shy-inhibited temperament in childhood lead to anxiety problems in adolescence? *J Am Acad Child Adolesc Psychiatry* 2000;39:461–468.

85. Reiss S, Pererson RA, Gursky DM, et al. Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behav Res Ther* 1986;24:1–8.

86. McNally RJ. *Panic disorder: a critical analysis*. New York: Guilford, 1994.

87. McNally RJ. Psychological approaches to panic disorder: a review. *Psychol Bull* 1990;108:403–419.

88. Schmidt N, Lerew D, Jackson R. The role of anxiety sensitivity in the pathogenesis of panic: prospective evaluation of spontaneous panic attacks during acute stress. *J Abnorm Psychol* 1997;106:355–364.

89. Pollock RA, Carter AS, Dierker L, et al. Anxiety sensitivity in children at risk for psychopathology. *J Consulting and Clinical Psychology*; *in press*.

90. Stein MB, Jang KL, Livesley WJ. Heritability of anxiety sensitivity: a twin study. *Am J Psychiatry* 1999;156:246–251.

91. Hayward C, Killen JD, Kraemer HC, et al. Predictors of panic attacks in adolescents. *J Am Acad Child Adolesc Psychiatry* 2000;39:207–214.

92. Bulik CM, Sullivan PF, Fear JL, et al. Eating disorders and antecedent anxiety disorders: a controlled study. *Acta Psychiatr Scand* 1997;96:101–107.

93. Rao U, Ryan ND, Dahl RE, et al. Factors associated with the development of substance use disorder in depressed adolescents. *J Am Acad Child Adolesc Psychiatry* 1999;38:1109–1117.

94. Patton GC, Carlin JB, Coffey C, et al. Depression, anxiety, and smoking initiation: a prospective study over 3 years. *Am J Public Health* 1998;88:1518–1522.

95. Giaconia RM, Reinherz HZ, Hauf A, et al. Comorbidity of substance use and post-traumatic stress disorders in a community sampler of adolescents. *Am J Orthopsychiatry* 2000;70:253–262.

96. Johnson J, Cohen P, Pine DS, et al. Association between cigarette smoking and anxiety disorders during adolescence and early adulthood. *JAMA* 2000;284:2348–51 *in press*.

97. Kagan J, Reznich SJ, Clarke C, et al. Behavioral inhibition to the unfamiliar. *Child Dev* 1984;55:2212–2225.

98. Allen TJ, Moeller FG, Rhoades HM, et al. Impulsivity and history of drug dependence. *Drug Alcohol Depend* 1998;50:137–145.

99. Taylor BK, Casto R, Printz MP. Dissociation of tactile and acoustic components in air puff startle. *Physiol Behav* 1991;49:527–532.

100. Cohen H, Kotler M, Matar M, et al. Analysis of heart rate variability in posttraumatic stress disorder patients in response to a trauma-related reminder. *Biol Psychiatry* 1998;44:1054–1059.

101. Kagan J. *Galen's prophecy: temperament in human nature*. New York: Basic Books, 1994:261–262.

102. Allen M, Matthews K. Hemodynamic responses to laboratory stressors in children and adolescents: the influences of age, race, and gender. *Psychophysiology* 1997;34:329–339.

103. Haines AP, Imeson JD, Meade TW. Phobic anxiety and ischaemic heart disease. *Br Med J* 1987;295:297–299.

104. Bovasso G, Eaton W. Types of panic attacks and their association with psychiatric disorder and physical illness. *Compr Psychiatry* 1999;40:469–477.

105. Merikangas KR, Isler H, Angst J. Comorbidity of migraine and psychiatric disorders: results of the Zurich cohort study of young adults. *Cephalalgia* 1991;(suppl 11):308–310.

106. Swartz KL, Pratt LA, Armenian HK, et al. Mental disorders and the incidence of migraine headaches in a community sample. *Arch Gen Psychiatry* 2000;57:945–950.

107. Gorman JM, Sloan RP. Heart rate variability in depressive and anxiety disorders. *Am Heart J* 2000;140:77–83.

108. Davidson RJ, Abercrombie H, Nitschke JB, et al. Regional brain function, emotion and disorders of emotion. *Curr Opin Neurobiol* 1999;9:228–234.

109. Lang P, Bradley M, Fitzsimmons J, et al. Emotional arousal and activation of the visual cortex: an fMRI analysis. *Psychophysiology* 1998;35:199–210.

110. Ekman P, Levenson RW, Friesen WV. Autonomic nervous system activity distinguishes among emotions. *Science* 1983;223:1208–1210.

111. Kagan J. *Galen's prophecy*. New York: Basic Books, 1995.

112. Bellodi L, Battaglia M, Diaferia G, et al. Lifetime prevalence of depression and family history of patients with panic disorder and social phobia. *Eur Psychiatry* 1993;8:147–152.

113. Shalev AY, Sahar T, Freedman S, et al. A prospective study of heart rate response following trauma and the subsequent development of posttraumatic stress disorder. *Arch Gen Psychiatry* 1998;55:553–559.

114. Davis M. Are different parts of the extended amygdala involved in fear versus anxiety? *Biol Psychiatry* 1998;44:1239–1247.

115. Eysenck HJ, The conditioning model of neurosis. *Behav Brain Sci* 1979;2:155–199.

116. Raine A, Reynolds C, Venables PH, et al. Fearlessness, stimulation-seeking, and large body size at age 3 years as early predisposition to childhood aggression at age 11 years. *Arch Gen Psychiatry* 1998;56:283–284.

117. Davis M, Walker DL, Lee Y. Amgydala and bed nucleus of the stria terminals: differential roles in the fear and anxiety measured with the acoustic startle reflex. *Philos Trans R Soc Lond* 1997;352:1675–1687.

118. Grillon C, Ameli R, Woods SW, et al. Fear-potentiated startle in humans: effects of anticipatory anxiety on the acoustic blink reflex. *Psychophysiology* 1991;28:588–595.

119. Grillon C, Dierker L, Merikangas KR. Startle modulation in children at risk for anxiety disorders and/or alcoholism. *J Am Acad Child Adolesc Psychiatry* 1997;36:925–932.

120. Grillon C, Pellowski M, Merikangas KR, et al. Darkness facilitates the acoustic startle in humans. *Biol Psychiatry* 1997;42:453–460.

121. Grillon C, Dierker L, Merikangas KR. Fear-potentiated startle in adolescent offspring of parents with anxiety disorders. *Biol Psychiatry* 1998;44:990–997.

122. Grillon C, Morgan CA, Davis M, et al. Effect of darkness on acoustic startle in Vietnam veterans with PTSD. *Am J Psychiatry* 1998;155:812–817.

123. Grillon C, Merikangas KR, Dierker LC, et al. Startle potentiation by threat of aversive stimuli and darkness in adolescents: a multi-site study. *Int J Psychophysiol* 1999;32:63–73.

124. Grillon C, Morgan CA. Fear-potentiated startle conditioning to explicit and contextual cues in Gulf war veterans with posttraumatic stress disorder. *J Abnorm Psychol* 1999;108:134–142.

125. Maren S. Long-term potentiation in the amygdala: a mechanism for emotional learning and memory. *Trends Neurosci* 1999;22:561–567.

126. Crestani F, Lorez M, Baer K, et al. Decreased GABA-receptor clustering results in enhanced anxiety and a bias for threat cues. *Nature Neurosci* 1999;2:833–839.

127. Aiba A, Chen C, Herrup K, et al. Reduced hippocampal long-term potentiation and context-specific deficit in associative learning in mGluR1 mutant mice. *Cell* 1994;79:365–375.

128. Anagnostaras SG, Craske MG, Fanselow MS. Anxiety: at the intersection of genes and experience. *Nature Neurosci* 1999;2:780–782.

129. Caldarone B, Saavedra C, Tartaglia K, et al. Quantitative trait loci analysis affecting contextual conditioning in mice. *Nature Genet* 1997;17:335–337.

130. Gray JA, Flint J, Dawson GR, et al. A strategy to home-in on polygenes influencing susceptibility to anxiety. *Hum Psychopharmacol Clin Exposure* 1999;14:S3–S10.

131. Mayford M, Bach ME, Huang YY, et al. Control of memory formation through regulated expression of a caMKII transgene. *Science* 1996;274:1678–1683.

132. Lang PL, Bradley MM, Cuthbert BN. Emotion, motivation and anxiety: Brain mechanisms and psychophysiology. *Biol Psychiatry* 1998;44:1248–1263.

133. Allen NB, Trinder J, Brennan C. Affective startle modulation in clinical depression: preliminary findings. *Biol Psychiatry* 1999;46:542–550.

134. Cook EW, III, Hawk LW, Davis TL, et al. Affective individual differences and startle reflex modulation. *J Abnorm Psychol* 1991;100:5–13.

135. Grillon C, Ameli R, Goddard A, et al. Baseline and fear-potentiated startle in panic disorder patients. *Biol Psychiatry* 1994;35:431–439.

136. Watson D, Clark LA, Carey G. Positive and negative affectivity and their relation to anxiety and depressive disorders. *J Abnorm Psychol* 1988;97:346–353.

137. Cowley DS, Arana GW. The diagnostic utility of lactate sensitivity in panic disorder. *Arch Gen Psychiatry* 1990;47:277–284.

138. Coryell W, Arndt S. The 35% CO2 inhalation procedure: test-retest reliability. *Biol Psychiatry* 1999;45:923–927.

139. van Beek N, Griez E. Reactivity to a 35% CO2 challenge in healthy first-degree relatives of patients with panic disorder. *Biol Psychiatry* 2000;47:830–835.

140. Horwath E, Wolk SI, Goldstein RB, et al. Is the comorbidity between social phobia and panic disorder due to familial cotransmission or other factors? *Arch Gen Psychiatry* 1995;47:21–26.

141. Pine DS, Weese-Mayer D, Silvestri JM, et al. Anxiety and congenital central hypoventilation syndrome. *Am J Psychiatry* 1994;151:864–870.

142. Breslau N, Klein DF. Smoking and panic attacks: an epidemiologic investigation. *Arch Gen Psychiatry* 1999;56:1141–1147.

143. Schmidt NB, Lerew DR, Jackson RJ. Prospective evaluation of anxiety sensitivity in the pathogenesis of panic: replication and extension. *J Abnorm Psychol* 1999;108:532–537.

144. Sallee FR, Sethuraman G, Sine L, et al. Yohimbine challenge in children with anxiety disorders. *Am J Psychiatry* 2000;157:1236–1242.

145. Birmaher B, Ryan ND, Williamson DE, et al. Childhood and adolescent depression: a review of the past 10 years, part II. *J Am Acad Child Adolesc Psychiatry* 1996;35:1575–1583.

146. Cowley DS, Hyde TS, Dager SR, et al. Lactate infusions: The role of baseline anxiety. *Psychiatry Res* 1988;21:169.

147. Balon R, Pohl R, Yeragani V, et al. Comparison of lactate-induced panic attacks in panic disorder patients and controls. *Psychiatry Res* 1988;22:1.

148. Reschke AH, Mannuzza S, Chapman TF, et al. Sodium lactate response and familial risk for panic disorder. *Am J Psychiatry* 1995;152:277–279.

149. McNally RJ. Cognitive bias in the anxiety disorder. *Nebr Symp Motiv* 1996;43:211–250.

150. McNally RJ. Memory and anxiety disorders. *Philos Trans R Soc Lond [B]* 1997;352:1755–1759.

151. LeDoux JE. Fear and the brain: where have we been, and where are we going? *Biol Psychiatry* 1998;44:1229–1238.

152. Williams JMG, Mathews A, MacLeod C. The emotional Stroop task and psychopathology. *Psychol Bull* 1996;120:3–24.

153. Weinstein AM, Nutt DJ. A cognitive dysfunction in anxiety and its amelioration by effective treatment with SSRis. *J Psychopharmacol* 1995;9:83–89.

154. Faravelli C, Guerrini Degl'Innocenti B, Giardnelli L. Epidemiology of anxiety disorders in Florence. *Acta Psychiatr Scand* 1989;79:308–312.

155. Roy-Byrne PP, Mellman TA, Uhde TW. Biologic findings in panic disorder: neuroendocrine and sleep-related abnormalities. *J Anxiety Dis* 1988;2:17–29.

156. DeLoof C, Zandbergen J, Lousberg H, et al. The role of life events in the onset of panic disorder. *Behav Res Ther* 1989;27:461–463.

157. Last C, Barlow D, O'Brien G. Precipitants of agoraphobia: role of stressful life events. *Psychol Rep* 1984;54:567–570.

158. Lteif GN, Mavissakalian MR. Life events and panic disorder/agoraphobia: a comparison at two time periods. *Compr Psychiatry* 1996;37:241–244.

159. Manfro GG, Otto MW, McArdle ET, et al. Relationship of antecedent stressful life events to childhood and family history of anxiety and the course of panic disorder. *J Affect Disord* 1996;41:135–139.

160. Pine DS, Klein RG, Coplan JD, et al. Differential CO2 sensitivity in childhood anxiety disorders and non-ill comparisons. *Arch Gen Psychiatry* 2000;51:960–962.

161. Bowlby J. The making and breaking of affectional bonds. *Br J Psychiatry* 1960;130:201–210.

162. Parker G, Tupling H, Brown L. A parental bonding instrument. *Br J Med Psychology* 1979;52:1–10.

163. Silove D, Parker G, Hadzipavlovic D, et al. Parental representations of patients with panic disorder and generalised anxiety disorder. *Br J Psychiatry* 1991;159:835–841.

164. Bennet A, Stirling J. Vulnerability factors in the anxiety disorders. *Br J Med Psychology* 1998;71:31–321.

165. Lieb R, Wittchen H-U, Hofler M, et al. Parental psychopathology, parenting styles and the risk of social phobia in offspring: a prospective-longitudinal community study. *Arch Gen Psychiatry* 2000;57:859–866.

166. Stein MB, Walker JR, Anderson G, et al. Childhood physical and sexual abuse in patients with anxiety disorders and in a community sample. *Am J Psychiatry* 1996;153:275–276.

167. Lewis MH, Gluck JP, Petitto JM, et al. Early social deprivation in nonhumhan primates: long-term effects on survival and cell-meditated immunity. *Biol Psychiatry* 2000;47:119–126.

168. Heim C, Nemeroff CB. The impact of early adverse experiences on brain systems involved in the pathophysiology of anxiety and affective disorders. *Biol Psychiatry* 1999;46:1509–1522.

169. Pynoos RS, Steinberg AM, Piacentini JC. A developmental psychopathology model of childhood traumatic stress and intersection with anxiety disorders. *Biol Psychiatry* 1999;1999:1542–1554.

170. Spear LP. The adolescent brain and age-related behavioral manifestations. *Neurosci Biobehav Rev* 2000;24:417–463.

171. Sroufe LA. Psychopathology as an outcome of development. *Dev Psychopathol* 1997;9:251–268.

62

Animal Models and Endophenotypes of Anxiety and Stress Disorders

Vaishali P. Bakshi

Ned H. Kalin

Vaishali P. Bakshi and Ned H. Kalin: Department of Psychiatry, University of Wisconsin at Madison, Wisconsin Psychiatric Institute and Clinics, Madison, Wisconsin.

ANIMAL MODELS OF PSYCHIATRIC ILLNESS

Part of "62 - Animal Models and Endophenotypes of Anxiety and Stress Disorders "

Animal Models: Types Of Validity

An important criterion for developing animal models to study psychopathology involves establishing the validity of the model as a true representation of the process being studied. Generally, three types of validity are applied to animal models: face validity, construct validity, and predictive validity (1 ,2 and 3). Face validity refers to the outward similarity in appearance between the model and the illness. Construct validity, on the other hand, does not exclusively involve outward tangible signs of the modeled illness. Rather, it refers to the internal mechanism or state that underlies the illness. Finally, predictive validity refers to the ability of the animal model to identify therapeutic treatments for the illness. It should be noted that the different types of validity can be independent of each other; an animal model can possess predictive and construct validity without possessing face validity. Ideally, an animal model should possess both construct and predictive validity so that it may be used to understand the mechanisms and etiology of the disorder and also to identify promising treatments for the disorder.

Endophenotype Approach

Species differences in the manifestation of a particular internal state can cloud the usefulness of face validity in animal models. In addition, when considering a complex psychiatric illness, it is likely that several different symptom clusters contribute to the final pathologic condition; these different sets of symptoms may have different underlying substrates and thus may be ameliorated by different treatments. Therefore, it is difficult to come up with an animal model for an illness that meets the aforementioned criteria and also models the pathologic syndrome in its entirety. An alternative approach that has been used involves the modeling of discrete symptom clusters and physiologic alterations rather than the whole syndrome, with the assumption that what causes the symptoms contributes mechanistically to the illness. This general approach has involved the use of endophenotypes that may be related to a particular psychiatric disorder. The term *endophenotype* refers to a set of behavioral and/or physiologic characteristics that accompany a basic process that is altered in relation to the illness that is being studied (4). It is important to note that this more narrowly defined endophenotype approach does not necessarily have to capture specific symptoms that are a part of the clinical diagnosis, but rather may focus on a core process or function that is abnormal in the clinical population under study and that is thought to be related to the manifestation of the illness. For example, in the case of anxiety-related disorders, investigators have focused on studying the genetic, physiologic, and neurochemical correlates of fearful or anxious endophenotypes because a core aspect of anxiety-related disorders involves the aberrant expression of fearful responses to neutral or mildly stressful contexts (5). Thus, by identifying animals that display fearful endophenotypes, it is possible to study the neural substrates that contribute to this basic process that may underlie the development and expression of anxiety-related psychopathology.

Using endophenotypes that are based on core and basic processes rather than the entire illness offers certain advantages. Because the whole illness is not being modeled, the endophenotype approach affords greater possibility for construct and predictive validity in the model, and can incorporate species-specific manifestations of the core process being modeled. This approach may also make screening for genetic abnormalities associated with the disorder more fruitful, because the genetic factors associated with a very discrete

process (which could be mediated by a small number of genes) rather than an entire syndrome (which is likely caused by a complex set of interactions between multiple genes) is being studied (4 ,6). Moreover, heterogeneity within a diagnostic category could potentially dilute the strength of a sample population (i.e., not all patients with anxiety disorders are identical in their clinical presentation), and diminish the chances of identifying genes that contribute to the illness. Ideally, one might be able to generate several different endophenotypes for a particular disorder, and then study the genetic underpinnings of each of these separate core processes in order to identify a set of genes that might be implicated in that particular disorder. The definition and use of endophenotypes in animal models of psychiatric illness is a developing area. This chapter presents some promising candidates of animal models of fearful and anxious endophenotypes, and outlines some of the preliminary genetic factors that have been identified to contribute to the manifestation of these endophenotypes.

PUTATIVE ANIMAL ENDOPHENOTYPES OF STRESS AND ANXIETY: STUDIES OF FEARFUL TEMPERAMENT

Part of "62 - Animal Models and Endophenotypes of Anxiety and Stress Disorders "

Defensive Behaviors

In an attempt to understand the basic neural mechanisms underlying psychiatric conditions involving fear and anxiety, several groups have focused on identifying the neural substrates of defensive behaviors in animals. Defensive behaviors are exhibited by a wide array of species including rats, nonhuman primates, and humans in response to perceived threats from the environment, and are essential components of an organism's behavioral repertoire that ensure its protection and survival. Because organisms display defensive behaviors in reaction to threat, it is thought that the aberrant expression of defensive behaviors may represent a good example of a fearful endophenotype that would have relevance to stress and anxiety-related disorders. Although the specific behavioral responses that compose defensive behaviors are dependent on the environmental context and vary from species to species, a common element that unites this cross-species phenomenon is that defensive behaviors represent an organism's behavioral response to fear. Because defensive behaviors are expressed in response to an immediate threat, they characteristically supersede and interrupt the expression of other normal homeostatic behaviors such as feeding and reproduction that may be ongoing at the time of the perceived threat (7 ,8 ,9 ,10 and 11). One defensive response pattern expressed by many species is to inhibit all body movements and assume an immobile or freezing posture. This phenomenon of behavioral inhibition is effective in preventing detection and attack by predators (12 ,13), and may have special relevance for understanding psychopathology.

In nonhuman primates, defensive behaviors are composed of a constellation of responses that include vocalizations, freezing, fleeing, or defensive hostility and aggression. The particular set of responses that is emitted depends on, among other variables, the nature of the perceived threat (14 ,15). Studies of defensive behaviors in rhesus monkeys may provide valuable information that could aid in the understanding of fear and anxiety-related psychopathology in humans, because extreme fearful or defensive responses occur in dispositionally fearful humans who have an increased risk to develop psychopathology (16).

Psychiatric illnesses such as anxiety disorders and depression might involve the aberrant expression of defensive behaviors. In other words, pathologic anxiety could be conceptualized as the inappropriate expression of defensive or fear-related behaviors, consisting of either an exaggerated or overly fearful response to an appropriate context, or a fearful response to an inappropriate or neutral context. Although appropriate levels of defensive behaviors in response to environmental threats are adaptive and ensure survival, the overly intense or context-inappropriate display of fearrelated defensive behaviors may represent a liability that interferes with normal behavior and would likely contribute to certain forms of fear-related psychopathology. Thus, inappropriate or exaggerated expression of defensive behaviors may represent an important animal endophenotype of anxiety. An understanding of the specific neural substrates underlying the expression and regulation of defensive behaviors may therefore ultimately shed insight into the processes that become dysregulated in stress-related psychopathology. In defining animal endophenotypes relevant to anxiety, specific symptoms of a particular type of anxiety disorder are not being modeled, but rather the general phenomenon of hyperreactivity to mildly stressful stimuli is studied. The approach of modeling anxiety by studying defensive behaviors in animals has been described previously for rodent models (17 ,18). In the following sections, both primate and rodent analogues of stress hyperresponsiveness are described, with a particular emphasis on models of either the overly intense but context-appropriate expression of defensive behaviors or the normal but context inappropriate expression of defensive behaviors. Initially, various behavioral paradigms that have been used to measure an animal's level of defensive behavior are described, and subsequently, specific examples of fearful endophenotypes that have been identified using these tests are discussed.

Measuring Defensive Behaviors in Nonhuman Primates: Human Intruder Paradigm

One laboratory paradigm that has been developed to identify animals with fearful dispositions characterizes monkeys' fearful behavioral responses to a human intruder. In the human intruder paradigm (HIP), the monkey is placed by itself in a test cage where it remains for 30 to 40 minutes

while its behavior is recorded on videotape. A human intruder then enters the test area, representing a potential predatorial threat to the animal (14 ,15 ,19). The test session consists of three consecutive brief conditions: alone (“A,” animal left alone in cage); no eye contact (“NEC,” animal presented with the facial profile of a human standing 2.5 m away); stare (“ST,” animal presented with a human who faces it and engages it in direct eye contact). Typically, animals respond to the A condition by increasing their levels of locomotion and by emitting frequent coo vocalizations, which have been likened to the human cry and function to signal the infant's location and facilitate maternal retrieval (20 ,21). The NEC condition causes a reduction in cooing and an increase in behavioral inhibition, which functions to help the monkey remain inconspicuous in the face of a predator and is often manifested as hiding behind the food bin and freezing. The ST condition elicits aggressive (open-mouth threats, lunges, cage shaking, barking vocalizations) and submissive (lip smacking, fear-grimacing) behaviors that represent adaptive responses to the perceived threat of the staring experimenter. The different test conditions (A, NEC, ST) reliably elicit responses in young or adult laboratory-reared monkeys or in feral animals (14 ,19). Moreover, these context-specific defensive responses are not dependent on the gender of the intruder, and can also be elicited by showing the animal a videotape of the intruder (Kalin et al., unpublished data).

Behavioral Tests Used to Measure Fearful Endophenotype in Rodents

To identify fearful endophenotypes in rodents, a variety of behavioral paradigms have been employed. All of the paradigms in some manner provide an assessment of the rodent's level of defensive behavior, which is essentially thought to be an index of its level of fearfulness or anxiousness. The behavioral tests measure one of four general categories of stress-related behavior: approach-avoidance conflicts, conditioned fear, aggression, and punished responding conflicts. Detailed descriptions and protocols for these tests can be found in a recent review by File and colleagues (22).

Approach-Avoidance Conflicts

Briefly, all of the paradigms presented in this section measure the animal's ratio of approach versus avoidance behaviors by presenting a choice between an environment that is safe (usually a dark, enclosed, small space) and an environment that seems novel but risky (usually bright, wide open, large spaces). The entries into and amount of time spent in the safe environment relative to the risky environment are used as an index of the animal's stress level (an increase in exploratory behaviors toward and into the risky environment indicate a relatively low level of stress). A number of paradigms including the *elevated plus maze* (composed of safer closed, dark arms versus riskier open bright arms), the *open field* (consisting of a darker wall-bordered peripheral portion versus a brighter open center section), a *light-dark transition box* (consisting of an exploratorium divided into two halves, one that is dark and one that is bright), and a *defensive withdrawal* apparatus (composed of a small dark chamber that is inside of a brightly lit open field) have been frequently used and validated as paradigms that are sensitive to detecting shifts in an animal's approach-avoidance–based conflict (23).

Conditioned Fear

Behavioral tests that measure conditioned fear utilize basic principles of Skinnerian conditioning. Two frequently used paradigms to assess fear conditioning are conditioned freezing and fear-potentiated startle. *Conditioned freezing* is evaluated using a two-step procedure. First, during the training or conditioning phase, a stressful unconditioned stimulus (UCS, such as a foot shock) that elicits freezing is paired with a neutral stimulus that subsequently becomes a conditioned stimulus (CS). On the test day, the amount of freezing in response to the CS is assessed; animals that have not undergone the CS-UCS pairing do not normally freeze when the CS is presented, but animals that have learned to associate the CS with a foot shock show marked levels of freezing simply in response to this stimulus. The CS can either be a context (i.e., the environment in which the shock is delivered) or a discrete cue (e.g., a tone or light). The level of conditioned freezing is thought to correspond to the level of fear or anxiety that the animal is experiencing due to anticipation of a threat (24). In the case of *fear-potentiated startle*, the unconditioned startle response to a sudden stimulus (e.g., a loud noise burst) is measured in the presence and in the absence of a CS that has been paired previously with shock. The startle response is markedly increased when the startling stimulus occurs in the presence of the CS; this relative increase in startle magnitude is quantified, and serves as an index of the level of fear (thought to be elicited by a discrete cue as the CS) or anxiety (thought to be elicited by a contextual CS) that the animal may be experiencing (25).

Aggression and Social Behavior

Aggressive behaviors are emitted as part of the behavioral repertoire an animal displays when it encounters a threatening situation. The study of defensive aggressive behaviors has been summarized and reviewed by a number of investigators (26 ,27 and 28). Briefly, aggressive behaviors can be studied using a *resident intruder paradigm*, in which the offensive/agonistic responses (e.g., upright postures, attacks) of a male resident or the defensive responses (e.g., submissive posture, flight, freezing) of a male intruder are measured. Other

stress-related paradigms involve the study of affiliative behaviors and include the *social interaction* test in which approach toward and contact between two rats is measured (e.g., sniffing or grooming each other).

Punished Responding Conflict

The basic principle of punished responding tests is to present the animal with a situation in which a particular behavioral response results in both a rewarding outcome and an aversive outcome. The extent to which the animal exhibits the behavioral response during the conflict schedule is used as an index of its level of stress. For example, in the classic *Geller-Seifter* conflict test (29), rats are trained to press a lever for a food reward. Gradually, the bar press is also paired with a mild foot shock, and a stable rate of responding is established under the conflict schedule. Drugs can then be administered and evaluated for their ability to increase responding under the punished schedule. For example, benzodiazepines have been found to increase bar-pressing during the conflict schedule, putatively by decreasing the stress or anxiety induced by the aversive stimulus. Similarly, in the *Vogel punished drinking* paradigm (30), thirsty rats with access to a water bottle are periodically given mild electric shocks through the spout of the bottle; the extent to which licking is decreased is used as an index of stress.

INDIVIDUAL DIFFERENCES IN DEFENSIVE BEHAVIORS: NATURALLY OCCURRING FEARFUL ENDOPHENOTYPES

Part of "62 - Animal Models and Endophenotypes of Anxiety and Stress Disorders "

Primates

Several lines of evidence support the notion that an individual's level of defensive responding is a relatively stable trait characteristic (which in part may be derived from the nature of early postnatal maternal interactions, see below). Extreme individual differences detected early in life may be predictive of future psychopathology. For example, extremely inhibited children are at greater risk to develop anxiety and depressive disorders and are more likely to have parents that suffer from anxiety disorders (31 ,32 ,33 and 34). Moreover, behavioral inhibition in childhood (based on retrospective self-reports) is highly associated with anxiety in adulthood (35). Some of the physiologic correlates that have been observed in extremely inhibited children are elevated levels of the stress-related hormone cortisol (36) and greater sympathetic nervous system activity (37). In nonhuman primates, individual differences in defensive behaviors have been studied in an attempt to elucidate the neuroendocrine and neurobiological concomitants of extreme behavioral inhibition and to characterize a primate analogue of an anxiety-related endophenotype.

Marked individual differences among rhesus monkeys have been noted with regard to the intensity of context-specific defensive responses. These defensive responses have been characterized using the HIP (see previous section). For example, some monkeys tend to coo frequently during the A condition (in which the animal is isolated), whereas other same-aged animals engage in little or no cooing. Large individual differences have also been observed in the duration of NEC-induced freezing (in the presence of a human profile) and ST-induced hostility (in response to direct eye contact with the human intruder). Some animals freeze the entire length of the test period, whereas at the other extreme some never freeze and act relatively undisturbed by the human intruder. These individual differences in fear-related responses seen in the laboratory are similar to those that have been observed in rhesus monkeys who inhabit Cayo Santiago, a 45-acre island with approximately 1,000 free-ranging monkeys (Kalin et al., unpublished data). Importantly, it has been found that monkeys' individual differences in defensive responses are relatively stable over time, suggesting that the intensity of defensive behavior that is displayed reflects a trait rather than a state characteristic. It was initially demonstrated that the duration of NECinduced freezing behavior remained stable in 12 animals tested twice with an interval of 4 months (r = .94). Using a larger sample size, the stability of NEC-induced freezing was confirmed; ST-induced hostility was also found to be relatively stable (Kalin et al., unpublished data). Interestingly, significant correlations between the magnitude of the different types of defensive responses were not observed within an animal. Thus, monkeys that exhibited extreme levels of NEC-induced freezing did not necessarily display extreme levels of ST-induced hostility. This lack of correlation between different types of defensive responses suggests that cooing, freezing, and defensive hostility represent different and somewhat unrelated characteristics of animals' defensive styles. Pharmacologic data also support this notion. For example, manipulations of the opiate system affect A (alone condition)–induced cooing without affecting threat-induced freezing or hostility. Conversely, benzodiazepines reduce the threat-related behaviors, but have little effect on A-induced cooing (14).

Finally, to identify some of the mechanisms underlying these individual differences in defensive responding, the relationships between the stress-related hormone cortisol or asymmetric frontal EEG activity and individual differences in fearful behavior were examined. Thus, in 28 mother-infant pairs, it was found that in both mothers and infants freezing duration was significantly and positively correlated with baseline (nonstressed) cortisol levels (38). These data are consistent with findings from human studies demonstrating that extremely inhibited children have elevated levels of salivary cortisol (36 ,37), and is also consistent with findings in rodents that corticosterone (the rodent analogue of cortisol) is required for rat pups to develop the ability to freeze when threatened (39).

Extremely fearful monkeys (as identified by the HIP) also exhibit characteristic EEG patterns. In adult humans,

asymmetric right frontal brain activity has been associated with negative emotional responses (40). Our studies in rhesus monkeys have demonstrated similarities in this measure between monkeys and humans (41). Thus, it has been found that dispositionally fearful monkeys have extreme right frontal brain activity, paralleling the pattern of extreme right frontal activity in humans who suffer from anxiety-related disorders. In addition, it was found that individual differences in asymmetric frontal activity in nonhuman primates in the 4- to 8-Hz range are a stable characteristic of an animal (41 ,42). Furthermore, a significant positive correlation between relative right asymmetric frontal activity and basal cortisol levels in 50 one-year-old animals was found. As predicted, the more right frontal an animal was, the higher was its cortisol level. An extreme groups analysis revealed that extreme right compared to extreme left frontal animals had greater cortisol concentrations as well as increased defensive responses, such as freezing and hostility. The association between extreme right frontal activity and increased cortisol appeared to be long-lasting because the right frontal animals continued to demonstrate elevated cortisol levels at 3 years of age. These results are the first to link individual differences in asymmetric frontal activity with circulating levels of cortisol. This finding is important because both factors have been independently associated with fearful temperamental styles.

It has recently been found that cerebrospinal fluid (CSF) levels of corticotropin-releasing hormone (CRH), a peptide that mediates stress responses, are significantly elevated in monkeys that display exaggerated defensive responses to threatening stimuli (5). As stated before, these extreme individual differences in defensive behaviors are stable over time. Moreover, it was found that CSF CRH levels are also stable over time in rhesus monkeys. Finally, when comparing monkeys with extreme right frontal activity (that display exaggerated fearful responses) to those with extreme left frontal activity (that display low levels of fearful behaviors), the right frontal group was found to consistently have increased CSF CRH levels over a period of 4 years (5). Thus, it appears that extreme fearful behavioral responses in nonhuman primates are associated with increased levels of stress hormones such as cortisol and brain CRH, and also with extreme right frontal brain activity versus left frontal brain activity, a profile that has been found in humans suffering from stress-related psychopathology (43). Taken together, these findings suggest that in primates, a fearful endophenotype can be conceptualized as a constellation of hormonal, electrophysiologic, and behavioral characteristics. Studying species-specific defensive behaviors and their neuroendocrine and physiologic correlates offers a powerful approach for identifying animal correlates of anxiety.

Rodents

Extreme individual differences in the expression of stress-related defensive behaviors have also been noted in rodent species. The examination of naturally occurring genetic variations with regard to stress reactivity may have important implications for the elucidation of individual differences in sensitivity to stressful situations. One example of naturally occurring individual differences comes from the study of different rodent strains with regard to their level of stress-like behavioral responding to environmental stimuli. Because of the important role of the CRH system in regulating defensive behaviors induced by stressful or threatening situations, attention has been focused on identifying rat or mouse strains that display differential stress reactivity and different baseline levels of CRH gene expression. For example, it has been found that baseline levels of CRH messenger RNA (mRNA) are significantly higher in the amygdala of fawn-hooded rats compared to either Sprague-Dawleys or Wistars (44 ,45). Fawn-hooded rats have also been reported to exhibit exaggerated behavioral responses to stress such as enhanced freezing, leading to the suggestion that this strain may have utility as a model for endogenous stress-related CRH overexpression and anxiety. Strain differences, which essentially reflect differential genetic makeups, have also been found to influence the effects of acute environmental stressors on regulating CRH system gene expression. Thus, the stress of whole-body restraint produces a much larger increase in CRH mRNA levels within the hypothalamus of Fisher rats than in Wistars or Sprague-Dawleys (46 ,47). Similarly, the spontaneously hypertensive and borderline hypertensive strains of rats have increased basal and stress-induced levels of hypothalamic CRH mRNA compared to the Wistar and Sprague-Dawley strains (48 ,49 and 50).

In mice, it has been shown that the BALB/c strain is hyperresponsive to a variety of stressors compared to the C57BL/6 strain; BALB/c mice exhibit significantly higher avoidance of aversive areas in a light-dark transition test and an open field (51 ,52). These mice also show high levels of neophobia (53). Recent genetic mapping studies in these strains have revealed that these behavioral differences may be associated with differential levels of γ-aminobutyric acid receptor A ($GABA_A$) expression between the strains. For example, it has been found that BALB/c mice have significantly lower levels of benzodiazepine binding sites in the amygdala compared to C57BL/6 mice (54). As described below, alterations in the expression of $GABA_A$ receptors have been found to lead to increased anxiety-like behaviors in genetically modified mice (see CRH System Transgenic Mice).

Taken together, these findings indicate that different rodent strains, as a consequence of their distinct genetic makeups, display different baseline levels of gene expression within various systems that are known to regulate the expression of stress-induced defensive behaviors. The study of various rodent strains may thus help to identify the neurogenetic differences that contribute to individual differences in stress susceptibility, and thereby further characterize the interaction between genes and environmental

conditions in the etiology of anxiety. Although such information is useful, it remains to be determined whether or not the specific genetic differences identified above actually underlie the different behavioral effects. It is probable that a number of genes in addition to those described above are differentially expressed across different rodent strains. Which other genes differ across strains, and of these, which ones contribute to the behavioral profile? It is also unclear whether the differential gene expression patterns are the cause or the result of the different phenotypes observed in the separate strains. Future studies in which behavioral phenotypes are assessed after the application of novel gene targeting techniques to selectively disrupt or restore gene function in these rodent strains will aid in clarifying these issues.

MATERNAL DEPRIVATION: AN ENVIRONMENTAL MANIPULATION THAT CAN LEAD TO FEARFUL ENDOPHENOTYPES IN PRIMATES AND RODENTS

Part of "62 - Animal Models and Endophenotypes of Anxiety and Stress Disorders "

Converging lines of evidence from a number of species point to the importance of the early postnatal period, and in particular the bond between mother and infant, in the development of normal defensive behaviors and the putative emotional states underlying these behaviors. It has been observed that children who were placed in nurseries that lacked adequate social stimulation developed a syndrome of “protest, despair, and detachment” that may be analogous to an increase in defensive responses (55). Furthermore, recent reports suggest that children reared without appropriate nurturance can display neuroendocrinologic alterations and may develop long-term behavioral and emotional difficulties including an increased risk for stress-related psychiatric illness (56 ,57).

Perhaps the most significant environmental factor during the early development of mammals is the interaction between the infant and its mother. As described above, separation of an infant from its mother during this early developmental phase represents a significant stressor that markedly and negatively affects the subsequent emotional development of the infant (55 ,57). In fact, disruption of normal attachment behavior at critical developmental phases can, in a number of species, lead to marked and persistent disturbances in behaviors and brain systems that are thought to participate in the regulation of fear-related responses; this disruption may ultimately contribute to an individual's propensity to develop exaggerated or inappropriate defensive responses.

Altered maternal-infant interactions can lead to anxiety endophenotypes in nonhuman primates and rodents, thus identifying an environmental manipulation that can be used to create animal models of increased stress-related functioning. Indeed, a large body of work in monkeys and rats indicates that a number of deleterious and long-lasting effects are produced as a result of separating infants from their mothers prior to weaning. The notion that perturbations in the early postnatal environment might have enduring neuroendocrine, neurochemical, and behavioral effects was originally put forth several decades ago by Levine (58). It has since been demonstrated that a likely source of these alterations is a disruption of the interaction between mothers and pups (59 ,60).

Nonhuman Primates

The classic studies by Harlow and colleagues (20 ,61 ,62) of the effects of maternal separation in primates found that in addition to life-supporting nourishment, physical contact and comfort are necessary for primates' normal social and emotional development. During the first months of life, the attachment between mother and infant is intense, and as a consequence the infant remains in close proximity to its mother (61 ,63). Long-term maternal separation can result in profound alterations in stress-related behavioral responses in the separated offspring. Monkeys that have been separated from their mothers for prolonged periods during this time exhibit symptoms of enhanced defensive or fear-related behavioral responses into adulthood and appear socially withdrawn, a phenomenon that has led to the suggestion that the behavioral and neuroendocrine sequelae of maternal separation might provide a model for some of the dysfunction that is observed in anxiety disorders and depression (64 ,65 ,66 ,67 and 68).

Furthermore, neuroendocrine studies in rhesus monkeys indicate that an infant's stress hormone levels are negatively correlated with the number of offspring the mother had, suggesting that when mothers are less experienced, cortisol levels in their (early born) infants are high; elevated cortisol levels also correspond to increased fearful behavioral responses in the infants (38). Cortisol has been found to play an important role in mediating the development of defensive responses (69); thus, factors that were expected to affect infant primate cortisol concentrations were examined. It was found that maternal cortisol levels were moderately correlated with those of their infants (38). Interestingly, it was also found that maternal parity was negatively correlated with infant cortisol levels such that the current infants of mothers that previously had more offspring were likely to have lower cortisol levels. This finding indicates that a mother's past infant rearing and/or pregnancy experience may contribute to individual differences in infant baseline cortisol levels, and provides further support for the notion that the mother-infant interactions may be a critical factor in determining the future fearful disposition of the offspring (38). Although the precise mechanism for this interaction remains to be determined, it is likely that mothers with little rearing experience would interact differently with their infants than mothers with more experience.

Evidence for the notion that long-lasting dysregulation

of the CRH system may in part underlie the harmful consequences of early developmental stressors has been provided in a study of nonhuman primates that were exposed to adverse rearing conditions during infancy. Coplan and colleagues (70 ,71) found that CSF levels of CRH are basally and chronically elevated in adult bonnet macaques whose mothers were exposed for 3 months to an unpredictable variable foraging demand (VFD), in comparison to mothers confronted with either a high but predictable or low but predictable foraging demand. Infants reared by VFD-exposed mothers have been found to subsequently display abnormal affiliative social behaviors in adulthood (72). These findings are consistent with the recent results from this lab that indicate that CSF CRH levels are elevated in dispositionally fearful monkeys, and that this CRH elevation is a stable trait-like characteristic of fearful endophenotype (5).

Rats

These aforementioned findings in nonhuman primates support the notion that mother-infant interactions may be a critical factor in determining the future fearful disposition of the offspring. Maternal separation has also been found to produce long-term changes in defensive behaviors into adulthood in rats. Using the maternal separation paradigm in rats, investigators have also been able to begin to elucidate some of the alterations in gene expression that take place in response to this early life stressor.

Interestingly, the nature of the separation determines the direction of the long-term changes, as has been reviewed in detail recently (73 ,74 and 75). Thus, brief periods of separation (3 to 15 minutes per bout, once a day, for roughly 2 weeks) from the mother result in a profile indicative of diminished anxiety, whereas more protracted separations (3 hours or more) have the opposite effect, resulting in increased stress-like responses. In an elegant series of studies by Plotsky and Meaney (76), the long-term effects of these different types of maternal separation have been described, and the behavioral and neuroendocrine mechanisms underlying these long-term effects have been characterized. It was initially found that rat pups that underwent very short periods of separation (termed “handling”) from their mothers had decreased basal levels of hypothalamic CRH mRNA and median eminence CRH immunoreactivity as adults compared to undisturbed control rats. As adults, these “handled” pups also displayed significantly lower elevations of stress-induced corticosterone levels and blunted CRH release from the median eminence relative to controls. It has since been found that the mechanism underlying this reduction in stress-related functioning in handled rat pups involves the type of maternal behavior that is displayed after the pups are returned to the mother (77), confirming earlier hypotheses that maternal behavior is the critical component in the developmental milieu of the infant (58). A brief removal of rat pups from the dam results in a significant increase in the amount of licking, grooming, and arched-back nursing (LG-ABN) that the mother lavishes on the pups when they are returned; the total amount of time spent nursing and contacting the offspring is not affected, but rather the quality of the interaction between mother and pup is altered. In nonseparated pups, individual differences in LG-ABN predict hypothalamic-pituitary-adrenal (HPA) axis responsivity in adulthood such that mothers that engage in high levels of LG-ABN have offspring that, as adults, show reduced HPA axis activation in response to stress and have decreased levels of CRH mRNA in the paraventricular nucleus (PVN) of the hypothalamus (77). Pups that are born to mothers that naturally exhibit high levels of LG-ABN grow up into adults that display low-anxiety–like behaviors (increased exploration of novel environments) and compared to low–LG-ABN offspring, have decreased levels of CRH receptors in brain regions such as the locus coeruleus that are thought to mediate stress responses (78). Taken together, these findings indicate that increased nurturing physical contact from the mother can lead to a toned-down stress-responsive system in the offspring.

In contrast, longer periods of maternal separation seem to have the opposite effect on stress-related functioning later in life. Rat pups that are separated from the mother for 3 hours or longer (investigators have often used a 24-hour separation) show in adulthood increased CRH system gene expression, exaggerated HPA axis responses to stress, and increased stress-like behaviors in paradigms such as the elevated plus maze (76 ,79 ,80). Other intense stressors such as an endotoxin insult during the perinatal stage are also able to produce marked elevations in basal CRH gene expression and lead to an exaggerated stress-induced HPA axis response in adulthood (81). It has accordingly been hypothesized that the perinatal environment plays a critical role in “programming” or “setting” the animal's stress coping system (perhaps through alterations in CRH system gene expression) for the remainder of its life (73 ,74 and 75). Maternally separated rats also show alterations in other systems that are known to regulate stress-related behaviors and that are consistent with an increased fearful endophenotype. For example, maternal separation increases the release of norepinephrine into the PVN of the hypothalamus in response to restraint stress; stress-induced plasma adrenocorticotropic hormone (ACTH) levels were also elevated in maternally deprived rats (82). Early life stressors such as maternal separation may therefore play an important role in determining the eventual stress-related endophenotype that is exhibited in adulthood. Moreover, the aforementioned studies provide an example of how the animal endophenotype approach can be applied to investigating molecular correlates of anxiety-related conditions.

It should be mentioned that prenatal stress can also produce alterations in indices of stress-induced responding in adulthood. For example, in rats, disturbing the prenatal environment by stressing the mother can lead to increases in

CRH gene expression in the fetal PVN, increases in CRH content in the amygdala of adult offspring, and potentiation of stress-like behavioral responses in these rats whose mothers had undergone stress during pregnancy (83 ,84 and 85). These findings further support the notion that mother-infant interactions may be a critical factor in determining the future fearful disposition of the offspring.

TARGETED MUTATIONS LEADING TO ANXIETY-LIKE ENDOPHENOTYPES: STUDIES OF TRANSGENIC MICE

Part of "62 - Animal Models and Endophenotypes of Anxiety and Stress Disorders "

A perhaps more direct approach for studying the genetic underpinnings of a particular animal endophenotype is to characterize the change in an organism's interaction with its environment following either overexpression or underexpression of a particular gene product. Transgenic and knockout mice are thus now widely used in the ongoing effort to understand the contributions of specific genes to psychopathology. The detailed methodology for the generation of these animals and their use in neuroscience research has been reviewed (86). Briefly, genetic alterations are introduced in the embryonic stage such that the mouse develops with the mutation, thereby putatively providing a model for congenital abnormalities that may contribute to anomalous functioning and the expression of a particular endophenotype.

Using this strategy, a variety of components within the CRH, serotonin (5-hydroxytryptamine, 5-HT), and GABA systems have been successfully targeted and studied for their roles in mediating stress-related behavioral effects (87 ,88 and 89). It should be noted that in addition to the aforementioned systems, there are several other important central regulators of stress and anxiety-related processes. The norepinephrine system has long been implicated in the modulation of anxiety states. Several recent reviews detail the preclinical and clinical evidence for the involvement of norepinephrine (NE) in anxiety-related disorders such as panic and posttraumatic stress disorder (90 ,91). Indeed, ß-adrenergic receptor antagonists and α_2-receptor agonists are effective in the treatment of certain anxiety-related symptoms in humans. In addition, recent preclinical evidence indicates that a variety of central nervous system (CNS) peptides including cholecystokinin (CCK), neuropeptide Y (NPY), and substance P/neurokinins participate in the regulation of anxiety-like behaviors (92 ,93 and 94).

Although these peptide and neurotransmitter systems undoubtedly play a role in stress- and anxiety-related behaviors, the focus of the following section will be the CRH, 5-HT, and GABA systems because these three systems are perhaps the most thoroughly studied with transgenic models. Although the transgenic/knockout approach has provided valuable new information about the genetic regulation of stress-related fearful endophenotypes, it should be kept in mind that there are a number of important caveats regarding the interpretation of findings from transgenic animal studies (described at the end of this section). As outlined below, alterations of discrete genes within each of these systems results in an anxiety-like murine endophenotype, characterized by the increased expression of certain aspects of rodent defensive behaviors.

The CRH System

CRH and the related endogenous peptide agonist urocortin (95) bind to the two cloned CRH receptors, designated CRH1 and CRH2 (96 ,97 and 98), and to the CRH binding protein (CRH-BP) (99). The CRH-BP has been postulated to function as an endogenous buffer for the actions of the CRH family of ligands at their receptors (100). CRH, its receptors, and its binding-protein are expressed in key structures of the HPA axis, and thereby participate in mounting the neuroendocrine response to environmental perturbations. The various elements of the CRH system are widely and heterogeneously expressed in cortical, limbic, and brainstem structures and these regions are thought to regulate behavioral responses to stress.

CRH System Transgenic Mice

Given that central infusion of CRH results in enhanced fear-related defensive behaviors, CRH-overexpressing mice were predicted to display increased stress-like behavioral responses. Indeed, these mice have been found to have several behavioral effects associated with acute CRH administration. CRH overexpressing mice exhibit a profile that is consistent with increased levels of stress, such as reduced baseline and stress-induced exploration of a novel environment, and decreased activity and time spent in the open arms of an elevated plus maze (101 ,102). These effects are potently blocked by administration of the CRH receptor antagonist α-helical CRH. CRH transgenic mice also show a profound decrease in sexual behaviors and significant deficits in learning; higher order functions such as these are typically abolished when a situation is found threatening and defensive behaviors are recruited (103 ,104). Thus, CRH overexpressing transgenic mice may represent a genetically engineered model of a murine anxiety-like endophenotype.

Consistent with the notion that heightened CRH transmission elicits stress-like behaviors is the finding that CRH-BP knockout mice show increased stress-like behaviors. CRH-BP knockout mice displayed decreases in open arm entries and open arm time in an elevated plus maze, and showed a decrease in the number of exits from a safe box in a defensive withdrawal/open field paradigm (105). These results indicate a heightened level of neophobia in these mice. Moreover, CRH-BP knockouts have reduced body weight gain over several weeks (105 ,106), which is also syntonic

with increased basal CRH activity. The behavioral profile of CRH-BP knockouts is similar to that which is seen with exogenous CRH administration (107), and supports the notion that removal of the CRH-BP may lead to increased basal stress responses due to increased CRH tone. Thus, as with the CRH-overexpressing mice, CRH-BP knockouts may represent another rodent anxiety-like endophenotype. In terms of the predictive validity of these models, CRH receptor antagonists have been found to block the stress-like behavioral profile observed in these animals; one recent clinical study indicates that CRH1 receptor antagonists may indeed prove to be effective anxiolytics or antidepressants (108).

Interestingly, deletion of the CRH gene does not appear to decrease stress-like behaviors, as might be predicted from the aforementioned work. Although certain endocrinologic deficits are observed in CRH knockout mice, stress-related behavioral function in these animals remains relatively unaffected as assessed by multiple stress-related paradigms (109 ,110 ,111 ,112 ,113 and 114). This sparing of normal stress responsivity may be due to compensatory increases in the expression of other CRH system ligands such as urocortin. Deletion of the CRH1 receptor gene, however, does appear to consistently result in a putative reduction in anxiety (115 ,116 and 117). For example, CRH1 knockout mice show increased exploration of the open arms on an elevated plus maze and spend more time in the brightly lit compartment of a dark-light transition box than do wild-type controls. Moreover, CRH1 knockout mice appear to be immune to the anxiogenic effects of ethanol withdrawal (117). Studies of CRH2 receptor knockout mice, on the other hand, indicate that these mice display a less consistent behavioral profile than the CRH1 knockout mice (118 ,119 and 120). Part of the behavioral profile of CRH2 knockouts is suggestive of increased stress-like responding, but other aspects of the behavioral profile indicate either no alteration of stress-related responding (118 ,119), or a decrease in anxiety-like behaviors (120). The observed increases in anxiety-like behaviors in these genetically altered mice may be due to increased levels of brain CRH and/or urocortin; in two of the three studies, an elevation of baseline CRH or urocortin mRNA levels in the CNS was seen in CRH2 knockout mice (118 ,119). Thus, the endophenotype displayed by CRH2 knockout mice may actually be indirectly due to a compensatory alteration induced by the mutation rather than simply due to a lack of CRH2 receptor expression. It should be noted that acute blockade of CRH2 receptors results in a decrease in stress-induced defensive behaviors; thus the behavioral profile of these animals is opposite to that of mice that are missing the CRH2 receptor (121 ,122). Thus, the timing of the gene deletion may critically influence the nature of the behavioral phenotype that ensues. Future studies utilizing novel inducible-knockout technologies may help in clarifying the developmental versus acute role of various genes in the development of anxiety-related endophenotypes (123).

Clinically Effective CRH System Drugs for Stress-Related Disorders

A large body of preclinical literature indicates that CRH is a critical modulator of stress and anxiety-like behaviors in nonhuman primates and rodents (107 ,124). Based on the ability of CRH1 receptor-selective antagonists to block many of the behavioral effects of stress or CRH administration, these antagonists have been proposed as potentially therapeutic agents for the treatment of stress-related psychiatric conditions including anxiety and depression (125). Pharmacologic analysis of stress-induced primate defensive responses has also revealed that the CRH system is a critical modulator of this index of anxiety-related behavior. For example, administration of CRH into the cerebral ventricles of nonhuman primates results in a constellation of behavioral responses that closely resemble the defensive responses that are exhibited upon presentation of a stressor (124). Consistent with the notion that increased levels of CRH are associated with increased anxiety-like responding are the recent findings that small-molecule CRH1 receptor antagonists block the expression of some behavioral, physiologic, and neuroendocrine responses to stressors in rhesus monkeys (126 ,127). The first report of an open-label clinical trial with a CRH1 antagonist was recently published, and revealed a significant effect of this compound in ameliorating symptoms of depression and anxiety (108). Although further research is needed to firmly establish the utility of CRH1 antagonists as psychotherapeutic agents and also to determine the possible side effects associated with their use, these preliminary data support the notion that these compounds represent an important new class of drugs that may offer great promise for the treatment of illnesses associated with increased anxiety and stress.

The 5-HT System

Serotonin is a member of the monoamine family of transmitters that also include dopamine and norepinephrine. As is typical for the monoamines, cell bodies for this neurotransmitter are found in discrete nuclei within the midbrain (dorsal and medial raphe nuclei) and send widespread 5-HT–containing projections throughout the brain (128). 5-HT produces its effects through at least 15 different 5-HT receptors that are differentially distributed throughout the CNS; the principal mode of 5-HT inactivation is cellular reuptake via terminal transporter proteins (129). The 5-HT system has long been implicated in the regulation of mood states and anxiety, and selective serotonin reuptake inhibitors (SSRIs) constitute a major class of antidepressants that have anxiolytic effects. As outlined below, 5-HT transmission also plays a critical role in the regulation of anxiety-like

behaviors. Given the plethora of 5-HT receptors and the paucity of highly selective ligands for these multiple target sites, several investigators have employed murine gene targeting strategies to elucidate the roles of specific 5-HT receptors in the regulation of stress and anxiety.

5-HT System Transgenic Mice

Studies of targeted gene deletions within the 5-HT system have revealed an important role for this system in the regulation of stress and anxiety-related behaviors in mice. The behavioral sequelae of disrupting 5-HT receptor gene expression have been elegantly summarized in several review articles (89 ,130 ,131 and 132). Perhaps the best-characterized 5-HT mutant mice are the 5-HT_{1A} and 5-HT_{1B} receptor knockouts. Mice with a mutation in the 5-HT_{1A} receptor gene have been found to display increased stress-like behaviors in multiple tests of approach-avoidance conflicts. These animals show decreased entries into and time spent in the more aversive region in paradigms such as the open field, elevated plus maze, and the elevated zero maze; thus 5-HT_{1A} knockout mice avoid the center of an open field, the open arms of a plus maze, and the unenclosed regions of a zero maze (133 ,134 and 135). It is worth noting that this “increased anxiety” pattern of results was found consistently across three different research labs, indicating its robustness and reproducibility. Consistent with this profile is the finding that these mice also exhibit decreased activity in the presence of and approach toward a novel object (135). This increase in stress-like responding is not accompanied by changes in overall locomotor activity or motor and spatial coordination, as assessed in photocell cages and a RotaRod apparatus. Curiously, 5-HT_{1A} knockout mice display increased mobility in response to an acute stressor such as forced swimming or tail suspension (133 ,134 and 135). Taken together, these findings indicate that 5-HT_{1A} knockout mice may represent another animal endophenotype of increased anxiety.

The constitutive knockout of the 5-HT_{1A} receptor does not seem to lead to compensatory alterations in the expression of serotonin or its transporter, or to changes in catecholamine levels in several brain regions (135). Interestingly, a recent report indicates that this mutation alters GABA system expression and function (136). It has been found that anxiety-like behaviors in 5-HT_{1A} knockout mice are relatively unaffected by benzodiazepine treatment. Analysis of brain tissue from these animals indicates that $GABA_A$ receptor binding is reduced and that the expression of α_1 and α_2 subunits of the $GABA_A$ receptor are decreased in the amygdala. The anxiolytic actions of benzodiazepines may in part be mediated by $GABA_A$ receptors within the amygdala; the profile of results in 5-HT_{1A} knockout mice has led to the intriguing speculation that the anxiety-like endophenotype in these mice may actually in part derive from a decrease in the expression and function of the GABA*A* receptor (136). This proposed mechanism is consistent with the increase in stress-related behaviors that are seen in certain transgenic mice with mutations in the GABA*A* receptor (see below). Further work is necessary to determine the precise mechanisms through which the developmental interruption of 5-HT_{1A} gene expression results in the observed anxiety-like endophenotype.

In contrast to 5-HT_{1A} knockout mice, mice that lack the 5-HT_{1B} receptor show decreased anxiety-like behaviors in several tests of approach-avoidance conflicts. 5-HT_{1B} knockout mice spend more time in the center of an open field and more readily explore novel objects than their wild-type controls; this profile is opposite from that of 5-HT_{1A} knockout mice, and is suggestive of diminished neophobia (89 ,137). Consistent with this pattern of results is the finding that as pups, 5-HT_{1B} mice emit fewer ultrasonic vocalizations when separated from their mothers; separationinduced vocalizations are thought to provide a measure of anxiety and distress in pups (138 ,139). It is interesting to note, however, that no changes in contextual or cue-induced conditioned freezing are observed in 5-HT_{1B} mutant mice, suggesting that approach-avoidance conflicts and conditioned fear may be differentially modulated by the 5-HT system. The other main behavioral effect of constitutive 5-HT_{1B} receptor deletion is a marked increase in aggressive behavior (89 ,140 ,141). Given that aggressive behaviors represent an important part of an organism's response to threat, 5-HT_{1B} knockout mice may also provide valuable information on the neural and genetic factors associated with stress and anxiety-related functioning (89 ,142).

It should be noted that mice with null mutations of other 5-HT receptor subtypes have also been generated, but these animals have not been found to display as robust an anxiety-related behavioral profile as the 5-HT_{1A} or 5-HT_{1B} knockout mice. It has been found that 5-HT_{5A} receptor knockout mice show increased exploratory activity in the presence of novelty, but do not differ from wild-type controls with regard to avoidance behaviors from an aversive environment such as the open arms of a plus maze, or the center of an open field (143). These knockout mice also do not respond differently from control subjects in tests of startle reactivity or in burying a probe that delivered a brief electric shock. Thus, these animals appear to have yet a different behavioral profile from that of the 5-HT_{1A} or 5-HT_{1B} knockout mice. An initial report indicates that 5-HT_6 receptor deficient mice may exhibit increased avoidance of aversive environments; although these preliminary findings are interesting, further work is needed to fully characterize the phenotype of these mutant mice (144 ,145). Mice lacking either the 5-HT_{2A} or 5-HT_{2C} receptors have also been created; to the best of our knowledge, the stress-related behavioral functioning of these animals has yet to be reported (146 ,147).

Clinically Effective 5-HT System Drugs for Stress-Related Disorders

As mentioned above, one of the most commonly prescribed and effective classes of drugs that is used in the treatment of depression and anxiety is the SSRIs, which block the reuptake of 5-HT by its transporter and thereby increase serotoninergic transmission. Based on the findings of preclinical studies including those obtained from 5-HT receptor knockout mice, 5-HT_{1A} agonists have been developed for the treatment of anxiety. The clinical utility of this class of compounds, however, remains to be determined. As these transgenic approaches develop and become more refined, they will undoubtedly aid in clarifying the roles of the many other 5-HT receptor subtypes in processes related to stress and anxiety and will aid in drug development.

The GABA System

The primary inhibitory neurotransmitter in the CNS is GABA; GABA-synthesizing cells are distributed throughout the brain (128). The actions of GABA are mediated by two major classes of receptors, $GABA_A$ and $GABA_B$, both of which modulate the activity of ion channels. The principal mode of inactivation of GABA transmission is the presynaptic reuptake of GABA by its transporter protein. Although both types of GABA receptors are widely distributed through the CNS, several important differences exist between the two. Relevant to psychopharmacology is the finding that traditional anxiolytics (benzodiazepines) do not bind to $GABA_B$ receptors, but rather mediate their effects through $GABA_A$ receptors. $GABA_A$ receptors consist of a chloride channel formed by the pentameric arrangement of at least 18 different protein subunits (α_{1-6}, β_{1-4}, γ_{1-3}, δ, ε, π, ρ_{1-3}), thus allowing for considerable heterogeneity of the $GABA_A$ receptor isoforms (148). Typically, benzodiazepine-responsive $GABA_A$ receptors consist of α, ß, and γ subunits; in addition to the benzodiazepine site, these receptors also contain distinct sites for the binding of GABA, barbiturates, and ethanol. These various regions act as allosteric regulators of GABA-induced chloride channel opening. Although psychotherapeutic effects such as anxiolysis are achieved through facilitation of GABA transmission at this receptor, drugs that act as $GABA_A$ receptor agonists also produce several deleterious side effects. The extent to which differences in $GABA_A$ receptor subunit composition might contribute to possible dissociations between the beneficial and negative effects of these compounds is currently being investigated.

GABA System Transgenic Mice

The synthesis of GABA is regulated by two isoforms of the enzyme glutamate decarboxylase (GAD), GAD67, and the shorter form GAD65 (149). Whereas GAD67 is thought to maintain basal GABA levels, GAD65 is thought to regulate the synthesis of GABA at nerve terminals in response to high GABA demand (150). Given the important role of GABA in inhibitory neurotransmission associated with anxiolysis, several investigators have evaluated the behavioral profile of genetically altered mice that lack the GAD65 gene. Two separate groups have reported that $GAD65^{-/-}$ mice display an increase in stress-like behaviors in numerous behavioral paradigms (151 ,152). GAD65 knockout mice had fewer entries into and time spent in the center of an open field or the open areas of an elevated zero maze (similar to an elevated plus maze), indicating that they were more avoidant of inherently aversive areas. Similarly, these mice had lower levels of activity in the bright portion of a light-dark transition box. It should be mentioned that $GAD65^{-/-}$ mice also displayed an elevation in the occurrence of spontaneous and stress-induced seizures, and that these mice had a dramatically increased mortality rate starting at 4 to 5 weeks after birth (151). Thus, although the behavioral profile of GAD65 knockout mice is suggestive of increased anxiety-like responses, it is possible that these effects are secondary to the occurrence of seizures and to the factors leading to early lethality. The usefulness of this knockout as a model for anxiety-related deficits may therefore be limited. Given that benzodiazepines and barbiturates act as positive modulators of GABA transmission at the $GABA_A$ receptor by enhancing GABA-induced chloride channel opening, it is of interest to note that $GAD65^{-/-}$ mice were not sensitive to the effects of either benzodiazepines or barbiturates, but did respond to the direct $GABA_A$ agonist muscimol, which binds directly to the GABA site of the $GABA_A$ receptor and increases opening of the chloride channel in the absence of GABA (152). This pharmacologic profile is consistent with the finding that GABA synthesis is blocked by the GAD65 null mutation, but that $GABA_A$ receptor binding is unaffected by this change. Furthermore, this mutation does not seem to alter the functioning of GABA receptors because direct agonists stimulate the receptor but indirect modulators of GABA do not.

In an attempt to delineate the roles of the various $GABA_A$ receptor subunits in the regulation of stress- and anxiety-related behaviors, investigators have generated mutant mice with alterations in the expression of specific $GABA_A$ receptor subunits. It was initially reported that deletion of the γ_2 subunit led to a selective (94%) reduction in the expression of benzodiazepine sites in the CNS without alterations in the level of GABA sites or changes in the expression of other $GABA_A$ receptor subunits (153). Thus, γ_2 knockout mice possessed functional $GABA_A$ receptors that responded normally to GABA site ligands or barbiturates, but did not respond to benzodiazepines; these findings led to the conclusion that the γ_2 subunit is not necessary for the formation of functional $GABA_A$ receptors, but is required to create

the benzodiazepine-responsive site of those receptors. Mice that were homozygous for the mutation, however, did not live past weaning in this study. In mice carrying only one copy of the functional γ_2 gene, a 20% reduction in benzodiazepine sites was observed, but these mice did not show overt developmental deficits. In a recent study, a detailed characterization of the behavioral profile of these animals was carried out. Heterozygotes displayed a decrease in the number of entries into and amount of time spent in the open arms of an elevated plus maze and the bright compartment of a light-dark box. These animals also exhibited a decrease in the exploration of novel areas, and an increase in certain forms of fear conditioning that are thought to be mediated by the hippocampus. Finally, γ_2 heterozygotes were found to react to partially conditioned stimuli (only weakly paired with aversive consequences) as if they were full and potent predictors of threat; compared to wild-types, which showed low levels of defensive behaviors to the partially conditioned stimulus, heterozygotes displayed high levels of conditioned freezing to the partial conditioned stimulus that were identical to those displayed by all animals in response to the full conditioned stimulus. This profile has been proposed to be a model for the tendency to interpret neutral situations as threatening that is seen in anxiety patients. Taken together, the results from this extensive behavioral profile indicate that $\gamma_2^{+/-}$ mice have increased neophobia and stress-like responses and may thus provide a model for increased anxiety-like behaviors (154 ,155 and 156). Interestingly, all of the elevations in stress-like behaviors in γ_2 heterozygotes were blocked by the benzodiazepine diazepam, suggesting that this animal model may also have good predictive validity for identifying clinically effective anxiolytics.

It is also extremely important to mention the α_1 subunit transgenic mice, whose behavioral profiles have been thoroughly and insightfully reviewed in recent articles (157 ,158). In these mice, a single amino acid is altered (histidine replaced by arginine at the 101 position of the peptide) in the α_1 subunit of the $GABA_A$ receptor complex. This subtle change does not produce any overt alterations in baseline responses to stress in the genetically altered mice; these animals behave similarly to wild-type controls in tests such as the elevated plus maze and the fear-potentiated startle paradigm, a measure of conditioned fear (159 ,160). Thus, under drug-free, normal conditions, these animals do not display a behavioral pattern that is consistent with an anxiety-like endophenotype. When these mice are treated with conventional benzodiazepines, however, they react very differently to the drug than their wild-type counterparts. Mice with the mutation in the α_1 subunit display a normal reduction of stress-induced anxiety-like behaviors after benzodiazepine treatment, but fail to display some of the more deleterious side effects associated with this class of drugs such as sedation, amnesia, and ataxia. These results indicate that the anxiolytic effects of benzodiazepines can be separated from the negative side effects of these compounds, and that the α_1 subunit of the $GABA_A$ receptor is likely to mediate some of these potentially harmful properties of benzodiazepines. Interestingly, McKernan and colleagues (160) demonstrate that a novel benzodiazepine-site ligand that binds to $GABA_A$ receptors containing α_2, α_3, or α_5 subunits but avoids receptors with the α_1 subunit produces a behavioral profile that is identical to that of the α_1 subunit knockout mice; in normal mice, this compound decreases murine anxiety-like behaviors without eliciting sedation or ataxia (160).

Clinically Effective GABA System Drugs for Stress-Related Disorders

As stated above, the most widely used GABA system-based drugs for the treatment of anxiety are the benzodiazepines, which facilitate GABA transmission through the $GABA_A$ receptor. As outlined in the previous section, the search for novel compounds that may act selectively at specific $GABA_A$ subunits is ongoing, with the ultimate hope of discovering ligands that produce anxiolysis but do not cause some of the serious side effects that are commonly associated with benzodiazepines. As demonstrated by McKernan and colleagues (160), drugs that selectively target certain $GABA_A$ receptor subunits may hold great promise for the treatment of anxiety without harmful side effects. This development would represent a major breakthrough in the pharmacotherapy of anxiety-related disorders. The use of targeted genetic alterations in identifying the roles of various $GABA_A$ subunits will undoubtedly aid in this effort to create “designer drugs” for the treatment of anxiety (158).

General Issues and Caveats of Transgenic Animal Studies

As mentioned above, mice carrying certain mutations within either the CRH, the 5-HT, or the GABA system display an anxiety-like endophenotype. It appears that these genetically engineered mouse models also have some predictive validity; the stress-like endophenotype observed in at least two of the aforementioned models is normalized by administration of a clinically effective antianxiety agent that acts within the system that was genetically targeted. It remains to be determined, however, the extent to which these genetically altered models serve to identify potential antianxiety agents from different chemical classes. For example, do benzodiazepines reduce stress-like effects of CRH overexpressers? The extent to which the stress-like endophenotype in these animals is altered by compounds that act on systems that were not directly targeted by the genetic mutation will aid in determining the generalizability and utility of these models as predictors of novel anxiolytic agents. If one assumes that these animals provide a model of inherent trait-like anxiety, they can serve as a powerful tool for

screening new potential anxiolytics. These models do provide a sound approach to study the long-term effects of congenital abnormalities in these neurotransmitter and neuropeptide systems.

Several broad issues should be considered when interpreting studies utilizing genetically altered mice. Generally, the hypotheses regarding the behavioral profiles of transgenic mice are based on earlier findings from psychopharmacologic studies. For example, within the CRH field, the prediction that CRH overexpressers would display increased anxiety-like behaviors was based on the observation that CRH administration produces stress-like behaviors in rodents and primates (107 ,124). When the outcome of the transgenic studies agrees with the psychopharmacology-based prediction, the findings are taken as a confirmation of that hypothesized mechanism of action. When the outcome of the transgenic studies disagrees with the predicted phenotype, however, concerns about possible developmental confounds are raised. One of the most commonly cited drawbacks of the transgenic/knockout strategy is that the gene of interest is altered from the embryonic stage, therefore possibly influencing other genes involved in the normal development of the animal. Thus, it is difficult to tease apart the effects of under- or overexpression of that gene on the endpoints under study from effects due to compensatory or downstream developmental changes that may have occurred as a result of the mutation (86 ,87 ,161). Therefore, the transgenic/knockout approach provides an excellent method for modeling a congenital abnormality that leads to a disease state, but this approach may be less useful for identifying the discrete functions of a specific gene product because of the problems of interpretation that arise from the developmental confound. Indeed, with regard to all of the studies discussed in this section on genetically altered mice, it will be important in future studies to delineate the compensatory alterations that occur in response to the congenital mutation, and that may indirectly contribute to the adult endophenotypes that are reported for these animals. Future studies utilizing novel inducible-knockout strategies will circumvent the developmental issue; inducible knockouts may thus become a valuable tool for exploring the functions of discrete gene products for which no selective ligands are available (123).

It should also be noted that there is a large literature concerning the use of antisense oligonucleotide infusions to knock down the expression of particular gene products that may be related to fearful endophenotypes. The antisense oligonucleotide approach, however, has been plagued with a number of issues regarding toxicity, and may therefore not represent the optimal method for studying gene function *in vivo* (162).

FUTURE DIRECTIONS

Although the studies summarized in this chapter have contributed a great deal of knowledge about some of the genetic contributions to the development of stress and anxiety-like endophenotypes in animals, further information is needed to understand the precise nature of gene–environment interactions in stress regulation. It is likely that a particular stressor results in alterations of gene expression in myriad systems and that the overall response to stress involves the coordination of gene activation and/or suppression within these various systems. Novel high-throughput technologies have recently been developed that enable the expression of thousands of genes to be assayed at once. “Gene chips” and “DNA arrays” are two powerful new tools for analyzing complex multilocus genetic interactions associated with a particular environmental perturbation or disease state (163 ,164). This approach and its application to psychiatry research have been discussed comprehensively in a recent review article (165). Briefly, gene chip and DNA array technology involve the hybridization of gene transcripts from a tissue sample onto a glass slide or filter that contains up to 10,000 different nucleotide sequences. The amount and pattern of the signal hybridized to the array are then assessed; this method thus permits a rapid analysis of changes in the expression of multiple genes. This technology can also be used to identify single nucleotide polymorphisms in a particular gene by comparing the hybridization patterns of samples from different candidate populations on chips that contain multiple copies of the gene of interest, each copy differing from the previous one by just one base in the sequence. Theoretically, depending on the size of the gene, it would be possible to carry out a base-by-base examination of the entire gene on a single gene chip. However, it is important to realize that although a broad approach can be taken with this technology, it may not be sensitive enough to detect small but functionally important changes in gene expression. This technology can be applied to preclinial and clinical questions regarding the complex genetic control of stress and anxiety by examining event-related gene expression changes and also baseline differences in gene sequences (polymorphisms) that might contribute to differential stress responsivity (165). This technique, along with the recent completion of the Human Genome Project, not only raises the potential to simultaneously profile multiple gene expression systems at once, but also holds great promise for the identification of completely novel genes in stress regulation and anxiety.

A greater challenge, however, is the elucidation of the functional role of these new genes in processes related to stress and anxiety. Given this daunting task, methods for more specific and long-term gene targeting will increasingly gain importance in neuroscience research aimed at uncovering genetic dysregulation relating to psychopathology. One technique that is likely to be helpful is that of virally mediated gene transfer. In this method, a gene of interest is cloned into viral vector (with most of the viral genome removed to reduce toxicity and infection) and the modified vector is then infused into a particular brain region using

standard stereotaxic procedures (see ref. 166 for review). Depending on the gene insertion and the selection of the promoter to drive the expression of the gene, it is possible to obtain either an increase or decrease in the amount of protein resulting from the gene of interest. This method allows for highly selective gene regulation and thus provides a valuable new tool with which to study the effects of a particular gene product on stress-related functioning. The virally mediated gene transfer approach also has certain advantages over the current transgenic and antisense oligonucleotide strategies: it can be administered to the animal at any time or into any brain region, it results in a fairly robust and long-lasting up- or down-regulation of the gene, and it can be used to insert several genes at once in the same animal. Thus, the viral gene transfer approach completely avoids the issue of developmental confounds, which are perhaps the most commonly cited problems that plague current transgenic and knockout approaches. A few groups have already reported successful long-term up-regulation or down-regulation of discrete gene products related to neuroscience research applications; the behavioral effects associated with this technique appear to be quite robust and do not appear to be associated with the high level of toxicity that has been reported with antisense oligonucleotides (167 ,168 and 169). Thus, these methods may provide valuable new strategies to more rapidly uncover the neurogenetic basis for stress-related psychopathology.

On the clinical side, human genomic studies are indicating the existence of polymorphisms in the regulatory region of the gene encoding CRH (170 ,171 and 172). As careful analysis of genes for the other elements of the CRH system progresses, it will be interesting to see if particular mutations can be associated with stress-related disease states. This method has been applied successfully to study the role of the serotonin (5-HT) system in anxiety disorders; reports of polymorphisms in the gene encoding the 5-HT transporter have been made in patients with anxiety-related traits (173 ,174 ,175 and 176). Clinically, one challenge will be to develop more discrete definitions of anxiety-related dysfunction that will optimize the screening of patient populations for abnormalities in genes that are believed to be related to stress and anxiety (177). Moreover, gene chip technology applied to animal analogues of stress endophenotypes may provide a rapid and comprehensive method for identifying novel gene candidates for stress-related disorders. Using these methods, it may be possible in the near future to have even greater crosstalk between animal studies and clinical findings. These combined efforts will undoubtedly facilitate our understanding of the interactions between environmental and genetic contributions to anxiety and stress-related disorders.

REFERENCES

1. Swerdlow NR, Braff DL, Taaid N, et al. Assessing the validity of an animal model of deficient sensorimotor gating deficits in schizophrenic patients. *Arch Gen Psychiatry* 1994;51:139–154.

2. D'Mello GD, Steckler T. Animal models in cognitive behavioural pharmacology: an overview. *Brain Res Cogn Brain Res* 1996;3:345–352.

3. Martin P. Animal models sensitive to anti-anxiety agents. *Acta Psychiatr Scand Suppl* 1998;393:74–80.

4. Freedman R, Adler LE, Leonard S. Alternative phenotypes for the complex genetics of schizophrenia. *Biol Psychiatry* 1999;45:551–558.

5. Kalin NH, Shelton SE, Davidson RJ. Cerebrospinal fluid corticotropin-releasing hormone levels are elevated in monkeys with patterns of brain activity associated with fearful temperament. *Biol Psychiatry* 2000;47:579–585.

6. Freedman R, Coon H, Myles-Worsley M, et al. Linkage of a neurophysiologic deficit in schizophrenia to a chromosome 15 locus. *Proc Natl Acad Sci USA* 1997;94:587–592.

7. Schaller GB. *The Serengeti lion. Chicago: University of Chicago Press*, 1972.

8. Ficken MS, Witkin SR. Responses of black-capped chickadees to predators. *Auk* 1977;94:156–157.

9. Magurran AF, Girling SL. Predator model recognition and response habituation in shoaling minnows. *Anim Behav* 1986;34:510–518.

10. Ficken MS. Acoustic characteristics of alarm calls associated with predation risk in chickadees. *Anim Behav* 1990;39:400–401.

11. Morse DH. Interactions between tit flocks and sparrowhawks *Accipiter nisus. Ibis* 1993;115:591–593.

12. Palmer W. Instinctive stillness in birds. *Auk* 1909;26:23–36.

13. Curio E. *The ethology of predation. New York: Springer-Verlag*, 1976.

14. Kalin NH, Shelton SE. Defensive behaviors in infant rhesus monkeys: environmental cues and neurochemical regulation. *Science* 1989;243:1718–1721.

15. Kalin NH. The neurobiology of fear. *Sci Am* 1993;268:94–101.

16. Bakshi VP, Shelton SE, Kalin NH. Neurobiological correlates of defensive behaviors. *Prog Brain Res* 2000;122:105–115.

17. Rodgers RJ. Animal models of anxiety: where next? *Behav Pharmacol* 1997;8:477–496.

18. Weiss SM, Lightowler S, Stanhope KJ, et al. Measurement of anxiety in transgenic mice. *Rev Neurosci* 2000;11:59–74.

19. Kalin NH, Shelton SE, Takahashi LK. Defensive behaviors in infant rhesus monkeys: ontogeny and context-dependent selective expression. *Child Dev* 1991;62:1175–1183.

20. Harlow HF, Harlow MK. The affectional systems. In: Schrier A, Harlow H, Stollnitz F, eds. *Behavior of nonhuman primates*. New York: Academic Press, 1965.

21. Newman JD. The infant cry of primates: an evolutionary perspective. In: Lester BM, Zachariah Boukydis CF, eds. *Infant crying*. New York: Plenum, 1985.

22. File SE, Lippa AS, Beer B, et al. Animal tests of anxiety. *Curr Protocols Neurosci* 2000;2:8.3.1–8.3.19.

23. File SE. New strategies in the search for anxiolytics. *Drug Design Deliv* 1990;5:195–201.

24. LeDoux J. Fear and the brain: where have we been, and where are we going? *Biol Psychiatry* 1998;44:1229–1238.

25. Davis M. Are different parts of the extended amygdala involved in fear versus anxiety? *Biol Psychiatry* 1998;44:1239–1247.

26. Miczek KA, Weerts E, Haney M, et al. Neurobiological mechanisms controlling aggression: preclinical developments for pharmacotherapeutic interventions. *Neurosci Biobehav Rev* 1994;18:97–110.

27. Miczek KA, Weerts EM, Vivian JA, et al. Aggression, anxiety and vocalizations in animals: $GABA_A$ and 5-HT anxiolytics. *Psychopharmacology* 1995;121:38–56.

28. Olivier B, Mos J, Miczek KA. Ethopharmacological studies of anxiolytics and aggression. *Eur Neuropsychopharmacol* 1991;1:97–100.

29. Geller I, Hartmann RJ. Effects of buspirone on operant behavior of laboratory rats and cynomolgus monkeys. *J Clin Psychiatry* 1982;43:25–33.

30. Vogel RA, Frye GD, Wilson JH, et al. Attenuation of the effects of punishment by ethanol: comparisons with chlordiazepoxide. *Psychopharmacology* 1980;71:123–129.

31. Biederman J, Rosenbaum JF, Hirshfeld DR, et al. Psychiatric correlates of behavioral inhibition in young children of parents with and without psychiatric disorders. *Arch Gen Psychiatry* 1990;47:21–26.

32. Hirshfeld DR, Rosenbaum JF, Biederman J, et al. A 3-year follow-up of children with and without behavioral inhibition. *J Am Acad Child Adolesc Psychiatry* 1993;32:814–821.

33. Rosenbaum JF, Biederman J, Bolduc-Murphy EA, et al. Behavioral inhibition in childhood: a risk factor for anxiety disorders. *Harvard Rev Psychiatry* 1993;1:2–16.

34. Pollock RA, Rosenbaum JF, Marrs A, et al. Anxiety disorders of childhood: implications for adult psychopathology. *Psychiatr Clin North Am* 1995;18:745–766.

35. Mick MA, Telch MJ. Social anxiety and history of behavioral inhibition in young adults. *J Anxiety Disord* 1998;12:1–20.

36. Kagan J, Reznick JS, Snidman N. Biological bases of childhood shyness. *Science* 1988;240:167–171.

37. Kagan J, Reznick JS, Snidman N. The physiology and psychology of behavioral inhibition in children. *Child Dev* 1987;58:1459–1473.

38. Kalin NH, Shelton SE, Rickman M, et al. Individual differences in freezing and cortisol in infant and mother rhesus monkeys. *Behav Neurosci* 1998;112:251–254.

39. Takahashi LK, Rubin WW. Corticosteroid induction of threat-induced behavioral inhibition in preweanling rats. *Behav Neurosci* 1993;107:860–866.

40. Tomarken AJ, Davidson RJ, Wheeler RE, et al. Individual differences in anterior brain asymmetry and fundamental dimensions of emotion. *J Personality Soc Psychol* 1992;62:676–687.

41. Davidson RJ, Kalin NH, Shelton SE. Lateralized response to diazepam predicts temperamental style in rhesus monkeys. *Behav Neurosci* 1993;107:1106–1110.

42. Kalin NH, Larson C, Shelton SE, et al. Asymmetric frontal brain activity, cortisol, and behavior associated with fearful temperaments in rhesus monkeys. *Behav Neurosci* 1998;112:286–292.

43. Henriques JB, Davidson RJ. Left frontal hypoactivation in depression. *J Abnormal Psychol* 1991;100:535–545.

44. Gomez F, Grauges P, Lopez-Calderon A, et al. Abnormalities of hypothalamic-pituitary-adrenal and hypothalamic-somatotrophic axes in fawn-hooded rats. *Eur J Endocrinol* 1999;141:290–296.

45. Altemus M, Smith AM, Diep V, et al. Increased mRNA for corticotropin-releasing hormone in the amygdala of fawn-hooded rats: a potential animal model of anxiety. *Anxiety* 1994–95;1:251–257.

46. Redei E, Pare WP, Aird F, et al. Strain differences in hypothalamic-pituitary-adrenal activity and stress ulcer. *Am J Physiol* 1994;266:R353–R360.

47. Sternberg EM, Glowa, JR, Smith MA, et al. Corticotropin-releasing hormone related behavioral and neuroendocrine responses to stress in Lewis and Fischer rats. *Brain Res* 1992;570:54–60.

48. Krukoff TL, Mactavish D, Jhamandas JH. Hypertensive rats exhibit heightened expression of corticotropin-releasing factor in activated central neurons in response to restraint stress. *Mol Brain Res* 1999;65:70–79.

49. Mansi JA, Rivest S, Drolet G. Effect of immobilization stress on transcriptional activity of inducible immediate-early genes, corticotropin-releasing factor, its type I receptor, and enkephalin in the hypothalamus of borderline hypertensive rats. *J Neurochem* 1998;70:1556–1566.

50. Imaki T, Naruse M, Harada S, et al. Stress-induced changes of gene expression in the paraventricular nucleus are enhanced in spontaneously hypertensive rats. *J Endocrinol* 1998;10:635–645.

51. Beuzen A, Belzung C. Link between emotional memory and anxiety states: a study by principal component analysis. *Physiol Behav* 1995;58:111–118.

52. Griebel G, Belzung C, Perrault G, et al. Differences in anxiety-related behaviours and in sensitivity to diazepam in inbred and outbred strains of mice. *Psychopharmacology* 2000;148:164–170.

53. Griebel G, Belzung C, Misslin R, et al. The free-exploratory paradigm: an effective method for measuring neophobic behavior in mice and testing potential neophobia-inducing drugs. *Behav Pharmacol* 1993;4:637–644.

54. Hode Y, Ratomponirina C, Gobaille S, et al. Hypoexpression of benzodiazepine receptors in the amygdala of neophobic BALB/c mice compared to C57BL/6 mice. *Pharmacol Biochem Behav* 2000;65:35–38.

55. Bowlby J. *Attachment and loss, vol II: separation*. New York: Basic Books, 1973.

56. Frank DA, Klass PE, Earls F, et al. Infants and young children in orphanages—one view from pediatrics and child psychiatry. *Pediatrics* 1996;97:569–578.

57. Carlson M, Earls F. Psychological and neuroendocrinological sequelae of early social deprivation in institutionalized children in Romania. *Ann NY Acad Sci* 1997;807:419–428.

58. Levine S. Infantile experience and resistance to physiological stress. *Science* 1957;126:405–406.

59. Sapolsky RM. The importance of a well-groomed child. *Science* 1997;277:1620–1621.

60. Kuhn CM, Schanberg SM. Responses to maternal separation-mechanisms and mediators. *Int J Dev Neurosci* 1998;16:261–270.

61. Harlow HF. Love in infant monkeys. *Sci Am* 1959;200:68–74.

62. Harlow HF, Zimmerman RR. Affectional responses in the infant monkey. *Science* 1959;130:421–432.

63. Berman CM. Mother-infant relationships among free-ranging rhesus monkeys on Cayo Santiago: a comparison with captive pairs. *Anim Behav* 1980;28:860–873.

64. Harlow HF, Rowland GL, Griffin GA. The effect of total social deprivation on the development of monkey behavior. *Psychiatr Res Rep* 1964;19:116–135.

65. McKinney WT, Bunney WE. Animal model of depression. I. Review of evidence: implications for research. *Arch Gen Psychiatry* 1969;21:240–248.

66. Suomi SJ, Harlow HF, McKinney WT. Monkey psychiatrists. *Am J Psychiatry* 1972;128:927–932.

67. Reite M. Maternal separation in monkey infants: a model of depression. In: Hanin I, Usdin E, eds. *Animal models in psychiatry and neurology*. Oxford, UK: Pergamon Press, 1977:127–140.

68. Rosenblum LA, Paully GS. Primate models of separation-induced depression. *Psychiatr Clin North Am* 1987;10:437–447.

69. Takahashi LK, Kim H. Intracranial action of corticosterone facilitates the development of behavioral inhibition in the adrenalectomized preweanling rat. *Neurosci Lett* 1994;176:272–276.

70. Coplan JD, Andrews MW, Rosenblum LA, et al. Persistent elevations of cerebrospinal fluid concentrations of corticotropin-releasing factor in adult nonhuman primates exposed to early-life stressors: implications for the pathophysiology of mood and anxiety disorders. *Proc Natl Acad Sci USA* 1996;93:1619–1623.

71. Coplan JD, Smith ELP, Trost RC, et al. Growth hormone response to clonidine in adversely reared young adult primates: relationship to serial cerebrospinal fluid corticotropin-releasing factor concentrations. *Psychiatry Res* 2000;95:93–102.

72. Andrews MW, Rosenblum LA. The development of affiliative and agonistic social patterns in differentially reared monkeys. *Child Dev* 1994;65:1398–1404.

73. Anisman H, Zaharia MD, Meaney MJ, et al. Do early-life events permanently alter behavioral and hormonal responses to stressors? *Int J Dev Neurosci* 1998;16:149–164.

74. Heim C, Nemeroff CB. The impact of early adverse experiences on brain systems involved in the pathophysiology of anxiety and affective disorders. *Biol Psychiatry* 1999;46:1509–1522.

75. Francis DD, Caldji C, Champagne F, et al. The role of corticotropin-releasing factor-norepinephrine systems in mediating the effects of early experience on the development of behavioral and endocrine responses to stress. *Biol Psychiatry* 1999;46:1153–1166.

76. Plotsky PM, Meaney MJ. Early, postnatal experience alters hypothalamic corticotropin-releasing factor (CRF) mRNA, median eminence CRF content and stress-induced release in adult rats. *Mol Brain Res* 1993;18:195–200.

77. Liu D, Diorio J, Tannenbaum B, et al. Maternal care, hippocampal glucocorticoid receptors, and hypothalamic-pituitary-adrenal responses to stress. *Science* 1997;277:1659–1662.

78. Caldji C, Tannenbaum B, Sharma S, et al. Maternal care during infancy regulates the development of neural systems mediating the expression of fearfulness in the rat. *Proc Natl Acad Sci USA* 1998;95:5335–5340.

79. Rots NY, de Jong J, Workel JO, et al. Neonatal maternally deprived rats have as adults elevated basal pituitary-adrenal activity and enhanced susceptibility to apomorphine. *J Neuroendocrinol* 1996;8:501–506.

80. Wigger A, Neumann ID. Periodic maternal deprivation induces gender-dependent alterations in behavioral and neuroendocrine responses to emotional stress in adult rats. *Physiol Behav* 1999;66:293–302.

81. Shanks N, Larcoque S, Meaney MJ. Neonatal endotoxin exposure alters the development of the hypothalamic-pituitaryadrenal axis: early illness and later responsivity to stress. *J Neurosci* 1995;15:376–384.

82. Liu D, Caldji C, Sharma S, et al. Influence of neonatal rearing conditions on stress-induced adrenocorticotropin responses and norepinephrine release in the hypothalamic paraventricular nucleus. *J Neuroendocrinol* 2000;12:5–12.

83. Takahashi LK, Turner JG, Kalin NH. Prenatal stress alters brain catecholaminergic activity and potentiates stress-induced behavior in adult rats. *Brain Res* 1992;574:131–137.

84. Cratty MS, Ward HE, Johnson EA, et al. Prenatal stress increases corticotropin-releasing factor (CRF) content and release in rat amygdala minces. *Brain Res* 1995;675:297–302.

85. Fujioka T, Sakata Y, Yamaguchi K, et al. The effects of prenatal stress on the development of hypothalamic paraventricular neurons in fetal rats. *Neuroscience* 1999;92:1079–1088.

86. Picciotto MR, Wickman K. Using knockout and transgenic mice to study neurophysiology and behavior. *Physiol Rev* 1998;78:1131–1163.

87. Contarino A, Heinrichs SC, Gold LH. Understanding corticotropin-releasing factor neurobiology: contributions from mutant mice. *Neuropeptides* 1999;33:1–12.

88. Heinrichs, SC. Stress-axis, coping and dementia: gene manipulation studies. *Trends Pharmacol Sci* 20:311–315.

89. Zhuang X, Gross C, Santarelli L, et al. Altered emotional states in knockout mice lacking 5-HT1A or 5-HT1B receptors. *Neuropsychopharmacol* 1999;21:52S–60S.

90. Southwick SM, Paige S, Morgan CA, et al. Neurotransmitter alterations in PTSD: catecholamines and serotonin. *Semin Clin Neuropsychiatry* 1999;4:242–248.

91. Sullivan GM, Coplan JD, Kent JM, et al. The noradrenergic system in pathological anxiety: a focus on panic with relevance to generalized anxiety and phobias. *Biol Psychiatry* 1999;46:1205–1218.

92. Fink H, Rex A, Voits M, et al. Major biological actions of CCK—a critical evaluation of research findings. *Exp Brain Res* 1998;123:77--83.

93. Heilig M, Soderpalm B, Engel JA, et al. Centrally administered neuropeptide Y (NPY) produces anxiolytic-like effects in animal anxiety models. *Psychopharmacology* 1989;98:524–529.

94. Saria A. The tachykinin NK1 receptor in the brain: pharmacology and putative functions. *Eur J Pharmacol* 1999;375:51–60.

95. Vaughan J, Donaldson C, Bittencourt J, et al. Urocortin, a mammalian neuropeptide related to fish urotensin I and to corticotropin-releasing factor. *Nature* 1995;378:287–292.

96. Chen R, Lewis KA, Perrin MG, et al. Expression cloning of a human corticotropin-releasing factor receptor. *Proc Natl Acad Sci USA* 1993;90:896–871.

97. Perrin M, Donaldson C, Chen R, et al. Identification of a second corticotropin-releasing factor receptor gene and characterization of a cDNA expressed in heart. *Proc Natl Acad Sci USA* 1995;92:2969–2973.

98. Lovenberg TW, Liaw CW, Grigoriadis DE, et al. Cloning and characterization of a functionally distinct corticotropin-releasing factor receptor subtype from rat brain. *Proc Natl Acad Sci USA* 1995;92:836–840.

99. Potter E, Behan DP, Fischer WH, et al. Cloning and characterization of the cDNAs for human and rat corticotropin-releasing factor-binding proteins. *Nature* 1991;349:423–426.

100. Potter E, Behan DP, Linton EA, et al. The central distribution of a corticotropin-releasing factor (CRF)-binding protein predicts multiple sites and modes of interaction with CRF. *Proc Natl Acad Sci USA* 1992;89:4192–4196.

101. Stenzel-Poore MP, Cameron VA, Vaughan J, et al. Development of Cushing's syndrome in corticotropin-releasing factor transgenic mice. *Endocrinology* 1992;130:3378–3386.

102. Stenzel-Poore MP, Heinrichs SC, Rivest S, et al. Overproduction of corticotropin-releasing factor in transgenic mice: a genetic model of anxiogenic behavior. *J Neurosci* 1994;14:2579–2584.

103. Heinrichs SC, Stenzel-Poore MP, Gold LH, et al. Learning impairment in transgenic mice with central overexpression of corticotropin-releasing factor. *Neuroscience* 1996;74:303–311.

104. Heinrichs SC, Min H, Tamraz S, et al. Anti-sexual and anxiogenic behavioral consequences of corticotropin-releasing factor overexpression are centrally mediated. *Psychoneuroendocrinology* 1997;22:215–224.

105. Karolyi IJ, Burrows HL, Ramesh TM, et al. Altered anxiety and weight gain in corticotropin-releasing hormone-binding protein-deficient mice. *Proc Natl Acad Sci USA* 1999;96:11595–11600.

106. Lovejoy DA, Aubry JM, Turnbull A, et al. Ectopic expression of the CRF-binding protein: minor impact on HPA axis regulation but induction of sexually dimorphic weight gain. *J Neuroendocrinol* 1998;10:483–491.

107. Koob GF, Heinrichs SC. A role for corticotropin-releasing factor and urocortin in behavioral responses to stressors. *Brain Res* 1999;848:141–152.

108. Zobel AW, Nickel T, Kunzel HE, et al. Effects of the high-affinity corticotropin-releasing hormone receptor 1 antagonist R121919 in major depression: the first 20 patients treated. *J Psychiatr Res* 2000;34:171–181.

109. Muglia LJ, Jacobson L, Dikkes P, et al. Corticotropin-releasing hormone deficiency reveals major fetal but not adult glucocorticoid need. *Nature* 1995;373:427–432.

110. Muglia LJ, Bae DS, Brown TT, et al. Proliferation and differentiation defects during lung development in corticotropin-releasing hormone-deficient mice. *Am J Respir Cell Mol Biol* 1999;20:181–188.

111. Dunn AJ, Swiergiel AH. Behavioral responses to stress are intact in CRF-deficient mice. *Brain Res* 1999;845:14–20.

112. Weninger SC, Dunn AJ, Muglia LJ, et al. Stress-induced behaviors require the corticotropin-releasing hormone (CRH) receptor, but not CRH. *Proc Natl Acad Sci USA* 1999;96:8283–8288.

113. Weninger SC, Muglia LJ, Jacobson L, et al. CRH-deficient mice have a normal anorectic response to chronic stress. *Regul Pept* 1999;84:69–74.

114. Weninger SC, Peters LL, Majzoub JA. Urocortin expression in the Edinger-Westphal nucleus is up-regulated by stress and corticotropin-releasing hormone deficiency. *Endocrinology* 2000;141:256–263.

115. Contarino A, Dellu F, Koob FG, et al. Reduced anxiety-like and cognitive performance in mice lacking the corticotropin-releasing factor receptor 1. *Brain Res* 1999;835:1–9.

116. Smith GW, Aubry JM, Dellu F, et al. Corticotropin-releasing factor receptor 1-deficient mice display decreased anxiety, impaired stress response, and aberrant neuroendocrine development. *Neuron* 1998;20:1093–1102.

117. Timpl P, Spanagel R, Sillaber I, et al. Impaired stress response and reduced anxiety in mice lacking a functional corticotropin-releasing hormone receptor 1. *Nature Genet* 1998;19:162–166.

118. Bale TL, Contarino A, Smith GW, et al. Mice deficient for corticotropin-releasing hormone receptor-2 display anxiety-like behaviour and are hypersensitive to stress. *Nature Genetics* 2000;24:410–414.

119. Coste SC, Kesterson RA, Heldwein KA, et al. Abnormal adaptations to stress and impaired cardiovascular function in mice lacking corticotropin-releasing hormone receptor-2. *Nature Genet* 2000;24:403–409.

120. Kishimoto T, Radulovic J, Radulovic M, et al. Deletion of Crhr2 reveals an anxiolytic role for corticotropin-releasing hormone receptor-2. *Nature Genet* 2000;24:415–419.

121. Bakshi VP, Smith-Roe SL, Yang LW, et al. Antagonism of CRF receptors within the lateral septum decreases stress-induced behavioral inhibition in rats. *Soc Neurosci Abst* 1999;25:62.

122. Radulovic J, Ruhmann A, Liepold T, et al. Modulation of learning and anxiety by corticotropin-releasing factor (CRF) and stress: differential roles of CRF receptors 1 and 2. *J Neurosci* 1999;19:5016–5025.

123. Stark KL, Oosting RS, Hen R. Inducible knockout strategies to probe functions of 5-HT receptors. *Ann NY Acad Sci* 1998;861:57–66.

124. Kalin NH. Behavioral effects of ovine corticotropin-releasing factor administered to rhesus monkeys. *Fed Proc* 1985;44:249–253.

125. McCarthy J, Heinrichs SC, Grigoriadis DE. Recent advances with the CRF_1 receptor: design of small molecule inhibitors, receptor subtypes and clinical indications. *Curr Pharmaceut Design* 1999;5:289–315.

126. Habib KE, Weld KP, Rice KC, et al. Oral administration of a corticotropin-releasing hormone receptor antagonist significantly attenuates behavioral, neuroendocrine, and autonomic responses to stress in primates. *Proc Natl Acad Sci USA* 2000;97:6079–6084.

127. He L, Gilligan PJ, Zaczek R, et al. 4-(1,2-dimethoxyprop-2-ylamino)-2,7-dimethyl-8-(2,4-dichlorophenyl)pyrazolo-[1,2-*a*]-1,3,5-triazine: a potent, orally bioavailable CRF1 receptor antagonist. *J Med Chem* 2000;43:449–456.

128. Cooper JR, Bloom FE, Roth RH. *The biochemical basis of neuropharmacology*. New York: Oxford University Press, 1996.

129. Martin GR, Eglen RM, Hamblin MW, et al. The structure and signalling properties of 5-HT receptors: an endless diversity? *Trends Pharmacol Sci* 1998;19:2–4.

130. Hen R. Mean genes. *Neuron* 1996;16:17–21.

131. Brunner D, Hen R. Insights into the neurobiology of impulsive behavior from serotonin receptor knockout mice. *Ann NY Acad Sci* 1997;836:81–105.

132. Julius D. Serotonin receptor knockouts: a moody subject. *Proc Nat Acad Sci USA* 1998;95:15153–15154.

133. Parks CL, Robinson PS, Sibille E, et al. Increased anxiety of mice lacking the serotonin1A receptor. *Proc Natl Acad Sci USA* 1998;95:10734–10739.

134. Ramboz S, Oosting R, Amara DA, et al. Serotonin receptor 1A knockout: an animal model of anxiety-related disorder. *Proc Natl Acad Sci USA* 1998;95:14476–14481.

135. Heisler LK, Chu HM, Brennan TJ, et al. Elevated anxiety and antidepressant-like responses in serotonin 5-HT1A receptor mutant mice. *Proc Natl Acad Sci USA* 1998:15049–15054.

136. Sibille E, Pavlides C, Benke D, et al. Genetic inactivation of the serotonin1A receptor in mice results in downregulation of major $GABA_A$ receptor α subunits, reduction of $GABA_A$ receptor binding, and benzodiazepine-resistant anxiety. *J Neurosci* 2000;20:2758–2765.

137. Malleret G, Hen R, Guillou JL, et al. 5-HT1B receptor knock-out mice exhibit increased exploratory activity and enhanced spatial memory performance in the Morris water maze. *J Neurosci* 1999;19:6157–6168.

138. Brunner D, Buhot MC, Hen R, et al. Anxiety, motor activation, and maternal-infant interactions in 5-HT1B knockout mice. *Behav Neurosci* 1999;113:587–601.

139. Olivier B, Molewijk HE, van der Heyden JA, et al. Ultrasonic vocalizations in rat pups: effects of serotonergic ligands. *Neurosci Biobehav Rev* 1998;23:215–227.

140. Ramboz S, Saudou F, Amara DA, et al. 5-HT1B receptor knock out—behavioral consequences. *Behav Brain Res* 1996;73:305–312.

141. Sadou F, Amara DA, Dierich A, et al. Enhanced aggressive behavior in mice lacking the 5-HT1B receptor. *Science* 1994;265:1875–1878.

142. Olivier B, Mos J, van Oorschot R, et al. Serotonin receptors and animal models of aggressive behavior. *Pharmacopsychiatry* 1995;28(suppl 2):80–90.

143. Grailhe R, Waeber C, Dulawa SC, et al. Increased exploratory activity and altered response to LSD in mice lacking the 5-HT5A receptor. *Neuron* 1999;22:581–591.

144. Tecott LH, Chu HM, Brennan TJ. Neurobehavioral analysis of 5-HT6 receptor null mutant mice. In: *IUPHAR satellite meeting on serotonin*, 4th ed. Rotterdam: 1998:abstr S1.2.

145. Branchek TA, Blackburn TP. 5-HT6 receptors as emerging targets for drug discovery. *Annu Rev Pharmacol Toxicol* 2000;40:319–334.

146. Rioux A, Fabre V, Lesch KP, et al. Adaptive changes of serotonin 5-HT2A receptors in mice lacking the serotonin transporter. *Neurosci Lett* 1999;262:113–116.

147. Applegate CD, Tecott LH. Global increases in seizure susceptibility in mice lacking 5-HT2C receptors: a behavioral analysis. *Exp Neurol* 1998;154:522–530.

148. Mehta AK, Ticku MK. An update on $GABA_A$ receptors. *Brain Res Rev* 1999;29:196–217.

149. Erlander MG, Tillakaratne NJ, Feldblum S, et al. Two genes encode distinct glutamate decarboxylases. *Neuron* 1991;7:91–100.

150. Martin DL, Martin SB, Wu SJ, et al. Regulatory properties of brain glutamate decarboxylase (GAD): the apoenzyme of GAD is present principally as the smaller of two molecular forms of GAD in brain. *J Neurosci* 1991;11:2725–2731.

151. Stork O, Feng-Yun J, Koichi K, et al. Postnatal development of a GABA deficit and disturbance of neural functions in mice lacking GAD65. *Brain Res* 2000;865:45–58.

152. Kash SF, Tecott LH, Hodge C, et al. Increased anxiety and altered responses to anxiolytics in mice deficient in the 65-kDa isoform of glutamic acid decarboxylase. *Proc Natl Acad Sci USA* 1999;96:1698–1703.

153. Gunther U, Benson J, Benke D, et al. Benzodiazepine-insensitive mice generated by targeted disruption of the γ_2 subunit gene of γ-aminobutyric acid type A receptors. *Proc Natl Acad Sci USA* 1995;92:7749–7753.

154. Crestani F, Lorenz M, Baer K, et al. Decreased $GABA_A$-receptor clustering results in enhanced anxiety and a bias for threat cues. *Nature Neurosci* 1999;2:833–839.

155. McNaughton N. A gene promotes anxiety in mice—and also in scientists. *Nature Med* 1999;5:1131–1132.

156. Anagnostaras SG, Craske MG, Fanselow MS. Anxiety: at the intersection of genes and experience. *Nature Neurosci* 1999;2:780–782.

157. Wisden W, Stephens DN. Towards better benzodiazepines. *Nature* 1999;401:751–752.

158. Tecott LH. Designer genes and anti-anxiety drugs. *Nature Neurosci* 2000;3:529–530

159. Rudolph U, Crestani F, Benke D, et al. Benzodiazepine actions mediated by specific γ-aminobutyric acid A receptor subtypes. *Nature* 1999;401:796–800.

160. McKernan RM, Rosahl TW, Reynolds DS, et al. Sedative but not anxiolytic properties of benzodiazepines are mediated by the $GABA_A$ receptor γ_1 subtype. *Nature Neurosci* 2000;3:587–592.

161. Gingrich JA, Hen R. The broken mouse: the role of development, plasticity and environment in the interpretation of phenotypic changes in knockout mice. *Curr Opin Neurobiol* 2000;10:146–152.

162. Bakshi VP, Kalin NH. Corticotropin-releasing hormone and animal models of anxiety: gene-environment interactions. *Biol Psychiatry* 2000;48:1175–1198.

163. Chee M, Yang R, Hubell E, et al. Accessing genetic information with high-density DNA arrays. *Science* 1996;274:610–614.

164. Schena M, Shalon D, Heller R, et al. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci USA* 1996;93:10614–10619.

165. Watson SJ, Akil H. Gene chips and arrays revealed: a primer on their power and their uses. *Biol Psychiatry* 1999;45:533–543.

166. Simonato M, Manservigi R, Marconi P, et al. Gene transfer into neurones for the molecular analysis of behaviour: focus on herpes simplex vectors. *Trends Neurosci* 2000;23:183–190.

167. Carlezon WA, Boundy VA, Haile CN, et al. Sensitization to morphine induced by viral-mediated gene transfer. *Science* 1997;277:812–814.

168. Kang W, Wilson MA, Bender MA, et al. Herpes virus-mediated preproenkephalin gene transfer to the amygdala is antinociceptive. *Brain Res* 1998;792:133–135.

169. Szczypka MS, Mandel RJ, Donahue BA, et al. Viral gene delivery selectively restores feeding and prevents lethality of dopamine-deficient mice. *Neuron* 1999;22:167–178.

170. Baerwald CG, Panayi GS, Lanchbury JS. A new X*mn*l polymorphism in the regulatory region of the corticotropin-releasing hormone gene. *Hum Genet* 1996;97:697–698.

171. Baerwald CG, Panayi GS, Lanchbury JS. Corticotropin-releasing hormone promoter region polymorphisms in rheumatoid arthritis. *J Rheumatol* 1997;24:215–216.

172. Gu J, Sadler L, Daiger S, et al. Dinucleotide repeat polymorphism at the CRH gene. *Hum Mol Genet* 1993;2:85.

173. Lesch KP, Bengel D, Heils A, et al. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 1996;274:1527–1531.

174. Ohara K, Nagai M, Suzuki Y, et al. Association between anxiety disorders and a functional polymorphism in the serotonin transporter gene. *Psychiatry Res* 1998;81:277–279.

175. Mazzanti CM, Lappalainen J, Long JC, et al. Role of the serotonin transporter promoter polymorphism in anxiety-related traits. *Arch Gen Psychiatry* 1998;55:936–940.

176. Flory JD, Manuck SB, Ferrell RE, et al. Neuroticism is not associated with the serotonin transporter (5-HTTLPR) polymorphism. *Mol Psychiatry* 1999;4:93–96.

177. Smoller JW, Tsuang MT. Panic and phobic anxiety: defining phenotypes for genetic studies. *Am J Psychiatry* 1998;155:1152–1162.

63

Neurobiological Basis of Anxiety Disorders

Dennis S. Charney

Wayne C. Drevets

Dennis S. Charney: Mood and Anxiety Disorder Research Program, National Institute of Mental Health, Bethesda, Maryland.

W. C. Drevets: Section on Mood and Anxiety Disorders Imaging, Molecular Imaging Branch, National Institute of Mental Health, Bethesda, Maryland.

The 1990s witnessed tremendous progress in the acquisition of knowledge about the molecular, cellular, and anatomic correlates of fear and anxiety. Advances in neuropharmacology and molecular biology have enabled elucidation of multiple chemical neurotransmitter systems that play roles in fear and anxiety behavior. The anatomic circuits where these transmitters participate in mediating and modulating fear and anxiety are also being illuminated through improvements in neurotoxic techniques, which have enhanced the selectivity of lesion analyses in experimental animals, and by advances in neuroimaging technology, which have permitted mapping of the neurophysiologic correlates of emotion in humans. The findings of these investigations have informed the design and interpretation of clinical neuroscience approaches aimed at investigating how dysfunction within these neurochemical and anatomic systems may result in psychiatric conditions such as panic, posttraumatic stress, and phobic disorders. This chapter reviews the preclinical and clinical data regarding the neural mechanisms underlying normal and pathologic anxiety and discusses their implications for guiding development of novel treatments for anxiety disorders.

NEUROANATOMIC CIRCUITS SUPPORTING FEAR AND ANXIETY

Part of "63 - Neurobiological Basis of Anxiety Disorders "

Fear and anxiety normally comprise adaptive responses to threat or stress. These emotional-behavioral sets may arise in response to exteroceptive visual, auditory, olfactory, or somatosensory stimuli or to interoceptive input through the viscera and the endocrine and autonomic nervous systems. Anxiety may also be produced by cognitive processes mediating the anticipation, interpretation, or recollection of perceived stressors and threats.

Emotional processing in general can be divided into evaluative, expressive, and experiential components (1). Evaluation of the emotional salience of a stimulus involves appraisal of its valence (e.g., appetitive versus aversive), its relationship with previous conditioning and behavioral reinforcement experiences, and the context in which it arises (2 ,3). Emotional expression conveys the range of behavioral, endocrine, and autonomic manifestations of the emotional response, whereas emotional experience describes the subjective feeling accompanying the response. To optimize their capacity for guiding behavior, all these aspects of emotional processing are modulated by complex neurobiological systems that prevent them from becoming persistent, excessive, inappropriate to reinforcement contingencies, or otherwise maladaptive.

The emotional processes pertaining to fear and anxiety that have been most extensively studied (largely because of their amenability to experimental manipulation) have involved pavlovian fear conditioning and fear-potentiated startle (4 ,5). These types of “fear learning” have been shown to comprise experience-dependent forms of neural plasticity in an extended anatomic network that centers around the critical involvement of the amygdala (1 ,6). The structures that function in concert with the amygdala during fear learning include other mesiotemporal cortical structures, the sensory thalamus and cortices, the orbital and medial prefrontal cortex (mPFC), the anterior insula, the hypothalamus, and multiple brainstem nuclei (1 ,5 ,7). Much of this network appears to participate in the general process of associating a conditioned stimulus (CS) or operant behavior with an emotionally salient unconditioned stimulus (US) (see Fig. 63.1 on p. 905) (5 ,8 ,9 ,10 and 11).

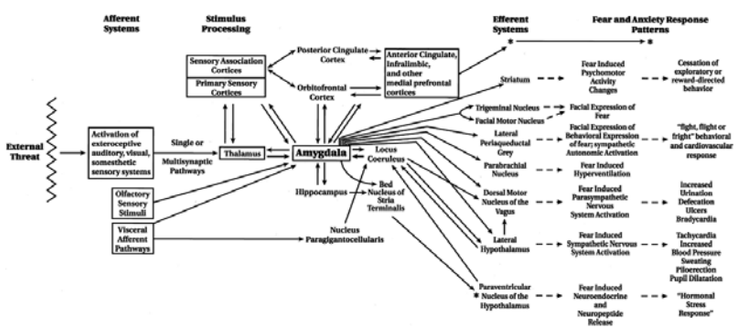


FIGURE 63.1. The innervation of the parasympathetic nervous system from limbic structures is thought to mediate visceral symptoms associated with anxiety.

Role of the Amygdala in Fear Learning and Expression

The anatomic systems supporting fear learning are organized to permit both rapid responses to simple perceptual

elements of potentially threatening stimuli and longerlatency responses to more highly processed information about complex sensory stimuli and environmental contexts. The former processes depend on monosynaptic projections from the sensory thalamus to the amygdala, whereas the latter involve projections from sensory association cortices and mesiotemporal cortical structures to the amygdala (1 ,12). These neural networks also respond to visceral input received both directly through the nucleus paragigantocellularis and the nucleus tractus solitarius (NTS) of the vagus nerve and indirectly through the locus ceruleus (LC), the anterior insula, and the infralimbic and prelimbic cortices (4 ,7 ,13). Finally, neural activity within the amygdala is modulated by cortisol, norepinephrine (NE), and other neurotransmitters and by mnemonic input related to previous conditioning and reinforcement experiences conveyed by projections from mesiotemporal and prefrontal cortical structures (14 ,15 ,16 ,17 and 18).

The lateral nucleus of the amygdala (LA) comprises the primary sensory interface of the amygdala and receives synaptic input representing sensory information from the sensory thalamus and cortex (4). Single neurons within the LA are responsive to auditory, visual, and somatic stimuli, thus enabling the LA to serve as a locus of convergence for information about CS and US (19). Olfactory input, in contrast, directly projects to the periamygdaloid cortex from the olfactory bulb through the olfactory tract (20). The olfactory tract also sends projections to the pyriform cortex and the entorhinal cortex, areas with reciprocal connections to the amygdala (20). Although the periamygdaloid cortex neurons project to deeper amygdaloid nuclei, the specific pathways conveying olfactory information through the amygdala have not been delineated.

In addition to its role in conditioning to explicit sensory stimuli, the amygdala is involved in the development of emotional responses to environmental context. The projections from the hippocampal formation to the amygdala through the fornix have been specifically implicated in spatial contextual conditioning (21 ,22). Thus, lesioning these projections specifically prevents fear conditioning to the chamber or the position within a maze in which aversive stimulation previously occurred (22 ,23 ,24 and 25). Other structures that participate in the modulation of contextual fear include the rostral perirhinal cortex and the ventrolateral PFC/ anterior (agranular) insula. Lesions of the latter regions reduce fear *reactivity* to contextual stimuli, but they do not affect CS acquisition or response extinction (26). In contrast, lesions placed in the rostral perirhinal cortex *after* fear conditioning interfere with the expression of conditioned fear responses elicited by visual and auditory stimuli when these stimuli are presented in contexts that differ from the initial conditioning context (27). Notably, genetic studies in mice identified a quantitative trait locus for contextual conditioning (28 ,29) that was associated with mouse “emotionality” in another study (30), although the molecular genetic, neurochemical, and functional anatomic correlates of this trait have not been established.

The projections from sensory thalamus to the LA are thought to support rapid conditioning to simple visual and auditory features, presumably accounting for fear responses below the level of conscious awareness (31). Thus, lesioning the *auditory cortex* before conditioning does not prevent conditioning to single auditory tones. In contrast, projections to the LA from the primary sensory and sensory association cortices appear to be essential for some aspects of conditioned responding to more complex sensory stimuli (4 ,32). These relationships are modality specific. For example, disruption of the projections from the auditory thalamus and auditory cortex to the LA specifically prevents acquisition of fear conditioning to auditory stimuli and fear-conditioned responses to previous auditory CSs (33 ,34 and 35).

After sensory input enters the LA, the neural representation of the stimulus is distributed in parallel to various amygdaloid nuclei, where it may be modulated by diverse functional systems, such as those mediating memories from past experiences or knowledge about ongoing homeostatic states (36). The most extensive extranuclear projections of the LA are composed of reciprocal projections to the basal and accessory basal nuclei and the central nucleus of the amygdala (CE) (37 ,38). Lesions of either the LA or the CE—but not of other amygdala nuclei—disrupt fear conditioning to a tone CS, a finding suggesting that this direct projection from LA to CE is sufficient to mediate conditioning to simple sensory features (4).

The projections from LA to the basal amygdaloid nuclei also participate in forming long-lasting memory traces for fear conditioning (2 ,15 ,39). Functional inactivation of the lateral and basal amygdaloid nuclei before pavlovian fear conditioning interferes with acquisition of learning, whereas inactivation immediately after conditioning has no effect on memory consolidation (40). The basal nuclei have widespread intranuclear connections and also project to other amygdalar nuclei, including the CE and the LA (41). They also share extensive, reciprocal projections with the orbital and mPFC (43). The basal nuclei are thus anatomically positioned to modulate neuronal responses in both the LA and the PFC (42 ,43).

The plasticity within the amygdala that constitutes memory for conditioning experiences has been shown to involve long-term potentiation–like associative processes (6). Plasticity related to fear learning also occurs in cortical areas, presumably making possible the establishment of explicit or declarative memories about the fear-related event through interactions with the medial temporal lobe memory system (44 ,45). The influence of the amygdala on cortically based memories has been most clearly characterized with respect to *late* plastic components of the auditory cortex neuronal responses to a CS. Single-unit recordings during fear conditioning indicate that some auditory cortex neurons, which before conditioning did not respond to the CS tone, develop

late-conditioned responses (i.e., 500 to 1,500 milliseconds after CS onset) that anticipate the US and show extinction-resistant memory storage (46). These late-conditioned auditory cortical neuronal responses take more trials to learn and respond more slowly than LA neurons within trials, and their late development is prevented by amygdala lesions. Thus, whereas rapid conditioning of fear responses to potentially dangerous stimuli depends on plasticity in the amygdala, learning involving higher cognitive (i.e., mnemonic and attentional) processing of fear experiences may depend on plasticity involving cortical neurons that is influenced by neural transmission from the amygdala to the cortex.

Other auditory cortex neurons show an early (less than 50 milliseconds of stimulus onset) plastic component during fear conditioning, in which the preexisting electrophysiologic responses of auditory cortex neurons to the CS become enhanced by conditioning (46). This short-latency plasticity within the auditory cortex appears to depend on input from the auditory thalamus and is unaffected by amygdala lesions. Nevertheless, such short-latency responses are extinguished more quickly (during repeated exposure to the CS alone) in animals with amygdala lesions, a finding implying that the amygdala is involved in preventing extinction of these responses.

In human neuroimaging studies, hemodynamic activity in the amygdala increases during initial exposures to fear-conditioned stimuli (47 ,48). However, during repeated, unreinforced exposures to the same stimulus, single-trial functional magnetic resonance imaging (fMRI) studies show that this initial elevation of hemodynamic activity attenuates and subsequently decreases to less than baseline (47). This observation suggests that synaptic input into the amygdala may be actively reduced during the extinction process (49), although the level at which this suppression of afferent synaptic activity into (or within) the amygdala is being suppressed during nonreinforced exposures to the CS has not been established.

Activation of the amygdala during an emotional event enhances the strength of long-term memory for emotional stimuli represented in other cortical memory circuits as well (16 ,50 ,51). These circuits presumably involve the medial temporal lobe memory system, which has extensive anatomic connections with the amygdala and presumably provides a neuroanatomic substrate for the interaction between storage and explicit recall of affectively salient memories (16). For example, as healthy humans read stories, the magnitude of physiologic activation in the amygdala correlates both with the negative emotional intensity and with the subsequent recall performance of the story's content (52 ,53). Physiologic activity in the amygdala and the hippocampus measured during memory encoding reportedly correlates with enhanced episodic memory for pleasant as well as aversive visual stimuli (54), and the amygdala's role in modulating emotional memory may depend more generally on the degree of arousal or the behavioral salience associated with verbally conveyed information (9 ,16).

Human neuroimaging and electrophysiologic and lesion analysis studies have also demonstrated that the amygdala is involved in the recall of emotional or arousing memories (4 ,53 ,55). In humans, bursts of electroencephalographic activity have been recorded in the amygdala during recollection of specific emotional events (56). Moreover, electrical stimulation of the amygdala can evoke emotional experiences (especially fear or anxiety) (57) and the recollection of emotionally charged life events from remote memory (58).

Role of the Amygdala in Organizing Emotional Expression

The amygdaloid output nuclei, especially the CE, receive convergent information from multiple amygdala regions and generate behavioral responses that are thought to reflect the sum of neuronal activity produced by different amygdaloid nuclei (36). The CE comprises the interface between the amygdala and the motor, autonomic, and neuroendocrine systems involved in expressing fear behavior (4 ,5). The CE projects to nuclei in the hypothalamus, midbrain, and medulla that mediate the neuroendocrine, autonomic, and behavioral responses associated with fear and anxiety. For example, the amygdala facilitates stress-related corticotropin-releasing hormone (CRH) release by both intrinsic CRH-containing neurons and bisynaptic (double γ-aminobutyric acid–ergic [GABAergic]) anatomic projections to the paraventricular nucleus (PVN) of the hypothalamus (59). Electrical stimulation of the CE produces responses similar to those elicited by fear-conditioned stimuli (60 ,61), and lesions of the CE prevent the expression of fear responses of various types (4 ,62 ,63). In contrast, lesioning of specific structures efferent to the CE, such as the lateral hypothalamus or periaqueductal gray (PAG), produces selective deficits in cardiovascular or somatomotor behavioral fear responses, respectively (1 ,64).

The amygdala also sends projections to the thalamus, the nucleus accumbens, the ventromedial caudate, and parts of the ventral putamen that participate in organizing motor responses to threatening stimuli (65). For example, activation of the amygdalar projections to the ventral striatum arrests goal-directed behavior in experimental animals (66), a finding suggesting a possible neural mechanism for the cessation of motivated or reward-directed behavior during anxiety and panic. The amygdala may also influence motor behavior by projections through the hypothalamus and PAG (1). For example, in experimental animals, stimulation of the lateral PAG produces defensive behaviors, sympathetic autonomic arousal, and hypoalgesia, whereas stimulation of the ventrolateral PAG produces social withdrawal and behavioral quiescence, as in response to deep injury or visceral pain (67).

Other Roles of the Amygdala in Fear Processing

The amygdala also appears to play important roles in mediating innate fear and in processing affective elements of social interactions (68). Amygdala lesions cause rats to lose their fear of cats and monkeys to lose their fear of snakes (reviewed in ref. 4). In monkeys, amygdala lesions reduce aggression as well as fear and cause animals to become more submissive to dominant animals (69). In humans, blood flow increases in the amygdala as subjects view faces expressing fear or sadness (70 ,71), and amygdala lesions impair the ability to recognize fear or sadness in facial expression (55 ,72) and fear and anger in spoken language (73).

Bed Nucleus of the Stria Terminalis: Hypothesized Role in Anxiety

The hypothalamic and brainstem structures that mediate the expression of emotional behavior can also be activated directly by the bed nucleus of the stria terminalis (BNST) (5). Anxiety-like responses elicited either by exposure to a threatening environment for several minutes or by intraventricular administration of CRH appear to be specifically mediated by the BNST, rather than the CE (5). This system is thus hypothesized to play a role in mediating anxiety during exposure to less explicit, or less well defined, sensory cues or to contexts that occur over a longer duration.

Other Temporal Cortical Structures

The perirhinal cortex shares reciprocal anatomic connections with the amygdala (74), and it is thought to play a role in conveying information about complex visual stimuli to the amygdala during presentation of fear-conditioned visual stimuli. Lesions of the anterior perirhinal cortex, the basolateral nucleus of the amygdala, or the CE can each completely eliminate fear-potentiated startle during exposure to some conditioned visual stimuli (75 ,76). In contrast, complete removal of the entire visual cortex, insular cortex, mPFC, and posterior perirhinal cortex produces no significant effect on the magnitude of fear-potentiated startle, and lesions of the frontal cortex only partly attenuate fear-potentiated startle. The perirhinal cortex receives input regarding conditioned visual stimuli from the lateral geniculate nucleus, and lesions of this structure can also block fear-potentiated startle (77). Finally, the anterior perirhinal cortex receives afferent projections from the visual cortices as well as from the anterior cingulate cortex (ACC), the infralimbic cortex, and the parietal cortex (74), structures implicated in modulating behavioral responses to fear-conditioned stimuli.

The temporopolar cortex has been implicated in modulating autonomic aspects of emotional responses and in processing emotionally provocative visual stimuli. Electrical stimulation of various sites within the temporopolar cortex can alter a variety of autonomic responses (reviewed in ref. 1). In humans with simple phobias or posttraumatic stress disorder (PTSD), physiologic activity increases in the anterior temporopolar cortex during experimentally induced exacerbations of anxiety involving visual exposure to phobic stimuli or word scripts describing traumatic events, respectively (78 ,79). Blood flow also increases in the anterior temporopolar cortex of healthy humans during exposure to emotionally provocative visual stimuli, whether the stimuli convey “sad,” “disgusting,” or “happy” content, relative both to conditions involving exposure to emotionally “neutral” visual stimuli and to conditions in which corresponding emotional states are elicited by recall of autobiographic information (80 ,81). Portions of the temporopolar cortex may thus function as sensory association areas that participate in evaluating the emotional salience of actual or anticipated stimuli and in modulating autonomic responses to such stimuli.

Neuroendocrine and Autonomic Responses during Fear or Stress

The peripheral hormonal and autonomic responses to threat mediated by the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic and parasympathetic autonomic nervous systems also play adaptive roles in responding to threat or stress (5). Stimulation of the lateral nucleus of the hypothalamus by afferent projections from the CE of the amygdala, the BNST, or the ventral striatum (82) activates the sympathetic system and produces increases in blood pressure and heart rate, sweating, piloerection, and pupillary dilation. Stress stimulates release of CRH from the PVN of the hypothalamus and amygdala. The CRH secretion from the PVN, in turn, increases peripheral adrenocorticotropic hormone (ACTH) levels, and this stimulates the adrenal glands to secrete cortisol. The ACC, anterior insula, and posterior orbital cortex send anatomic projections to the hypothalamus that participate in modulating or inhibiting cardiovascular and endocrine responses to threat and stress (1 ,43 ,83).

The vagus and splanchnic nerves constitute the major efferent projections of the parasympathetic nervous system to the viscera. The vagal nuclei receive afferent projections from the lateral hypothalamus, the PVN, the LC, the amygdala, the infralimbic cortex, and the prelimbic portion of the ACC (43 ,84). The splanchnic nerves receive afferent connections from the LC. The innervation of the parasympathetic nervous system from these limbic structures is thought to mediate visceral symptoms associated with anxiety, such as gastrointestinal and genitourinary disturbances (Fig. 63.1).

Role of Prefrontal Cortical Structures in Modulating Fear and Anxiety Behavior

Multiple areas of the medial and orbital PFC appear to play roles in modulating anxiety and other emotional behaviors.

These PFC structures are thought to participate in interpreting the higher-order significance of experiential stimuli, in modifying behavioral responses based on competing reward versus punishment contingencies, and in predicting social outcomes of behavioral responses to emotional events (8 ,11 ,85 ,86). These areas share extensive, reciprocal projections with the amygdala, through which the amygdala can modulate PFC neuronal activity and the PFC can modulate amygdala-mediated responses to emotionally salient stimuli (17 ,18 ,42 ,43).

Areas within the orbital and mPFC and the anterior insula also participate in modulating peripheral responses to stress, including heart rate, blood pressure, and glucocorticoid secretion (13 ,17 ,43 ,87). The neuronal activities within these areas are, in turn, modulated by various neurotransmitter systems that are activated in response to stressors and threats. For example, the noradrenergic, dopaminergic, and serotonergic systems play roles in enhancing vigilance, modulating goal-directed behavior, and facilitating decision making about probabilities of punishment versus reward by modulating neuronal activity in the PFC (86 ,88 ,89 and 90).

Medial Prefrontal Cortex

The mPFC areas implicated in anxiety and fear-related behavior in humans and experimental animals *include* the infralimbic cortex, the ACC located ventral (“subgenual”) and anterior (“pregenual”) to the genu of the corpus callosum, and a more dorsal mPFC region that extends from the rostral ACC (BA 24, 32) toward the frontal pole (91). The reciprocal projections between the amygdala and the mPFC are hypothesized to play critical roles in attenuating fear responses and extinguishing behavioral responses to fear-conditioned stimuli that are no longer reinforced (17 ,18). Lesions of the ACC in rats resulted in enhanced freezing to a fear-conditioned tone, a finding suggesting that this mPFC region may be involved in fear reduction (17). In addition, neurons in the rat prelimbic cortex (thought to be homologous to subgenual PFC) reduce their spontaneous firing activity in the presence of a conditioned, aversive tone to an extent that is inversely proportional to the magnitude of fear (42). This suppression of prelimbic cortex neuronal firing activity is inversely correlated with increases in amygdala neuronal activity. Finally, lesions of the infralimbic cortex specifically interfere with the recall of extinction processes after long delays between the acquisition of extinction learning and reexposure to the initial CS (18). Extinction does not appear to occur by erasing memory traces of the CS-US association, but rather by new learning through which the behavioral response to the CS is actively inhibited (31).

In humans, the pregenual ACC shows areas of elevated hemodynamic activity during a variety of anxiety states elicited in healthy or anxiety-disordered subjects (reviewed in ref. 49). Electrical stimulation of this region elicits fear, panic, or a sense of foreboding in humans and vocalization in experimental animals (reviewed in ref. 7). Nevertheless, physiologic activity also increases in the ACC during the generation of positive emotions in healthy humans (92 ,93) and during depressive episodes in some subtypes of major depressive disorder (MDD) (94 ,95).

The subgenual ACC has been implicated in healthy sadness, MDD, mania, and PTSD (90 ,96 ,97). In patients with familial unipolar and bipolar depression, reductions in cerebral blood flow (CBF) and metabolism were associated with left-lateralized reductions in the volume of the corresponding cortex (96 ,98 ,99). The subgenual PFC activity shows a mood state dependency in which the metabolism is higher in the depressed than the remitted phase of MDD, consistent with the findings that blood flow increases in this region in healthy, nondepressed humans during experimentally induced sadness (85 ,100 ,101) and in persons with PTSD during internally generated imagery of past trauma (97).

Both the subgenual and the pregenual ACC share reciprocal anatomic connections with areas implicated in emotional behavior such as the posterior orbital cortex, amygdala, hypothalamus, nucleus accumbens, PAG, ventral tegmental area (VTA), raphe, LC, and NTS (Fig. 63.1) (102 ,103). Humans with mPFC lesions that include the pregenual and subgenual ACC show abnormal autonomic responses to emotionally provocative stimuli, inability to experience emotion related to concepts, and inability to use information regarding the probability of aversive social consequences versus reward in guiding social behavior (104). In rats, bilateral or *right*-lateralized lesions of the ventral mPFC composed of infralimbic, prelimbic, and ACC cortices *attenuate* corticosterone secretion, sympathetic autonomic responses, and gastric stress disorders during restraint stress or exposure to fear-conditioned stimuli (17 ,83 ,105). In contrast, *left*-sided lesions of this cortical strip *increase* sympathetic arousal and corticosterone responses to restraint stress (105). Finally, the ventral ACC contains glucocorticoid receptors that, when stimulated, inhibit stress-induced corticosterone release in rats (87).

Physiologic activity also increases in more dorsal mPFC areas in healthy humans as they perform tasks that elicit emotional responses or require emotional evaluations (81 ,106 ,107). During anxious anticipation of an electrical shock, CBF increases in the rostral mPFC (vicinity of anterior BA24, BA32, and rostral BA9), and the magnitude of ΔCBF correlates inversely with changes in anxiety ratings and heart rate (107). In rats, lesions of the rostral mPFC result in exaggerated heart rate responses to fear-conditioned stimuli, and stimulation of these sites attenuates defensive behavior and cardiovascular responses evoked by amygdala stimulation (83). In primates, whereas BA24 and 32 have extensive reciprocal connections with the amygdala through which they may modulate emotional expression, the BA9 cortex has only sparse projections to the amygdala. Nevertheless, all three areas send extensive efferent projections to the PAG and the hypothalamus through which cardiovascular

responses associated with emotional behavior can be modulated (43 ,108).

In the depressed phase of MDD and bipolar disorder, metabolic activity is abnormal in the dorsomedial and dorsal anterolateral PFC (in the vicinity of rostral BA9) (91 ,109). Postmortem studies of these regions have shown abnormal reductions in the size of glia and neurons in MDD (110). Given the preclinical and neuroimaging evidence presented earlier, indicating that this area may modulate anxiety, it may be hypothesized that dysfunction of this mPFC area contributes to the development of anxiety symptoms in mood disorders.

Orbital and Anterior Insular Cortex

Other areas of the PFC that are implicated in studies of fear or anxiety in human and nonhuman primates are the posterior and lateral orbital cortex, the anterior (agranular) insula, and the ventrolateral PFC (1 ,43). Physiologic activity increases in these areas during experimentally induced anxiety states in healthy subjects and in subjects with obsessive-compulsive disorder (OCD), simple phobia, and panic disorder (PD) (49 ,111). (See Chapter 65) The baseline metabolic activity is also abnormally elevated in these regions in unmedicated study subjects with primary MDD (91) and OCD (112) scanned while resting with eyes closed. The elevated activity in these areas in both MDD and OCD appears state dependent, and effective antidepressant or antiobsessional treatment results in decreases in CBF and metabolism in the medicated-improved relative to the unmedicated-symptomatic phase (112 ,113 and 114).

A complex relationship exists between anxiety-depressive symptoms and physiologic activity in the orbital cortex and the ventrolateral PFC. In MDD, whereas CBF and metabolism increase in these areas in the depressed relative to the remitted phase, the magnitude of these measures correlates inversely with ratings of depressive ideation and severity (115 ,116). Similarly, posterior orbital cortex flow increases in OCD and animal phobic subjects during exposure to phobic stimuli and in healthy subjects during induced sadness, but this change in CBF correlates inversely with changes in obsessive thinking, anxiety, and sadness, respectively (114 ,117 ,118).

These data appear consistent with electrophysiologic and lesion analysis data showing that the orbital cortex participates in modulating behavioral and visceral responses associated with fearful, defensive, and reward-directed behavior as reinforcement contingencies change. Nearly one-half of the orbital cortex pyramidal neurons alter their firing rates during the delay period between stimulus and response, and this firing activity relates to the presence or absence of reinforcement (11). These cells are thought to play roles in extinguishing unreinforced responses to aversive or appetitive stimuli (7 ,11 ,66). The posterior and lateral orbital cortex and the amygdala send projections to each other and to overlapping portions of the striatum, hypothalamus, and PAG through which these structures modulate each other's neural transmission (Fig. 63.1) (42 ,66 ,108 ,119). For example, the defensive behaviors and cardiovascular responses evoked by electrical stimulation of the amygdala are attenuated or ablated by concomitant stimulation of orbital sites, which, when stimulated alone, exert no autonomic effects (120).

Humans with orbital cortex lesions show impaired performance on tasks requiring application of information related to punishment or reward, perseverate in behavioral strategies that are unreinforced, and exhibit difficulty in shifting intellectual strategies in response to changing task demands (11 ,121). Likewise, monkeys with surgical lesions of the lateral orbital cortex and ventrolateral PFC demonstrate “perseverative interference,” characterized by difficulty in learning to withhold prepotent responses to nonreinforced stimuli as reinforcement contingencies change (122). Activation of the orbital cortex during anxiety or obsessional states may thus reflect endogenous attempts to attenuate emotional expression or to interrupt unreinforced aversive thought and emotion (91). Conversely, dysfunction of the orbital cortex may contribute to pathologic anxiety and obsessional states by impairing the ability to inhibit nonreinforced or maladaptive emotional, cognitive, and behavioral responses to social interactions and sensory or visceral stimuli.

Posterior Cingulate Cortex

Many functional imaging studies report that exposure to aversive stimuli of various types increases physiologic activity in the retrosplenial cortex and other portions of the posterior cingulate gyrus (reviewed in ref. 123). Posterior cingulate cortical flow and metabolism have also been found abnormally elevated in some studies of depressed subjects with MDD (reviewed in ref. 91). In contrast, Mayberg et al. reported that script-driven sadness resulted in decreased posterior cingulate activity in healthy subjects, and flow was decreased in depressed relative to remitted subjects with MDD, findings raising the possibility that this large region is functionally heterogenous with respect to emotional behavior (101). The posterior cingulate cortex appears to serve as a sensory association cortex and may participate in processing the affective salience of sensory stimuli. The posterior cingulate cortex sends a major anatomic projection to the ACC, through which it may relay such information into the limbic circuitry (124).

FUNCTIONAL ANATOMIC CORRELATES OF SPECIFIC ANXIETY DISORDERS

Neuroimaging studies have assessed neurophysiologic abnormalities in anxiety-disordered samples in the baseline,

“resting” condition and during symptom provocation. These data converge with those obtained from studies of healthy subjects and of experimental animals to implicate the limbic, paralimbic, and sensory association areas reviewed earlier in the functional anatomy of emotional behavior. Nevertheless, the results of most of the imaging studies reviewed herein await replication, and the data they provide do not clearly establish whether differences between anxiety-disordered and control subjects reflect physiologic correlates of anxiety symptoms or traitlike biological abnormalities underlying the vulnerability to anxiety syndromes.

Panic Disorder

The baseline state in PD is characterized by mild to moderate levels of chronic anxiety (termed *anticipatory anxiety*). In this state, abnormalities of CBF and glucose metabolism have been reported in the vicinity of the hippocampus and parahippocampal gyrus. Reiman et al. initially reported an abnormal resting asymmetry (left less than right) of blood flow and oxygen metabolism in a region of interest placed over the parahippocampal gyrus (125). Nordahl et al. similarly found that glucose metabolism measured over the hippocampus-parahippocampal gyrus was asymmetric and concluded that this abnormality reflected an abnormal metabolic elevation on the right side (126). Bisaga et al. also found abnormal metabolism in this vicinity, but with the opposite laterality (i.e., elevated metabolism in the left hippocampal-parahippocampal area) in lactate-sensitive PD study subjects relative to healthy controls (127). In contrast, De Cristofaro et al. reported that resting perfusion, measured using single photon emission computed tomography (SPECT) and [^{99m}Tc]hexamethylpropyleneamineoxime (HMPAO), was abnormally decreased in the hippocampus, bilaterally, in lactate-sensitive PD study subjects relative to controls (128).

Each of these studies employed region-of-interest based approaches that were incapable of localizing the center of mass of the abnormality in this region. Reanalysis of some of these data using a voxel-by-voxel approach suggested that the abnormal radioactivity in the vicinity of the mesiotemporal cortex may actually reflect elevated metabolism in the adjacent midbrain (111). This midbrain region, which may reflect the lateral PAG, has been implicated in lactate-induced panic (129), other acute anxiety states (130), and animal models of panic attacks (67).

Study subjects with PD have also been imaged during panic elicited using a variety of chemical challenges. Panic attacks induced by intravenous sodium lactate infusion were associated with regional CBF increases in the anterior insula, the anteromedial cerebellum, and the midbrain (129); areas of increased CBF may also exist in the temporal polar cortex, but these findings were confounded by corresponding increases in the adjacent facial muscles during severe anxiety (115). Blood flow also increased in these regions in animal phobic subjects during exposure to phobic stimuli and in healthy subjects during the threat of a painful electrical shock, findings suggesting that these CBF changes reflect the neurophysiologic correlates of fear processing in general (111 ,130). Consistent with this hypothesis, anxiety attacks induced in healthy humans using cholecystokinin tetrapeptide (CCK-4) were also associated with CBF increases in the insular-amygdala region and the anteromedial cerebellum (131).

Indirect evidence suggests that the neurophysiologic responses in the PFC during panicogen challenge may differ between PD subjects and healthy controls. For example, panic attacks induced using CCK-4 were associated with CBF increases in the ACC in healthy humans (131), but flow did not significantly change in the ACC in subjects with PD during lactate-induced panic (129). The ACC was also a region where flow significantly increased in healthy subjects but not in subjects with PD during fenfluramine challenge in a study in which fenfluramine induced panic attacks in 56% of subjects with PD but in only 11% of control subjects (132). Finally, Cameron et al. found that normalized medial frontal CBF increased in healthy controls after yohimbine administration (i.e., after normalizing to remove effects on whole brain CBF) (133), whereas Woods et al. found that the relative prefrontal cortical flow was decreased in PD relative to control subjects following yohimbine challenge (134).

Structural MRI studies have begun to investigate whether morphometric or morphologic abnormalities may exist in PD. Ontiveros et al. reported qualitative abnormalities of temporal lobe structure in PD (135), although these findings have not been replicated. Vythilingam et al. reported that hippocampal volume did not differ between PD and healthy control subjects (136).

Phobias

In simple animal phobias, phobic anxiety was imaged by acquiring blood flow scans during exposures to the feared animal. During the initial fearful scans, flow increased in the lateral orbital-anterior insular cortex, bilaterally, the pregenual ACC, and the anteromedial cerebellum (78 ,111), areas where CBF also increases in other anxiety states (see earlier). During the development of habituation to phobic stimuli, the magnitude of the hemodynamic responses to the phobic stimulus diminished in the anterior insula and the medial cerebellum, but it increased in the left posterior orbital cortex in an area where flow had not changed during exposures that preceded habituation (117). The magnitude of the CBF increase in this latter region was inversely correlated with the corresponding changes in heart rate and anxiety

ratings. As discussed earlier, the posterior orbital cortex was a site where CBF increased in subjects with OCD during exposure to phobic stimuli, with the increase in flow inversely correlated with obsessional ratings (114).

In social anxiety disorder, an aversive conditioning paradigm (in which the US was an aversive odor and the CS was a picture of a human face) showed that hemodynamic activity decreased in the amygdala and the hippocampus during presentations of the CS in healthy controls, but it increased in social phobic subjects (137). Interpretation of these data was confounded by the problem that both human faces and aversively CSs normally activate the amygdala, so it remained unclear which of the stimuli produced abnormal responses in social phobia. Nevertheless, these data appear conceptually intriguing, given the role of hippocampal-amygdalar projections in mediating contextual fear and the possibility that deficits in the transmission of information regarding context may be involved in the pathogenesis of phobias (21).

Posttraumatic Stress Disorder

PTSD is hypothesized to involve the emotional-learning circuitry associated with the amygdala, because the traumatic event constitutes a fear-conditioning experience, and subsequent exposure to sensory, contextual, or mnemonic stimuli that recall aspects of the event elicits psychological distress and sympathetic arousal. Potentially consistent with this expectation, some studies demonstrated activation of the amygdala as patients with PTSD listened to auditory scripts describing the traumatic event (79) or to combat sounds (in combat-related PTSD) (138) or generated imagery related to the traumatic event without sensory cues (139). However, other studies found no significant changes in amygdala CBF as patients with PTSD listened to scripts describing the traumatic event or viewed trauma-related pictures, and studies comparing CBF responses with trauma-related stimuli have not shown significant differences in the amygdala between patients with PTSD and trauma-matched, non-PTSD control subjects (97 ,139 ,140 and 141). The extent to which these negative findings reflect limitations in statistical sensitivity or in positron emission tomography (PET) temporal resolution must be addressed in provocation studies involving larger subject samples and employing fMRI instead of PET. In this regard, it is noteworthy that a preliminary fMRI study found exaggerated hemodynamic changes in the amygdala in patients with PTSD relative to trauma-matched, non-PTSD control subjects during exposure to pictures of fearful faces presented using a backward-masking technique (142). If replicated, this finding may suggest that the emotional dysregulation associated with PTSD may involve amygdalar responses to emotional stimuli of various types.

Other limbic and paralimbic cortical structures have also been implicated in provocation studies of PTSD. In both patients with PTSD and trauma-matched, non-PTSD control subjects, CBF increases in the posterior orbital cortex, anterior insula, and temporopolar cortex during exposure to trauma-related stimuli, but these changes have generally not differentiated PTSD and control samples (79 ,139 ,140). In contrast, the pattern of CBF changes elicited in the mPFC by traumatic stimuli may differ between PTSD and control subjects. During exposure to trauma-related sensory stimuli, flow decreased in the left (97 ,140) but increased in the right pregenual ACC in PTSD (79 ,138), a finding potentially consistent with the evidence reviewed earlier that the role of the mPFC in emotional behavior is lateralized (105). However, CBF in the right pregenual ACC increased significantly more in non-PTSD, trauma-matched control subjects than in patients with PTSD (139). Moreover, in the infralimbic cortex, CBF *decreased* in patients with combat-related PTSD but *increased* in combat-matched, non-PTSD

control subjects during exposure to combat-related visual and auditory stimuli (141).

Given evidence that the ACC and the infralimbic cortex play roles in extinguishing fear-conditioned responses (17 ,18), the observation that patients with PTSD fail to activate these structures to a similar extent as traumatized, non-PTSD control subjects during exposure to traumatic cues (139 ,141) suggests that neural processes mediating extinction to trauma-related stimuli may be impaired in PTSD. Compatible with this hypothesis, PTSD samples have been shown to acquire *de novo* conditioned responses more readily and to extinguish them more slowly than control samples (143 ,144). Such an impairment could conceivably be related to the vulnerability to developing PTSD, because PTSD occurs in only 5% to 20% of individuals exposed to similar traumatic events.

Structural MRI studies of PTSD have identified subtle reductions in the volume of the hippocampus in PTSD samples relative to healthy or traumatized, non-PTSD control samples (145 ,146 ,147 and 148). Although limitations existed in these studies in the matching of alcohol use or abuse between PTSD and control samples, the reductions in hippocampal volume did not correlate with the extent of alcohol exposure in the PTSD samples, and no volumetric differences were found between PTSD and control samples in the amygdala, entire temporal lobe, caudate, whole brain, or lateral ventricles. Although the magnitude of the reduction in hippocampal volume only ranged from 5% to 12% in the PTSD samples relative to trauma-matched controls, these abnormalities were associated with short-term memory deficits in some studies (145 ,147). It remains unclear whether the difference in hippocampal volume may reflect a result of the chronic stress associated with PTSD (e.g., from sustained exposure to elevated glucocorticoid concentrations) or a biological antecedent that may confer risk for developing PTSD (149 ,150).

Obsessive-Compulsive Disorder

The anatomic circuits involved in the production of obsessions and compulsions have been elucidated by converging evidence from functional neuroimaging studies of OCD, analysis of lesions resulting in obsessive-compulsive symptoms, and observations regarding the neurosurgical interventions that ameliorate OCD (113 ,114 ,151). PET studies of OCD have shown that “resting” CBF and glucose metabolism are abnormally increased in the orbital cortex and the caudate nucleus bilaterally in primary OCD (reviewed in ref. 112). With symptom provocation by exposure to relevant phobic stimuli (e.g., skin contact with “contaminated” objects for patients with OCD who have germ phobias), flow increased further in the orbital cortex, ACC, caudate, putamen, and thalamus (114). During effective pharmacotherapy, orbital metabolism decreased toward normal, and both drug treatment and behavioral therapy were associated with a reduction of caudate metabolism (112). The baseline areas of hypermetabolism in the orbital cortex and the caudate may thus reflect physiologic concomitants of obsessive thoughts or chronic anxiety, and, conversely, the reduction in caudate metabolism associated with effective (but not ineffective) treatment may reflect a physiologic correlate of symptom resolution rather than a primary mechanism of treatment.

Based on the evidence reviewed earlier from electrophysiologic and lesion analysis studies indicating that the orbital cortex participates in the correction of behavioral responses that become inappropriate as reinforcement contingencies change, posterior orbital areas may be specifically activated as an endogenous attempt to interrupt patterns of nonreinforced thought and behavior in OCD (11 ,91). Compatible with this hypothesis, the posterior orbital cortex CBF increases during symptom provocation in OCD, but the magnitude of this increase correlates *inversely* with the corresponding rise in obsession ratings (r = -0.83) (114). In contrast, flow also increases in an area of the right anterior orbital cortex implicated in a variety of types of mnemonic processing, and the change in CBF in this region correlates positively with changes in obsession ratings (114 ,152).

The neurologic conditions associated with the development of secondary obsessions and compulsions also provide evidence that dysfunction within circuits formed by the basal ganglia and the PFC may be related to the pathogenesis of OCD. Such conditions involve lesions of the globus pallidus and the adjacent putamen: Sydenham chorea (a poststreptococcal autoimmune disorder associated with neuronal atrophy in the caudate and putamen), Tourette syndrome (an idiopathic syndrome characterized by motoric and phonic tics that may have a genetic relationship with OCD), chronic motor tic disorder, and lesions of the ventromedial PFC (151 ,152 ,153 and 154). Several of these conditions are associated with complex motor tics (repetitive, coordinated, involuntary movements occurring in patterned sequences in a spontaneous and transient manner). It is conceivable that complex tics and obsessive thoughts may reflect homologous, aberrant neural processes manifested within the motor and cognitive-behavioral domains, respectively, because of their origination in distinct portions of the cortical-striatal-pallidal-thalamic circuitry (113 ,155).

In contrast to the regional metabolic abnormalities found in primary OCD, imaging studies of obsessive-compulsive syndromes arising in the setting of Tourette syndrome or basal ganglia lesions have not found elevated blood flow and metabolism in the caudate and in some cases have found reduced metabolism in the orbital cortex in such subjects relative to controls (111 ,151). The differences in the functional anatomic correlates of primary versus secondary OCD are consistent with a neural model in which dysfunction arising at various points within the ventral prefrontal cortical-striatal-pallidal-thalamic circuitry may result in pathologic obsessions and compulsions. This circuitry appears

to be generally involved in organizing internally guided behavior toward a reward, switching of response strategies, habit formation, and stereotypic behavior (66 ,155).

These circuits have also been implicated in the pathophysiology of MDD, another illness in which intrusive, distressing thoughts recur to an extent that the ability to switch to goal-oriented, rewarding cognitive-behavioral sets is impaired (91). Although MDD and OCD appear distinct in their course, prognosis, genetics, and neurochemical concomitants, substantial comorbidity exists across these syndromes. Major depressive episodes occur in about one-half of patients with OCD, pathologic obsessions can arise in primary MDD, and the pharmacologic interventions that ameliorate OCD can also effectively treat MDD. Moreover, the neurosurgical procedures that are effective at reducing both obsessive-compulsive and depressive symptoms in intractable cases of OCD and MDD interrupt white matter tracts carrying neural projections between the frontal lobe, the basal ganglia, and the thalamus (155). The clinical comorbidity across these two disorders may thus reflect involvement of an overlapping neural circuitry by otherwise distinct pathophysiologic processes.

NEUROCHEMICAL BASIS OF FEAR AND ANXIETY

Part of "63 - Neurobiological Basis of Anxiety Disorders "

The neuroanatomic circuits that support fear and anxiety behavior are modulated by a variety of chemical neurotransmitter systems. These include the peptidergic neurotransmitters, CRH, neuropeptide Y (NPY), and substance P, the monoaminergic transmitters, NE, serotonin (5-hydroxytryptamine or 5-HT), and dopamine (DA), and the amino acid transmitters, GABA and glutamate. The neurotransmitter systems that have been best studied in association with responses to stress or threat involve the HPA axis and the central noradrenergic system. These neurochemical systems subserve important adaptive functions in preparing the organism for responding to threat or stress, by increasing vigilance, modulating memory, mobilizing energy stores, and elevating cardiovascular function. Nevertheless, these biological responses to threat and stress can become maladaptive if they are chronically or inappropriately activated. Additional neurochemical systems that play important roles in modulating stress responses and emotional behavior include the central GABAergic, serotonergic, dopaminergic, opiate, and NPY systems. The preclinical and clinical literature regarding these neurochemical concomitants of stress and fear and their potential relevance to the pathophysiology of anxiety disorders are reviewed in the following sections.

Role of the Central Noradrenergic System in Fear and Anxiety

Exposure to stressful stimuli of various types increases central noradrenergic function. Thus, exposure to fear-conditioned stimuli, immobilization stress, foot shock, or tail pinch increases NE turnover in the LC, the hypothalamus, the hippocampus, the amygdala, and the cerebral cortex (156). The firing activity of LC neurons also increases during exposure to fear-conditioned stimuli and other stressors or threats (157 ,158 and 159). For example, the firing activity of NE neurons in the cat LC increases two- to threefold during confrontation with a dog or an aggressive cat, but it remains unchanged during exposure to other novel stimuli or to nonaggressive cats (160). However, repeated exposure to severe stressors from which the animal cannot escape results in the behavioral pattern termed *learned helplessness*, which is associated with depletion of NE, possibly reflecting a point at which NE synthesis cannot keep pace with NE release (161 ,162).

Acquisition of fear-conditioned responses requires an intact central noradrenergic system, a finding suggesting that NE release plays a critical role in fear learning (157 ,163 ,164). For at least some types of emotional learning, memory consolidation depends on noradrenergic stimulation of ß- and α_1-adrenoreceptors in the basolateral nucleus of the amygdala (15). The activation of NE release in such models may, in turn, depend on effects of stress hormones on noradrenergic neurons (15).

The responsiveness of LC neurons to future novel stressors can be enhanced by chronic exposure to some stressful experiences. In rats, the amount of NE synthesized and released in the hippocampus and the mPFC in response to a novel stressor or to local depolarization is increased after repeated exposure to chronic cold stress (165 ,166 and 167). This effect may result from a stress-mediated alteration in the sensitivity of presynaptic α_2-adrenoreceptors, which inhibit NE synthesis and release. In the native state, administration of the α_2-adrenoreceptor antagonists, idazoxan or yohimbine, increases the electrophysiologic response of LC neurons to stressful stimuli (without altering their basal firing rates) and increases NE release and synthesis, whereas administration of the α_2-adrenoreceptor agonist, clonidine, decreases NE release and synthesis (167 ,168). In chronically cold-stressed rats, idazoxan administration produces a greater increase in NE release and synthesis, and clonidine administration produces a blunted attenuation of NE release and synthesis relative to naive rats (167). Consistent with these observations, Torda et al. found that cold immobilization stress decreases the α_2-adrenoreceptor density in the hippocampus and the amygdala (169).

The effect of chronic stress on noradrenergic responses to subsequent, novel stressors may constitute a form of “behavioral sensitization,” a process by which single or repeated exposures to aversive stimuli or pharmacologic agents can

increase the behavioral sensitivity to subsequent stressors (reviewed in ref. 170). Such phenomena are hypothesized to account for clinical observations that patients with anxiety disorders report experiencing exaggerated sensitivity to psychosocial stress. Neural models for the pathogenesis of anxiety disorders built on sensitization phenomena thus hold that *repeated* exposure to traumatic stress comprises a risk factor for the subsequent development of anxiety disorders, particularly PTSD.

Noradrenergic Function in Anxiety Disorders

The recurrent symptoms of anxiety disorders, such as panic attacks, insomnia, exaggerated startle, and chronic sympathetic autonomic arousal, may conceivably reflect elevated noradrenergic function (171 ,172 and 173). Patients with PTSD and PD show evidence of heightened peripheral sympathetic nervous system arousal that, because of the correlation between peripheral sympathetic activity and central noradrenergic function, is compatible with the hypothesis of increased central NE activity in these disorders (174 ,175). Moreover, patients with PD, PTSD, and phobic disorders report that their hyperarousal symptoms and intrusive memories are attenuated by alcohol, benzodiazepines (BZDs), and opiates, agents known to decrease LC neuronal firing activity, but are exacerbated by cocaine, which increases LC neuronal firing. The risk of abuse of these substances appears increased in patients with anxiety disorders, a finding raising the possibility that such patients are "self-medicating" anxiety symptoms with these agents. It remains unclear, however, whether alterations in noradrenergic function play a primary, etiologic role in the pathogenesis of anxiety disorders, or instead reflect secondary, compensatory changes in response to disorders in other systems.

PD has been specifically associated with elevations of α_2-adrenoreceptor sensitivity and nocturnal urinary NE excretion (176), although ß-adrenoreceptor function, baseline heart rate and blood pressure, and other measures reflecting central NE secretion have not been consistently altered in PD (see Table 63.1) (177). Altered α_2-adrenoreceptor sensitivity is evidenced by findings that administration of the α_2-adrenoreceptor agonist, clonidine, results in greater hypotension and larger reductions in plasma 3-methoxy-4-hydroxyphenylethylene glycol (MHPG) in PD relative to control subjects (178 ,179 ,180 and 181). In addition, administration of the α_2-adrenoreceptor antagonist, yohimbine (which stimulates NE release by antagonizing presynaptic α_2-adrenoreceptors) produces exaggerated anxiogenic and cardiovascular responses and enhanced plasma MHPG and cortisol increases in PD relative to control subjects (133 ,172 ,173 ,182 ,183 ,184 ,185 and 186). Finally, yohimbine administration resulted in reduced relative frontal cortex flow in patients with PD that did not occur in control subjects, as measured with SPECT and [^{99m}Tc]HMPAO (134); it remains unclear, however, whether this difference reflected a differential physiologic sensitivity to yohimbine or an effect of greater anxiety in the patients with PD, because all the patients with PD but only one control subject developed increased anxiety in response to yohimbine. The sensitivity of α_2-adrenoreceptors also appears increased in PTSD. Patients with combat-related PTSD show increased behavioral, chemical, and cardiovascular responses to yohimbine, relative to healthy controls (187 ,188 and 189).

TABLE 63.1. EVIDENCE OF ALTERED CATECHOLAMINERGIC FUNCTION IN ANXIETY DISORDERS

	PTSD	Panic Disorder
Increased resting heart rate and blood pressure	+/–	+/–
Increased heart rate and blood pressure response to traumatic reminders/panic attacks	+++	++
Increased resting urinary NE and E	+	+/–
Increased resting plasma NE or MHPG	–	–
Increased plasma NE with traumatic reminders/panic attacks	+	+/–
Increased orthostatic heart rate response to exercise	+	+
Decreased binding to platelet α2 receptors	+	+/–
Decrease in basal and stimulated activity of cAMP	+/–	+
Decrease in platelet MAO activity	+	NS
Increased symptoms, heart rate and plasma MHPG with yohimbine noradrenergic challenge	++	+++
Differential brain metabolic response to yohimbine	+	+

–, One or more studies did not support this finding (with no positive studies), or the majority of studies do not support this finding; +/–, an equal number of studies support this finding and do not support this finding; +, at least one study supports this finding and no studies do not support the finding, or the majority of studies support the finding; ++, two or more studies support this finding, and no studies do not support the finding; +++, three or more studies support this finding, and no studies do not support the finding; cAMP, cyclic adenosine 3′, 5′-monophosphate; E, epinephrine; MAO, monoamine oxidase; MHPG, 3-methoxy-4-hydroxyphenylglycol; NE, norepinephrine; NS, not studied; PTSD, posttraumatic stress disorder.

Considerable evidence also indicates that noradrenergic function is abnormal in PTSD (see Table 63.1). Women with PTSD secondary to childhood sexual abuse showed elevated 24-hour urinary excretion of catecholamines and cortisol (190). In addition, men—but not women—with PTSD resulting from a motor vehicle accident exhibited elevated urinary levels of epinephrine, NE, and cortisol 1 month after the accident and still had higher epinephrine levels 5 months later (191). Similarly, maltreated children with PTSD excreted greater amounts of urinary DA, NE, and cortisol over 24 hours than controls, with the urinary catecholamine and cortisol output positively correlated with the duration of PTSD trauma and the severity of PTSD symptoms (192). Exposure to traumatic reminders (e.g., combat films or sounds) produced greater increases in plasma, epinephrine, NE, and cortisol in patients with PTSD than in control subjects (191 ,193 ,194), although baseline concentrations of catecholamines are not consistently altered in combat-related PTSD (188 ,189). Geracioti et al. found that cerebrospinal fluid (CSF) NE concentrations are abnormally elevated in PTSD (195). Finally, platelet α_2-adrenoreceptor density (196), platelet basal adenosine, isoproterenol, forskolin-stimulated cyclic adenosine monophosphate signal transduction (197), and basal platelet monoamine oxidase activity (198) were decreased in PTSD, findings hypothesized to reflect compensatory responses to chronically elevated NE release.

In study subjects with specific phobias, plasma NE and epinephrine concentrations, heart rate, blood pressure, and subjective anxiety ratings increase in response to exposure to phobic stimuli (199). Subjects with social anxiety disorder show greater increases in plasma NE during orthostatic challenge than healthy subjects or those with PD (200).The growth hormone response to intravenous clonidine (a marker of central α_2-adrenoreceptor function) is blunted in social anxiety disorder (201), although the density of lymphocyte ß-adrenoreceptors has not differed between social anxiety–disordered and control samples (202) (Table 63.1).

Finally, Gerra et al. reported that, plasma NE concentrations increased to a greater extent in male peripubertal patients with generalized anxiety disorder than in controls in response to a psychological stress test (203). However, the

pretest baseline NE concentrations did not differ between the anxious and control subjects.

Hypothalamic-Pituitary-Adrenal Axis and Corticotropin-Releasing Hormone

Exposure to acute stress of various types results in release of CRH, ACTH, and cortisol. This HPA-axis activation during acute stress can produce a transient elevation of the plasma cortisol concentration and partial resistance to feedback inhibition of cortisol release that persists during and shortly after the duration of the stressful stimulus. This phenomenon may involve a rapid down-regulation of glucocorticoid receptors, because elevated glucocorticoid levels such as those elicited by acute stress decrease the number of hippocampal glucocorticoid receptors, with a resulting increase in corticosterone secretion and feedback resistance (204). After stress termination, as glucocorticoid levels decrease (presumably because the limbic drive on CRH release diminishes), glucocorticoid-receptor density increases, and feedback sensitivity normalizes (204).

During some types of *chronic* stress, adaptive changes in ACTH and corticosterone secretion occur such that the plasma ACTH and corticosterone concentrations achieved are lower than those seen in response to acute stress (205). In contrast, other types of chronic stress are associated with enhanced corticosterone secretion in rats (206). Moreover, Dallman and Jones showed that the experience of prior stress can result in augmented corticosterone responses to subsequent stress exposure (207). The factors that determine whether adaptation or sensitization of glucocorticoid activity occurs after chronic stress remain poorly understood.

Some stressors experienced within critical periods of neurodevelopment exert long-term effects on HPA-axis function. In rats exposed to either severe prenatal *(in utero)* stress or early maternal deprivation stress (208 ,209), the plasma concentrations of corticosterone achieved in response to subsequent stressors are increased, and this tendency to show exaggerated glucocorticoid responses to stress persists into adulthood. Early postnatal adverse experiences such as maternal separation are associated with long-lasting alterations in the basal concentrations of hypothalamic CRH mRNA, hippocampal glucocorticoid-receptor mRNA, median eminence CRH, and in the magnitude of stress-induced CRH, corticosterone, and ACTH release (210 ,211 and 212). In nonhuman primates, adverse early experiences induced by variable maternal foraging requirements reportedly result in alterations in juvenile and adult social behavior, such that animals are more timid, less socially interactive, and more subordinate (213). Adult monkeys who were raised in such a maternal environment are also hyperresponsive to yohimbine and have elevated CRH concentrations and decreased cortisol levels in the CSF, findings that parallel those in humans with PTSD (213).

Conversely, positive early-life experiences during critical developmental periods may have beneficial long-term consequences on the ability to mount adaptive responses to stress or threat. For example, daily postnatal handling of rat pups by human experimenters within the first few weeks of life has been shown to produce persistent (throughout life) increases in the density of type II glucocorticoid receptors. This increase was associated with enhanced feedback sensitivity to glucocorticoid exposure and reduced glucocorticoid-mediated hippocampal damage in later life (214 ,215). These effects are hypothesized to comprise a type of “stress inoculation” induced by the mothers' repeated licking of the pups after they were handled by humans. Taken together with the data reviewed in the preceding paragraph, these data indicate that a high degree of plasticity exists in stress-responsive neural systems during the prenatal and early postnatal periods that “programs” future biological responses to stressful stimuli (210).

Regional differences in the regulation of CRH function by glucocorticoid-receptor stimulation and stress may play major roles in the mediation of fear and anxiety (216). The feedback inhibition of CRH function by glucocorticoids (to suppress HPA-axis activity) occurs at the level of the PVN of the hypothalamus, where systemically administered glucocorticoids reduce CRH expression, and the anterior pituitary, where glucocorticoids decrease CRH *receptor* expression (217 ,218 ,219 and 220). The regulation of CRH *receptor* mRNA expression shows a regional specificity that becomes altered when stress occurs concomitantly with elevated glucocorticoid concentrations. After both short-term and long-term corticosterone (CORT) administration, the CRH *receptor* RNA expression decreases in the PVN and the anterior pituitary (219). However, after acute or repeated immobilization stress sufficient to produce a large increase in plasma CORT levels, the CRH mRNA expression decreases in the anterior pituitary, but *increases* in the PVN. In contrast, neither CORT administration nor restraint stress alters the CRH *receptor* expression in the CE of the amygdala or the BNST. Furthermore, CRH secretion is not constrained by glucocorticoids in the CE or the lateral BNST, and CRH mRNA expression *increases* in these areas during systemic CORT administration (217 ,218 ,220). It is thus conceivable that the positive feedback of glucocorticoids on extrahypothalamic CRH function in the amygdala or the BNST may contribute to the production of anxiety symptoms (216 ,221).

Another level through which the CRH-glucocorticoid system maintains homeostasis and provides mechanisms for modulating mechanism over stress or anxiety responses involves functional differences between CRH-receptor subtypes. The CRH_1 and CRH_2 receptors appear to play reciprocal roles in mediating stress responsiveness and anxiety-like behaviors (221). Mice genetically deficient in CRH_1-receptor expression exhibit diminished anxiety and

stress responses to threat or stress (222 ,223). In contrast, mice deficient in CRH_2 receptors display heightened anxiety in response to stress (224 ,225). The affinity of CRH is higher for CRH_1 than CRH_2 receptors, a finding consistent with evidence that CRH elicits anxiogenic effects either when exogenously administered to native animals or when endogenously released in mice genetically altered to overexpress CRH (221). Also consistent with the hypothesis that CRH_1-receptor stimulation facilitates anxiety responses, oral administration of the CRH_1-receptor antagonist, antalarmin, inhibits the behavioral, sympathetic autonomic, and neuroendocrine responses (i.e., attenuating increases in the CSF CRH concentration and in the pituitary-adrenal and adrenal-medullary activity) to acute social stress in monkeys (226).

Regional differences in the anatomic distribution of CRH_1 and CRH_2 receptors likely play a role in balancing facilitatory versus modulatory effects of CRH-receptor stimulation on stress responses. In monkeys, the CRH_1-receptor density is high in most amygdaloid nuclei, the cingulate cortex, the PFC, the insular cortex, the parietal cortex, the dentate gyrus, and the entorhinal cortex, and it is moderate in the CE and the LC. The CRH_2-receptor density is high in the cingulate cortex, the mPFC, the CE, the CA-1 region of the hippocampus, and the PVN and supraoptic nucleus of the hypothalamus. An important avenue of future research will involve assessments of the homeostatic balance between CRH_1- and CRH_2-receptor systems in anxiety disorders.

HPA-Axis Function and CRH Release in Anxiety Disorders

The anxiety disorder for which abnormalities of CRH or HPA-axis function has been most commonly reported is PTSD. Nevertheless, the nature of such abnormalities has been inconsistent across studies, because basal plasma or 24-hour urine cortisol concentrations have been reported to be abnormally decreased (227 ,228 and 229), not different (230 ,231), or abnormally increased (190 ,192 ,232 ,233) in PTSD samples relative to healthy or trauma-matched control samples. Differences across these studies may reflect effects of gender, age of illness onset (i.e., childhood versus adult), trauma type or duration, or physiologic variation relative to illness phase. For example, Hawk et al. showed that 24-hour urine cortisol concentrations were elevated in males but not females with PTSD, and that this abnormality in males was evident at 1 month but not 6 months after the traumatic event (191).

The HPA-axis responses to behavioral or pharmacologic challenge have also been assessed in PTSD. During provocation of PTSD symptoms by exposure to combat sounds, the *changes* in plasma cortisol and ACTH concentrations did not differ between patients with combat-related PTSD and either healthy or combat-matched, non-PTSD control subjects (232). In response to dexamethasone administration, cortisol suppression was found to be normal (234) or enhanced (228 ,235 ,236) in PTSD, with the latter result particularly found in response to low-dose (0.25 and 0.5 mg) dexamethasone. Yehuda et al. also observed that patients with PTSD have an increased density of glucocorticoid receptors on peripheral lymphocytes (228). This finding, together with the observations that patients with PTSD show hypersensitivity to low-dose dexamethasone, led Yehuda et al. to hypothesize that an increase in hypothalamic glucocorticoid-receptor function results in enhanced feedback sensitivity to cortisol, leading to decreased peripheral cortisol levels (237). Preliminary data suggest that a reduced cortisol response after trauma exposure may predict PTSD development, a finding raising the possibility that enhanced feedback sensitivity to cortisol may arise acutely or may even antedate illness onset in some patients with PTSD (229 ,238).

The central release of CRH in PTSD was examined in two studies of CSF concentrations, both of which found abnormally *increased* in chronic, combat-related PTSD (239 ,240). Potentially consistent with this observation, PTSD samples show a blunted ACTH response to CRH relative to control samples (241 ,242). Although these observations would appear most consistent with findings that basal cortisol secretion and excretion are abnormally increased in PTSD (190 ,192 ,232 ,233), they do not clearly contradict the findings of normal or reduced peripheral cortisol concentrations in PTSD because hypothalamic and extrahypothalamic CRH secretion are independently regulated (216).

Nevertheless, the studies that either identified reductions or were unable to identify elevations in peripheral cortisol concentrations in PTSD present a challenge to the hypothesis that the reduced hippocampal volume found in MRI studies of PTSD (reviewed earlier) are accounted for by cortisol hypersecretion (150). This hypothesis may still be reconciled with the peripheral cortisol measures associated with chronic PTSD if the cortisol secretion was elevated near the time of the stressor (191 ,243). Longitudinal studies in male patients who developed PTSD after motor vehicle accidents suggest that cortisol secretion is elevated 1 month after the trauma, but it is normal when measured 6 months after the trauma (191). In rats, the atrophy of pyramidal cell apical dendrites that occurs in response to stress-induced corticosterone secretion is reversible when the exposure to elevated glucocorticoid concentrations is terminated early, but it can become irreversible if the elevated corticosterone concentration persists beyond a critical time period (149). Hippocampal damage may thus conceivably occur in PTSD during a period of excessive cortisol secretion that follows the traumatic event and is prolonged enough so that hippocampal neuronal atrophy becomes irreversible. An alternative hypothesis for the reduction of hippocampal volume in PTSD, however, is that this abnormality antedates the

development of PTSD and may comprise a risk factor for developing PTSD in response to traumatic stress.

In PD, the results of studies examining CRH-receptor and HPA-axis function have been less consistent (Table 63.2). Elevated plasma cortisol levels were reported in one study (244), but not in another (245), and the results of studies assessing urinary free cortisol have been similarly inconsistent (177 ,246). In a study of 24-hour secretion of ACTH and cortisol, PD subjects had subtle elevations of nocturnal cortisol secretion and greater amplitude of ultraradian secretory episodes relative to control subjects (247), but these findings await replication. Both normal and elevated rates of cortisol nonsuppression after dexamethasone administration have been reported in PD (248). After combined dexamethasone-CRH challenge, the HPA-axis response was higher in PD subjects than in healthy controls, but the magnitude of this abnormality was less than that seen in depressed samples (249 ,250). The ACTH response to CRH was blunted in some studies (249 ,250), but not in others (250), in PD relative to control samples, although CSF levels of CRH did not differ between PD and control samples (251). The extent to which pathophysiologic heterogeneity within PD samples may account for the inconsistency of these findings remains unclear.

TABLE 63.2. EVIDENCE OF ALTERATIONS IN CRF-HPA AXIS FUNCTION IN ANXIETY DISORDERS[a]

	PTSD	Panic Disorder
Alteration in urinary cortisol	+/–[a]	+/–
Altered plasma cortisol with 24-hour sampling	+ (dec.)	+ (inc.)/–
Supersuppression with DST	++[b]	–
Blunted ACTH response to CRF	++	+/–
Elevated CRF in CSF	++	–
Increased lymphocyte glucocorticoid receptors	++	NS

[a]Findings of decreased urinary cortisol in older male combat veterans and holocaust survivors and increased cortisol in younger female abuse survivors may be explainable by differences in gender, age, trauma type, developmental epoch at the time of the trauma, or timing within illness course.
[b]Pertains specifically to response to low-dose dexamethasone (0.25 and 0.5 mg).
–, One or more studies did not support this finding (with no positive studies), or the majority of studies do not support this finding; +/–, an equal number of studies support this finding and do not support this finding; +, atleast one study supports this finding and no studies do not, or the majority of studies support the finding; ++, two or more studies support this finding, and no studies do not support the finding; +++, three or more studies support this finding, and no studies do not; ACTH, adrenocorticotropic hormone; CRF, corticotropin-releasing factor; CSF, cerebrospinal fluid; dec., decrease; DST, dihydrostreptomycin; HPA, hypothalamic pituitary adrenal axis; inc., increase; NS, not studied; PTSD, posttraumatic stress disorder.

Functional Interactions among Noradrenergic, HPA, and CRH Systems

Coordinated functional interactions between the HPA axis and the noradrenergic systems play major roles in producing adaptive responses to stress, anxiety, or fear. The secretion of CRH increases LC neuronal firing activity and results in enhanced NE release in a variety of cortical and subcortical regions (252 ,253). Conversely, NE release stimulates CRH secretion in the PVN (the nucleus containing most of the CRH-synthesizing neurons in the hypothalamus). During chronic stress in particular, the LC is the brainstem noradrenergic nucleus that appears preferentially to mediate NE release in the PVN (254). Conversely, as CRH release in the PVN stimulates ACTH secretion from the pituitary and thereby increases cortisol secretion from the adrenal glands, the rise in plasma cortisol concentrations acts through a negative feedback pathway to decrease both CRH and NE synthesis at the level of the PVN. Glucocorticoid-mediated inhibition of NE-induced CRH stimulation may be evident primarily during stress, rather than under resting conditions, as an adaptive response that restrains stress-induced neuroendocrine and cardiovascular effects mediated by the PVN (254). NE, cortisol, and CRH thus appear tightly linked as a functional system that offers a homeostatic mechanism for responding to stress.

A clinical phenomenon of anxiety disorders that may be specifically regulated by interactions between NE and glucocorticoid secretion involves the acquisition and consolidation of traumatic memories. A characteristic feature of PTSD and PD is that memories of the traumatic experience or the original panic attack, respectively, persist for decades and are recalled in response to multiple stimuli or stressors. In experimental animals, alterations of both brain catecholamine and glucocorticoid levels affect the consolidation and retrieval of emotional memories (50 ,51). Glucocorticoids influence memory storage by activation of glucocorticoid receptors in the hippocampus, whereas NE effects are mediated in part through ß-adrenoreceptor stimulation in the amygdala (255). In humans, adrenocortical suppression blocks the memory-enhancing effects of amphetamine and epinephrine (256), and propranolol impairs memory for an emotionally provocative story, but not for an emotionally “neutral” story (257). These data suggest that the acute release of glucocorticoids and NE in response to trauma may modulate the encoding of traumatic memories. It is conceivable that long-term alterations in these systems may account for memory distortions seen in PTSD, such as the memory fragmentation, hypermnesia, and deficits in declarative memory.

Central Benzodiazepine-GABA–Receptor System

Several lines of preclinical and clinical evidence have established that BZD-receptor agonists exert anxiolytic effects

and have suggested that BZD-receptor function may be altered in anxiety disorders. Central BZD receptors are expressed are present throughout the brain, but they are most densely concentrated in the cortical gray matter. The BZD and $GABA_A$ receptors form parts of the same macromolecular complex, and although they constitute distinct binding sites, they are functionally coupled and regulate each other in an allosteric manner (258). Central BZD-receptor agonists potentiate and prolong the synaptic actions of the inhibitory neurotransmitter, GABA, by increasing the frequency of GABA-mediated chloride channel openings (258 ,259). Microinjection of BZD-receptor agonists in limbic and brainstem regions such as the amygdala and the PAG exert antianxiety effects in animal models of anxiety and fear (260). Conversely, administration of BZD-receptor *inverse* agonists, such as ß-carboline-3-carboxylic acid ethylester, produces behaviors and increases in heart rate, blood pressure, plasma cortisol, and catecholamines similar to those seen in anxiety and stress (261 ,262), effects that can be blocked by administration of BZD-receptor agonists (263).

Transgenic mouse studies have identified behavioral roles for specific $GABA_A$-receptor subunits. The anxiolytic action of diazepam appears absent in mice with α_2 subunit point mutations, but it is present in mice with α_1 or α_3 subunit point mutations (264 ,265). These data suggest that the anxiolytic effect of BZD agonists is at least partly mediated by the $GABA_A$-receptor α_2 subunit, which is largely expressed in the limbic system, but not by the α_3 subunit, which is predominately expressed in the reticular activating system, or the α_1 subunit, which is implicated in mediating the sedative, amnestic, and anticonvulsive effects of BZDs (265 ,266). These findings hold clear implications for investigations of the pathophysiology of anxiety disorders and for the development of anxioselective BZD-receptor agonists.

Some other agents with anxiolytic effects appear to modulate the function of the $GABA_A$/BZD-receptor–chloride ionophore complex by mechanisms distinct from those of the BZD agonists. The neurosteroid, allopregnenolone, exerts antianxiety effects in conflict paradigms that serve as putative animal models of anxiety. The anticonflict effects of allopregnenolone are reversed by either isopropylbicyclophosphate, which binds at the picrotoxinin site on the $GABA_A$ receptors, or RO15-4513 (ethyl-8-azido-5,6-dihydro-5-methyl-6-oxo-4H-imidazo[1,5-α]-[1,4]benzodiazepine-3-carboxylate), a BZD-receptor inverse agonist that inhibits $GABA_A$-activated chloride flux in neuronal membranes. In contrast, administration of the BZD-receptor antagonist flumazenil (ethyl-8-fluoro-5,6-dihydro-5-methyl-6-oxo-4H-imidazo[1,5-α]-[1,4]benzodiazepine-3-carboxylate) does not block allopregnenolone's anxiolytic-like effects, a finding indicating that allopregnenolone does not bind at the BZD site. Allopregnenolone may thus exert anxiolytic-like effects by stimulating the chloride channel in $GABA_A$ receptors by binding at the picrotoxinin site or at a site specific for RO15-4513.

The antianxiety effects of antidepressant drugs with primary effects on monoamine reuptake may also be partly mediated through a GABAergic mechanism. These agents are effective for the treatment of a spectrum of anxiety disorders including social anxiety disorder, generalized anxiety disorder, PD, and PTSD. One of the multiple secondary effects of these agents involves potentiation of GABAergic function. For example, in rats, the effective dose of phenelzine (15 mg/kg) on the elevated plus maze administered produces a more than twofold increase in whole-brain level GABA concentrations, whereas an ineffective dose of phenelzine (5.1 mg/kg) does not significantly alter GABA levels (267). Moreover, the *N*-acetylated metabolite of phenelzine, *N*-2-acetylphenelzine, which potently inhibits monoamine oxidase but does not change whole-brain GABA concentrations, does not produce anxiolytic effects in the elevated plus-maze test (267). Phenelzine's anxiolytic effects in the plus-maze model may thus depend on elevating brain GABA concentrations, in contrast to the mechanism of the classic BZDs, which instead increase the affinity of $GABA_A$ receptors for GABA.

Effects of Stress on Benzodiazepine-$GABA_A$ Receptors

BZD- and GABA-receptor function can be altered by exposure to stress in some brain regions. In experimental animals exposed to inescapable stress in the form of cold swim or foot shock, the BZD-receptor binding decreases in the frontal cortex, with less consistent reductions occurring in the hippocampus and hypothalamus, but no changes in the occipital cortex, striatum, midbrain, thalamus, cerebellum, or pons (268). Chronic stress in the form of repeated foot shock or cold water swim resulted in decreased BZD-receptor binding in the frontal cortex and hippocampus, and possibly in the cerebellum, midbrain, and striatum, but not in the occipital cortex or pons (268 ,269 and 270). These reductions in BZD-receptor binding were associated with deficits in maze escape behaviors that may have reflected alterations in mnemonic processing (269 ,270). Some of these stress effects may be mediated by glucocorticoids, because chronic exposure to stress levels of CORT alters mRNA levels of multiple $GABA_A$-receptor subunits (271). Consistent with the effects of chronic stress on BZD-receptor expression, the *Maudsley "genetically fearful" rat strain* shows decreased BZD-receptor density relative to other rats in several brain structures including the hippocampus (272).

Stressors arising early in life may also influence the development of the GABAergic system. In rats, early-life adverse experiences such as maternal separation result in decreased $GABA_A$-receptor concentrations in the LC and the NTS, reduced BZD-receptor sites in the LC, the NTS, the frontal cortex, and the CE and the LA of the amygdala, and reduced mRNA levels for the γ2 subunit of the $GABA_A$-receptor complex in the LC, the NTS, and the amygdala (273). The extent to which these developmental responses to early-life

stress may alter the expression of fear and anxiety in adulthood remains unclear.

Benzodiazepine-GABA–Receptor Function in Anxiety Disorders

The central BZD receptor has been implicated in anxiety disorders on the basis of the anxiolytic and anxiogenic properties of BZD agonists and inverse agonists, respectively, and by the evidence that the BZD-receptor sensitivity to BZD agonists is reduced in some anxiety-disordered subjects (21 ,274 ,275). Hypotheses advanced regarding the role of $GABA_A$-BZD–receptor function in anxiety disorders have proposed either that changes in the $GABA_A$-BZD macromolecular complex conformation or that alterations in the concentration or properties of an endogenous ligand account for the pathologic anxiety symptoms seen in anxiety disorders. However, these hypotheses have not been conclusively tested by *in vivo* or postmortem studies of anxiety-disordered humans.

In PD, oral (276) and intravenous (274) administration of the BZD-receptor antagonist, flumazenil, produces panic attacks and increases anticipatory anxiety in some subjects with PD, but not in healthy controls. In addition, the sensitivity to the effects of diazepam on saccadic eye movement velocity is abnormally reduced in PD, a finding implying that the functional sensitivity of the $GABA_A$-BZD supramolecular complex is attenuated in brainstem regions controlling saccadic eye movements (275). Subjects with PD also show abnormally reduced sensitivity to the suppressant effects of diazepam on plasma NE, epinephrine, and heart rate (see Table 63.3 on p. 920) (277).

TABLE 63.3. EVIDENCE OF ALTERATION IN OTHER NEUROTRANSMITTER SYSTEMS IN ANXIETY DISORDERS

	PTSD	Panic Disorder
Benzodiazepine		
Increased symptomatology with benzodiazepine antagonist	–	++
Decreased number of benzodiazepine receptors using SPECT-iomazenil or PET-flumazenil binding	+	+++/–
Opiate		
Naloxone-reversible analgesia	+	NS
Reduced plasma β-endorphin	+	NS
Elevated levels of CSF β-endorphin	+	–
Serotonin		
Decreased serotonin reuptake site binding in platelets	++	+/–
Decreased serotonin transmitter in platelets	–	+/–
Blunted endocrine response to 5-HT_{1A} probe	–	+
Altered serotonin effect on cAMP in platelets (5-HT_{1A} probe)	–	NS
Increased anxiogenic responses to 5-HT agonists	+	+/–
Thyroid		
Increased baseline indices of thyroid function	+	–
Increased TSH response to TRH	+	–
Somatostatin		
Increased somatostatin levels at baseline in CSF	+	–
Cholecystokinin		
Increased anxiogenic responses to CCK agonists	+	+++

–, One or more studies did not support this finding (with no positive studies), or the majority of studies do not support this finding; +/–, an equal number of studies support this finding and do not support this finding; +, at least one study supports this finding and no studies do not support the finding, or the majority of studies support the finding; ++, two or more studies support this finding, and no studies do not support the finding; +++, three or more studies support this finding, and no studies do not support the finding; +++/–, three or more studies support this finding, and one study does not support the finding; cAMP, cyclic adenosine 3', 5'-monophosphate; CCK, cholecystokinin; CSF, cerebrospinal fluid; NS, not studied; PTSD, posttraumatic stress disorder; SPECT, single photon emission computed tomography; TRH, thyrotropin-releasing hormone; TSH, thyroid-stimulating hormone.

Receptor imaging studies using PET and SPECT have assessed central BZD-receptor binding in anxiety disorders. SPECT studies have reported reduced uptake of the selective BZD-receptor radioligand, [^{123}I]iomazenil, in the frontal (278 ,279 and 280), temporal (278 ,279), and occipital (278) cortices in subjects with PD relative to control subjects. However, interpretation of these results was limited by the absence of medication-free PD study subjects and of healthy controls (278 ,279) or by the dependence on nonquantitative methods for estimating BZD-receptor binding. A SPECT-iomazenil study that quantitated BZD-receptor binding by derivation of distribution volumes found reduced binding in the left hippocampus and precuneus in unmedicated PD relative to healthy control samples and reported an inverse correlation between panic anxiety ratings and frontal cortex iomazenil binding (281). Another SPECT-iomazenil study reported lower distribution volumes for BZD receptors in the dorsomedial PFC in PTSD relative to control samples (281a). These findings appeared consistent with the evidence cited earlier that stress down-regulates BZD-receptor binding in the frontal cortex and the hippocampus of experimental animals.

Central BZD-receptor binding has also been assessed in PD using PET and [^{11}C]flumazenil. Malizia et al. reported a global reduction in BZD site binding in seven study subjects with PD relative to eight healthy controls, with the most prominent decreases evident in the right orbitofrontal cortex and the right insula (areas consistently activated during normal anxiety processing) (282). In contrast, Abadie et al. found no differences in the B_{max}, K_d or bound/free values for [^{11}C]flumazenil in any brain region in ten *unmedicated* PD study subjects relative to healthy controls (283).

Dopaminergic System

Acute stress increases DA release and turnover in multiple brain areas. The dopaminergic projections to the mPFC appear particularly sensitive to stress, because brief or low-intensity stressors (e.g., exposure to fear-conditioned stimuli) increase DA release and turnover in the mPFC in the absence of corresponding changes in other mesotelencephalic dopaminergic projections (284). For example, in rats, low-intensity electric foot shock increases tyrosine hydroxylase activity and DA turnover in the mPFC, but not in the nucleus accumbens or the caudate-putamen (285). In contrast, stress of greater intensity or longer duration additionally enhances DA release and metabolism in other areas as well (285). The regional sensitivity to stress appears to follow a pattern in which dopaminergic projections to the mPFC are more sensitive to stress than the mesoaccumbens and nigrostriatal projections, and the mesoaccumbens dopaminergic projections are more sensitive to stress than the nigrostriatal projections (284).

Thus far, there is little evidence that dopaminergic dysfunction plays a primary role in the pathophysiology of human anxiety disorders. In PD, Roy-Byrne et al. found a higher plasma concentration of the DA metabolite, homovanillic acid (HVA), in patients with high levels of anxiety and frequent panic attacks relative to controls (286). Patients with PD were also shown to have a greater growth hormone response to the DA-receptor agonist, apomorphine, than depressed controls (287). However, Eriksson et al. found no evidence of alterations in the CSF HVA concentrations in patients with PD or for correlations between CSF HVA and anxiety severity or panic attack frequency (288). In addition, genetic studies examining associations between PD and gene polymorphisms for the DA D4 receptor and the DA transporter have produced negative results (289).

In social phobia, two preliminary SPECT imaging studies involving small subject samples reported abnormal reductions in DA-receptor binding. Tiihonen et al. found a significant reduction in ß-CIT binding in the striatum in social phobic relative to healthy control samples (290), presumably reflecting a reduction in DA-transporter binding. Schneier et al. reported reduced uptake of the DA D2/D3-receptor radioligand, [^{123}I]IBZM, in social phobic subjects relative to healthy control subjects (291). Both findings await replication.

Serotonergic System

Exposure to various stressors including restraint stress, tail shock, tail pinch, and high-level (but not low-level) foot shock results in increased 5-HT turnover in the mPFC, nucleus accumbens, amygdala, and lateral hypothalamus in experimental animals (285). During exposure to fear-conditioned stimuli, the 5-HT turnover in the mPFC appears particularly sensitive to the severity of stress, increasing as the aversiveness of the US and the magnitude of the conditioned fear behavioral response increases (285). However, exposure to repeated electric shocks sufficient to produce learned helplessness is associated with reduced *in vivo* release of 5-HT in the frontal cortex (292), a finding possibly reflecting a state in which 5-HT synthesis is outpaced by release. Preadministration of BZD-receptor agonists or tricyclic antidepressant drugs prevents stress-induced reductions in 5-HT release and interferes with the acquisition of learned helplessness, whereas infusion of 5-HT into the frontal cortex after stress exposure *reverses* learned-helplessness behavior (292 ,293). Finally, administration of 5-HT–receptor antagonists produces behavioral deficits resembling those of the learned helplessness seen after inescapable shock during animal stress models that do not ordinarily result in learned helplessness (293).

The effect of stress in activating 5-HT turnover may stimulate both anxiogenic and anxiolytic pathways within the forebrain, depending on the region involved and the 5-HT–receptor subtype that is predominantly stimulated. For example, microinjection of 5-HT into the amygdala appears to enhance conditioned fear, whereas 5-HT injection into the PAG inhibits unconditioned fear (260). Graeff et al. hypothesized that the serotonergic innervation of the amygdala and the hippocampus mediates anxiogenic effects by 5-HT_{2A}–receptor stimulation (260), whereas serotonergic innervation of hippocampal 5-HT_{1A} receptors suppresses formation of new CS-US associations and provides resilience to aversive events. Potentially compatible with this hypothesis, 5-HT_{1A}-receptor knockout mice exhibit behaviors consistent with increased anxiety and fear, and long-term administration of 5-HT_{1A}–receptor partial agonists exerts anxiolytic effects in generalized anxiety disorder (295).

Notably, stress and glucocorticoids exert major effects on the genetic expression of 5-HT_{1A} and 5-HT_{2A} receptors. Postsynaptic 5-HT_{1A}–receptor gene expression is under tonic inhibition by adrenal steroids in the hippocampus and possibly other regions where mineralocorticoid receptors are expressed (reviewed in ref. 296). Thus, 5-HT_{1A}–receptor density and mRNA levels decrease in response to chronic stress or CORT administration and increase after adrenalectomy (296 ,297 ,298 and 299). The stress-induced down-regulation of 5-HT_{1A}–receptor expression is prevented by adrenalectomy, a finding showing the importance of circulating adrenal steroids in mediating this effect (296). Although both mineralocorticoid-receptor stimulation and glucocorticoid-receptor stimulation are involved in mediating this effect, the former is most potent, and 5-HT_{1A} mRNA levels markedly decrease within hours of mineralocorticoid-receptor stimulation (296). Conversely, 5-HT_{2A}–receptor expression is up-regulated during chronic stress and CORT administration, and it is down-regulated in response to adrenalectomy (298 ,300). In view of evidence that 5-HT_{1A} and 5-HT_{2A} receptors may play reciprocal roles in mediating anxiety, it is conceivable that these corticosteroid mediated effects on 5-HT_{1A} and 5-HT_{2A} expression may be relevant to the pathophysiology of anxiety.

Serotonergic Function in Anxiety Disorders

The literature regarding serotonergic function in anxiety disorders is in disagreement (see Table 63.3). In PD, platelet 5-HT uptake has been reported to be abnormally elevated (301), normal (302), or abnormally reduced (303). Platelet imipramine binding (to a site related to the 5-HT transporter site), did not differ in PD relative to control samples (304 ,305). Another study reported reduced concentrations of circulating 5-HT in PD relative to control samples (306), although this finding has not been replicated.

Pharmacologic challenge studies involving 5-HT have been similarly unable to establish a primary role for 5-HT in the pathophysiology in PD. Neuroendocrine responses to challenge with the 5-HT precursors, L-tryptophan and 5-hydroxytryptophan (5-HTP), did not differentiate PD study subjects from healthy controls (307 ,308). Moreover, tryptophan depletion did not prove anxiogenic in unmedicated PD study subjects (309). Nevertheless, challenge with the 5-HT releasing agent, fenfluramine, produced greater increases in anxiety, plasma prolactin, and cortisol in PD compared with control subjects (131 ,310). Fenfluramine challenge also resulted in reduced CBF in the left posterior parietal-superior temporal cortex in PD study subjects relative to healthy controls (131), although it was unclear whether this abnormality reflected an abnormality of serotonergic function or a physiologic correlate of fenfluramine-induced anxiety, because more PD study subjects (56%) developed panic attacks than did control subjects (11%).

Preliminary data regarding the sensitivity of specific 5-HT–receptor subtypes appear more promising, particularly because the elevation of plasma ACTH and cortisol and the hypothermic responses to the 5-HT_{1A} partial agonist, ipsapirone, were blunted in PD relative to healthy control samples (311). Finally, increases in anxiety and plasma cortisol in PD relative to control samples have been reported after oral (312), but not intravenous, administration of the 5-HT_2–receptor agonist, m-chloromethylpiperazine (mCPP) (313).

Samples with combat-related PTSD have been shown to have decreased paroxetine binding in platelets relative to controls, a finding suggesting alterations in the 5-HT transporter (314).

Southwick et al. observed that a subgroup (five of 14 subjects) with PTSD experienced panic anxiety and “flashbacks” after mCPP challenge (189). Thus, a subgroup of patients with PTSD may have abnormal sensitivity to serotonergic provocation.

Cholecystokinin

CCK is an anxiogenic neuropeptide present in both the brain and the gastrointestinal tract. CCK-containing neurons are found in high density in the cerebral cortex, amygdala, hippocampus, midbrain PAG, substantia nigra, and raphe. Iontophoretic administration of CCK has depolarizing effects on pyramidal neurons and stimulates action potential formation in the dentate gyrus of the hippocampus (reviewed in ref. 315).

The CCK-receptor agonist, CCK-4, is anxiogenic in a variety of animal models of anxiety, whereas CCK-receptor antagonists exert anxiolytic effects in the same models (315). CCK has important functional interactions with other systems implicated in anxiety and fear (noradrenergic, dopaminergic, BZD). For example, the panicogenic effect of CCK-4 in PD is attenuated by administration of the ß-adrenoreceptor antagonist, propranolol, and by long-term imipramine treatment, which down-regulates ß-adrenoreceptors (316).

Study subjects with PD or PTSD are more sensitive to the anxiogenic effects of CCK-4 than are control subjects (317 ,318). For example, Strohle et al. found that of 24 PD study subjects tested, 15 experienced a panic attack after CCK-4 administration (319). Although the mechanism underlying the enhanced sensitivity to CCK-4 has not been elucidated, it is noteworthy that CSF concentrations of CCK are lower in PD study subjects than in healthy controls (320).

The neuroendocrine effects associated with CCK-4 induced panic appear to differ between PD and PTSD. In PTSD, CCK-4–induced panic was associated with a *lower* ACTH response in the PTSD study subjects than in healthy controls, and cortisol concentrations increased in both the PTSD and control groups (318). The elevation in the cortisol concentrations attenuated more rapidly in the PTSD group than in the control group.

In contrast to the findings in PTSD, ACTH secretion was higher in subjects with PD who developed panic attacks in response to CCK-4 than in those who did not, although even the latter subjects showed brief, less pronounced increases in ACTH concentrations (319). Neither PD subgroup showed significant changes in the plasma cortisol concentration after CCK-4 administration. The elevation of ACTH concentrations suggested that CRH secretion increases in CCK-4–induced panic in PD (consistent with preclinical evidence regarding the role of CRH in stress and anxiety and the interaction of CRH and CCK in modulating anxiety) (221).

The CCK receptors are classified into CCK-A and CCK-B subtypes. Kennedy et al. reported a significant association between PD and a single nucleotide polymorphism found in the coding region of the CCK-B–receptor gene (321). In contrast, genetic polymorphisms for the CCK-A–receptor gene and the CCK-pre-pro hormone genes showed no association with PD (321). If confirmed by replication, these data would suggest that a CCK-B–receptor gene variation may be involved in the pathogenesis of PD.

Pande et al. assessed the efficacy of the selective CCK-B–receptor antagonist, CI-988, for preventing panic attacks in PD (322). No differences in the rate of panic attacks were seen between the active drug and placebo treatment groups. Nevertheless, because of the limited bioavailability of oral CI-988, studies involving this drug may not have sufficiently tested the hypothesis that CCK-B–receptor antagonism produces antipanic effects in PD.

Other Neuropeptides

Opioid Peptides

Acute, uncontrollable shock increases secretion of opiate peptides and decreases µ-opiate–receptor density (323 ,324). The elevation of opioid peptide secretion may contribute to the analgesia observed after uncontrollable stress and exposure to fear-conditioned stimuli (325). This analgesic effect shows evidence of sensitization, because subsequent exposure to less intense shock in rats previously exposed to uncontrollable shock also results in analgesia (326).

Potentially consistent with these data, Pitman et al. found that patients with PTSD showed reduced pain sensitivity compared with veterans without PTSD after exposure to a combat film (327), an effect that was reversed by the opiate antagonist naloxone (a finding suggesting mediation by endogenous opiate release during symptom provocation). In the baseline state, the CSF ß-endorphin levels were abnormally elevated in PTSD relative to control samples (328). However, Hoffman et al. found *lower* morning and evening *plasma* ß-endorphin levels in a PTSD group versus healthy control samples (329). Another study found no differences in plasma methionine-enkephalin concentrations between PTSD subjects and control subjects, although this compound's degradation half-life was higher in the PTSD group (330).

During opiate administration, Bremner et al. reported that some patients with combat-related PTSD experience an attenuation of their hyperarousal symptoms (331). Because preclinical studies in experimental animals have shown that opiates potently suppress central and peripheral noradrenergic activity, these data appear compatible with the hypothesis that some PTSD symptoms are mediated by noradrenergic hyperactivity (discussed earlier). Conversely, during opiate withdrawal noradrenergic activity increases, and it

has been noted that some symptoms of PTSD resemble those of opiate withdrawal (170).

Neuropeptide Y

NPY administered in low doses intraventricularly attenuates experimentally induced anxiety in a variety of animal models (332). Consistent with these data, transgenic rats that overexpress hippocampal NPY show behavioral insensitivity to restraint stress and absent fear suppression of behavior in a punished drinking task (333). In healthy humans subjected to uncontrollable stress during military training exercises, plasma NPY levels increased to a greater extent in persons rated as having greater stress resilience (334). During stress exposure, the NPY plasma levels were positively correlated with plasma cortisol concentrations and behavioral performance, and they were negatively correlated with dissociative symptoms (334).

In humans with PD, plasma NPY concentrations were abnormally elevated, and this finding, given NPY's putative anxiolytic effects, may reflect an adaptive response to anxiety symptoms (335). In contrast, patients with combat-related PTSD had *lower* plasma NPY concentrations both at baseline and in response to yohimbine challenge than healthy controls (336). In the PTSD group, the baseline NPY levels were inversely correlated with PTSD and panic symptoms and with yohimbine-induced increases in MHPG and systolic blood pressure (336). If this finding proves reproducible, it suggests that a deficit in endogenous NPY secretion may be involved in the generation of anxiety and sympathetic autonomic symptoms in PTSD.

Thyrotropin-Releasing Hormone and the Thyroid Axis

In the early twentieth century, Graves described cases in which thyroid hormone hypersecretion was associated with anxiety, palpitations, breathing difficulties, and rapid heart rate in persons recently exposed to traumatic stress. Nevertheless, systematic epidemiologic studies of the relationship between stress and thyroid disease have not been conducted. Although few studies have looked at thyroid function in anxiety disorders, Mason et al. found elevated levels of triiodothyronine in patients with combat-related PTSD (337) (Table 63.3), a finding consistent with evidence that stress results in long-lasting elevations of thyroid hormone secretion (338).

Respiratory System Dysfunction in Panic Disorder

Associations between respiratory perturbation and acute anxiety have been demonstrated in PD, in which various forms of respiratory stimulation consistently produce panic anxiety and alterations in parameters of respiratory physiology (339 ,340 ,341 and 342). The most straightforward forms of respiratory stimulation that produce panic anxiety produce elevations of carbon dioxide pressure (hypercapnia). Thus, panic attacks can be consistently induced in patients with PD by rebreathing air, inhaling 5% to 7% carbon dioxide in air (343 ,344), or inhaling a single deep breath of 35% carbon dioxide (345 ,346). Other panicogenic chemical challenges have also been hypothesized to induce anxiogenic effects

through respiratory stimulation (340 ,341 ,347). Although the panicogenic mechanism of intravenous administration of sodium lactate remains unclear, it may also involve respiratory stimulation (339 ,340).

The evidence that respiratory parameters index risk for panic anxiety includes data showing the following: (a) asymptomatic adult relatives of patients with PD have abnormally increased sensitivity to respiratory stimulation by carbon dioxide inhalation; (b) among PD samples, stronger family loading for PD is found among persons with evidence of respiratory dysregulation; and (c) the respiratory indices associated with PD are heritable, a finding suggesting a shared genetic vulnerability for panic attacks and respiratory dysregulation (reviewed in Chapter 61). Nevertheless, these data partly depend on subjective ratings of dyspnea during stress or respiratory stimulation, and the mechanisms underlying this sensitivity remain unclear. One possibility is that this hypersensitivity reflects an overall sensitivity to somatic sensations, because high degrees of anxiety sensitivity are linked to future panic attacks (348).

The associations between respiratory perturbation and acute anxiety are not specific to PD. Exaggerated sensitivity to respiratory perturbation has also been reported in anxiety-disordered patients with some simple phobias, limited symptom panic attacks, childhood separation anxiety disorder, or limited-symptom anxiety attacks and in nonpsychiatrically ill subjects with high ratings on anxiety sensitivity scales. (See Chapter 61) For example, children with separation anxiety disorder exhibit greater changes in somatic symptoms during carbon dioxide inhalation that positively correlate with increases in respiratory rate, tidal volume, minute ventilation, end-tidal carbon dioxide pressure, and irregularity in respiratory rate during room-air breathing (349).

CONCLUDING REMARKS

Part of "63 - Neurobiological Basis of Anxiety Disorders "

The inconsistency in the results of biological investigations of anxiety disorders highlights the importance of addressing the neurobiological heterogeneity inherent within criteria-based, psychiatric diagnoses. Understanding this heterogeneity will be facilitated by the continued development and application of genetic, neuroimaging, and neurochemical approaches that can refine anxiety disorder phenotypes and can elucidate the genotypes associated with these disorders. Application of these experimental approaches will also facilitate research aimed at elucidating the mechanisms of antianxiety therapies.

The knowledge reviewed herein regarding the neurobiology of fear and anxiety already suggests themes along which the development of new therapeutic approaches can be organized. In general, anxiolytic treatments appear to inhibit neuronal activity in the structures mediating fear expression and behavioral sensitization and facilitate endogenous mechanisms for modulating the neural transmission of information about aversive stimuli and responses to such stimuli. Novel treatments being developed to exploit the former type of mechanisms include pharmacologic agents that selectively target subcortical and brainstem pathways supporting specific components of emotional expression (e.g., CRH-receptor antagonists). In contrast, nonpharmacologic treatments for anxiety may augment the brain's systems for modulating anxiety responses, by facilitating the extinction of putative fear-conditioned responses or directing the reinterpretation of anxiety-related thoughts and somatic sensations (so they produce less subjective distress). Informed by increasingly detailed knowledge about the pathophysiology of specific anxiety disorders and the neural pathways involved in anxiety and fear processing, the development of therapeutic strategies that combine both types of approaches may ultimately provide the optimal means for reducing the morbidity of anxiety disorders.

REFERENCES

1. LeDoux JE. Emotion. In: Mills J, Mountcastle VB, Plum F, et al., eds. *Handbook of physiology: the nervous system V*. Baltimore: Williams & Wilkins, 1987:373–417.

2. LeDoux, Joseph E. Emotion: clues from the brain. *Annu Rev Psychol* 1995;46:209–235.

3. Izard CE. Four systems for emotion activation: cognitive and noncognitive processes. *Psychol Rev* 1993;100:68–90.

4. LeDoux J. Fear and the brain: where have we been, and where are we going? *Biol Psychiatry* 1998;44:1229–1238.

5. Davis M. Are different parts of the extended amygdala involved in fear versus anxiety? *Biol Psychiatry* 1998;44:1239–1247.

6. Rogan MT, Staubli UV, LeDoux JE. Fear conditioning induces associative long-term potentiation in the amygdala. *Nature* 1997;390:604–607.

7. Price JL, Carmichael ST, Drevets WC. Networks related to the orbital and medial prefrontal cortex: a substrate for emotional behavior? *Prog Brain Res* 1996;107:523–536.

8. Baxter MG, Parker A, Lindner CCC, et al. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J Neurosci* 2000;20:4311–4319.

9. Everitt BJ, Cador M, Robbins TW. Interactions between the amygdala and ventral striatum in stimulus-reward associations: studies using a second-order schedule of sexual reinforcement. *Neuroscience* 1989;30:63–75.

10. Nishijo H, Ono T, Nishino H. Single neuron responses in amygdala of alert monkey during complex sensory stimulation with affective significance. *J Neurosci* 1988;8:3570–3583.

11. Rolls ET. A theory of emotion and consciousness, and its application to understanding the neural basis of emotion. In: Gazzaniga MS, ed. *The cognitive neurosciences*, Cambridge, MA: MIT Press, 1995:1091–1106.

12. Doron NN, LeDoux JE. Organization of projections to the lateral amygdala from auditory and visual areas of the thalamus in the rat. *J Comp Neurol* 1999;412:383–409.

13. Neafsey EJ, Terreberry RR, Hurley KM, et al. Anterior cingulate cortex in rodents: connections, visceral control functions, and implications for emotion. In: Vogt BA, Gabriel M, eds. *Neurobiology of cingulate cortex and limbic thalamus*. Boston: Birkhauser, 1993:206–223.

14. Armony JL, Quirk GJ, LeDoux JE. Differential effects of amygdala lesions on early and late plastic components of auditory cortex spike trains during fear conditioning. *J Neurosci* 1998;18:2592–2601.

15. Ferry B, Roozendaal B, McGaugh JL. Role of norepinephrine in mediating stress hormone regulation of long-term memory storage: a critical involvement of the amygdala. *Biol Psychiatry* 1999;46:1140–1152.

16. Cahill L, McGaugh JL. Mechanisms of emotional arousal and lasting declarative memory. *Trends Neurosci* 1998;21:294–299.

17. Morgan MA, LeDoux JE. Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Behav Neurosci* 1995;109:681–688.

18. Quirk GJ, Russo GK, Barron JL, et al. The role of the ventromedial prefrontal cortex in the recovery of extinguished fear. *J Neurosci* 2000;20:6225–6231.

19. Romanski LM, Clugnet MC, Bordi F., et al. Somatosensory and auditory convergence in the lateral nucleus of the amygdala. *Behav Neurosci* 1993;107:444–450.

20. Carmichael ST, Clugnet M-C, Price JL. Central olfactory connections in the macaque monkey. *J Comp Neurol* 1994;346:403–434.

21. Gray JA, McNaughton N. The neuropsychology of anxiety: reprise. *Nebr Symp Motiv* 1996;43:61–134.

22. Phillips RG, LeDoux JE. Lesions of the fornix but not the entorhinal or perirhinal cortex interfere with contextual fear conditioning. *J Neurosci* 1995;15:5308–5315.

23. Kim JJ, Fanselow MD. Modality-specific retrograde amnesia of fear. *Science* 1992;256:675–677.

24. Maren S, Aharonov G, Fanselow MS. Neurotoxic lesions of the dorsal hippocampus and pavlovian fear conditioning in rats. *Behav Brain Res* 1997;88:261–274.

25. Phillips RG, LeDoux JE. Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behav Neurosci* 1992;106:274–285.

26. Morgan MA, LeDoux JE. Contribution of ventrolateral prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Neurobiol Learn Mem* 1999;72:244–251.

27. Corodimas KP, LeDoux JE. Disruptive effects of posttraining perirhinal cortex lesions on conditioned fear: contributions of contextual cue. *Behav Neurosci* 1995;109:613–619.

28. Caldarone B, Saavedra C, Tartaglia K, et al. Quantitative trait loci analysis affecting contextual conditioning in mice. *Nat Genet* 1997;17:335–337.

29. Wehner JM, Radcliffe RA, Rosmann ST, et al. Quantitative trait locus analysis of contextual fear conditioning in mice [see Comments]. *Nat Genet* 1997;17:331–334.

30. Flint J, Corley R, DeFries JC, et al. A simply genetic basis for a complex psychological trait in laboratory mice. *Science* 1995;269:1432–1435.

31. LeDoux JE, Romanski L, Xagoratis A. Indelibility of subcortical emotional memories. *J Cogn Neurosci* 1989;1:238–243.

32. Shi C, Davis M. Pain pathways involved in fear conditioning measured with fear-potentiated startle: lesion studies. *J Neurosci* 1999;19:420–430.

33. Campeau S, Davis M. Involvement of subcortical and cortical afferents to the lateral nucleus of the amygdala in fear conditioning measured with fear-potentiated startle in rats trained concurrently with auditory and visual conditioned stimuli. *J Neurosci* 1995;15:2312–2327.

34. LeDoux JE, Cicchetti P, Xagoraris A, et al. The lateral amygdaloid nucleus: sensory interface of the amygdala in fear conditioning. *J Neurosci* 1990;10:1062–1069.

35. Romanski LM, LeDoux JE. Equipotentiality of thalamo-amygdala and thalamo-cortico-amygdala circuits in auditory fear conditioning. *J Neurosci* 1992;12:4501–4509.

36. Pitkanen A, Savander V, LeDoux JE. Organization of intra-amygdaloid circuitries in the rat: an emerging framework for understanding functions of the amygdala. *Trends Neurosci* 1997;20:517–523.

37. Pitkanen A, Stefanacci L, Farb CR, et al. Intrinsic connections of the rat amygdaloid complex: projections originating in the lateral nucleus. *J Comp Neurol* 1995;356:288–310.

38. Savander V, Miettinen R, LeDoux JE, et al. Lateral nucleus of the rat amygdala is reciprocally connected with basal and accessory basal nuclei: a light and electron microscopy study. *Neuroscience* 1997;77:767–781.

39. Maren S. Long-term potentiation in the amygdala: a mechanism for emotional learning and memory. *Trends Neurosci* 1999;22:561–567.

40. Welinsky AE, Scafe GE, LeDoux JE. Functional inactivation of the amygdala before but not after auditory fear conditioning prevents memory consolidation. *J Neurosci* 1999;19:RC48.

41. Savander V, Go CG, LeDoux JE, et al. Intrinsic connections of the rat amygdaloid complex: projections originating in the basal nucleus. *J Comp Neurol* 1995;361:345–368.

42. Garcia R, Vouimba R-M, Baudry M, et al. The amygdala modulates prefrontal cortex activity relative to conditioned fear. *Nature* 1999;402:294–296.

43. Öngür D, Price JL. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* 2000;10:206–219.

44. Murray EA. Memory for objects in nonhuman primates. In: Gazzaniga MS, ed. *The new cognitive neurosciences*. Cambridge, MA: MIT Press, 2000:753–763.

45. Squire LR. Knowlton BJ. Learning about categories in the absence of memory. *Proc Natl Acad Sci USA* 1995;92:12470–12474.

46. Quirk GJ, Armony JL, LeDoux JE. Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. *Neuron* 1997;19:481–484.

47. Büchel C, Morris J, Dolan RJ, et al. Brain systems mediating aversive conditioning: an event related fMRI study. *Neuron* 1998;20:947–957.

48. LaBar KS, Gatenby JC, Gore JC, et al. Human amygdala activation during conditioned fear acquisition and extinction: a mixed trial fMRI study. *Neuron* 1998;20:937–945.

49. Drevets WC, Raichle ME. Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: implications for interactions between emotion and cognition. *Cogn Emot* 1998;12:353–385.

50. McGaugh JL. Involvement of hormonal and neuromodulatory systems in the regulation of memory storage. *Annu Rev Neurosci* 1989;2:255–287.

51. McGaugh JL. Significance and remembrance: the role of neuromodulatory systems. *Psychol Sci* 1991;1:15–45.

52. Cahill L, Haier RJ, Fallon J, et al. Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proc Nat Acad Sci USA* 1996;93:8016–8021.

53. Canli T, Zhao Z, Brewer J, et al. Event-related activation in the human amygdala associates with later memory for individual emotional experience. *J Neurosci* 2000;20:RC99.

54. Hamann SB, Ely TD, Grafton ST, et al. Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nat Neurosci* 1999;2:289–293.

55. Phelps EA, Anderson AK. Emotional memory: what does the amygdala do? *Curr Biol* 1997;7:R311–314.

56. Halgren E. The amygdala contribution to emotion and memory: current studies in humans. In: Ben-Ari Y, ed. *The amygdaloid complex*. Amsterdam: Elsevier/North Holland Biomedical Press, 1981:395–408.

57. Gloor P, Olivier A, Quesney LF, et al. The role of the limbic system in experiential phenomena of temporal lobe epilepsy. *Ann Neurol* 1982;12:129–144.

58. Brothers L. Neurphysiology of the perception of intentions by primates. In: Gazzaniga MS, ed. *The cognitive neurosciences*. Cambridge, MA: MIT Press, 1995:1107–1116.

59. Herman JP, Cullinan WE. Neurocircuitry of stress: central control of the hypothalamo-pituitary-adrenocortical axis. *Trends Neurosci* 1997;20:78–84.

60. Iwata J, Chida K, LeDoux JE. Cardiovascular responses elicited by stimulation of neurons in the central amygdaloid nucleus in awake but not anesthetized rats resemble conditioned emotional responses. *Brain Res* 1987;418:183–188.

61. Kapp BS, Gallagher M, Underwood MD, et al. Cardiovascular responses elicited by electrical stimulation of the amygdala central nucleus in the rabbit. *Brain Res* 1982;234:251–262.

62. Hitchcock JM, Davis M. Efferent pathway of the amygdala involved in conditioned fear as measured with the fear-potentiated startle paradigm. *Behav Neurosci* 1991;105:826–842.

63. Kim M, Davis M. Lack of a temporal gradient of retrograde amnesia in rats with amygdala lesions assessed with the fear-potentiated startle paradigm. *Behav Neurosci* 1993;107:1088–1092.

64. LeDoux JE, Iwata J, Cicchetti P, et al. Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear *J Neurosci* 1988;8:2517–2529.

65. Russchen FT, Bakst I, Amaral DG, et al. The amygdalostriatal projections in the monkey-an anterograde tracing study. *Brain Res* 1985;329:241–257.

66. Mogenson GJ, Brudzynski SM, Wu M, et al. From motivation to action: a review of dopaminergic regulation of limbic nucleus accumbens ventral pallidum pedunculopontine nucleus circuitries involved in limbic-motor integration. In: Kalivas PW, Barnes CD, eds. *Limbic motor circuits and neuropsychiatry*. London: CRC Press, 1993:193–236.

67. Bandler R, Shipley MT. Columnar organization in the midbrain periaqueductal grey: modules for emotional expression? *Trends Neurosci* 1994;17:379–389.

68. Weiskrantz L. Behavioral changes associated with ablation of the amygdaloid complex in monkeys. *JCPP* 1956;49:381–391.

69. Meunier M, Bachevalier J, Murray EA, et al. Effect of aspiration versus neurotoxic lesions of the amygdala on emotional responses in monkeys. *Eur J Neurosci* 1999;11:4403–4418.

70. Morris JS, Frith CD, Perrett DI, et al. A differential neural response in the human amygdala to fearful and happy facial expression. *Nature* 1996;383:812–815.

71. Blair RJR, Morris JS, Frith CD, et al. Neural responses to sad and angry expressions. *Brain* 1999;122:883–893.

72. Adolphs R, Tranel D, Damascio H, et al. Fear and the human amygdala. *J Neurosci* 1995;15:5879–5891.

73. Scott SK, Young AW, Calder AJ, et al. Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature* 1997;385:254–257.

74. Krettek JE, Price JL. Projections from the amygdaloid complex and adjacent olfactory structures to the entorhinal cortex and to the subiculum in the rat and cat. *J Comp Neurol* 1977;172:723–752.

75. Rosen, JB, Hitchcock JM, Miserendino MJ, et al. Lesions of the perirhinal cortex but not of the frontal, medial prefrontal, visual, or insular cortex block fear-potentiated startle using a visual conditioned stimulus. *J Neurosci* 1992;12:4624–4633.

76. Sananes CB, Davis M. *N*-methyl-D-aspartate lesions of the lateral and basolateral nuclei of the amygdaloid block fear-potentiated startle and shock sensitization of startle. *Behav Neurosci* 1992;106:72–80.

77. Tischler MD, Davis M. A visual pathway that mediates fear-conditioned enhancement of acoustic startle. *Brain Res* 1983;276:55–71.

78. Rauch SL, Savage CR, Alpert NM, et al. A positron emission tomographic study of simple phobic symptom provocation. *Arch Gen Psychiatry* 1995;52:20–28.

79. Rauch SI, van der Kolk BA, Fisler RF, et al. A symptom provocation study of posttraumatic stress disorder using positron emission tomography and script driven imagery. *Arch Gen Psychiatry* 1996;53:380–387.

80. Lane RD, Reiman EM, Ahern GL, et al. Neuroanatomical correlates of happiness, sadness, and disgust. *Am J Psychiatry* 1997;154:926–933.

81. Reiman EM, Lane RD, Ahern GL, et al. Neuroanatomical correlates of externally and internally generated human emotion. *Am J Psychiatry* 1997;154:918–925.

82. Sawchenko PE, Swanson LWW. Central noradrenergic pathways for the integration of hypothalamic neuroendocrine and autonomic responses. *Science* 1983;214:685–687.

83. Frysztak RJ, Neafsey EJ. The effect of medial frontal cortex lesions on cardiovascular conditioned emotional responses in the rat. *Brain Res* 1994;643:181–193.

84. Neafsey EJ, Hurley-Gius KM, Arvanitis D. The topographical organization of neurons in the rat medial frontal, insular and olfactory cortex projecting to the solitary nucleus, olfactory bulb, periaqueductal gray and superior colliculus. *Brain Res* 1986;377:261–270.

85. Damasio AR, Grabowski TJ, Bechara A, et al. Neural correlates of the experience of emotions. *Soc Neurosci Abstr* 1998;24:258.

86. Rogers RD, Everitt BJ, Baldacchino A, et al. Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-deleted normal volunteers: evidence for monoaminergic mechanisms. *Neuropsychopharmacology* 1999;20:322–339.

87. Dioro D, Viau V, Meaney MJ. The role of the medial prefrontal cortex (cingulate gyrus) in the regulation of hypothalamic-pituitary-adrenal responses to stress. *J Neurosci* 1993;3:3839–3847.

88. Schultz W. Dopamine neurons and their role in reward mechanisms. *Curr Opin Neurobiol* 1997;7:191–197.

89. Usher M, Cohen JD, Servan-Schreiber D, et al. The role of locus coeruleus in the regulation of cognitive performance. *Science* 1999;283:549–554.

90. Robertson IH, Mattingley JB, Rorden C, et al. Phasic alerting of neglect patients overcomes their spatial deficit in visual awareness. *Nature* 1998;395:169–172.

91. Drevets WC. Neuroimaging studies of mood disorders. *Biol Psychiatry* 2000;48:813–829.

92. Teasdale JD, Howard RJ, Cox SG, et al. Functional MRI study of the cognitive generation of affect. *Am J Psychiatry* 1999;156:79–88.

93. Northoff G, Richter A, Gessner, et al. Functional dissociation between medial and lateral prefrontal cortical spatiotemporal activation in negative and positive emotions: a combined fMRI/MEG study. *Cereb Cortex* 2000;10:93–107.

94. Drevets WC, Videen TO, Price JL, et al. A functional anatomical study of unipolar depression. *J Neurosci* 1992;12:3628–3641.

95. Mayberg HS, Brannan SK, Mahurin RK, et al. Cingulate function in depression: a potential predictor of treatment response. *Neuroreport* 1997;8:1057–1061.

96. Drevets WC, Price JL, Simpson JR, et al. Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* 1997;386:824–827.

97. Shin LM, Kosslyn SM, McNally RJ, et al. Visual imagery and perception in posttraumatic stress disorder: a positron emission tomographic investigation. *Arch Gen Psychiatry* 1997;54:233–237.

98. Hirayasu Y, Shenton ME, Salisbury DF, et al. Subgenual cingulate cortex volume in first-episode psychosis. *Am J Psychiatry* 1999;156:1091–1093.

99. Öngür D, Drevets WC, Price JL. Glial reduction in the subgenual prefrontal cortex in mood disorders. *Proc Natl Acad Sci USA* 1998;95:13290–13295.

100. George MS, Ketter TA, Parekh PI, et al. Brain activity during transient sadness and happiness in healthy women. *Am J Psychiatry* 1995;152:341–351.

101. Mayberg HS, Liotti M, Brannan SK, et al. Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *Am J Psychiatry* 1999;156:675–682.

102. Carmichael ST, Price JL. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys, *J Comp Neurol* 1995;363:615–641.

103. Leichnetz GR, Astruc J. The efferent projections of the medial prefrontal cortex in the squirrel monkey *(Saimiri sciureus). Brain Res* 1976;109:455–472.

104. Damasio AR, Tranel D, Damasio H. Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behav Brain Res* 1990;41:81–94.

105. Sullivan RM, Gratton A. Lateralized effects of medial prefrontal cortex lesions on neuroendocrine and autonomic stress responses in rats. *J Neurosci* 1999;19:2834–2840.

106. Dolan RJ, Fletcher P, Morris J, et al. Neural activation during covert processing of positive emotional facial expressions. *Neuroimage* 1996;4:194–200.

107. Drevets WC, Videen TO, Snyder AZ, et al. Regional cerebral blood flow changes during anticipatory anxiety. *Soc Neurosci Abstr* 1994;20:368.

108. Price JL. Networks within the orbital and medial prefrontal cortex. *Neurocase* 1999;5:231–241.

109. Baxter LR, Schwartz JM, Phelps ME, et al. Reduction of prefrontal cortex glucose metabolism common to three types of depression. *Arch Gen Psychiatry* 1989;46:243–250.

110. Rajkowska G, Miguel-Hidalgo JJ, Wei J, et al. Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. *Biol Psychiatry* 1999;45:1085–1098.

111. Drevets WC, Botteron, K. Neuroimaging in psychiatry. In: Guze SB, ed. *Adult psychiatry*. St. Louis: CV Mosby, 1997:53–81.

112. Baxter LR. Neuroimaging studies of human anxiety disorders. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven, 1995:921–932.

113. Drevets WC. Functional anatomical abnormalities in limbic and prefrontal cortical structures in major depression. *Prog Brain Res* 2000;126:413–431.

114. Rauch SL, Jenike MA, Alpert NM, et al. Regional cerebral blood flow measured during symptom provocation in obsessive-compulsive disorder using oxygen 15–labeled carbon dioxide and positron emission tomography. *Arch Gen Psychiatry* 1994;51:62–70.

115. Drevets WC, Videen TO, MacLeod AK, et al. PET images of blood flow changes during anxiety: correction. *Science* 1992;256:l696.

116. Drevets WC, Spitznagel E, Raichle ME. Functional anatomical differences between major depressive subtypes. *J Cereb Blood Flow Metab* 1995;15:S93.

117. Drevets WC, Simpson JR, Raichle ME. Regional blood flow changes in response to phobic anxiety and habituation. *J Cereb Blood Flow Metab* 1995;15:S856.

118. Schneider F, Gur RE, Mozley LH, et al. Mood effects on limbic blood flow correlate with emotional self-rating: a PET study with oxygen-15 labeled water. *Psychiatry Res* 1995;61:265–283.

119. Carmichael ST, Price JL. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol* 1995;363:615–641.

120. Timms RJ. Cortical inhibition and facilitation of the defense reaction. *J Physiol (Lond)* 1977;266:98–99.

121. Bechara A, Damasio H, Damasio AR. Emotion, decision making and the orbitofrontal cortex. *Cereb Cortex* 2000;10:295–307.

122. Iversen SD, Mishkin M. Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp Brain Res* 1970;11:376–386.

123. Maddock RJ. The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends Neurosci* 1999;22:310–316.

124. Van Hoesen GW, Morecraft RJ, Vogt B. Connections of the monkey cingulate cortex. In: Vogt BA, Gabriel M, eds. *Neurobiology of cingulate cortex and limbic thalamus* Boston: Birkhauser, 1993.

125. Reiman E, Raichle MF, Butler K, et al. A focal brain abnormality in panic disorder, a severe form of anxiety. *Nature* 1984;310:683–685.

126. Nordahl TE, Semple WE, Gross M, et al. Cerebral glucose metabolic differences in patients with panic disorder. *Neuropsychopharmacology* 1990;3:261–271.

127. Bisaga A, Katz JL, Antonini A, et al. Cerebral glucose metabolism in women with panic disorder. *Am J Psychiatry* 1998;155:1178–1183.

128. De Cristofaro MT, Sessarego A, Pupi A, et al. Brain perfusion abnormalities in drug-naive, lactate-sensitive panic patients: a SPECT study. *Biol Psychiatry* 1993;33:505–512.

129. Reiman EM, Raichle ME, Robins E, et al. Neuroanatomical correlates of a lactate-induced anxiety attack. *Arch Gen Psychiatry* l989;46:493–500.

130. Rauch SL, Savage CR, Alpert NM, et al. The functional neuroanatomy of anxiety: a study of three disorders using positron emission tomography and symptom provocation. *Biol Psychiatry* 1997;42:446–452.

131. Benkelfat C, Bradwejn J, Meyer E, et al. Functional neuroanatomy of CCK4–induced anxiety in normal healthy volunteers. *Am J Psychiatry* 1995;152:1180–1184.

132. Meyer JH, Swinson R, Kennedy SH, et al. Increased left posterior parietal-temporal cortex activation after D-fenfluramine in women with panic disorder. *Psychiatry Res* 2000;98:133–143.

133. Cameron OG, Zubieta JK, Grunhaus L, et al. Effects of yohimbine on cerebral blood flow, symptoms, and physiological functions in humans. *Psychosom Med* 2000;62:549–559.

134. Woods SW, Kosten K, Krystal JH, et al. Yohimbine alters regional cerebral blood flow in panic disorder [Letter]. *Lancet* 1988;2:678.

135. Ontiveros A, Fonaine R, Breton G. Correlation of severity of panic disorder and neuroanatomical changes on magnetic resonance imaging. *J Neuropsychiatry Clin Neurosci* 1989;1:404–408.

136. Vythilingam M, Anderson GM, Owens MJ, et al. Cerebrospinal fluid corticotropin-releasing hormone in healthy humans: effects of yohimbine and naloxone. *J Clin Endocrinol Metab* 2000;85:4138–4145.

137. Schneider F, Weiss U, Kessler C, et al. Subcortical correlates of differential classical conditioning of aversive emotional reactions in social phobia. *Biol Psychiatry* 1999;45:863–871.

138. Liberzon I, Taylor SF, Amdur R, et al. Brain activation in PTSD in response to trauma-related stimuli. *Biol Psychiatry* 1999;45:817–826.

139. Shin LM, McNally RJ, Kosslyn SM, et al. Regional cerebral blood flow during script-driven imagery in childhood sexual abuse-related PTSD: a PET investigation. *Am J Psychiatry* 1999;156:575–584.

140. Bremner JD, Narayan M, Staib LH, et al. Neural correlates of memories of childhood sexual abuse in women with and without posttraumatic stress disorder. *Am J Psychiatry* 1999;156:1787–1795.

141. Bremner JD, Staib LH, Kaloupek D, et al. Neural correlates of exposure to traumatic pictures and sounds in Vietnam combat veterans with and without posttraumatic stress disorder: a positron emission tomography study. *Biol Psychiatry* 1999;45:806–816.

142. Rauch SL, Whalen PJ, Shin LM, et al. Exaggerated amygdala response to masked facial stimuli in posttraumatic stress disorder: a functional MRI study. *Biol Psychiatry* 2000;47:769–776.

143. Orr SP, Metzger LJ, Lasko NB, et al. Pitman RK. *De novo* conditioning in trauma-exposed individuals with and without posttraumatic stress disorder. *J Abnorm Psychol* 2000;109:290–298.

144. Peri T, Ben-Shakhar G, Orr SP, et al. Psychophysiologic assessment of aversive conditioning in posttraumatic stress disorder. *Biol Psychiatry* 2000;47:512–519.

145. Bremner JD, Randall P, Scott TM, et al. MRI-based measurement of hippocampal volume in combat-related posttraumatic stress disorder. *Am J Psychiatry* 1995;152:973–981.

146. Bremner J.D, Randall P, Vermetten E, et al. MRI-based measurement of hippocampal volume in posttraumatic stress disorder related to childhood physical and sexual abuse: a preliminary report. *Biol Psychiatry* 1997;41:23–32.

147. Gurvits TG, Shenton MR, Hokama H, et al. Magnetic resonance imaging study of hippocampal volume in chronic combat-related posttraumatic stress disorder. *Biol Psychiatry* 1996;40:192–199.

148. Stein MB, Koverola C, Hanna C, et al. Hippocampal volume in women victimized by childhood sexual abuse. *Psychol Med* 1997;27:951–959.

149. McEwen BS. Stress and hippocampal plasticity. *Annu Rev Neurosci* 1999;22:105–122.

150. Sapolsky RM. Why stress is bad for your brain. *Science* 1996;273:749–750.

151. Laplane D, Levasseur M, Pillon B, et al. Obsessive-compulsive and other behavioural changes with bilateral basal ganglia lesions. *Brain* 1989;112:699–725.

152. Buckner RL, Petersen SE, Ojemann JG, et al. Functional anatomical studies of explicit and implicit memory retrieval tasks. *J Neurosci* 1995;15:12–29.

153. Swedo SE, Rapoport JL, Cheslow DL, et al. High prevalence of obsessive-compulsive symptoms in patients with Sydenham's chorea. *Am J Psychiatry* 1989;146:246–249.

154. Eslinger PJ, Damasio AR. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology* 1985;35:1731–1741.

155. Nauta WJH. Reciprocal links of the corpus striatum with the cerebral cortex and limbic system: a common substrate for movement and thought? In: Mueller J, ed.*Neurology and psychiatry: a meeting of minds*. New York: Karger, 1989.

156. Cassens G, Kuruc A, Roffman M, et al. Alterations in brain norepinephrine metabolism and behavior induced by environmental stimuli previously paired with inescapable shock. *Behav Brain Res* 1981;2:387–407.

157. Rasmussen K, Marilak DA, Jaclobs BI. Single unit activity of the locus coeruleus in the freely moving cat. I. During naturalistic behaviors and in response to simple and complex stimuli. *Brain Res* 1986;371:324–334.

158. Redmond DF Jr. Studies of the nucleus locus coeruleus in monkey and hypotheses for neuropsychopharmacology. In: Meltzer HY, ed. *Psychopharmacology: the third generation of progress*. New York: Raven, 1987:967–975.

159. Abercrombie ED, Jacobs BL. Single-unit response of noradrenergic neurons in the locus coeruleus of freely moving cats. I. Acutely presented stressful and nonstressful stimuli. *J. Neurosci* 1987;7:2837–2843.

160. Levine ES, Litto WJ, Jacobs BL. Activity of cat locus coeruleus noradrenergic neurons during the defense reaction. *Brain Res* 1990;531:189–195.

161. Bremner JD, Krystal JH, Southwick SM, et al. Noradrenergic mechanisms in stress and anxiety. I. Preclinical studies. *Synapse* 1996;23:28–38.

162. Bremner JD, Krystal JH, Southwick SM, et al. Noradrenergic mechanisms in stress and anxiety. II.Clinical studies. *Synapse* 1996;23:39–51.

163. Cose BJ, Robbins TW. Dissociable effects of lesions to dorsal and ventral noradrenergic bundle on the acquisition, performance, and extinction of aversive conditioning. *Behav Neurosci* 1987;101:476–488.

164. Charney DS, Deutch A. A functional neuroanatomy of anxiety and fear: implications for the pathophysiology and treatment of anxiety disorders. *Crit Rev Neurobiol* 1996;10:419–446.

165. Finlay JM, Abercrombie ED. Stress induced sensitization of norepinephrine release in the medial prefrontal cortex. *Soc Neurosci Abstr* 1991;17:151.

166. Nisenbaum LK, Zigmund MJ, Sved AF, et al. Prior exposure to chronic stress results in enhanced synthesis and release of hippocampal norepinephrine in response to a novel stressor. *J Neurosci* 1991;11:1473–1484.

167. Nisenbaum LK, Abercrombie ED. Presynaptic alterations associated with enhancement of evoked release and synthesis of NE in hippocampus of chronically cold stressed rats. *Brain Res* 1993;608:280–287.

168. Simson PE, Weiss JM. Altered activity of the locus coeruleus in an animal model of depression.*Neuropsychoparmacology* 1988;1:287–295.

169. Torda T, Kvetnansky R, Petrikova M. Effect of repeated immobilization stress on rat central and peripheral adrenoceptors. In: Usdin E, Kvetnansky R, Axelrod J, eds. *Stress: the role of catecholamines and other neurotransmitters*. New York: Gordon & Breach, 1984:691–701.

170. Charney DS, Deutch AY, Krystal JH, et al. Psychobiologic mechanisms of posttraumatic stress disorder. *Arch Gen Psychiatry* 1993;50:294–299.

171. Grillon C, Morgan CA. Fear-potentiated startle conditioning to explicit and contextual cues in Gulf War veterans with posttraumatic stress disorder. *Biol Psychiatry* 1998;44:990–997.

172. Charney DS, Heninger GR, Breier A. Noradrenergic function in panic anxiety: effects of yohimbine in healthy subjects and patients with agoraphobia and panic disorder. *Arch Gen Psychiatry* 1984;41:751–763.

173. Charney DS, Woods SW, Goodman WK, et al. Neurobiological mechanisms of panic anxiety: biochemical and behavioral correlates of yohimbine-induced panic attacks. *Am J Psychiatry* 1987;144:1030–1036.

174. Aston-Jones G, Shipley MT, Chouvet G, et al. Afferent regulation of locus coeruleus neurons: anatomy, physiology and pharmacology. *Prog Brain Res* 1991;88:47–75.

175. Prins A, Kaloupck DG, Keane TM. Psychophysiological evidence for autonomic arousal and startle in traumatized adult populations. In: Friedman MJ, Charney DS, Deutch AY, eds. *Neurobiological and clinical consequences of stress: from normal adaptation to PTSD* New York: Raven, 1995:291–314.

176. Bandelow B, Sengos G, Wedekind D, et al. Urinary excretion of cortisol, norepinephrine, testosterone, and melatonin in panic disorder. *Pharmacopsychiatry* 1997;30:113–117.

177. Wilkinson D, Thompson JM, Lambert GW, et al. Sympathetic activity in patients with panic disorder at rest, under laboratory mental stress, and during panic attacks. *Arch Gen Psychiatry* 1998;55:511–520.

178. Uhde T, Joffe RT, Jimerson DC, et al. Normal urinary free cortisol and plasma MHPG in panic disorder: clinical and theoretical implications. *Biol Psychiatry* 1988;23:575–585.

179. Nutt DJ. Altered alpha$_2$-adrenoceptor sensitivity in panic disorder. *Arch Gen Psychiatry* 1989;46:165–169.

180. Coplan JD, Pine D, Papp L, et al. Uncoupling of the noradrenergic-hypothalamus-pituitary adrenal axis in panic disorder patients. *Neuropsychopharmacology* 1995;13:65–73.

181. Coplan JD, Papp LA, Martinez MA, et al. Persistence of blunted human growth hormone response to clonidine in fluoxetine-treated patients with panic disorder. *Am J Psychiatry* 1995;152:619–622.

182. Gurguis GNM, Uhde TW. Plasma 3-methoxy-4hydroxyphenylethylene glycol (MHPG) and growth hormone responses in panic disorder patients and normal controls. *Psychoneuroendocrinology* 1990;15:217–227.

183. Gurguis GN, Vitton BJ, Uhde TW. Behavioral, sympathetic and adrenocortical responses to yohimbine in panic disorder patients and normal controls. *Psychiatry Res* 1997;71:27–39.

184. Albus M, Zahn TP, Brier A. Anxiogenic properties of yohimbine: behavioral, physiological and biochemical measures. *Eur Arch Psychiatry* 1992;241:337–344.

185. Charney DS, Woods SW, Krystal JH, et al. Noradrenergic neuronal dysregulation in panic disorder: the effects of intravenous yohimbine and clonidine in panic disorder patients. *Acta Psychiatr Scand* 1992;86:273–282.

186. Yeragani VK, Berger R, Pohl R, et al. Effects of yohimbine on heart rate variability in panic disorder patients and normal controls: a study of power spectral analysis of heart rate. *J Cardiovasc Pharmacol* 1992;20:609–618.

187. Bremner JD, Innis RB, Ng CK, et al. PET measurement of central metabolic correlates of yohimbine administration in posttraumatic stress disorder. *Arch Gen Psychiatry* 1997;54:246–256.

188. Southwick SM, Krystal JH, Morgan CA, et al. Abnormal noradrenergic function in posttraumatic stress disorder. *Arch Gen Psychiatry* 1993;50:266–274.

189. Southwick SM, Krystal JH, Bremner JD, et al. Noradrenergic and serotonergic function in posttraumatic stress disorder. *Arch Gen Psychiatry* 1997;54:749–758.

190. Lemieux AM, Coe CL. Abuse-related PTSD: evidence for chronic neuroendocrine activation in women. *Psychosom Med* 1995;57:105–115.

191. Hawk LW, Dougall AL, Ursano RJ, et al. Urinary catecholamines and cortisol in recent-onset posttraumatic stress disorder after motor vehicle accidents. *Psychosom Med* 2000;62:423–434.

192. DeBellis MD, Baum AS, Birmaher B, et al. A.E. Bennett Research Award: developmental traumatology. I. Biological stress systems. *Biol Psychiatry* 1999;45:1259–1270.

193. McFall ME, Veith RC, Murburg MM. Basal sympathoadrenal function in posttraumatic stress disorder. *Biol Psychiatry* 1992;31:1050–1056.

194. Blanchard EB, Kolb LC, Prins A, et al. Changes in plasma norepinephrine to combat-related stimuli among Vietnam veterans with post traumatic stress disorder. *J Nerv Ment Dis* 1991;179:371–373.

195. Geracioti TD, Baker DG, Ekhator NN, et al. Csf norepinephrine concentrations in posttraumatic stress disorder. *Am J Psychiatry* 2001;158:1227–30.

196. Perry GD, Giller EL, Southwick SM. Altered platelet alpha$_2$ adrenergic binding sites in posttraumatic stress disorder. *Am J Psychiatry* 1987;144:1511–1512.

197. Lerer B, Ebstein RP, Shestatsky M, et al. Cyclic AMP signal transduction in posttraumatic stress disorder. *Am J Psychiatry* 1987;144:1324–1327.

198. Davidson J, Lipper S, Kilts CD, et al. Platelet MAO activity in posttraumatic stress disorder. *Am J Psychiatry* 1985;142:1341–1343.

199. Nesse RM, Curtis GC, Thyer BA, et al. Endocrine and cardiovascular responses during phobic anxiety. *Psychosom Med* 1985;47:320–332.

200. Stein MB, Tancer ME, Uhde TW. Heart rate and plasma norepinephrine responsivity to orthostatic challenge in anxiety disorders: comparison of patients with panic disorder and social phobia and normal control subjects. *Arch Gen Psychiatry* 1992;49:311–317.

201. Tancer ME, Stein MB, Uhde TW. Effects of thyrotropin-releasing hormone on blood pressure and heart rate in phobic and panic patients: a pilot study. *Biol Psychiatry* 1990;27:781–783.

202. Stein MB, Huzel LL, Delaney SM. Lymphocyte α-adrenoreceptors in social phobia. *Biol Psychiatry* 1993;34:45–50.

203. Gerra G, Zaimovic A, Zambelli U, et al. Neuroendocrine responses to psychological stress in adolescents with anxiety disorder. *Neuropsychobiology* 2000;42:82–92.

204. Sapolsky RM, Plotsky PM. Hypercortisolism and its possible neural bases. *Biol Psychiatry* 1990;27:937–952.

205. Kant GJ, Leu JR, Anderson SM, et al. Effects of chronic stress on plasma corticosterone, ACTH and prolactin. *Physiol Behav* 1987;40:775–779.

206. Irwin J, Ahluwalia P, Zacharko RM, et al. Central norepinephrine and plasma corticosterone following acute and chronic stressors: influence of social isolation and handling. *Pharmacol Biochem Behav* 1986;24:1151–1154.

207. Dallman MF, Jones MT. Corticosteroid feedback control of ACTH secretion: effect of stress-induced corticosterone secretion on subsequent stress responses in the rat. *Endocrinology* 1973;92:1367–1375.

208. Stanton ME, Gutierrez YR, Levine S. Maternal deprivation potentiates pituitary-adrenal stress responses in infant rats. *Behav Neurosci* 1988;102:69–70.

209. Levine S, Wiener SG, Coe CL. Temporal and social factors influencing behavioral and hormonal responses to separation in mother and infant squirrel monkey. *Psychoneuroendocrinology* 1993;18:297–306.

210. Liu D, Diorio J, Tannenbaum B, et al. Maternal care, hippocampal glucocorticoid receptors and hypothalamic-pituitaryadrenal responses to stress. *Science* 1997;277:1654–1662.

211. Plotsky PM, Meaney MJ. Early postnatal experience alters hypothalamic corticotropin-releasing factor mRNA medial CRF context, and stress induced release in adult rats. *Mol Brain Res* 1993;18:195–200.

212. Heit C, Woens MJ, Plotsky PM, et al. Persistent changes in corticotropin releasing factor systems due to early life stress: relationship to pathophysiology of major depression and post-traumatic stress disorder. *Psychopharmacol Bull* 1997;33:185–192.

213. Coplan JD, Andrews MW, Rosenblum LA, et al. Persistent elevations of cerebrospinal fluid concentrations of corticotropin-releasing factor in adult nonhuman primates exposed to early-life stressors: implications for the pathophysiology of mood and anxiety disorders. *Proc Natl Acad Sci USA* 1996;93:1619–1623.

214. Meaney MJ, Airken DH, vanBerkel C, et al. Effect of neonatal handling on age-related impairments associated with the hippocampus. *Science* 1988;239:766–768.

215. Meaney MJ, Aitken DH, Sharma S, et al. Neonatal handling alters adrenocortical negative feedback sensitivity and hippocampal type II glucocorticoid receptor binding in the rat. *Neuroendocrinology* 1989;50:597–604.

216. Shulkin J, Gold PW, McEwen BS. Induction of corticotropin-releasing hormone gene expression by glucocorticoids: implication for understanding the states of fear and anxiety and allostatic load. *Psychoneuroendocrinology* 1998;23:219–243.

217. Makino S, Gold PW, Schulkin J. Corticosterone effects on corticotropin-releasing hormone mRNA in the central nucleus of the amygdala and the parvocellular region of the paraventricular nucleus of the hypothalamus. *Brain Res* 1994;640:105–112.

218. Makino S, Gold PW, Schulkin J. Effects of corticosterone on CRH mRNA and content in the bed nucleus of the stria terminalis; comparison with the effects in the central nucleus of the amygdala and the paraventricular nucleus of the hypothalamus. *Brain Res* 1994;657:141–149.

219. Makino S, Schulkin J, Smith MA, et al. Regulation of corticotropin-releasing hormone receptor messenger ribonucleic acid in the rat brain and pituitary by glucocorticoids and stress. *Endocrinology* 1995;136:4517–4525.

220. Swanson LW, Simmons DM. Differential steroid hormone and neural influences on peptide mRNA levels in CRH cells of the paraventricular nucleus: a hybridization histochemical study in the rat. *J Comp Neurol* 1989;285:413–435.

221. Koob GF, Heinrichs SC. A role for corticotropin releasing factor and urocortin in behavioral responses to stressors. *Brain Res* 1999;848:141–152.

222. Smith GW, Aubry JM, Dellu F, et al. Corticotropin releasing factor receptor 1-—deficient mice display decreased anxiety, impaired stress response, and aberrant neuroendocrine development. *Neuron* 1998;20:1093–1102.

223. Timpl P, Spanagel R, Sillaber I, et al. Impaired stress response and reduced anxiety in mice lacking a functional corticotropin-releasing hormone receptor. *Nat Genet* 1998;19:162–166.

224. Bale TL, Contarino A, Smith GW, et al. Mice deficient for corticotropin-releasing hormone receptor-2 display anxiety-like behaviour and are hypersensitive to stress. *Nat Genet* 2000;24:410–414.

225. Kishimoto T, Radulovic J, Radulovic M, et al. Delection of *Crhr2* reveals an anxiolytic role for corticotropin-releasing hormone receptor-2. *Nat Genet* 2000;24;415–419.

226. Habib KE, Weld KP, Rice KC, et al. Oral administration of a corticotropin-releasing hormone receptor antagonist significantly attenuates behavioral, neuroendocrine, and autonomic responses to stress in primates. *Proc Natl Acad Sci USA* 2000;97:6079–6084.

227. Mason JW, Giller EL, Kosten TR, et al. Urinary free-cortisol levels in post-traumatic stress disorder patients. *J Nerv Ment Dis* 1986;174:145–149.

228. Yehuda R, Boisoneau D, Lowy MT, et al. Dose-response changes in plasma cortisol and lymphocyte glucocorticoid receptors following dexamethasone administration in combat veterans with and without posttraumatic stress disorder. *Arch Gen Psychiatry* 1995;52:583–593.

229. Yehuda R, Bierer LM, Schmeidler J, et al. Low cortisol and risk for PTSD in adult offspring of holocaust survivors. *Am J Psychiatry* 2000;157:1252–1259.

230. Pitman RK, Orr SP. Twenty-four-hour urinary cortisol and catecholamine excretion in combat-related post traumatic stress disorder. *Biol Psychiatry* 1990;27:245–247.

231. Yehuda R, Lowy MT, Southwick SM, et al. Lymphocyte glucocorticoid receptor number in posttraumatic stress disorder. *Am J Psychiatry* 1991;148:499–504.

232. Liberzon I, Abelson JL, Flagel SB, et al. Neuroendocrine and psychophysiological responses in PTSD: a symptom provocation study. *Neuropsychopharmacology* 1999;21:40–50.

233. Maes M, Lin A, Bonaccorso S, et al. Increased 24-hour urinary cortisol excretion in patients with post-traumatic stress disorder and patients with major depression, but not in patients with fibromyalgia. *Acta Psychiatr Scand* 1998;90:328–325.

234. Kosten TR, Wahby V, Giller EL, et al. The dexamethasone suppression test and thyrotropin-releasing hormone stimulation test in posttraumatic stress disorder. *Biol Psychiatry* 1990;28:657–664.

235. Yehuda R, Southwick SM, Krystal JH, et al. Enhanced suppression of cortisol following dexamethasone administration in posttraumatic stress disorder. *Am J Psychiatry* 1993;150:83–86.

236. Stein MB, Yehuda R, Koverola C, et al. Enhanced dexamethasone suppression of plasma cortisol in adult women traumatized by childhood sexual abuse. *Biol Psychiatry* 1997;42:680–686.

237. Yehuda R, Teicher MH, Trestman RL, et al. Cortisol regulation in posttraumatic stress disorder and major depression: a chronobiological analysis. *Biol Psychiatry* 1996;40:79–88.

238. Delahanty DL, Raimonde AJ, Spoonster E. Initial posttraumatic urinary cortisol levels predict subsequent PTSD symptoms in motor vehicle accident victims. *Biol Psychiatry* 2000;48:940–947.

239. Bremner JD, Licinio J, Darnell A, et al. Elevated CSF corticotropin-releasing factor concentrations in posttraumatric stress disorder. *Am J Psychiatry* 1997;154:624–629.

240. Baker D, West SA, Orth DN, et al. Cerebrospinal fluid corticotropin-releasing hormone and adrenal cortical activity in post traumatic stress disorder. *Am J Psychiatry* 1999;156:585–588.

241. Smith MA, Davidson J, Ritchie JC, et al. The corticotropin-releasing hormone test in patients with PTSD. *Biol Psychiatry* 1989;26:349–355.

242. Heim C, Newport DJ, Heit S, et al. Pituitary-adrenal and autonomic responses to stress in women after sexual and physical abuse in childhood. *JAMA* 2000;284:592–597.

243. Sapolsky RM, Uno H, Rebert CS, Finch CE. Hippocampal damage associated with prolonged glucocorticoid exposure in primates. *J. Neurosci* 1990;10:2897–2902.

244. Goldstein S, Halbreich U, Asnis G, et al. The hypothalamic pituitary-adrenal system in panic disorder. *Am J Psychiatry* 1987;144:1320–1323.

245. Holsboer F, vonBardeleben U, Buller R, et al. Stimulation response to corticotropin-releasing hormone (CRH) in patients with depression, alcoholism an panic disorder. *Horm Metab Res* 1987;16[Suppl]:80–88.

246. Kathol RG, Anton R, Noyes R, et al. Relationship of urinary free cortisol levels in patients with panic disorder to symptoms of depression and agoraphobia. *Psychiatry Res* 1988;24:211–221.

247. Abelson JL, Curtis GC. Hypothalamic-pituitary-adrenal axis activity in panic disorder. *Arch Gen Psychiatry* 1996;53:323–332.

248. Coryell W, Noyes R. HPA axis disturbance and treatment outcome in panic disorder. *Biol Psychiatry* 1988;24:762–755.

249. Roy-Byrne PP, Uhde TW, Post RM, et al. The corticotropin-releasing hormone stimulation test in patients with panic disorder. *Am J Psychiatry* 1986;143:896–899.

250. Rapaport MH, Risch SC, Golshan S, et al. Neuroendocrine effects of ovine corticotropin-releasing hormone in panic disorder patients. *Biol Psychiatry* 1989;26:344–348.

251. Jolkkonen J, Lepola V, Bissette G, et al. CSF corticotropin-releasing factor is not affected in panic disorder. *Biol Psychiatry* 1993;33:136–138.

252. Curtis AL, Lechner SM, Pavcovich LA, et al. Activation of the locus coeculeus noradrenergic system by intracoerulear microinfusion of corticotropin releasing factor: effects on discharge, rate, cortical norepinephrine levels, and cortical electroencephalographic activity. *J Pharmacol Exp Ther* 1997;281:163–172.

253. Smagin GN, Swiergiel AH, Dunn AJ. Corticotropin releasing factor administered in the locus coeruleus, but not the parabrachial nucleus, stimulates norepinephrine release in the prefrontal cortex. *Brain Res Bull* 1995;36:71–76.

254. Pacak K, Palkovits M, Kopin IJ, et al. Stress induced NE release in the hypothalamic PVN and pituitary-adrenal and sympathoadrenal activity: *in vivo* microdialysis studies. *Front Neuroendocrinol* 1995;16:89–150.

255. Roozendaal B. Glucocorticoids and the regulation of memory consolidation. *Psychoneuroendocrinology* 2000;25:213–238.

256. Roozendaal B, Carmi O, McGaugh JL. Adrencortical suppression blocks the memory-enhancing effects of amphetamine and epinephrine. *Proc Natl Acad Sci USA* 1996;93:1429–1433.

257. Cahill L, Prins B, Weber M, et al. Beta-adrenergic activation and memory for emotional events. *Nature* 1994;371:702–703.

258. Choi DW, Farb DH, Fischbach GD. Chlordiazepoxide selectively potentiates GABA conductance of spinal cord and sensory neurons in cell culture. *J Neurophysiol* 1981;45:621–631.

259. Study RE, Barker JL. Cellular mechanisms of benzodiazepine action. *JAMA* 1982;247:2147–2151.

260. Graeff FG, Silveira MC, Nogueira RL, et al. Role of the amygdala and periaqueductal gray in anxiety and panic. *Behav Brain Res* 1993;58:123–131.

261. Dorow R, Horowski R, Paschelke G, et al. Severe anxiety induced by FG7142, a beta-carboline ligand for benzodiazepine receptors. *Lancet* 1983;2:98–99.

262. Braestrup C, Schmiechen R, Neef G, et al. Interaction of convulsive ligands with benzodiazepine receptors. *Science* 1982;216:1241–1243.

263. Ninan PT, Insel TM, Cohen RM, et al. Benzodiazepine receptor mediated experimental "anxiety" in primates. *Science* 1982;218:1332–1334.

264. Low K, Crestani F, Keist R, et al. Molecular and neuronal substrate for the selective attenuation of anxiety. *Science* 2000;290:131–134.

265. McKernan RM, Rosahl TW, Reynolds DS, et al. Sedative but not anxiolytic properties of benzodiazepines are mediated by the $GABA_A$ receptor alpha$_1$ subtype. *Nat Neurosci* 2000;3:587–592.

266. Rudolph U, et al. Benzodiazepine actions mediated by specific γ-aminobutyric acid a receptor subtypes. *Nature* 1999;410:796–800.

267. Paslawski T, Treit D, Baker GB, et al. The antidepressant drug phenelzine produces antianxiety effects in the plus-maze and increases in rat brain GABA. *Psychopharmacology (Berl)* 1996;127:19–24.

268. Weizman A, Weisman R, Kook KA, et al. Adrenalectomy prevents the stress-induced decrease in *in vivo* [^{3}H]Ro 15–1788 binding to $GABA_A$ benzodiazepine receptors in the mouse. *Brain Res* 1990;519:347–350.

269. Drugan RC, Morrow AL, Weizman R, et al. Stress-induced behavioral depression in the rat is associated with a decrease in GABA receptor-mediated chloride ion flux and brain benzodiazepine receptor occupancy.*Brain Res* 1989;487:45–51.

270. Weizman R, Weizman A, Kook KA, et al. Repeated swim stress alters brain benzodiazepine receptors measured *in vivo. J Pharmacol Exp Ther* 1989;249:701–707.

271. Orchinik M, Weiland NG, McEwen BS. Chronic exposure to stress levels of corticosterone alters GABAA receptor subunit mRNA levels in rat hippocampus. *Brain Res Mol Brain Res* 1995;34:29–37.

272. Robertson HA, Martin IL, Candy JM. Differences in benzodiazepine receptor binding in Maudsley-reactive and non-reactive rats. *Eur J Pharmacol* 1978;50:455–457.

273. Caldji C, Francis D, Sharma S, et al. The effects of early rearing environment on the development of GABAA and central benzodiazepine receptor levels and novelty-induced fearfulness in the rat. *Neuropsychopharmacology* 2000;22:219–229.

274. Nutt DJ, Glue P, Lawson C, et al. Flumazenil provocation of panic attacks: evidence for altered benzodiazepine receptor sensitivity in panic disorder. *Arch Gen Psychiatry* 1990;47:917–925.

275. Roy-Byrne P, Wingerson DK, Radant A, et al. Reduced benzodiazepine sensitivity in patients with panic disorder: comparison with patients with obsessive compulsive disorder and normal subjects. *Am J Psychol* 1996;153:1444–1449.

276. Woods SW, Charney DS, Silver JM, et al. Behavioral, biochemical, and cardiovascular responses to the benzodiazepine receptor antagonist flumazenil in panic disorder. *Psychiatry Res* 1991;36:115–124.

277. Roy-Byrne PP, Lewis N, Villacres E, et al. Preliminary evidence of benzodiazepine subsensitivity in panic disorder. *Biol Psychiatry* 1989;26:744–748.

278. Schlegel S, Teinert H, Bockisch A, et al. Decreased benzodiazepine receptor binding in panic disorder measured by iomazenil SPECT: a preliminary report. *Eur Arch Psychiatry Clin Neurosci* 1994;244:49–51.

279. Kascka W, Feistel H, Ebert D. Reduced benzodiazepine receptor binding in panic disorders measured by iomazenil SPECT. *J Psychiatry Res* 1995;29:427–434.

280. Kuikka JT, Pitkanen A, Lepola U, et al. Abnormal regional benzodiazepine receptor uptake in the prefrontal cortex in patients with panic disorder. *Nucl Med Commun* 1995;16:273–280.

281. Bremner JD, Innis RB, White T, et al. SPECT [I-123] Iomazenil measurement of the benzodiazepine receptor in panic disorder. *Biol Psychiatry* 2000;47:96–106.

281a. Bremner JD, Innis RB, Southwick SM, et al. Decreased benzodiazepine receptor binding in prefrontal cortex in combat-related posttraumatic stress disorder. *Am J Psychiatry* 2000;157:1120–1126.

282. Malizia AL, Cunningham VJ, Bell CJ, et al. Decreased brain $GABA_A$ -benzodiazepine receptor binding in panic disorder: preliminary results from a quantitative PET study. *Arch Gen Psychiatry* 1998;55:715–720.

283. Abadie P, Boulenger JP, Benail K, et al. Relationships between trait and state anxiety and the central benzodiazepine receptor: a PET study. *Eur J Neurosci* 1999;11:1470–1478.

284. Deutch AY, Young CD. A model of the stress-induced activation of prefrontal cortical dopamine systems: coping and the development of post-traumatic stress disorder. In: Friedman MJ, Charney DS, Deutch AY, eds. *Neurobiological and clinical consequences of stress*. Philadelphia: Lippincott-Raven, 1995:163–175.

285. Inoue T, Tsuchiya K, Koyama T. Regional changes in dopamine and serotonin activation with various intensity of physical and psychological stress in the rat brain. *Pharmacol Biochem Behav* 1994;49:911–920.

286. Roy-Byrne PP, Uhde TW, Sack DA, et al. Plasma HVA and anxiety in patients with panic disorder. *Biol Psychiatry* 1986;21:847–849.

287. Pichot W, Annsseau M, Moreno AG, et al. Dopaminergic function in panic disorder: comparison with major and minor depression. *Biol Psychiatry* 1992;32:1004–1011.

288. Eriksson E, Westberg P, Alling C, et al. Cerebrospinal fluid levels of monoamine metabolites in panic disorder. *Psychiatry Res* 1991;36:243–251.

289. Hamilton SP, Haghighi F, Heiman GA, et al. Investigation of dopamine receptor (DRD4) and dopamine transporter (DAT) polymorphisms for genetic linkage or association to panic disorder. *Am J Med Genet* 2000;96:324–330.

290. Tiihonen J, Kuikka J, Bergstrom K, et al. Dopamine reuptake site densities in patients with social phobia. *Am J Psychiatry* 1997;154:239–242.

291. Schneier FR, Liebowitz MR, Abi-Dargham A, et al. Low dopamine D(2) receptor binding potential in social phobia. *Am J Psychiatry* 2000;157:457–459.

292. Petty F, Kramer G, Wilson L. Prevention of learned helplessness: *in vivo* correlation with serotonin. *Pharmacol Biochem Behav* 1992;43:361–367.

293. Petty F, Kramer GL, Wu J. Serotonergic modulation of learned helplessness. *Ann NY Acad Sci* 1997;821:538–541.

294. Graeff F. Role of 5-HT in defensive behavior and anxiety. *Rev Neurosci* 1993;4:181–211.

295. Ramboz S, Oosting R, Amara DA, et al. Serotonin receptor 1A knockout: an animal model of anxiety-related disorder. *Proc Natl Acad Sci USA* 1998;95:14476–14481.

296. López JF, Chalmers DT, Little KY, et al. Regulation of serotonin$_{1A}$, glucocorticoid, and mineralocorticoid receptor in rat and human hippocampus: implications for the neurobiology of depression. *Biol Psychiatry* 1998;43:547–573.

297. Meijer OC, Van Oosten RV, de Kloet ER. Elevated basal trough levels of corticosterone suppress hippocampal 5-HT1A receptor expression in adrenally intact rats: implication for the pathogenesis of depression. *Neuroscience* 1997;80:419–426.

298. Mendelson SD, McEwen BS. Autoradiographic analyses of the effects of restraint-induced stress on 5-HT1A, 5-HT1C and 5-HT2 receptors in the dorsal hippocampus of male and female rats. *Neuroendocrinology* 1991;54:454–461.

299. Zhono P, Ciaranello R. Transcriptional regulation of hippocampal 5-HT1A receptors by glucocorticoid hormones. *Neuroscience* 1994;20:1161.

300. Watanabe Y, Sakai RR, McEwen BS, et al. Stress and antidepressant effects on hippocampal and cortical 5-HT$_{1a}$ and 5-HT$_{2}$ receptors and transport sites for serotonin. *Brain Res* 1993;615:87–94.

301. Norman TR, Judd FK, Gregory M, et al. Platelet serotonin uptake in panic disorder. *J Affect Disord* 1986;11:69–72.

302. Balon R, Poh R, Yeragani V, et al. Platelet serotonin levels in panic disorder. *Acta Psychiatr Scand* 1987;75:315.

303. Pecknold JC, Suranyi-Cadotte B, Chang H, et al. Serotonin uptake in panic disorder and agoraphobia. *Neuropsychopharmacology* 1988;1:173–176.

304. Uhde TW, Berrettini WH, Roy-Byrne PP, et al. Platelet 3H-imipramine binding in patients with panic disorder. *Biol Psychiatry* 1987;22:52–58.

305. Innis RB, Charney DS, Heninger GR. Differential 3H imipramine platelet binding in patients with panic disorder and depression. *Psychiatry Res* 1987;21:33–41.

306. Schneider LS, Munjack D, Severson JA, et al. Platelet H3 imipramine binding in generalized anxiety disorder, panic disorder, and agoraphobia with panic attacks. *Biol Psychiatry* 1987;21:3–41.

307. Charney DS, Heninger GR. Serotonin function in panic disorders: the effects of intravenous tryptophan in healthy subjects and panic disorder patients before and after alprazolam treatment. *Arch Gen Psychiatry* 1986;43:1059–1065.

308. DenBoer JA, Westenberg HGM. Behavioral, neuroendocrine, and biochemical effects of 5-hydroxytryptophan administration in panic disorder. *Psychiatry Res* 1990;31:367–378.

309. Goddard AW, Sholomskas DE, Augeri FM, et al. Effects of tryptophan depletion in panic disorders. *Biol Psychiatry* 1994;36:775–777.

310. Targum SD, Marshall LE. Fenfluramine provocation of anxiety in patients with panic disorder. *Psychiatry Res* 1989;28:295–306.

311. Lesch KP, Wiesmann M, Hoh A, et al. 5-HT1A receptor-effector system responsivity in panic disorder. *Psychopharmacology (Berl)* 1992;106:111–117.

312. Kahn RS, Asnis GM, Wetzler S, et al. Serotonin and anxiety revisited. *Biol Psychiatry* 1988;23:189–208.

313. Charney DS, Woods SW, Goodman WK, et al. Serotonin function in anxiety. II. Effects of the serotonin agonist MCPP in panic disorder patients and healthy subjects. *Psychopharmacology (Berl)* 1987;92:14–24.

314. Arora RC, Fichtner CG, O'Connor F, et al. Paroxetine binding in the blood platelets of post-traumatic stress disorder patients. *Life Sci* 1993;53:919–928.

315. Hano J, Vasar E, Bradwejn J. Cholecystokinin in animal and human research on anxiety. *Trends Pharmacol Sci* 1993;14:244–249.

316. Bradwejn J, Koszycki D. Imipramine antagonism of the panicogenic effects of CCK-4 in panic disorder patients.*Am J Psychiatry* 1994;151:261–263.

317. Bradwejn J, Koszycki D, Couetoux du Tetre A, et al. The panicogenic effects of CCK-4 are antagonized by L-365–260, a CCK receptor antagonist, in patients with panic disorder. *Arch Gen Psychiatry* 1994;51:486–493.

318. Kellner M, Wiedemann K, Yassouridis A, et al. Behavioral and endocrine response to cholecystokinin tetrapeptide in patients with posttraumatic stress disorder. *Biol Psychiatry* 2000;47:107–111.

319. Strohle A, Holsboer F, Rupprecht R. Increased ACTH concentrations associated with cholecystokinin tetrapeptide–induced panic attacks in patients with panic disorder. *Neuropsychopharmacology* 2000;22:251–256.

320. Lydiard RB, Ballenger JC, Laraia MT, et al. CSF cholecystokinan concentrations in patients with panic disorder and normal comparison subjects. *Am J Psychiatry* 1992;149:691–693.

321. Kennedy JL, Bradwejn J, Koszycki D, et al. Investigation of cholecystokinin system genes in panic disorder. *Mol Psychiatry* 1999;4:284–285.

322. Pande AC, Greiner M, Adams JB, et al. Placebo-controlled trial of the CCK-B antagonist, CI-988 in panic disorder. *Biol Psychiatry* 1999;46:860–862.

323. Madden J, Akil H, Patrick RI, et al. Stress induced parallel changes in central opioid levels and pain responsiveness in the rat. *Nature* 1977;265:358–360.

324. Stuckey J, Marra S, Minor T, et al. Changes in mu opiate receptors following inescapable shock. *Brain Res* 1989;476:167–169.

325. Fanselow MS. Conditioned fear-induced opiate analgesia: a competing motivational state theory of stress analgesia. *Ann NY Acad Sci* 1986;467:40–54.

326. Maier SF. Stressor controllability and stress induced analgesia. *Ann NY Acad Sci* 1986;467:55–72.

327. Pitman RK, van der Kolk BA, Orr SP, et al. Naloxone-reversible analgesic response to combat-related stimuli in posttraumatic stress disorder. *Arch Gen Psychiatry* 1990;47:541–544.

328. Baker DG, West SA, Orrth DN, et al. Cerebrospinal fluid and plasma beta endorphin in combat veterans with post traumatic stress disorder. *Psychoneuroendocrinology* 1997;22:517–529.

329. Hoffman I, Watsgon PD, Wilson G, et al. Low plasma endorphin in posttraumatic stress disorder. *Aust NZ J Psychiatry* 1989;23:268–273.

330. Wolf M. Plasma methionine enkephalin in PTSD. *Biol Psychiatry* 1991;29:295.

331. Bremner JD, Southwick SM, Darnell A, et al. Chronic PTSD in Vietnam combat veterans: course of illness and substance abuse. *Am J Psychiatry* 1996;153:369–375.

332. Heilig M, Koob GF, Ekman R, et al. Corticotropin-releasing factor and neuropeptide Y: role in emotional integration. *Trends Pharmacol Sci* 1994;17:80–85.

333. Thorsell A, Michalkiewicz M, Dumont Y, et al. Behavioral insensitivity to restraint stress, absent fear suppression of behavior and impaired spatial learning in transgenic rats with hippocampal neuropeptide Y overexpression. *Proc Natl Acad Sci USA* 2000;97:12852–12857.

334. Morgan CA III, Wang S, Southwick SM, et al. Plasma neuropeptide-Y concentrations in humans exposed to military survival training. *Biol Psychiatry* 2000;47:902–909.

335. Boulenger J, Jerabek I, Jolicoeur FB. Elevated plasma levels of neuropeptide Y in patients with panic disorder. *Am J Psychiatry* 1996;153:114–116.

336. Rasmusson AM, Hauger RL, Morgan CA, et al. Low baseline and yohimbine-stimulated plasma neuropeptide Y (NPY) levels in combat related PTSD. *Biol Psychiatry* 2000;47:526–539.

337. Mason J, Southwick S, Yehuda R, et al. Elevation of serum free triodothyronine, total triiodothronine, Thyroxine-binding globulin, and total thyroxine levels in combat-related posttraumatic stress disorder. *Arch Gen Psychiatry* 1994;51: 629–641.

338. Mason JW, Mougey FH, Brady JV, et al. Thyroid (plasma butanol-extractable iodine) responses to 72-hr avoidance sessions in the monkey. *Psychosom Med* 1968;30:682–696.

339. Gorman JM, Goetz RR, Dillon E, et al. Sodium d-lactate infusion in panic disorder patients. *Neuropsychopharmacology* 1990;3:181–190.

340. Klein DF. False suffocation alarms, spontaneous panics, and related conditions.*Arch Gen Psychiatry* 1993;50:306–317.

341. Papp LA, Klein DF, Gorman JM. Carbon dioxide hypersensitivity, hyperventilation, and panic disorder. *Am J Psychiatry* 1993;150:1149–1155.

342. Pine DS, Coplan JD, Lazlo AP, et al. Ventilatory physiology of children and adolescents with anxiety disorders. *Arch Gen Psychiatry* 1998;55:123–129.

343. Gorman JM, Fyer MR, Goetz R, et al. Ventilatory physiology of patients with panic disorder. *Arch Gen Psychiatry* 1988;45:31–39.

344. Gorman JM, Battista D, Goetz R, et al. A comparison of sodium bicarbonate and sodium lactate infusion in the induction of panic attacks. *Arch Gen Psychiatry* 1989;46:145–150.

345. Van Den Hout MA, Griez E. Panic symptoms after inhalation of carbon dioxide. *Br J Psychiatry* 1984;144:503–507.

346. Greiz E, Lousberg H, Van Den Hout MA, et al. Carbon dioxide vulnerability in panic disorder. *Psychiatry Res* 1987;20:87–95.

347. Pitts LN, McClure JN. Lactate metabolism in anxiety neuroses. *N Engl J Med* 1967;277:1329–1336.

348. Schmidt NB. Lerew DR. Jackson RJ. Prospective evaluation of anxiety sensitivity in the pathogenesis of panic: replication and extension. *J Abnormal Psychol* 1999;108:532–537.

349. Pine DS, Klein RG, Coplan JD, et al. differential carbon dioxide sensitivity in childhood anxiety disorders and nonill comparison group. *Arch Gen Psychiatry* 2000;57:960–967.

64

Neural Circuitry of Anxiety and Stress Disorders

Michael Davis

Michael Davis: Department of Psychiatry, Emory University School of Medicine, Atlanta, Georgia.

BRAIN SYSTEMS IN THE GENERATION OF FEAR AND ANXIETY

Part of "64 - Neural Circuitry of Anxiety and Stress Disorders "

Role of the Amygdala

Many data now indicate that the amygdala, along with its many efferent connections, is critically involved in emotion. Although the amygdala complex is generally defined by several distinct groups of cells, including the lateral, basal, accessory basal central, medial, and cortical nuclei, new data indicate that it is more useful to think of the amygdala as the *basolateral amygdala* (Bla) and to think of its several target areas as parts of a broader network that subserve more specialized functions (Fig. 64.1). The Bla receives sensory information from the thalamus, cortex (169), and ventral hippocampus (54) and then activates or modulates synaptic transmission in target areas appropriate for the reinforcement signal with which the sensory information has been associated. This involves both positive and negative associations. However, because most of the literature on the amygdala has analyzed the role of the Bla and its adjacent target, the central nucleus of the amygdala (CeA), in aversive conditioning, this work serves as the main focus of this chapter. Brief summaries of the role of Bla outputs to other targets shown in Fig. 64.1 follow. Because the periaqueductal gray (PAG) has received considerable attention in the study of defensive behavior and the hippocampus in the study of contextual fear conditioning, these data are reviewed next. Finally, brain systems and neurotransmitters involved in the inhibition of fear are reviewed, given the clinical significance of this information.

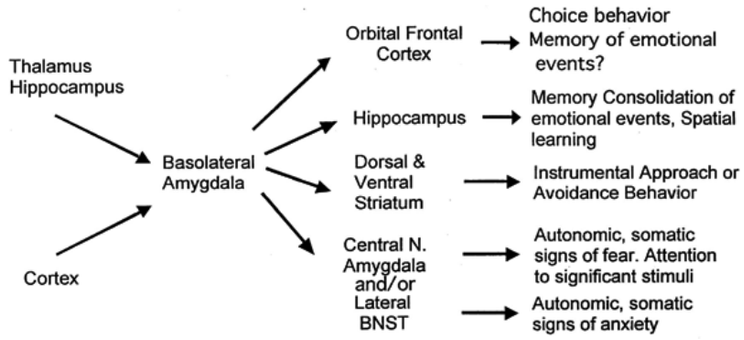


FIGURE 64.1. Schematic diagram of the outputs of the basolateral amygdala to various target areas and how these connections may be involved in fear and anxiety.

Basolateral Nucleus of the Amygdala to CeA or BNST Pathway as It Relates to Conditioned and Unconditioned Fear

Figure 64.1 shows that the Bla projects directly to the CeA, as well as to a related structure, the lateral division of the bed nucleus of the stria terminalis (BNST), to form part of the *lateral extended amygdala* (6). Figure 64.2 summarizes work done in many different laboratories indicating that the CeA and BNST have direct projections to various anatomic areas that may be expected to be involved in many of the symptoms of fear or anxiety (65). The CeA and BNST have been grouped together because fibers from the Bla that project to the BNST pass through the CeA and cells in the lateral division of the CeA project to the BNST. Thus, many effects previously attributed to the CeA may really depend on the BNST.

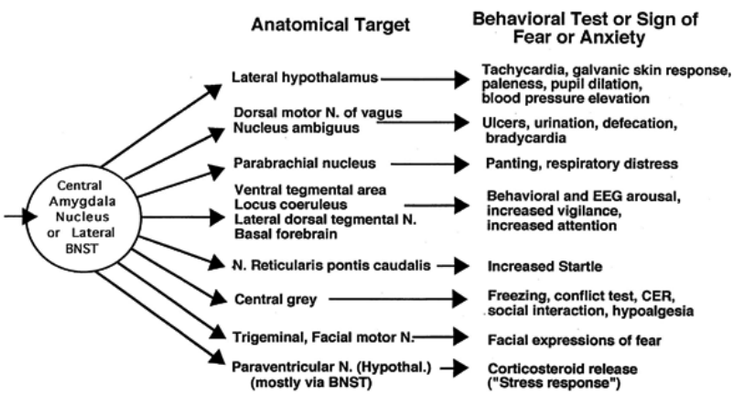


FIGURE 64.2. Schematic diagram of the outputs of the central nucleus of the amygdala and the lateral division of the bed nucleus of the stria terminalis to various target areas and how these connections may be related to specific aspects of fear and anxiety. BNST, bed nucleus of the stria terminalis; CER, conditioned emotional response; EEG, electroencephlographic; N, nucleus.

Autonomic and Hormonal Measures

Anatomically, the CeA and the BNST are well situated to mediate the various components of the fear response. Both structures send prominent projections to areas such as the lateral hypothalamus, which is involved in activation of the sympathetic autonomic nervous system seen during fear and anxiety (155). Direct projections to the dorsal motor nucleus of the vagus, nucleus of the solitary tract, and ventrolateral medulla may be involved in lateral extended amygdala modulation of heart rate and blood pressure, which are known to be regulated by these brainstem nuclei (222). Projections to the parabrachial nucleus may be involved in respiratory (as well as cardiovascular changes) during fear, because electrical stimulation and lesions of this nucleus are known to alter various measures of respiration. Indirect projections of the CeA to the paraventricular nucleus through the BNST and preoptic area may mediate the prominent neuroendocrine responses to fearful or stressful stimuli.

Attention and Vigilance

Projections from the lateral extended amygdala to the ventral tegmental area may mediate stress-induced increases in dopamine metabolites in the prefrontal cortex (101). Direct projections to the dendritic field of the locus ceruleus or indirect projections through the paragigantocellularis nucleus may mediate the increase in firing rates of cells in the locus ceruleus in the presence of a fearful stimulus. Direct

projections to the lateral dorsal tegmental nucleus and parabrachial nuclei, which have cholinergic neurons that project to the thalamus, may mediate increases in synaptic transmission in thalamic sensory relay neurons during states of fear. This cholinergic activation, along with increases in thalamic transmission accompanying activation of the locus ceruleus, may thus lead to increased vigilance and superior signal detection in a state of fear or anxiety.

As emphasized by Kapp et al. (141), in addition to its direct connections to the hypothalamus and brainstem, the CeA has the potential for indirect widespread effects on the cortex through its projections to cholinergic neurons that project to the cortex. The rapid development of conditioned bradycardia during pavlovian aversive conditioning, critically dependent on the amygdala, may reflect a general increase in attention.

Motor Behavior

Release of norepinephrine onto motor neurons by lateral extended amygdala activation of the locus ceruleus, or through projections to serotonin containing raphe neurons, could lead to enhanced motor performance during a state of fear, because both norepinephrine and serotonin facilitate excitation of motor neurons. Direct projections to the nucleus reticularis pontis caudalis, as well as indirect projections to this nucleus through the central gray, probably are involved in fear potentiation of the startle reflex. Direct projections to the lateral tegmental field, including parts of the trigeminal and facial motor nuclei, may mediate some of the facial expressions of fear as well as potentiation of the eyeblink reflex. The lateral extended amygdala also projects to regions of the central gray that appear to be a critical part of a general defense system and that have been implicated in conditioned fear in certain behavioral tests including freezing, sonic and ultrasonic vocalization, and stress-induced hypalgesia (20 ,33 ,78 ,103 ,121 ,155).

Elicitation of Fear Responses by Electrical or Chemical Stimulation of the Extended Amygdala

Electrical stimulation or abnormal electrical activation of the amygdala (i.e., by temporal lobe seizures) can produce a complex pattern of behavioral and autonomic changes that, taken together, highly resembles a state of fear.

Autonomic and Hormonal Measures

As outlined by Gloor: “The most common affect produced by temporal lobe epileptic discharge is fear…. It arises ‘out

of the blue.' Ictal fear may range from mild anxiety to intense terror. It is frequently, but not invariably, associated with a rising epigastric sensation, palpitation, mydriasis, and pallor and may be associated with a fearful hallucination, a frightful memory flashback, or both" (98). In humans, electrical stimulation of the amygdala elicits feelings of fear or anxiety as well as autonomic reactions indicative of fear (57 ,99). Although other emotional reactions occasionally are produced, the major reaction is one of fear or apprehension.

Electrical stimulation of the CeA or chemical activation by the cholinergic agonist carbachol or the neurotransmitter glutamate produces prominent cardiovascular effects that depend on the species, site of stimulation, and state of the animal. CeA stimulation can also produce gastric ulceration and can increase gastric acid, and these features can be associated with chronic fear or anxiety. It can also alter respiration, a prominent symptom of fear, especially in panic disorder.

Using very small infusion cannulas and very low doses, Sanders and Shekhar found increases in blood pressure and heart rate when the γ-aminobutyric $acid_A$ ($GABA_A$) antagonist bicuculline was infused into the Bla but not the CeA (215). Local infusion of *N*-methyl-D-aspartate (NMDA) or AMPA into the basolateral nucleus also increased blood pressure and heart rate (230). Repeated infusion of initially subthreshold doses of bicuculline into the anterior basolateral nucleus led to a "priming" effect in which increases in heart rate and blood pressure were observed after three to five infusions (216). This change in threshold lasted at least 6 weeks and could not be ascribed to mechanical damage or generalized seizure activity based on EEG measurements. Similar changes in excitability were produced by repetitive infusion of very low doses of corticotropin-releasing hormone (CRH) or the related peptide, urocortin (210). Once primed, these animals exhibited behavioral and cardiovascular responses to intravenous sodium lactate, a panic-inducing treatment in certain types of psychiatric patients.

In general, electrical stimulation of the amygdala causes an increase in plasma levels of corticosterone. The effect of electrical stimulation appears to depend on both norepinephrine and serotonin in the paraventricular nucleus. Depletion of these transmitters through local infusions of 6-hydroxydopamine or 5,7-DHT, or local infusion of the norepinephrine or serotonin antagonists prazosin or ketanserin, into the paraventricular nucleus attenuated the effects of electrical stimulation (80).

Attention and Vigilance

Studies in several species indicate that electrical stimulation of the CeA increases attention or processes associated with increased attention. For example, stimulation of sites in the CeA that produce bradycardia (142) also produce low-voltage fast EEG activity (140). In fact, an attention or orienting reflex was the most common response elicited by electrical stimulation of the amygdala (16 ,241). These and other observations led Kapp et al. to hypothesize that the "central nucleus and its associated structures function, at least in part, in the acquisition of an increased state of nonspecific attention or arousal manifested in a variety of CRs which function to enhance sensory processing. This mechanism is rapidly acquired, perhaps via an inherent plasticity within the nucleus and associated structures in situations of uncertainty but of potential import; for example, when a neutral stimulus (CS) precedes either a positive or negative reinforcing, unexpected event (US)" (141). Electrical stimulation of the amygdala can also activate cholinergic cells that are involved in arousal-like effects depending on the state of sleep and perhaps the species.

Motor Behavior

Electrical or chemical stimulation of the CeA produces a reduction of prepotent, ongoing behavior, a critical component in several animal models such as freezing, the operant conflict test, the conditioned emotional response, and the social interaction test. Electrical stimulation of the amygdala also elicits jaw movements and activation of facial motoneurons, which may mimic components of the facial expressions seen during the fear reaction. These motor effects may be indicative of a more general effect of amygdala stimulation, namely, that of modulating brainstem reflexes such as the massenteric, baroreceptor nictitating membrane, eyeblink, and the startle reflex.

Summary of the Effects of Stimulation of the Amygdala

Viewed in this way, the pattern of behaviors seen during fear may result from activation of a single area of the brain (the extended amygdala), which then projects to various target areas that themselves are critical for each of the specific symptoms of fear (the expression of fear), as well as the experience of fear. Moreover, it must be assumed that all these connections are already formed in an adult organism, because electrical stimulation produces these effects in the absence of prior explicit fear conditioning. Thus, much of the complex behavioral pattern seen during a state of "conditioned fear" has already been "hard wired" during evolution.

For a formerly neutral stimulus to produce the constellation of behavioral effects used to define a state of fear or anxiety, it is only necessary for that stimulus to activate the amygdala, which, in turn, will produce the complex pattern of behavioral changes by virtue of its innate connections to different brain target sites. Viewed in this way, plasticity during fear conditioning probably results from a change in synaptic inputs before or in the Bla (173 ,192 ,204), rather than from a change in its efferent target areas. The ability to produce long-term potentiation (LTP) in the Bla (55 ,56 ,58 ,91 ,129 ,226) that can lead to an increase in responsiveness to a physiologic stimulus (203) and the finding

that local infusion of NMDA antagonists into the amygdala block the acquisition of fear conditioning (65) are consistent with this hypothesis.

Effects of Lesions of the Amygdala on Conditioned and Unconditioned Fear in Rodents and Other Species

Many studies in rodents and other species indicate that lesions of the Bla or CeA block many different measures of conditioned fear, as well as unconditioned fear. Table 64.1 and Table 64.2 show selected examples of such studies in animals, which have been extensively reviewed elsewhere (64 ,65). More recent studies in humans also point to the amygdala in fear and anxiety.

Method	Species	Site	Effect of Lesion	Reference
Aspiration (pre)	Human	AC	Decrease galvanic skin response during classic fear conditioning unilateral lesions	153
Aspiration (pre)	Human	AC	Decrease galvanic skin response during classic fear conditioning	26
Electrolytic	Rabbit	Ce	Decrease bradycardia to cue paired with shock	139
Ibotenic acid (pre)	Rabbit	Ce	Decrease bradycardia to cue paired with shock	168
Ibotenic acid (Ce)	Rat	Ce	Decrease blood pressure rise to cue paired with shock	137
Electrolytic (MG)		MG	(Ce, unilateral, MG, contralateral)	
Cooling (pre)	Cat	Ce	Decrease bradycardia, respiratory increases and blood pressure changes to cue paired with shock	259
Electrolytic (pre)	Rat	L	Decrease blood pressure rise during classic conditioning	154,206
NMDA (pre or post)	Rat	C	Decrease secretion of corticosterone and defecation to cue paired with shock; decrease rise in dopamine, serotonin, or norepinephrine metabolites in prefrontal cortex to cue paired with shock	101
Electrolytic (pre)	Rat	Ce	Decrease freezing to cue paired with shock	157
Ibotenic acid (pre)	Rat	Ce and Bla	Decrease freezing to context paired with shock	118
Electrolytic (pre)	Rat	L	Decrease freezing to cue paired with shock	154
Radiofrequency (pre)	Rat	AC	Decrease freezing to cues paired with shock; decrease shock probe avoidance	30
Ibotenic acid (amygdala) M (medial geniculate) (pre)	Rat	Ce, MG	Decrease freezing to cue paired with shock (Ce, unilateral; MG, contralateral), some damage to L and Bla	137
NMDA (pre or post)	Rat	Bla	Decrease acquisition of freezing to odor or context or expression of both when lesions made 1 or 15 days after conditioning	61
NMDA (pre and post)	Rat	Bla	Decrease acquisition or expression of freezing to context or cue paired with shock even when lesions made 1 month after training	166
Lidocaine (pre or pretesting or both)	Rat	Ce and Bla	Decrease expression of freezing to context paired with shock; weaker effect when inactivation given before training	119
Muscimol (pre or post)	Rat	Bla	Decrease freezing when given either before testing or training but not immediately after training	183
NMDA (pre or post)	Rat	Bla	Decrease freezing or high-frequency vocalizations to shock or cues paired with shock	101
Ibotenic acid (pre)	Rat	Ce not Bla	Anticonflict (licking), decrease effects of restraint stress on plus maze	175
Electrolytic (pre)	Rat	Ce not Bla	Anticonflict effect	225,256
Quinolinic acid or 6-OHDA (pre)	Rat	Bla	Decrease reduction in licking during conditioned emotional response test	224
NMDA (pre)	Rat	AC	Decrease avoidance of water spout paired with shock but not with quinine	48
Ibotenic acid (pre)	Rat	Ce, not Bla	Decrease disruption of bar pressing to cue paired with shock; could still avoid shock bar	146
Quinolinic acid (pre)	Rat	Bla not Ce	Decrease avoidance of shocked bar, no reduction of disruption of bar pressing to cue paired with shock	146
NMDA (pre or post)	Rat	Bla	Decrease expression or acquisition of fear-potentiated startle to visual conditioned stimulus	214
NMDA (post)	Rat	Bla	Decrease expression of fear-potentiated startle even when lesions made 1 month after training	159
Ibotenic acid (post)	Rat	Ce	Decrease expression of fear-potentiated startle to visual or auditory conditioned stimulus	52
Ibotenic acid (pre)	Rat	Ce and Bla	Decrease expression of hypoalgesia to context paired with shock (formalin test)	118

AC, amygdala complex; Bla, basolateral complex; central nucleus; MG, medial geniculate; post, posttraining; pre, pretraining.

TABLE 64.1. EFFECTS OF LESIONS OF THE AMYGDALA ON CONDITIONED FEAR

Method	Species	Site	Effect of Lesion	Reference
Electrolytic	Rat	Ce	Decrease secretion of ACTH to immobilization stress	23,24
Ibotenic acid	Rat	Ce	Decrease rise corticosterone to immobilization stress	242
Radiofrequency	Rat	**	Decrease the compensatory hypersecretion of ACTH that normally occurs following adrenalectomy	7,8
Electrolytic (pretraining but not posttraining)	Rat	Ce	Decrease secretion of corticosterone and prolactin to shock; no effect on epinephrine, norepinephrine	207
Radiofrequency	Rat	Ce	Decrease ulceration produced by restraint	122,123
Radiofrequency	Rat	Ce	Decrease ulceration produced by shock stress	124
Electrolytic	Rat	Ce	Decrease gastric ulcers to water restraint	59
Ibotenic acid	Rat	Ce	No effect on ulcers to water restraint	59
Electrolytic	Rat	Ce	Decrease noise-elicited hypertension	88
Electrolytic	Wild Rat	CO, Ce	Decrease emotionality in measured in terms of flight and defensive behaviors	144,145
Electrolytic	Rat	CO, Me Ce	Increase the number of contacts a rat will make with a sedated cat	30
Electrolytic	Many species	AC	General taming effect	For review, see 100
Electrolytic	Wild Rat	CO, Ce	Decrease emotionality in measured in terms of flight and defensive behaviors	144,145
Electrolytic	Rat	CO, Me Ce	Increase the number of contacts a rat will make with a sedated cat	30
Radiofrequency	Rat	Ce	Decrease jump withdrawal sign in morphine dependent rats after intraperitoned naloxone	51
Radiofrequency	Rat	Ce, L and Bla	Decrease analgesia produced by exposure to cat or shock (tail flick test)	85
Electrolytic	Rat	L or Ce	Decrease loud noise-induced hypoalgesia (tail fick test)	28
Electrolytic	Rat	Ce, not Bla	Decrease analgesic effects of systemic flumazenil (hot plate test)	106
NMDA	Rat	Ce not Bla or Me	Decrease morphine (low dose) induced antinociception (formalin or tail flick tests)	164,165
Lidocaine	Rat	Ce	Decrease morphine induced antinociception (tail flick test)	165

AC, amygdala complex; Bla, basolateral complex; Ce, central nucleus; Me, medial nucleus.

TABLE 64.2. EFFECTS OF LESIONS OF THE AMYGDALA ON UNCONDITIONED FEAR OR STRESS

Effects of Lesions of the Amygdala in Humans and Nonhuman Primates

In nonhuman primates (160 ,184 ,205) and in humans (9 ,117), cells have been found in the amygdala that respond selectively to faces or direction of gaze (42). In humans, removal of the amygdala has been associated with an impairment of memory for faces (4 ,138 ,240 ,257) and deficits in recognition of emotion in people's faces and interpretation of gaze angle (41 ,50 ,257). In a very rare case involving bilateral calcification confined to the amygdala (Urbach-Wiethe disease), patient SM046 could not identify the emotion of fear in pictures of human faces and could not draw a fearful face, even though other emotions such as happy, sad, angry, and disgusted were identified and drawn within the normal range. The deficit in recognizing facial expressions of fear

only seemed to occur after bilateral amygdala damage (3). This patient and two others also tended to view even the most threatening faces as trustworthy and approachable (2). A more detailed evaluation of patient SM046 showed that she correctly identified valence (e.g., pleasant versus unpleasant) in faces displaying happy, surprised, afraid, angry, disgusted, or sad emotion, but she was highly abnormal in rating the level of arousal to the afraid, angry, disgusted, and sad faces (1). Another patient (SP) with extensive bilateral amygdala damage also showed a major deficit in her ability to rate levels of fear in human faces, yet was perfectly normal in generating a fearful facial expression in comparison with neurologically normal subjects, based on the ratings of three judges (13).

Patients with unilateral (153) or bilateral (26) lesions of the amygdala also have been reported to have deficits in classic fear conditioning using the galvanic skin response as a measure of fear. In monkeys, removal of the amygdala decreases reactivity to sensory stimuli measured with the galvanic skin response (17 ,18).

Effects of Local Infusion of Drugs into the Amygdala on Measures of Fear and Anxiety

If the amygdala is critically involved in fear and anxiety, then drugs that reduce fear or anxiety clinically may well act within the amygdala. It is also probable that certain neurotransmitters within the amygdala may be involved in fear and anxiety. In fact, many studies indicate that local infusions of GABA or GABA agonists, benzodiazepines, CRH antagonists, opiate agonists, neuropeptide Y, dopamine antagonists, or glutamate antagonists decrease measures of fear and anxiety in several animal species. Table 64.3 gives selected examples of some of these studies, which have been extensively reviewed (65). Conversely, local infusions of GABA antagonists, CRH or CRH analogues, vasopressin, thyrotropin-releasing hormone, opiate antagonists, cholecystokinin (CCK) or CCK analogues tend to have anxiogenic effects. Table 64.4 shows selected examples of such studies that also have been reviewed (65).

Substance	Species	Site	Effect of Substance Infused	Reference
GABA or chlordiazepoxide	Rat	Ce	Decrease stress-induced gastric ulcers	232
GABA or benzodiazepines	Rat	Bla	Increase punished responding in operant conflict test (anticonflict effect)	105,126,189, 218,239
Benzodiazepines	Rat	Ce	Increase punished responding in operant conflict test (anticonflict effect)	225,234
Midazolam	Rat	Bla	More time on open arms in plus-maze, no effect on shock probe avoidance	188
Diazepam	Rat	Ce or Bla	Decrease freezing to footshock	120,258
Diazepam	Mice	AC	More time in light side in light-dark box test (anxiolytic effect)	60
Muscimol	Rat	Bla	Anxiolytic effect in the social interaction test; no effect in Ce	216
Muscimol	Rat	Bla	Increase punished responding in operant conflict test (anticonflict effect); no effect in Ce	218
a-CRH	Rat	Ce	Block noise-elicited increase in tryptophan hydroxylase in cortex	35
a-CRH	Rat	Ce	Anxiolytic effect (plus maze) in socially defeated rat	116
a-CRH	Rat	Ce	Anxiolytic effect in plus maze during ethanol withdrawal in ethanol-dependent rats; no effect in plus maze in nondependent rats	194
a-CRH	Rat	Ce	Decrease behavioral effects of opiate withdrawal	115
CRH receptor antisense	Rat	Ce	Anxiolytic effect in the plus maze in rats that previously experienced defeat stress	161
a-CRH	Rat	Ce	Decrease duration of freezing to an initial shock treatment or to re-exposure to shock box 24 h later	233
a-CRH	Rat	Ce	No effect on grooming and exploration activity under stress-free conditions	247
Enkephalin analogue	Rat	Ce	Decrease stress-induced gastric ulcers, prevented by 6-OHDA or clozapine	195,196,198
Opiate agonists	Rb	Ce	Block acquisition of conditioned bradycardia	89,90
Morphine	Rat	Ce	Anxiolytic effect in social interaction test	83
Neuropeptide Y	Rat	Bla, not Ce	Anxiolytic effect in social interaction test, blocked by Y-1 antagonist	213
Neuropeptide Y1 agonist	Rat	Ce	Anxiolytic effects in conflict test; NPY-Y2 agonist much less potent	114
Oxytocin	Rat	Ce	Decrease stress-induced bradycardia and immobility responses	209
SCH 23390	Rat	AC	Decrease expression of fear-potentiated startle	151
SCH 23390	Rat	AC	Decrease acquisition and expression of freezing to tone or context; not due to state-dependent learning	107
CNQX	Rat	Bla	Blocks expression of fear-potentiated startle (visual, auditory CS)	149
NBQX	Rat	Bla or Ce	Blocks expression of fear-potentiated startle (visual CS)	243
AP5	Rat	Bla	Block facilitation of eyeblink conditioning by prior stress when given prior to stressor session	227
AP5 or CNQX	Rat	Bla	Anxiolytic effect in social interaction test	212
CNQX	Rat	Ce	Decrease naloxone precipitated withdrawal signs in morphine-dependent rats	235

AC, amygdala complex; Bla, basolateral complex; Ce, central nucleus.

TABLE 64.3. EFFECTS OF NEUROTRANSMITTER AGONISTS INFUSED INTO THE AMYGDALA ON FEAR AND ANXIETY

Substance	Species	Site	Effect of Substance Infused	Reference
Bicuculline, picrotoxin	Rat	Bla	Anxiogenic effects in the social interaction test; repeated infusion led to sensitization	216
Bicuculline	Rat	Bla	Anxiogenic effects in social interaction, blocked by either NMDA or non-NMDA antagonists into the amygdala	211
Bicuculline (un)	Rat	Bla not Ce	Increases in blood pressure, heart rate, and locomotor activity; bigger effect with repeated infusions	215,216
Bicuculline (un)	Rat	Bla	Increases in blood pressure, heart rate; blocked by infusion of either NMDA or non-NMDA antagonists into the amygdala	211
Bicuculline NMDA, AMPA (un)	Rat	Bla	Increases in blood pressure, heart rate blocked by either NMDA or non-NMDA antagonists infused into Bla or the dorsomedial hypothalamus	229,230
CRH	Rat	Ce	Increase heart rate; effect blocked by a-CRH into Ce	248
CRH, TRH, or CGRP	Rat	Ce	Increase in blood pressure, heart rate, and plasma catecholamines	43
Urocortin or CRH	Rat	Bla	After repeated subthreshold doses get increase in blood pressure to systemic lactate	210
CRH	Rat	Ce, not Bla	Increased grooming and exploration in animals tested under stress-free conditions (i.e., in the home cage)	247,250
CRH	Rat	Ce	Increase defensive burying	249
CRH or Urocortin	Rat	Bla	Anxiogenic effect in plus maze, sensitization with repeated subthreshold doses; now get behavioral and cardiovascular effects to systemic lactate	210
Vasopressin	Rat	Ce	Increased stress-induced bradycardia and immobility responses in rats bred for low rates of avoidance behavior but not the more aggressive rats that show high avoidance rates	209
Vasopressin	Rat	Ce	Bradycardia (low doses) or tachycardia and release of corticosterone (high dose); tachycardia blocked by oxytocin antagonist	208
Vasopressin	Rat	Ce	Immobility, seizures second infusion	251
Vasopressin	Rat	Ce	Immobility in rats bred for low rates of avoidance but not bred for high avoidance rates	209
TRH	Rat	Ce	Increase stress-induced gastric ulcers	196,198
TRH or physostigmine	Rat	Ce	Increase stress-induced gastric ulcers, blocked by muscarinic or benzodiazepine agonists	197
TRH analogue	Rat	Ce	Increase gastric contractility, blocked by vagotomy	182
TRH	Rat	Ce	Produce gastric lesions and stimulated acid secretion	125
TRH analogue	Rat	AC	No effect on gastric secretion, whereas large effect after infusion into dorsal vagal complex or nucleus ambiguus	136
Naloxone	Rat	Ce	Increase stress-induced gastric ulcers	196,198
Naloxone	Rat	AC	Elicit certain signs of withdrawal (depending on site) in morphine-dependent rats (unilateral)	51
Methyl naloxonium	Rat	AC	Place aversion to context where injections given to morphine-dependent rats	231
Methyl naloxonium	Rat	AC	Weak withdrawal signs in morphine-dependent rats	163
Yohimbine	Rat	Ce	Facilitation of the startle reflex	82
CCK analogues	Rat	AC	Anxiogenic effect in plus maze but not clear because significant decrease in overall activity	27
Pentagastrin	Rat	AC	Increase acoustic startle, blocked by CCK B antagonist that also blocked effect of pentagastrin (ICV)	86

AC, amygdala complex; Bla, basolateral complex; Ce, central nucleus; un, unanesthetized.

TABLE 64.4. EFFECTS OF NEUROTRANSMITTER ANTAGONISTS INFUSED INTO THE AMYGDALA ON FEAR AND ANXIETY

In summary, connections between the Bla and the CeA or BNST are critically involved in various autonomic and motor responses seen during a state of fear or anxiety. However, it is also the case that connections between the basolateral nucleus and other target areas are involved in emotional behavior.

Basolateral Nucleus of the Amygdala to Dorsal Striatum Pathway as It Relates to Avoidance or Escape from Aversive Events

As emphasized by McGaugh, Packard, and others, the amygdala modulates memory in various tasks such as inhibitory avoidance and motor or spatial learning (49 ,170 ,171 ,185 ,186). For example, posttraining intracaudate injections of amphetamine enhanced memory in a visible platform water maze task but had no effect in the hidden platform, spatially guided task (185 ,186). Conversely, posttraining intrahippocampal infusion of amphetamine enhanced memory in the hidden platform water maze task but not in the visible platform task. However, posttraining intraamygdala injections of amphetamine enhanced memory in both water maze tasks (185 ,186).

Moreover, preretention intrahippocampal lidocaine injections blocked expression of the memory-enhancing effects of posttraining intrahippocampal amphetamine injections in the hidden platform task, and preretention intracaudate lidocaine injections blocked expression of the memory-enhancing effects of posttraining intracaudate amphetamine injections in the visible platform task. However, preretention intraamygdala lidocaine injections did not block the memory-enhancing effect of posttraining intraamygdala amphetamine injections on either task. Finally, in the hidden platform task, posttraining intrahippocampal, but not intracaudate, lidocaine injections blocked the memory-enhancing effects of posttraining intraamygdala the visible platform task, posttraining intracaudate, but not intrahippocampal, lidocaine injections blocked the memory-enhancing effects of posttraining intraamygdala amphetamine. The findings indicate a double dissociation between the roles of the hippocampus and caudate-putamen in memory and suggest that the amygdala exerts a modulatory influence on both the hippocampal and caudate-putamen memory systems.

Perhaps similarly, lesions of the CeA block freezing but not escape to a tone previously paired with shock, whereas lesions of the basal nucleus of the Bla have just the opposite effect (10). However, lesions of the lateral nucleus, which receive sensory information required by both measures, block both freezing and escape. Lesions of the Bla, but not the CeA, also block avoidance of a bar associated with shock (146). It is possible that basolateral outputs to the dorsal or the ventral striatum mediate the escape behavior, given the importance of the striatum in several measures of escape or avoidance learning. However, combined, unilateral lesions of each structure on opposite sides of the brain would be required to evaluate whether this results from serial transmission from the basolateral nucleus to the striatum.

Basolateral Nucleus of the Amygdala to Hippocampus Pathway as It Relates to Avoidance or Escape from Aversive Events

As mentioned earlier, posttraining intrahippocampal as well as intraamygdala injections of amphetamine selectively enhance memory in a hidden platform water maze task (185 ,186). Posttraining infusion of norepinephrine into the basolateral nucleus enhanced retention in the hidden platform water maze task, whereas posttraining infusion of propranolol had the opposite effect (113). These results suggest that

the amygdala exerts a modulatory influence on hippocampal-dependent memory systems, presumably by direct projections from the basolateral nucleus of the amygdala, perhaps by modulation of LTP in the hippocampus. Lesions (131), NMDA antagonists (132), or local anesthetics (134) infused into the Bla decrease LTP in the dentate gyrus of the hippocampus. Conversely, high-frequency stimulation of the amygdala facilitates induction of LTP in the dentate gyrus (130 ,133). However, combined, unilateral lesions of each structure on opposite sides of the brain would be required to evaluate whether this results from serial transmission from the basolateral nucleus to the hippocampus.

Basolateral Nucleus of the Amygdala to Frontal Cortex Pathway as It Relates to Emotion

Importance of the Bla in US Representation

After pavlovian conditioning, presentation of a conditioned stimulus (CS) elicits some neural representation of the unconditioned stimulus (US) with which it was paired. For example, the sound of a refrigerator door opening or an electric can opener may bring the family cat into the kitchen in expectation of dinner. Several studies suggest that the Bla, perhaps by connections with cortical areas such as the perirhinal cortex (93), is critical for these US representations based on studies using a procedure called *US devaluation*. In these experiments, a neutral stimulus (e.g., a light) is first paired with food so a conditioned response can be measured. Some animals then have the food paired with something that makes them sick (US devaluation). After such treatment, these animals show a reduction in the conditioned response to the light compared with animals that did not experience US devaluation. This result suggests that, after conditioning, animals have a representation of the value of a reinforcement that is elicited by the cue paired with that US. When that representation is changed, then the behavior elicited by the cue also is changed in the same direction. Lesions of the basolateral, but not the CeA, block US devaluation (112). In a related paradigm, rats are trained to be fearful of a weak shock in the presence of a tone. When this is followed by presentation of a stronger shock, without further tone-shock pairing, more freezing occurs to the tone. Temporary inactivation of the Bla during this *inflation* procedure blocks this effect when testing subsequently occurs with a normal, unlesioned, amygdala (5).

Second-order conditioning also depends on a US representation elicited by a CS. In this procedure, cue 1 is paired with a particular US (e.g., shock or food), and cue 2 is paired with cue 1. After such training, cue 2 elicits a similar behavior as that elicited by cue 1, depending on the US with which cue 1 was paired. Thus, it may elicit approach behavior if cue 1 was formerly paired with food and avoidance if cue 1 was paired with shock. This indicates that cue 1 elicits a representation of the US that then becomes associated with cue 2. Lesions of the Bla, but not the CeA, block second-order conditioning (72 ,73 ,112), as do local infusions of NMDA antagonists into the Bla (92).

Importance of the Bla Projection to the Frontal Cortex in Using US Representations to Guide Behavior

Converging evidence now suggests that the connection between the Bla and the prefrontal cortex is critically involved in the way in which a US representation (e.g., very good, pretty good, very bad, pretty bad) guides approach or avoidance behavior. Patients with late- or early-onset lesions of the orbital regions of the prefrontal cortex fail to use important information to guide their actions and decision making (14 ,25 ,63). For example, on a gambling task, they choose high, immediate reward associated with long-term loss rather than low, immediate reward associated with positive long-term gains. They also show severe deficits in social behavior and make poor life decisions.

Studies using single-unit recording techniques in rats indicate that cells in both the Bla and the orbitofrontal cortex fire differentially to an odor, depending on whether the odor predicts a positive (e.g., sucrose) or negative (e.g., quinine) US. These differential responses emerge before the development of consistent approach or avoidance behavior elicited by that odor (220). Many cells in the Bla reverse their firing pattern during reversal training (i.e., the cue that used to predict sucrose now predicts quinine and vice versa) (221), although this has not always been observed (217). In contrast, many fewer cells in the orbitofrontal cortex showed selectivity before the behavioral criterion was reached, and many fewer reversed their selectivity during reversal training (221). These investigators suggest that cells in the Bla encode the associative significance of cues, whereas cells in the orbitofrontal cortex are active when that information, relayed from the Bla, is required to guide choice behavior.

Taken together, these data suggest that the connection between the Bla and the frontal cortex may be involved in determining choice behavior based on how an expected US is represented in memory. The necessity for communication between the amygdala and the frontal cortex was shown in monkeys using a “disconnection approach” in which the amygdala on one side of the brain and the frontal cortex on the other side were lesioned together (22). Because the reciprocal connections between the two structures are ipsilateral, this procedure completely eliminated activity of the network connections while preserving partial function of each structure. Using this approach in rhesus monkeys, Baxter et al. found a decrease in US devaluation after unilateral neurotoxic lesions of the basolateral nucleus in combination with unilateral aspiration of orbital prefrontal cortex (22). These monkeys continued to approach a food on which they had recently been satiated, whereas control monkeys consistently switched to the other food.

Neuroimaging Studies of the Amygdala in Humans

As reviewed by Whalen (244), neuroimaging studies in normal human subjects have shown activation of the amygdala by presentation of biologically relevant sensory stimuli that probably induce strong negative emotional states. For example, the functional magnetic resonance imaging (fMRI) signal intensity within the amygdala is greater when subjects view graphic photographs of negative material (e.g., mutilated human bodies) compared with when they view neutral pictures (135). Positron emission tomography metabolic activity within the amygdala increased to negative material presented by film clips (199), and the amount of amygdala activity during film clips predicted later recall (47). In addition, fMRI signal intensities in humans during classic fear conditioning increased in response to stimuli that predicted an aversive event (45 ,150 ,179).

Amygdala activation also seems to be greater during presentations of fearful faces compared with neutral facial expressions (40 ,180), happy facial expressions (180 ,246), or when subjects looked at a fixation point on an otherwise blank screen (246). Whalen et al. used a backward masking technique in which very brief presentations of fearful and happy facial expressions (33 milliseconds) were followed immediately by presentations (167 milliseconds) of neutral faces (245). Most study subjects reported seeing neutral “expressionless” faces, but not any afraid or smiling faces. Nonetheless, the amygdala demonstrated greater fMRI signal intensity to masked fearful faces compared with masked happy faces. In addition, subjects reported that these masked stimuli did not induce any noticeable changes in their state of emotional arousal. As suggested by Whalen (244), “this study offers preliminary support for the notion that the amygdala constantly monitors the environment for such signals. More than functioning primarily for the production of strong emotional states, the amygdala would be poised to modulate the moment-to-moment vigilance level of an organism.”

Role of the Periaqueductal (Central) Gray

Outputs from the CeA to the ventral central gray appear to mediate several components of the fear response including freezing, conditioned analgesia, and fear-related vocalizations, but, surprisingly, maybe not cessation of operant behavior (11). More dorsal regions of the PAG play a role in active defensive responses (33), depending on whether the threat is distal or proximal (31 ,71). Fanselow (79), for example, showed that dorsal, but not ventral, PAG lesions eliminate activity bursts elicited by foot shock, whereas ventral, but not dorsal, PAG lesions diminish freezing responses elicited by cues previously paired with foot shock. Fanselow suggested that these stimuli (i.e., foot shock versus stimuli predicting foot shock) access different points on a “predatory imminence” continuum in which proximal threats activate the dorsal PAG to generate active defensive behaviors. More distal threats activate the ventral PAG and generate passive or preparatory defensive behaviors such as freezing and analgesia. From a similar perspective, Deakin and Graeff et al. proposed that moderately threatening stimuli inhibit the dorsal PAG (66 ,104), but this inhibition is overcome with more extreme danger, thus allowing active defense or panic behaviors to emerge.

Results from stimulation studies have suggested an anatomic division of function within PAG. In particular, Depaulis and colleagues showed that chemical or electrical stimulation of PAG regions lateral to the aqueduct produces active behaviors such as forward avoidance, defensive aggression, and cardiovascular activation (67 ,68), whereas stimulation of more ventral regions of the PAG elicits passive responses such as behavioral arrest and decreased cardiovascular output (21 ,69). Electrical stimulation of the dorsal PAG in humans produces a pattern of cardiovascular effects that resemble those seen during a natural panic attack, and patients often experience fear, anxiety, and the desire to terminate stimulation (162). Exposure of rats to a cat or high-frequency vocalizations of conspecifics that often signal a predator in the environment increases neuronal firing in the dorsal PAG inferred from an increase in the immediate early gene c-*fos* (162).

Based on these and other data, several investigators have suggested that the dorsal PAG may be involved in panic attacks in humans, perhaps resulting from a dysregulation of various transmitters systems within this structure (162). The dorsal PAG has heavy innervation of the panicogenic peptide CCK, which has been shown to excite the majority of cells in this region. CCK antagonists functionally decrease the effects of electrical stimulation of the dorsal PAG, as does elevating serotonin, perhaps relevant to the use of serotonin reuptake inhibitors in the treatment of panic disorder. Whether these effects depend on connections between the amygdala and the PAG or whether they represent examples where the PAG can function autonomously remains to be determined.

Role of the Hippocampus in Contextual Fear Conditioning

Rats given cue-shock pairings learn to be afraid of the cue as well as the place where cue-shock pairings occurred *(context conditioning)*. In 1992, two seminal articles reported that the hippocampus was necessary for context but not explicit cue conditioning (148 ,190). Both studies found that electrolytic lesions of the dorsal hippocampus blocked freezing in the presence of a fearful context but not in the presence of a cue paired with shock in that context. Kim and Fanselow found such effects when lesions were made shortly after training (1 or 7 days) but not 28 days later (148), whereas Phillips and LeDoux used pretraining lesions (190).

Although the role of the hippocampus in contextual fear conditioning had been discovered earlier using a place aversion measure (223), these articles were more influential because they integrated the well-known role of the hippocampus in spatial learning with a simple, yet powerful measure of classic fear conditioning. Contextual freezing was quickly adopted by investigators interested in the role of hippocampal LTP in learning because contextual fear conditioning was rapid and long lasting, like LTP, and it was easy to measure without complex or expensive equipment. The idea was that the hippocampus was required to form a representation of the context and that this representation was then associated with shock, perhaps in the amygdala. The hippocampus was not needed when an explicit cue, such as a tone, was used because this could be relayed directly to the amygdala without having to be processed by the hippocampus.

As attractive as this hypothesis is, there are problems with concluding that the hippocampus is involved in all forms of contextual conditioning (96 ,97 ,202). Hippocampal lesions often produce substantial behavioral activation, which may interact with the expression of freezing and lead to a disruption of the freezing response itself, rather than of contextual fear. In fact, hippocampal lesions disrupt not only conditioned freezing responses, but also *unconditioned* freezing responses, such as the response elicited by a rat when confronted by a cat (32 ,34 ,147). The finding that hippocampus lesions did not block freezing to an explicit cue makes this competing response interpretation more difficult to accept, but some studies have found that hippocampal lesions disrupt freezing to an explicit cue (167). However, increases in activity cannot account for disruption of contextual freezing by hippocampal lesions in all instances. In an elegantly designed study, Anagnostaras et al. showed that hippocampal lesions disrupted freezing to a context that had been paired with shock shortly before surgery (12). In the same subjects, however, freezing to a second context, that had been paired with shock 28 days preoperatively, was not impaired. Thus, the freezing deficit to the recently conditioned context could not have resulted from an inability to freeze.

Although pretraining electrolytic lesions of the dorsal hippocampus (167) block contextual fear conditioning, neurotoxic lesions fail to do so (97 ,167 ,202), as does local infusion of muscimol (128). To explain this difference, Maren et al. suggested that rats with damage to cells in the hippocampus pick out salient explicit cues in the context and use these as elemental cues for fear conditioning (167). However, these investigators suggested that rats with electrolytic lesions do not do this because the lesion disrupts fibers that connect the ventral subiculum to the nucleus accumbens, which decreases exploration and thus sampling of the context to pick out salient explicit cues to associate with shock. In fact, experiments found a blockade of the acquisition but not the expression of contextual fear conditioning measured by freezing using infusion into the nucleus accumbens of a local anesthetic (108). This effect did not occur using tone-shock pairings, even using a weaker trace conditioning procedure that produced relatively low levels of freezing to the tone.

Another possibility is that fibers from the dorsal to the ventral hippocampus are important in these anterograde amnestic effects of electrolytic lesions of the dorsal hippocampus because neurotoxic lesions of either the entire hippocampus (102 ,202) or just the ventral hippocampus blocked contextual freezing, whereas neurotoxic lesions of the dorsal hippocampus again failed to block contextual conditioning (202). However, in contrast to the hypothesis that contextual fear conditioning involves processes similar to spatial learning, lesions of the ventral hippocampus did not block but instead actually facilitated spatial learning in a water maze task (202). As these investigators concluded, these data directly contradict the “widely held notion that spatial and contextual forms of learning are essentially different manifestations of the same basic underlying process” (202). Because these lesions also impaired freezing to a tone, these authors suggested that the ventral lesions disrupted freezing by increasing activity.

Because all these studies have relied on freezing as the measure of conditioned fear, it is important to assess the effects of hippocampal lesions on other behavioral or autonomic responses associated with fear. If hippocampal lesions disrupted multiple measures of contextual fear, it would provide further support for the hippocampal theory of context conditioning. However, posttraining hippocampal lesions were found not to disrupt context-specific potentiated startle, even though context-elicited freezing was disrupted in the same animals (174). This could not be explained by an excitatory effect of hippocampal lesions on startle amplitude itself (96). In contrast, lesions of the CeA completely blocked both freezing and startle.

In another experimental design (95), lesions of the dorsal hippocampus failed to block a phenomenon called *contextual blocking*, whereby prior contextual fear conditioning retards subsequent cue conditioning. However, as in other studies, freezing to the fearful context was blocked by hippocampal lesions. These data, along with several other reports in the literature (96 ,97 ,202), severely limit the general impression that the hippocampus is required for contextual fear conditioning. However, it does seem to be involved in certain situations, so further work is needed to predict those occasions in which it is and is not involved.

BRAIN SYSTEMS IN THE INHIBITION OR SUPPRESSION OF FEAR AND ANXIETY

Part of "64 - Neural Circuitry of Anxiety and Stress Disorders "

Extinction

The inability to suppress unwanted fear memories or irrational worry is a major problem in many psychiatric disorders, yet very little is known about brain systems involved

in the inhibition of fear. One way to study this important problem is to analyze brain systems involved in *extinction*, defined as a reduction in conditioned fear when the CS is presented many times in the absence of the US. Although such a procedure can decrease the conditioned response, this does not result from an erasure of the original fear memory. Instead, something new is learned that overcomes or competes with the original fear memory. For example, an extinguished conditioned response can return with the passage of time (spontaneous recovery, 187), after a subsequent stressor (reinstatement, 201), or when testing occurs in a different context (renewal, 36). Such results indicate that extinction (but see discussion in ref. 74) may involve a form of active inhibition that is fragile compared with conditioned fear itself.

Conditioned Inhibition

In a conditioned inhibition procedure, cue 1 predicts food or shock, and a compound stimulus (cue 1 plus cue 2) predicts the absence of these USs. There is general agreement that conditioned inhibition, closely related to extinction, does involve active inhibition. In fact, it has been argued that extinction is a special case of conditioned inhibition (38). The *summation test* is the basic method for observing conditioned inhibition (200). In this procedure, the putative conditioned inhibitor (e.g., a light) is presented in compound with an excitatory CS (e.g., a tone). If the combination produces a decrease in the conditioned response below the level observed when the CS is presented alone, then that stimulus is said to act as a conditioned inhibitor. When the conditioned inhibitor is removed, excitation returns to its original level. Various control procedures indicate that a stimulus trained in this way is in fact acting by inhibition.

Because psychotherapy often involves procedures to rid patients of unwanted fear memories, a behavioral analysis of extinction or conditioned inhibition has certain clinical implications, as suggested by Bouton and Swartzentruber (39). As they pointed out, “performance after extinction is inherently unstable” (39). Phenomena such as spontaneous recovery and reinstatement may explain why conditioned fears and phobias in humans sometimes seem to return spontaneously without any obvious cause. The renewal effect may explain why fears reduced successfully in the therapist's office reappear when the patient returns home or to work. If a drug is used as an adjunct to therapy, renewal of fear could occur when the fearful stimulus is encountered in the absence of the drug. In fact, animal experiments show that when benzodiazepines are given during extinction, fear of the CS returns when testing occurs in the absence of the drug (37).

Brain Areas in Extinction or Conditioned Inhibition

Sensory Cortex

Assuming that extinction results from active inhibition (see earlier), one could expect that lesions of various brain areas would disrupt either the development or expression of extinction. LeDoux, Romanski, and Xagoraris reported that rats given ablations of visual cortex before light–foot shock pairings failed to show extinction of lick suppression relative to sham controls over days (156). In a similar study employing heart rate conditioning in the rabbit, Teich et al. showed that although bilateral lesions of either auditory or visual cortex did not disrupt acquisition of fear conditioning to a tone CS, auditory cortex lesions, but not visual cortex lesions, blocked extinction of conditioned heart rate responses to the tone (237). Based on known anatomic connections between sensory cortex and thalamic structures, the authors of both experiments concluded that, during extinction, sensory cortices exert a modality specific inhibition of the thalamic structures important for the performance of conditioned responses.

However, my colleague and I found no effect of complete visual cortex lesions on extinction of fear-potentiated startle using a visual CS when the lesions were made either before light-shock pairings or after light-shock pairings and extinction (77). Although there were procedural differences between these reports, the conclusion that sensory cortex is universally involved in extinction of conditioned fear is not supported.

Frontal Cortex

Rats with lesions of the ventral medial prefrontal cortex made before fear conditioning required more days to reach an extinction criterion using an auditory CS and freezing as the measure of fear (178). However, in these same animals, extinction of conditioned fear to contextual cues was not impaired. In an extensive series of experiments, my colleagues and I found normal rates of extinction to both explicit and contextual cues after total removal of the ventral medial prefrontal cortex using both freezing and fear-potentiated startle as measures of conditioned fear (94). Because the lesions in the study by Morgan et al. were performed before fear conditioning (178), the apparent blockade of extinction after ventral medial prefrontal cortex lesions may have resulted from an increase in the strength of original fear conditioning, rather than from interfering with the process of extinction. Although the lesions and shams groups did not differ significantly in their level of freezing before the extinction sessions, freezing to explicit cues often becomes maximal after a very few training trials, so “ceiling effects” may well have been operating. Because extinction rate can be a more sensitive index of the strength of original

conditioning than the terminal level of performance before the initiation of extinction (15), the slower rate of extinction in the lesioned animals may have reflected a stronger degree of original learning. Although these authors do not believe their effects can be explained in this way (177), the finding that the lesions had no effect on the rate of extinction of context conditioning, which clearly was not at the ceiling of the freezing scale, is consistent with this interpretation. Similarly, we did not find any effect of pretraining ventral prefrontal cortex lesions on extinction of contextual fear-potentiated startle or freezing (94). In addition, we did not find any effect of ventral medial prefrontal cortex lesions on extinction when lesions were made after fear conditioning but before extinction (94). Morgan and LeDoux also found no effect on the rate of extinction when ventral prefrontal cortex lesions were made after fear conditioning, but before extinction (176). If the frontal cortex is required for the development of extinction or for the inhibition of fear after extinction, one would expect lesions to block the development of extinction, irrespective of whether the lesions were made before or after the initial phase of fear conditioning.

Similarly complex effects on extinction have been reported regarding depletion of dopamine in the prefrontal cortex (181). Preconditioning lesions of dopamine terminals in the medial prefrontal cortex retarded the rate of extinction when a 0.8-mA shock was used but not when a 0.4-mA shock was used. Inspection of the results strongly suggests that the 0.8-mA group was at the ceiling of the measurement scale at the beginning of the extinction session, whereas the 0.4-mA group was not. Conversely, 6-hydroxydopamine lesions of the frontal cortex substantially retarded extinction after 0.8-mA tone-shock pairings, even when the lesions were made after training (181). Thus, it is possible that dopamine levels in the prefrontal cortex are important for extinction when conditioning has produced high, but not more moderate, levels of fear, although further studies using posttraining lesions are required to verify this.

Quirk et al. found that pretraining lesions of the ventral medial prefrontal cortex did not block the development of conditioned freezing or the rate of within session extinction (191). However, the lesioned rats showed much more spontaneous recovery measured 24 hours later. Similar results were found in rats given systemic administration of an NMDA antagonist (193). In contrast, we found no change in the rate or final level of extinction, measured with fear-potentiated startle and freezing, when extinction was assessed over 18 daily sessions using a small number of CS presentations each day (94). There also were no differences in the degree of spontaneous recovery measured 5 days later or in shock-induced reinstatement measured 24 hours after a single foot shock. Hence, the findings of Quirk et al. may depend critically on the use of a relatively small amount extinction training (191 ,193). Moreover, their lesions were generally more ventral than ours, and this may have contributed to the differences. Clearly, more work needs to be done to examine the limits of the role of the prefrontal cortex in extinction of conditioned fear, given the clinical importance of these data.

Hippocampus

Although a complete review of the hippocampal literature on extinction is beyond the scope of this chapter, this brain area has received a great deal of experimental attention and was once widely believed to be involved in extinction. Theories of extinction confront the problem of designing a mechanism capable of discriminating occasions of reinforcement from nonreinforcement. Douglas suggested that the hippocampus is a nonreinforcement detector providing the organism with the means to “tune out” information that is of no motivational consequence (70). It is possible that the hippocampus recognizes that the CS is no longer followed by the US and inhibits relevant sensory or conditioned response production centers.

Various conditioning paradigms have been used to assess the role of the hippocampus in extinction, including the rabbit nictitating membrane response (29 ,219 ,228), conditioned heart rate (44), and conditioned suppression (152). Although some of these experiments have found that hippocampal lesions attenuate extinction (219), others have found no effect (228), and still another has shown facilitated extinction (152).

Because extinction is context specific (see earlier), one could expect that lesions of the hippocampus would disrupt this contextual control of extinction. However, direct tests of this hypothesis have not found a disruption of context specific extinction using pretraining lesions. Hence, neither fimbria-fornix lesions (252) nor excitotoxic lesions of the entire hippocampus (87) had any effect on the rate of extinction or on renewal of conditioned fear, although both types of lesions disrupted reinstatement. In contrast, large hippocampal lesions were not found to disrupt reinstatement of appetitively conditioned behavior (84). Overall, therefore, the role of the hippocampus in extinction remains uncertain.

Neurotransmitters in Extinction

Role of NMDA Receptors in Extinction

Local infusion of NMDA antagonists into the amygdala blocked the development of extinction measured with fear-potentiated startle (76). Intraamygdala infusion of AP5 also blocked extinction using an auditory CS and freezing as a measure of conditioned fear (158). Perhaps similarly, systemic administration of the NMDA antagonist MK-801 fully blocked extinction of conditioned analgesia, a reliable

measure of conditioned fear (62), and this effect could not be explained by state-dependent context extinction. A similar blockade of extinction was reported using a lick-suppression paradigm (19), as well as extinction of the rabbit nictitating membrane preparation (143). Taken together, these data indicate that NMDA antagonists can block the development of extinction measured on subsequent test sessions. This may even occur under conditions in which the antagonist does not block the development of short-term extinction. Thus, systemic injection of the NMDA antagonist CPP before extinction blocked the expression of conditioned freezing by about 40% but did not block the development of extinction. However, the CPP group showed substantial recovery of conditioned freezing measured 24 hours later, a finding suggesting that CPP blocked the long-term development of extinction. (193). Interestingly, these investigators found a similar effect with preconditioning lesions of the ventral prefrontal cortex (191), although the connection between these two sets of data remains to be made.

Role of GABA in Extinction

Several studies have suggested that GABA agonists given before nonreinforced CS presentations interfere with the development of extinction (37 ,74). However, many of these effects may be attributable to state-dependent learning rather than to a blockade of learning during nonreinforced CS exposure. For example, Bouton, using lick suppression as a measure of fear, showed a blockade of extinction when rats were given chordiazepoxide during nonreinforced CS presentations and were then tested in the absence of the drug (37). However, when chordiazepoxide also was given before testing, extinction was still evident. This finding suggests that the benzodiazepine did not actually block the learning that was occurring during extinction, but, instead, the change in drug state between extinction and testing was the factor that produced renewal (36).

Interestingly, a series of experiments by Harris and Westbrook found similar effects on excitatory conditioning. For example, rats given fear conditioning after injection with benzodiazepines showed an impairment in conditioned freezing measured 24 hours later in the same context compared with rats trained under the drug but tested in a different context (111) or rats given a stressor before testing (109). Thus, the benzodiazepines did not actually prevent original learning, but instead produced a state during conditioning that interfered with retrieval during testing.

Hence, it would seem that GABA agonists do not directly interfere with either excitatory or inhibitory learning, but, instead, act on processes that are important for retrieval of prior learning. However, if extinction is a form of active inhibition, it is possible that GABA may be one of the neurotransmitters necessary for the expression of extinction. In fact, in an elegant set of experiments, Harris and Westbrook provided evidence that extinction is mediated by GABA release (110). Pretraining or pretesting administration of the inverse agonist FG 7142, which decreases GABA transmission, blocked both development and expression of extinction to an auditory CS paired with foot shock, using freezing as a measure. This effect could not be ascribed to state dependency or to a ceiling effect. Pretest administration of FG 7142 reinstated freezing when assessed in the context where extinction took place, but not in a novel context, which itself reinstated freezing, and the two effects were not additive statistically. However, the disruption of extinction by FG 7142 was not complete, a finding leaving open the possibility that other mechanisms and neurotransmitters also may be involved.

Role of Adrenocorticotropic Hormone and Vasopressin in Extinction

Work by DeWied, Van Wiersima, Izquierdo, and Richardson and their co-workers indicated that administration of various peptides such as adrenocorticotropic hormone or vasopressin either before or after extinction training attenuates subsequent extinction performance (for review see ref. 74).

Neural Systems in Conditioned Inhibition

Using fluorodeoxyglucose autoradiography to measure neural activity, Mcintosh and Gonzalez-Lima compared region-specific activity in parallel auditory pathways in two groups presented with a tone-light compound (172). In both groups, the tone was a fear-eliciting CS. The groups differed with respect to the significance of the light, which had been trained as a conditioned inhibitor in one group and as a neutral stimulus in the other. Thus, the experiment allowed for an analysis of whether a conditioned inhibitor modulates activity in sensory areas normally activated by an auditory fear-eliciting CS. Interestingly, in only one area, the ventral medial geniculate nucleus, was there a significant difference in activation between the two groups. This structure showed less activation in the conditioned inhibition group, a finding suggesting that a conditioned inhibitor may act at this locus in the auditory pathway to inhibit conditioned fear normally produced by an auditory CS.

We found normal conditioned inhibition of fear-potentiated startle, using a visual excitatory stimulus and an auditory conditioned inhibitor after lesions of either the medial prefrontal cortex (94) or the nucleus accumbens (75). In addition, local infusion into the nucleus accumbens of either amphetamine or glutamate antagonists did not alter the magnitude of conditioned inhibition, as they alter responding to conditioned reinforcers trained in an operant situation (46 ,236).

In an appetitive learning situation, Holland et al. reported that lesions of the hippocampus appeared to block

feature negative conditional discrimination (127), a phenomenon closely related to conditioned inhibition.

Several studies suggested that the lateral septum may play an important role in conditioned inhibition. Using pavlovian discriminative fear conditioning, single-unit firing rates in the dorsal lateral septal nucleus increased in the presence of a conditioned inhibitor and decreased in the presence of a conditioned excitor (238 ,254). This finding was not seen when recordings were made in the medial septal nucleus (253). More recently, Yadin and Thomas reported that stimulation of the same area of dorsolateral septal nucleus inhibited restraint stress-induced ulcers (255). Using c-*fos* mRNA as a measured of neuronal activation, we found a unique increase in c-*fos* in a ventral part of the lateral septum, the so-called septohypothalamic nucleus, when a conditioned inhibitor of fear was presented (53). Curiously, however, lesions of the lateral septal nucleus did not block the expression of conditioned inhibition in preliminary pilot studies, although further work certainly is required to evaluate the role of the lateral septum, perhaps using acute inactivation techniques rather than lesions.

One study suggests that the dorsal central gray may play an important role in conditioned inhibition of fear. Fendt reported that posttest infusions of 5 ng of picrotoxin (a GABA chloride channel blocker) into the dorsal central gray, but not the lateral or ventrolateral central gray, reduced the expression of conditioned inhibition without affecting the expression of conditioned fear (81). Although this result is complicated by the finding that neither 2.5-ng doses nor 10-ng doses affected conditioned inhibition, it raises the intriguing possibility that a conditioned inhibitor of fear releases GABA into the dorsal central gray. Alternatively, because low doses of picrotoxin would be expected to activate the dorsal central gray by removing tonic inhibition, these results could be interpreted as indicating that the dorsal central gray is involved in inhibiting an unknown brain structure mediating conditioned inhibition (81). Given the prominent role of the central gray in the expression of fear (162), more work is needed to investigate the role of the dorsal central gray in conditioned inhibition of fear.

REFERENCES

1. Adolphs R, Russell JA, Tranel D. A role for the human amygdala in recognizing emotion arousal from unpleasant stimuli. *Psychol Sci* 1999;10:167–171.

2. Adolphs R, Trane D, Damasio AR. The human amygdala in social judgment. *Nature* 1998;393:470–474.

3. Adolphs R, Trane D, Damasio H, et al. Fear and the human amygdala. *J Neurosci* 1995;15:5879–5891.

4. Aggleton JP. The functional effects of amygdala lesions in humans: a comparison with findings from monkeys. In: Aggleton JP, ed. *The amygdala: neurobiological aspects of emotion, memory and mental dysfunction*. New York: Wiley-Liss, 1992:485–503.

5. Akbari Y, Mongeau R, Maren S, et al. Reversible inactivation of the basolateral amygdala prevents inflation of fear conditioning in rats. Society for Neuroscience, New Orleans, 1997.

6. Alheid G, deOlmos JS, Beltramino CA. Amygdala and extended amygdala. In: Paxinos G, ed. *The rat nervous system*. New York: Academic Press, 1995:495–578.

7. Allen JP, Allen CF. Role of the amygdaloid complexes in the stress-induced release of ACTH in the rat. *Neuroendocrinology* 1974;15:220–230.

8. Allen JP, Allen CF. Amygdalar participation in tonic ACTH secretion in the rat. *Neuroendocrinology* 1975;19:115–125.

9. Allison T, McCarthy G, Nobre A, et al. Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cereb Cortex* 1994;4:544–554.

10. Amorapanth P, LeDoux JE, Nader K. Different lateral amygdala outputs mediate reactions and actions elicited by a fear-arousing stimulus. *Nat Neurosci* 2000;3:74–79.

11. Amorapanth P, Nader K, LeDoux JE. Lesions of periaqueductal gray dissociate-conditioned freezing from conditioned suppression behavior in rats. *Learn Mem* 1999;6:491–499.

12. Anagnostaras SG, Maren S, Fanselow MS. Temporally graded retrograde amnesia of contextual fear after hippocampal damage in rats: within-subjects examination. *J Neurosci* 1999;19:1106–1114.

13. Anderson AK, Phelps EA. Expression without recognition: contributions of the human amygdala to emotional communication. *Psychol Sci* 2000;11:106–111.

14. Anderson SW, Bechara A, Damasio H, et al. Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nat Neurosci* 1999;2:1032–1037.

15. Annau Z, Kamin LJ. The conditioned emotional response as a function of US intensity. *J Comp Physiol Psychol* 1961;54:428–432.

16. Applegate CD, Kapp BS, Underwood MD, et al. Autonomic and somatomotor effects of amygdala central n. stimulation in awake rabbits. *Physiol Behav* 1983;31:353–360.

17. Bagshaw MH, Benzies S. Multiple measures of the orienting reaction and their dissociation after amygdalectomy in monkeys. *Exp Neurol* 1968;20:175–187.

18. Bagshaw MH, Kimble DP, Pribram KH. The GSR of monkeys during orienting and habituation and after ablation of the amygdala, hippocampus and inferotemporal cortex. *Neuropsychologia* 1965;3:111–119.

19. Baker JD, Azorlosa JL. The NMDA antagonist MK-801 blocks the extinction of pavlovian fear conditioning. *Behav Neurosci* 1996;110:618–620.

20. Bandler R, Carrive P. Integrated defence reaction elicted by excitatory amino acid microinjection in the midbrain periaqueductal grey region of the unrestrained cat. *Brain Res* 1988;439:95–106.

21. Bandler R, Shipley MT. Columnar organization in the midbrain periaqueductal gray: modules for emotional expression? *Trends Neurosci* 1994;17:379–389.

22. Baxter MG, Parker A, Lindner CCC, et al. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J Neurosci* 2000;20:4311–4319.

23. Beaulieu S, DiPaolo T, Barden N. Control of ACTH secretion by central nucleus of the amygdala: implication of the serotonergic system and its relevance to the glucocorticoid delayed negative feed-back mechanism. *Neuroendocrinology* 1986;44:247–254.

24. Beaulieu S, DiPaolo T, Cote J, et al. Participation of the central amygdaloid nucleus in the response of adrenocorticotropin secretion to immobilization stress: opposing roles of the noradrenergic and dopaminergic systems. *Neuroendocrinology* 1987;45:37–46.

25. Bechara A, Damasio H, Tranel D, et al. Deciding advantageously before knowing the advantageous strategy. *Science* 1997;275:1293–1294.

26. Bechara A, Tranel D, Damasio H, et al. Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* 1995;269:1115–1118.

27. Belcheva I, Belcheva S, Petkov VV, et al. Asymmetry in behavioral responses to cholecystokinin microinjected into rat nucleus accumbens and amygdala. *Neuropharmacology* 1994;33:995–1002.

28. Bellgowan PSF, Helmstetter FJ. Neural systems for the expression of hypoalgesia during nonassociative fear. *Behav Neurosci* 1996;110:727–736.

29. Berger TW, Weikert CL, Basset JL, et al. Lesions of the retrosplinal cortex produce deficits in reversal learning of the rabbit nictitating membrane response: implications for potential interactions between hippocampal and cerebellar brain systems. *Behav Neurosci* 1986;100:802–809.

30. Blanchard DC, Blanchard RJ. Innate and conditioned reactions to threat in rats with amygdaloid lesions. *J Comp PhysioI Psychol* 1972;81:281–290.

31. Blanchard DC, Blanchard RJ. Ethoexperimental approaches to the biology of emotion. *Annu Rev Psychol* 1988;39:43–68.

32. Blanchard DC, Blanchard RJ, Lee MC, et al. Movement arrest and the hippocampus. *PhysioI Psychol* 1977;5:312–324.

33. Blanchard DC, Williams G, Lee EMC, et al. Taming of wild *Rattus norvegicus* by lesions of the mesencephalic central gray. *PhysioI Psychol* 1981;9:157–163.

34. Blanchard RJ, Blanchard DC. The effects of hippocampal lesions on the rat's reaction to a cat. *J Comp PhysioI Psychol* 1972;78:77–82.

35. Boadle-Biber MC, Singh VB, Corley KC, et al. Evidence that corticotropin-releasing factor within the extended amygdala mediates the activation of tryptophan hydroxylase produced by sound stress in the rat. *Brain Res* 1993;628:105–114.

36. Bouton ME, Bolles RC. Contextual control of the extinction of conditioned fear. *Learn Motiv* 1979;10:455–466.

37. Bouton ME, Kenney FA, Rosengard C. State-dependent fear extinction with two benzodiazepine tranquilizers. *Behav Neurosci* 1990;104:44–55.

38. Bouton ME, Nelson JB. Context specificity of target versus feature inhibition in a feature-negative discrimination. *J Exp Psychol Anim Behav Process* 1994;20:51–65.

39. Bouton ME, Swartzentruber D. Sources of relapse after extinction in pavlovian instrumental learning. *Clin Psychol Rev* 1991;11:123–140.

40. Breiter HC, Etcoff NL, Whalen PJ, et al. Response and habituation of the human amygdala during vissual processing of facial expression. *Neuron* 1996;17:875–887.

41. Broks P, Young AW, Maratos EJ, et al. Face processing impairments after encephalitis: amygdala damage and recognition of fear. *Neuropsychologia* 1998;36:59–70.

42. Brothers L, Ring B. Mesial temporal neurons in the macaque monkey with responses selective for aspects of social stimuli. *Behav Brain Res* 1993;57:53–61.

43. Brown MR, Gray TS. Peptide injections into the amygdala of conscious rats: effects on blood pressure, heart rate and plasma catecholamines. *Regul Pept* 1988;21:95–106.

44. Buchanan SL, Powell DA. Divergencies in pavlovian conditioned heart rate and eyeblink responses produced by hippocampectomy in the rabbit *(Oryctolagus cuniculus). Behav Neural BioI* 1980;30:20–38.

45. Buchel C, Morris J, Dolan RJ, et al. Brain systems mediating aversive conditioning: an event-related fMRI study. *Neuron* 1998;20:947–957.

46. Burns LH, Everitt BJ, Kelly AE, et al. Glutamate-dopamine interactions in the ventral striatum: role in locomotor activity and responding with conditioned reinforcement. *Psychopharmacology* 1994;115:516–528.

47. Cahill L, Haier RJ, Fallon J, et al. Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proc Natl Acad Sci USA* 1996;93:8016–8021.

48. Cahill L, McGaugh JL. Amygdaloid complex lesions differentially affect retention of tasks using appetitive and aversive reinforcement. *Behav Neurosci* 1990;104:532–543.

49. Cahill L, McGaugh JL. Mechanisms of emotional arousal and lasting declarative memory. *Trends Neurosci* 1998;21:294–299.

50. Calder AJ, Young AW, Rowland D, et al. Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. *Cogn Neuropsychol* 1996;13:699–745.

51. Calvino B, Lagowska J, Ben-Ari Y. Morphine withdrawal syndrome: differential participation of structures located within the amygdaloid complex and striatum of the rat. *Brain Res* 1979;177:19–34.

52. Campeau S, Davis M. Involvement of the central nucleus and basolateral complex of the amygdala in fear conditioning measured with fear-potentiated startle in rats trained concurrently with auditory and visual conditioned stimuli. *J Neurosci* 1995;15:2301–2311.

53. Campeau S, Falls WA, Cullinan WE, et al. Elicitation and reduction of fear: behavioral and neuroendocrine indices and brain induction of the immediate-early gene *c-fos. Neuroscience* 1997;78:1087–1104.

54. Canteras N, Swanson L. Projections of the ventral subiculum to the amygdala, septum, and hypothalamus: a PHAL anterograde tract-tracing study in the rat. *J Comp Neurol* 1992;324:180–194.

55. Chapman PF, Bellavance LL. Induction of long-term potentiation in the basolateral amygdala does not depend on NMDA receptor activation. *Synapse* 1992;11:310–318.

56. Chapman PF, Kairiss EW, Keenan CL, et al. Long-term synaptic potentiation in the amygdala. *Synapse* 1990;6:271–278.

57. Chapman WP, Schroeder HR, Guyer G, et al. Physiological evidence concerning the importance of the amygdaloid nuclear region in the integration of circulating function and emotion in man. *Science* 1954;129:949–950.

58. Clugnet MC, LeDoux JE. Synaptic plasticity in fear conditioning circuits: induction of LTP in the lateral nucleus of the amygdala by stimulation of the medial geniculate body. *J Neurosci* 1990;10:2818–2824.

59. Coover GD, Murison R, Jellestad FK. Subtotal lesions of the amygdala: the rostral central nucleus in passive avoidance and ulceration. *PhysioI Behav* 1992;51:795–803.

60. Costall B, Kelly ME, Naylor RJ, et al. Neuroanatomical sites of action of 5-HT3 receptor agonist and antagonists for alteration of aversive behaviour in the mouse. *Br J PharmacoI* 1989;96:325–332.

61. Cousens G, Otto T. Both pre- and posttraining excitotoxic lesions of the basolateral amygdala abolish the expression of olfactory and contextual fear conditioning. *Behav Neurosci* 1998;112:1092–1103.

62. Cox J, Westbrook RF. The NMDA receptor antagonist MK-801 blocks acquisition and extinction of conditioned hypoalgesia responses in the rat. *Q J Exp Psychol* 1994;47B:187–210.

63. Damasio AR. *Descartes' error*. New York: Grosset/Putnam, 1994.

64. Davis M. Neurobiology of fear responses: the role of the amygdala. *J Neuropsychiatry Clin Neurosci* 1997;9:382–402.

65. Davis M, Shi C-J. The amygdala. *Curr BioI* 2000;10:R131.

66. Deakin JWF, Graeff FG. 5-HT and mechanisms of defence. *J Psychopharmacol* 1991;5:305–315.

67. Depaulis A, Bandler R, Vergnes M. Characterization of pretentorial periaqueductal gray matter neurons mediating intraspecific defensive behaviors in the rat by microinjections of kainic acid. *Brain Res* 1989;486:121–132.

68. Depaulis A, Keay KA, Bandler R. Longitudinal neuronal organization of defensive reactions in the midbrain periaqueductal gray region of the rat. *Exp Brain Res* 1992;90:307–318.

69. Depaulis A, Keay KA, Bandler R. Quiescence and hyporeactivity evoked by activation of cell bodies in the ventrolateral midbrain periaqueductal gray of the rat. *Exp Brain Res* 1994;99:75–83.

70. Douglas RJ. *Inhibition and learning: pavlovian conditioning in the brain*. London: Academic Press, 1972.

71. Edmunds M. *Defense in animals*. Essex, UK: Longmans, 1974.

72. Everitt BJ, Cador M, Robbins TW. Interactions between the amygdala and ventral striatum in stimulus-reward associations: studies using a second-order schedule of sexual reinforcement. *Neuroscience* 1989;30:63–75.

73. Everitt BJ, Morris KA, O'Brien A, et al. The basolateral amygdala-ventral striatal system and conditioned place preference: further evidence of limbic-striatal interactions underlying reward-related processes. *Neuroscience* 1991;42:1–18.

74. Falls WA, Davis M. Behavioral and physiological analysis of fear inhibition. In: Friedman MJ, Charney DS, Deutch AY, eds. *Neurobiological and clinical consequences of stress: from normal adaptation to PTSD*. Philadelphia: Lippincott–Raven, 1995:177–202.

75. Falls WA, Josselyn SA, Gewirtz JC, et al. The nucleus accumbens if not critical for condtioned inhibition of fear as measured with fear-potentiated startle. *Soc Neurosci Abstr* 1998;28.

76. Falls WA, Miserendino MJ D, Davis M. Extinction of fear-potentiated startle: blockade by infusion of an NMDA antagonist into the amygdala. *J Neurosci* 1992;12:854–863.

77. Falls WF, Davis M. Visual cortex ablations do not prevent extinction of fear-potentiated startle using a visual conditioned stimulus. *Behav Neural Biol* 1994;60:259–270.

78. Fanselow MS. The midbrain periaqueductal gray as a coordinator of action in response to fear and anxiety. In: Depaulis A, Bandler R, eds. *The midbrain periaqueductal gray matter: functional, anatomical and neurochemical organization*. New York: Plenum, 1991:151–173.

79. Fanselow MS. Neural organization of the defensive behavior system responsible for fear. *Psychonom Bull Rev* 1994;1:429–438.

80. Feldman S, Weidenfeld J. The excitatory effects of the amygdala on hypothalamo-pituitary-adrenocortical responses are mediated by hypothalamic norepinephrine, serotonin, and CRF-41. *Brain Res Bull* 1998;45:389–393.

81. Fendt M. Different regions of the periaqueductal grey are involved differently in the expression and conditioned inhibition of fear-potentiated startle. *Eur J Neurosci* 1998;10:3876–3884.

82. Fendt M, Koch M, Schnitzler HU. Amygdaloid noradrenaline is involved in the sensitization of the acoustic startle response in rats. *Pharmacol Biochem Behav* 1994;48:307–314.

83. File SE, Rodgers RJ. Partial anxiolytic actions of morphine sulphate following microinjection into the central nucleus of the amygdala in rats. *Pharmacol Biochem Behav* 1979;11:313–318.

84. Fox GD, Holland PC. Neurotoxic hippocampal lesions fail to impair reinstatement of an appetitively conditioned response. *Behav Neurosci* 1998;112:255–260.

85. Fox RJ, Sorenson CA. Bilateral lesions of the amygdala attenuate analgesia induced by diverse environmental challenges. *Brain Res* 1994;648:215–221.

86. Frankland PW, Josselyn SA, Bradwejn J, et al. Activation of amygdala cholecystokinin B receptors potentiates the acoustic startle response in the rat. *J Neurosci* 1997;17:1838–1847.

87. Frohardt R, Guarraci FA, Bouton ME. The effects of neurotoxic hippocampal lesions on two effects of context following fear extinction. *Behav Neurosci* 2000;114:227–240.

88. Galeno TM, VanHoesen GW, Brody MJ. Central amygdaloid nucleus lesion attenuates exaggerated hemodynamic responses to noise stress in the spontaneously hypertensive rat. *Brain Res* 1984;291:249–259.

89. Gallagher M, Kapp BS, McNall CL, et al. Opiate effects in the amygdala central nucleus on heart rate conditioning in rabbits. *Pharmacol Biochem Behav* 1981;14:497–505.

90. Gallagher M, Kapp BS, Pascoe JP. Enkephalin analogue effects in the amygdala central nucleus on conditioned heart rate. *Pharmacol Biochem Behav* 1982;17:217–222.

91. Gean PW, Chang FC, Huang CC, et al. Long-term enhancement of EPSP and NMDA receptor-mediated synaptic transmission in the amygdala. *Brain Res Bull* 1993;31:7–11.

92. Gewirtz J, Davis M. Second order fear conditioning prevented by blocking NMDA receptors in the amygdala. *Nature* 1997;388:471–474.

93. Gewirtz JC, Davis M. Application of pavlovian higher-order conditioning to the analysis of the neural substrates of learning and memory. *Neuropharmacology* 1998;37:453–460.

94. Gewirtz JC, Falls WA, Davis M. Normal conditioned inhibition and extinction of freezing and fear potentiated startle following electrolytic lesions of medial prefrontal cortex. *Behav Neurosci* 1997;111:712–726.

95. Gewirtz JC, McNish KA, Davis M. Disruption of contextual freezing, but not contextual “blocking” of fear-potentiated startle after lesions of the dorsal hippocampus. *Behav Neurosci* 2000;114:64–76.

96. Gewirtz JC, McNish KA, Davis M. Is the hippocampus necessary for contextual fear conditioning? *Behav Brain Res* 2000;110:83–95.

97. Gisquet-Verrier P, Dutrieux G, Richer P, et al. Effects of lesions to the hippocampus on contextual fear: evidence for a disruption of freezing and avoidance behavior but not context conditioning. *Behav Neurosci* 1999;113:507–522.

98. Gloor P. Role of the amygdala in temporal lobe epilepsy. In: Aggleton JP, ed. *The amygdala: neurobiological aspects of emotion, memory and mental dysfunction*. New York: Wiley–Liss, 1992:505–538.

99. Gloor P, Olivier A, Quesney LF. The role of the amygdala in the expression of psychic phenomena in temporal lobe seizures. In: Ben-Ari Y, ed. *The amygdaloid complex*. New York: Elsevier/North Holland, 1981:489–507.

100. Goddard GV. Functions of the amygdala. *Psychol Bull* 1964;62:89–109.

101. Goldstein LE, Rasmusson AM, Bunney BS, et al. Role of the amygdala in the coordination of behavioral, neuroendocrine, and prefrontal cortical monoamine responses to psychological stress in the rat. *J Neurosci* 1996;16:4787–4798.

102. Good M, Honey RC. Dissociable effects of selective lesions to hippocampal subsytems on exploratory behavior, contextual learning and spatial learning. *Behav Neurosci* 1997;111:487–493.

103. Graeff FG. Animal models of aversion. In: Simon P, Soubrie P, Wildlocher D, eds. *Selected models of anxiety, depression and psychosis*. vol 1. Basel: Karger, 1988:115–141.

104. Graeff FG, Silveira MCL, Nogueira RL, et al. Role of the amygdala and periaqueductal gray in anxiety and panic. *Behav Brain Res* 1993;58:123–131.

105. Green S, Vale AL. Role of amygdaloid nuclei in the anxiolytic effects of benzodiazepines in rats. *Behav Pharmacol* 1992;3:261–264.

106. Grijalva CV, Levin ED, Morgan M, et al. Contrasting effects of centromedial and basolateral amygdaloid lesions on stress-related responses in the rat. *Physiol Behav* 1990;48:495–500.

107. Guarraci FA, Frohardt RJ, Kapp BS. Amygdaloid D1 dopamine receptor involvement in pavlovian fear conditioning. *Brain Res* 1999;827:28–40.

108. Haralambous T, Westbrook RF. An infusion of bupivacaine into the nucleus accumbens disrupts the acquisition but not the expression of contextual fear conditioning. *Behav Neurosci* 1999;113:925–940.

109. Harris JA, Westbrook RF. Benzodiazepine-induced amnesia in rats: reinstatement of conditioned performance by noxious stimulation on test. *Behav Neurosci* 1998;112:183–192.

110. Harris JA, Westbrook RF. Evidence that GABA transmission mediates context-specific extinction of learned fear. *Psychopharmacology* 1998;140:105–115.

111. Harris JA, Westbrook RF. The benzodiazepine midazolam does not impair pavlovian fear conditioning but regulates when and where fear is expressed. *J Exp Psychol Anim Behav Process* 1999;25:236–246.

112. Hatfield T, Han J-S, Conley M, et al. Neurotoxic lesions of basolateral, but not central, amygdala interfere with pavlovian second-order conditioning and reinforcer devaluation effects. *J Neurosci* 1996;16:5256–5265.

113. Hatfield T, McGaugh JL. Norepinephrine infused into the basolateral amygdala posttraining enhances retention in the spatial water maze task. *Neurobiol Learn Mem* 1999;71:232–239.

114. Heilig M, McLeod S, Brot M, et al. Anxiolytic-like action of neuropeptide Y: mediation by Y1 receptors in amygdala, and dissociation from food intake effects. *Neuropsychopharmacology* 1993;8:357–363.

115. Heinrichs SC, Menzaghi F, Schulteis G, et al. Suppression of corticotropoin-releasing factor in the amygdala attenuates aversive consequences of morphine withdrawal. *Behav Pharmacol* 1995;6:74–80.

116. Heinrichs SC, Pich EM, Miczek KA, et al. Corticotropin-releasing factor antagonist reduces emotionality in socially defeated rats via direct neurotropic action. *Brain Res* 1992;581:190–197.

117. Heit G, Smith ME, Halgren E. Neural encoding of individual words and faces by the human hippocampus and amygdala. *Nature* 1988;333:773–775.

118. Helmstetter FJ. The amygdala is essential for the expression of conditioned hypoalgesia. *Behav Neurosci* 1992;106:518–528.

119. Helmstetter FJ. Contribution of the amygdala to learning and performance of conditional fear. *Physiol Behav* 1992;51:1271–1276.

120. Helmstetter FJ. Stress-induced hypoalgesia and defensive freezing are attenuated by application of diazepam to the amygdala. *Pharmacol Biochem Behav* 1993;44:433–438.

121. Helmstetter FJ, Tershner SA, Poore LH, et al. Antinociception following opioid stimulation of the basolateral amygdala is expressed through the periaqueductal gray and rostral ventromedial medulla. *Brain Res* 1998;779:104–118.

122. Henke PG. The amygdala and restraint ulcers in rats. *J Comp Physiol Psychol* 1980;94:313–323.

123. Henke PG. Facilitation and inhibition of gastric pathology after lesions in the amygdala in rats. *Physiol Behav* 1980;25:575–579.

124. Henke PG. Attenuation of shock-induced ulcers after lesions in the medial amygdala. *Physiol Behav* 1981;27:143–146.

125. Hernandez DE, Salaiz AB, Morin P, et al. Administration of thyrotropin-releasing hormone into the central nucleus of the amygdala induces gastric lesions in rats. *Brain Res Bull* 1990;24:697–699.

126. Hodges H, Green S, Glenn B. Evidence that the amygdala is involved in benzodiazepine and serotonergic effects on punished responding but not on discrimination. *Psychopharmacology* 1987;92:491–504.

127. Holland PC, Lamoureux JA, Han J-S, et al. Hippocampal lesions interfere with pavlovian negative ossasion setting. *Hippocampus* 1999;9:143–157.

128. Holt W, Maren S. Muscimol inactivation of the dorsal hippocampus impairs contextual retrieval of fear memory. *J Neurosci* 1999;19:9054–9062.

129. Huang YY, Kandel ER. Postsynaptic induction and PKAdependent expression of LTP in the lateral amygdala. *Neuron* 1998;21:169–178.

130. Ikegaya Y, Abe K, Saito H, et al. Medial amygdala enhances synaptic transmission and synaptic plasticity in the dentate gyrus of rats *in vivo. J Neurophysiology* 1995;74:2201–2203.

131. Ikegaya Y, Saito H, Abe K. Attenuated hippocampal long-term potentiation in basolateral amygdala-lesioned rats. *Brain Res* 1994;1994:157–164.

132. Ikegaya Y, Saito H, Abe K. Amygdala *N*-methyl-D-aspartate receptors participate in the induction of long-term potentiation in the dentate gyrus *in vivo. Neurosci Lett* 1995;192:193–196.

133. Ikegaya Y, Saito H, Abe K. High-frequency stimulation of the basolateral amygdala facilitates the induction of long-term potentiation in the dentate gyrus *in vivo. Neurosci Res* 1995;22:203–207.

134. Ikegaya Y, Saito H, Abe K. Requirement of basolateral amygdala neuron activity for the induction of long-term potentiation in the dentate gyrus *in vivo. Brain Res* 1995;671:351–354.

135. Irwin W, Davidson RJ, Lowe MJ, et al. Human amygdala activation detected with echo-planar functional magnetic resonance imaging. *Neuroreport* 1996;7:1765–1769.

136. Ishikawa T, Yang H, Tache Y. Medullary sites of action of the TRH analogue, RX 77368, for stimulation of gastric acid secretion in the rat. *Gastroenterology* 1988;95:1470–1476.

137. Iwata J, LeDoux JE, Meeley MP, et al. Intrinsic neurons in the amygdala field projected to by the medial geniculate body mediate emotional responses conditioned to acoustic stimuli. *Brain Res* 1986;383:195–214.

138. Jacobson R. Disorders of facial recognition, social behaviour and affect after combined bilateral amygdalotomy and subcaudate tractotomy: a clinical and experimental study. *Psychol Med* 1986;16:439–450.

139. Kapp BS, Frysinger RC, Gallagher M, et al. Amygdala central nucleus lesions: effect on heart rate conditioning in the rabbit. *Physiol Behav* 1979;23:1109–1117.

140. Kapp BS, Supple WF, Whalen PJ. Effects of electrical stimulation of the amygdaloid central nucleus on neocortical arousal in the rabbit. *Behav Neurosci* 1994;108:81–93.

141. Kapp BS, Whalen PJ, Supple WF, et al. Amygdaloid contributions to conditioned arousal and sensory information processing. In: Aggleton JP, ed. *The amygdala: neurobiological aspects of emotion, memory and mental dysfunction*. New York: Wiley-Liss, 1992:229–254.

142. Kapp BS, Wilson A, Pascoe JP, et al. A neuroanatomical systems analysis of conditioned bradycardia in the rabbit. In: Gabriel M, Moore J, eds. *Neurocomputation and learning: foundations of adaptive networks*. New York: Bradford Books, 1990:55–90.

143. Kehoe EJ, Macrae M, Hutchinson CL. MK-801 protects conditioned response from extinction in the rabbit nictitating membrane preparation. *Psychobiology* 1996;24:127–135.

144. Kemble ED, Blanchard DC, Blanchard RJ. Effects of regional amygdaloid lesions on flight and defensive behaviors of wild black rats *(Rattus rattus). Physiol Behav* 1990;48:1–5.

145. Kemble ED, Blanchard DC, Blanchard RJ, et al. Taming in wild rats following medial amygdaloid lesions. *Physiol Behav* 1984;32:131–134.

146. Killcross S, Robbins TW, Everitt BJ. Different types of fear-conditioned behaviour mediated by separate nuclei within amygdala. *Nature* 1997;388:377–380.

147. Kim C, Kim CC, Kim JK, et al. Fear response and aggressive behavior in hippocampectomized house rats. *Brain Res* 1971;29:237–251.

148. Kim JJ, Fanselow MS. Modality-specific retrograde amnesia of fear. *Science* 1992;256:675–677.

149. Kim M, Campeau S, Falls WA, et al. Infusion of the non-NMDA receptor antagonist CNQX into the amygdala blocks the expression of fear-potentiated startle. *Behav Neural Biol* 1993;59:5–8.

150. LaBar KS, Gatenby JC, Gore JC, et al. Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron* 1998;20:937–945.

151. Lamont EW, Kokkinidis L. Infusion of the dopamine D1 receptor antagonist SCH 23390 into the amygdala blocks fear expression in a potentiated startle paradigm. *Brain Res* 1998;795:128–136.

152. Leaton RN, Borszcz GS. Hippocampal lesions and temporally chained conditioned stimuli in a conditioned suppression paradigm. *Psychobiology* 1990;18:81–88.

153. LeBar KS, LeDoux JE, Spencer DD, et al. Impaired fear conditioning following unilateral temporal lobectomy in humans. *J Neurosci* 1995;15:6846–6855.

154. LeDoux JE, Cicchetti P, Xagoraris A, et al. The lateral amygdaloid nucleus: sensory interface of the amygdala in fear conditioning. *J Neurosci* 1990;10:1062–1069.

155. LeDoux JE, Iwata J, Cicchetti P, et al. Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. *J Neurosci* 1988;8:2517–2529.

156. LeDoux JE, Romanski L, Xagoraris A. Indelibility of subcortical memories. *J Cogn Neurosci* 1989;1:238–243.

157. LeDoux JE, Sakaguchi A, Reis DJ. Interuption of projections from the medial geniculate mediate emotional responses conditioned to acoustic stimuli. *J Neurosci* 1986;17:615–627.

158. Lee H, Kim J. Amygdalar NMDA receptors are critical for new fear learning in previously fear-conditioned rats. *J Neurosci* 1998;18:8444–8454.

159. Lee Y, Walker D, Davis M.Lack of a temporal gradient of retrograde amnesia following NMDA-induced lesions of the basolateral amygdala assessed with the fear-potentiated paradigm. *Behav Neurosci*. 1996;110:836–839.

160. Leonard CM, Rolls ET, Wilson FAW, et al. Neurons in the amygdala of the monkey with responses selective for faces. *Behav Brain Res* 1985;15:159–176.

161. Liebsch G, Landgraf R, Gerstberger R, et al. Chronic infusion of a CRH1 receptor antisense oligodeoxynucleotide into the central nucleus of the amygdala reduced anxiety-related behavior in socially defeated rats. *Regul Pept* 1995;59:229–239.

162. Lovick TA. Panic disorder: a malfunction of multiple transmitter control systems with the midbrain periaqueductal gray matter. *Neuroscientist* 2000;6:48–59.

163. Maldonado R, Stinus L, Gold LH, et al. Role of different brain structures in the expression of the physical morphine withdrawal syndrome. *J Pharmacol Exp Ther* 1992;261:669–677.

164. Manning BH, Mayer DJ. The central nucleus of the amygdala contributes to the production of morphine antinociception in the formalin test. *Pain* 1995;63:141–152.

165. Manning BH, Mayer DJ. The central nucleus of the amygdala contributes to the production of morphine antinociception in the rat tail-flick test. *J Neurosci* 1995;15:8199–8213.

166. Maren S, Aharonov G, Fanselow MS. Retrograde abolition of conditioned fear after excitotoxic lesions in the basolateral amygdala of rats: absence of a temporal gradient. *Behav Neurosci* 1996;110:718–726.

167. Maren S, Aharonov G, Fanselow MS. Neurotoxic lesions of the dorsal hippocampus and pavlovian fear conditioning. *Behav Brain Res* 1997;88:261–274.

168. McCabe PM, Gentile CG, Markgraf CG, et al. Ibotenic acid lesions in the amygdaloid central nucleus but not in the lateral subthalamic area prevent the acquisition of differential pavlovian conditioning of bradycardia in rabbits. *Brain Res* 1992;580:155–163.

169. McDonald AJ. Cortical pathways to the mammalian amygdala. *Prog Neurobiol* 1998;55:257–332.

170. McGaugh JL, Introini-Collison IB, Cahill L, et al. Neuromodulatory systems and memory storage: role of the amygdala. *Behav Brain Res* 1993;58:81–90.

171. McGaugh JL, Introini-Collison IB, Cahill L, et al. Involvement of the amygdala in neuromodulatory influences on memory storage. In: Aggleton JP, ed. *The amygdala: neurobiological aspects of emotion, memory and mental dysfunction*. New York: Wiley-Liss, 1992:431–451.

172. McIntosh AR, Gonzalez-Lima F. Functional network interactions between parallel auditory pathways during pavlovian conditioned inhibition. *Brain Res* 1995;683:228–241.

173. McKernan MG, Shinnick-Gallagher P. Fear conditioning induces a lasting potentiation of synaptic currents *in vitro. Nature* 1997;390:607–611.

174. McNish KA, Gewirtz JC, Davis M. Evidence of contextual fear conditioning following lesions of the hippocampus: a disruption of freezing but not fear-potentiated startle. *J Neurosci* 1997;17:9353–9360.

175. Moller C, Wujkybdm K, Sommer W, et al. Decreased experimental anxiety and voluntary ethanol consumption in rats following central but not basolateral amygdala lesions. *Brain Res* 1997;760:94–101.

176. Morgan MA, LeDoux JE. Medial prefrontal cortex (mPFC) and the extinction of fear: Differential effecs of pre- or post-training lesions. *Soc Neurosci Abstr* 1996;22:1116.

177. Morgan MA, LeDoux JE. Contribution of ventrolateral prefrontal cortex to the acquistion and extinction of conditioned fear in rats. *Neurobiol Learn Mem* 1999;72:244–251.

178. Morgan MA, Romanski LM, LeDoux JE. Extinction of emotional learning: contribution of medial prefrontal cortex. *Neurosci Lett* 1993;163:109–113.

179. Morris JS, Friston KJ, Buchel C, et al. A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain* 1998;121:47–57.

180. Morris JS, Frith CD, Perrett DI, et al. A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* 1996;383:812–815.

181. Morrow BA, Elsworth JD, Rasmusson AM, et al. The role of mesoprefrontal dopamine neurons in the acquisition and expression of conditioned fear in the rat. *Neuroscience* 1999;92:553–564.

182. Morrow NS, Hodgson DM, Garrick T. Microinjection of thyrotropin-releasing hormone analogue into the central nucleus of the amygdala stimulates gastric contractility in rats. *Brain Res* 1996;735:141–148.

183. Muller J, Corodimas KP, Fridel Z, et al. Functional inactivation of the lateral and basal nuclei of the amygdala by muscimol infusion prevents fear conditioning to an explicit conditioned stimulus and to contextual stimuli. *Behav Neurosci* 1997;111:683–691.

184. Nakamura K, Mikami A, Kubota K. Activity of single neurons in the monkey amygdala during performance of a visual discrimination task. *J Neurophysiol* 1992;67:1447–1463.

185. Packard MG, Cahill L, McGaugh JL. Amygdala modulation of hippocampal-dependent and caudate nucleus-dependent memory processes. *Proc Natl Acad Sci USA* 1994;91:8477–8481.

186. Packard MG, Teather LA. Amygdala modulation of multiple memory systems: hippocampus and caudate-putamen. *Neurobiol Learn Mem* 1998;69:163–203.

187. Pavlov IP. *Conditioned reflexes*. Oxford: Oxford University Press, 1927.

188. Pesold C, Treit D. The central and basolateral amygdala differentially mediate the anxiolytic effects of benzodiazepines. *Brain Res* 1995;671:213–221.

189. Petersen EN, Braestrup C, Scheel-Kruger J. Evidence that the anticonflict effect of midazolam in amygdala is mediated by the specific benzodiazepine receptor. *Neurosci Lett* 1985;53:285–288.

190. Phillips RG, LeDoux JE. Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behav Neurosci* 1992;106:274–285.

191. Quirk GJ, Kohanski GJ, Ayala O. Lesions of medial prefrontal cortex retard extinction of fear conditioning between sessions, but not within a session. *Soc Neurosci Abstr* 1998;28:1683.

192. Quirk GJ, Repa JC, LeDoux JE. Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: parallel recordings in the freely behaving rat. *Neuron* 1995;15:1029–1039.

193. Quirk GJ, Rosaly E, Romero RV, et al. NMDA receptors are required for long-term but not short-term memory of extinction learning. *Soc Neurosci Abstr* 1999;25:1620.

194. Rassnick S, Heinrichs SC, Britton KT, et al. Microinjection of a corticotroin-releasing factor antagonist into the central nucleus of the amygdala reverses anxiogenic-like effects of ethanol withdrawal. *Brain Res* 1993;605:25–32.

195. Ray A, Henke PG. Enkephalin-dopamine interactions in the central amygdalar nucleus during gastric stress ulcer formation in rats. *Behav Brain Res* 1990;36:179–183.

196. Ray A, Henke PG. TRH-enkephalin interactions in the amygdaloid complex during gastric stress ulcer formation in rats. *Regul Pept* 1991;35:11–17.

197. Ray A, Henke PG, Sullivan RM. Effects of intra-amygdalar thyrotropin releasing hormone (TRH) and its antagonism by atropine and benzodiazepines during stress ulcer formation in rats. *Pharmacol Biochem Behav* 1990;36:597–601.

198. Ray A, Sullivan RM, Henke PG. Interactions of thyrotropin-releasing hormone (TRH) with neurotensin and dopamine in the central nucleus of the amygdala during stress ulcer formation in rats. *Neurosci Lett* 1988;91:95–100.

199. Reiman EM, Lane RD, Ahern GL, et al. Neuroanatomical correlates of externally and internally generated human emotion. *Am J Psychiatry* 1997;54:918–925.

200. Rescorla RA. Pavlovian conditioned inhibition. *Psychol Bull* 1969;72:77–94.

201. Rescorla RA, Heth CD. Reinstatement of fear to an extinguished conditioned stimulus. *J Exp Psychol Anim Behav Process* 1975;1:88–96.

202. Richmond MA, Pouzet B, Veenman L, et al. Dissociating context and space within the hippocampus: effects of complete, dorsal, and ventral excitotixic hippocampal lesions on conditioned freezing and spatial learning. *Behav Neurosci* 1999;113:1189–1203.

203. Rogan MT, LeDoux JE. LTP is accompanied by commensurate enhancement of auditory-evoked responses in a fear conditioning circuit. *Neuron* 1995;15:127–136.

204. Rogan MT, Staubli UV, LeDoux JE. Fear conditioning induces associative long-term potentiation in the amygdala. *Nature* 1997;390:604–607.

205. Rolls ET. Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Hum Neurobiol* 1984;3:209–222.

206. Romanski LM, Clugnet MC, Bordi F, et al. Somatosensory and auditory convergence in the lateral nucleus of the amygdala. *Behav Neurosci* 1993;107:444–450.

207. Roozendaal B, Koolhaas JM, Bohus B. Central amygdaloid involvement in neuroendocrine correlates of conditioned stress responses. *J Neuroendocrinol* 1992;4:483–489.

208. Roozendaal B, Schoorlemmer GH, Koolhaas JM, et al. Cardiac, neuroendocrine, and behavioral effects of central amygdaloid vasopressinergic and oxytocinergic mechanisms under stress-free conditions in rats. *Brain Res Bull* 1993;32:573–579.

209. Roozendaal B, Wiersma A, Driscoll P, et al. Vasopressinergic modulation of stress responses in the central amygdala of the Roman high-avoidance and low-avoidance rat. *Brain Res* 1992;596:35–40.

210. Sajdyk TJ, Schober DA, Gehlert DR, et al. Role of corticotropin-releasing factor and urocortin within the basolateral amygdala of rats in anxiety and panic responses. *Behav Brain Res* 1999;100:207–215.

211. Sajdyk TJ, Shekhar A. Excitatory amino acid receptor antagonists block the cardiovascular and anxiety responses elicited by γ-aminobutyric acid: a receptor blockade in the basolateral amygdala of rats. *J Pharmacol Exp Ther* 1997;283:969–977.

212. Sajdyk TJ, Shekhar A. Excitatory amino acid receptors in the bsolateral amygdala regulate anxiety responses in the social interaction test. *Brain Res* 1997;764:262–264.

213. Sajdyk TJ, Vandergriff MG, Gehlert DR. Amygdalar neuropeptide Y Y-1 receptors mediate the anxiolytic-like actions of neuropeptide Y in the social interaction test. *Eur J Pharmacol* 1999;368:143–147.

214. Sananes CB, Davis M. *N*-methyl-D-aspartate lesions of the lateral and basolateral nuclei of the amygdala block fear-potentiated startle and shock sensitization of startle. *Behav Neurosci* 1992;106:72–80.

215. Sanders SK, Shekhar A. Blockade of $GABA_A$ receptors in the region of the anterior basolateral amygdala of rats elicits increases in heart rate and blood pressure. *Brain Res* 1991;576:101–110.

216. Sanders SK, Shekhar A. Regulation of anxiety by $GABA_A$ receptors in the rat amygdala. *Pharmacol Biochem Behav* 1995;52:701–706.

217. Sanghera MK, Rolls ET, Roper-Hall A. Visual responses of neurons in the dorsolateral amygdala of the alert monkey. *Exp Neurol* 1979;63:610–626.

218. Scheel-Kruger J, Petersen EN. Anticonflict effect of the benzodiazepines mediated by a GABAergic mechanism in the amygdala. *Eur J Pharmacol* 1982;82:115–116.

219. Schmaltz LW, Theiosus J. Acquisition and extinction of a classically conditioned response in hippocampectomized rabbits *(Oryctolagus cuniculus). J Comp Physiol Psychol* 1972;79:328–333.

220. Schoenbaum G, Chiba AA, Gallagher M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat Neurosci* 1998;1:155–159.

221. Schoenbaum G, Chiba AA, Gallagher M. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. *J Neurosci* 1999;19:1876–1884.

222. Schwaber JS, Kapp BS, Higgins GA, et al. Amygdaloid basal forebrain direct connections with the nucleus of the solitary tract and the dorsal motor nucleus. *J Neurosci* 1982;2:1424–1438.

223. Selden NR, Everitt BJ, Jarrard LE, et al. Complementary roles for the amygdala and hippocampus in aversive conditioning to explicit and contextual cues. *Neuroscience* 1991;42:335–350.

224. Selden NRW, Everitt BJ, Jarrard LE, et al. Complementary roles for the amygdala and hippocampus in aversive conditioning to explicit and contextual cues. *Neuroscience* 1991;42:335–350.

225. Shibata K, Kataoka Y, Yamashita K, et al. An important role of the central amygdaloid nucleus and mammillary body in the mediation of conflict behavior in rats. *Brain Res* 1986;372:159–162.

226. Shindou T, Watanabe S, Yamamoto K, et al. NMDA receptor-dependent formation of long-term potentiation in the rat medial amygdala neuron in an in vitro slice prepartation. *Brain Res Bull* 1993;31:667–672.

227. Shors TJ, Mathew PR. NMDA receptor antagonism in the lateral/basolateral but not central nucleus of the amygdala prevents the induction of facilitated learning in response to stress. *Learn Mem* 1998;5:220–230.

228. Solomon PR. Role of hippocampus in blocking and conditioned inhibition of the rabbit's nictitating membrane response. *J Comp Physiol Psychiatry* 1977;91:407–417.

229. Soltis RP, Cook JC, Gregg AE, et al. EAA receptors in the dorsomedial hypothalamic area mediate the cardiovascular response to activation of the amygdala. *Am J Physiol* 1998;275:R624–R631.

230. Soltis RP, Cook JC, Gregg AE, et al. Interaction of GABA and excitatory amino acids in the basolateral amygdala: role in cardiovascular regulation. *J Neurosci* 1997;17:9367–9374.

231. Stinus L, LeMoal M, Koob GF. Nucleus accumbens and amygdala are possible substrates for the aversive stimulus effects of opiate withdrawal. *Neuroscience* 1990;37:767–773.

232. Sullivan RM, Henke PG, Ray A, et al. The GABA/benzodiazepine receptor complex in the central amygdalar nucleus and stress ulcers in rats. *Behav Neural Biol* 1989;51:262–269.

233. Swiergiel AH, Takahashi LK, Kalin NH. Attenuation of stress-induced behavior by antagonism of corticotropin-releasing factor in the central amygdala of the rat. *Brain Res* 1993;623:229–234.

234. Takao K, Nagatani T, Kasahara K-I, et al. Role of the central serotonergic system in the anticonflict effect of *d*-AP159. *Pharmacol Biochem Behav* 1992;43:503–508.

235. Taylor JR, Punch LJ, Elsworth JD. A comparison of the effects of clonidine and CNQX infusion into the locus coeruleus and the amygdala on naloxone-precipitated opiate withdrawal in the rat. *Psychopharmacology* 1998;138:133–142.

236. Taylor JR, Robbins TW. Enhanced behavioral control by conditioned reinforcers following microinjections of D-amphetamine into the nucleus accumbens. *Psychopharmacology* 1984;84:405–412.

237. Teich AH, McCabe PM, Gentile CG, et al. Role of auditory cortex in the acquisition of differential heart rate conditioning. *Physiol Behav* 1988;44:405–412.

238. Thomas E, Yadin E, Strickland CE. Septal unit activity during classical conditioning: a regional comparison. *Brain Res* 1991;547:303–308.

239. Thomas SR, Lewis ME, Iversen SD. Correlation of [3H]diazepam binding density with anxiolytic locus in the amygdaloid complex of the rat. *Brain Res* 1985;342:85–90.

240. Tranel D, Hyman BT. Neuropsychological correlates of bilateral amygdala damage. *Arch Neurol* 1990;47:349–355.

241. Ursin H, Kaada BR. Functional localization within the amygdaloid complex in the cat. *Electroencephalogr Clin Neurophysiol* 1960;12:109–122.

242. Van de Kar LD, Piechowski RA, Rittenhouse PA, et al. Amygdaloid lesions: differential effect on conditioned stress and immobilization-induced increases in cortiocosterone and renin secretion. *Neuroendocrinology* 1991;15:89–95.

243. Walker DL, Davis M. Double dissociation between the involvement of the bed nucleus of the stria terminalis and the central nucleus of the amygdala in light-enhanced versus fear-potentiated startle. *J Neurosci* 1997;17:9375–9383.

244. Whalen PJ. Fear, vigilance, and ambiguity: initial neuroimaging studies of the human amygdala. *Curr Dir Psychol Sci* 1999;7:177–188.

245. Whalen PJ, Rauch SL, Etcoff NL, et al. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *J Neurosci* 1998;18:411–418.

246. Whalen PJ, Shin LM, McInerney SC, et al. Greater fMRI activation to fearful vs. angry facial expression in the amygdaloid region. *Neurosci Abstr* 1998;24:692.

247. Wiersma A, Baauw AD, Bohus B, et al. Behavioural activation produced by CRH but not α-helical CRH (CRH-receptor antagonist) when microinfused into the central nucleus of the amygdala under stress-free conditions. *Psychoneuroendocrinology* 1995;20:423–432.

248. Wiersma A, Bohus B, Koolhaas JM. Corticotropin-releasing hormone microinfusion in the central amygdala diminishes a cardiac parasympathetic outflow under stress-free conditions. *Brain Res* 1993;625:219–227.

249. Wiersma A, Bohus B, Koolhaas JM. Corticotropin-releasing hormone microinfusion in the central amygdala enhances active behaviour responses in the conditioned burying paradigm. *Stress* 1997;1:113–122.

250. Wiersma A, Tuinstra T, Koolhaas JM. Corticotropin-releasing hormone microinfusion into the basolateral nucleus of the amygdala does not induce any changes in cardiovascular, neuroendocrine or behavioural output in a stress-free condition. *Brain Res* 1993;625:219–227.

251. Willcox BJ, Poulin P, Veale WL, et al. Vasopressin-induced motor effects: localization of a sensitive site in the amygdala. *Brain Res* 1992;596:58–64.

252. Wilson A, Brooks D, Bouton ME. The role of the rat hippocampal system in several effects of context extincition. *Behav Neurosci* 1995;109:828–836.

253. Yadin E. Unit activity in the medial septum during differential appetitive conditioning. *Behav Brain Res* 1989;33:45–50.

254. Yadin E, Thomas E. Septal correlates of conditioned inhibition. *J Comp Physiol Psychol* 1981;95:331–340.

255. Yadin E, Thomas E. Stimulation of the lateral septum attenuates immobilization-induced stress ulcers. *Physiol Behav* 1996;59:883–886.

256. Yamashita K, Kataoka Y, Shibata K, et al. Neuroanatomical substrates regulating rat conflict behavior evidenced by brain lesioning. *Neurosci Lett* 1989;104:195–200.

257. Young AW, Aggleton JP, Hellawell DJ, et al. Face processing impairments after amygdalotomy. *Brain* 1995;118:15–24.

258. Young BJ, Helmstetter FJ, Rabchenuk SA, et al. Effects of systemic and intra-amygdaloid diazepam on long-term habituation of acoustic startle in rats. *Pharmacol Biochem Behav* 1991;39:903–909.

259. Zhang JX, Harper RM, Ni H. Cryogenic blockade of the central nucleus of the amygdala attenuates aversively conditioned blood pressure and respiratory responses. *Brain Res* 1986;386:136–145.

65

Structural and Functional Imaging of Anxiety and Stress Disorders

Scott L. Rauch

Lisa M. Shin

Scott L. Rauch: Department of Psychiatry, Harvard Medical School; Departments of Psychiatry and Radiology, Massachusetts General Hospital, Boston, Massachusetts.

Lisa M. Shin: Department of Psychiatry, Harvard Medical School; Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts; Department of Psychology, Tufts University, Medford, Massachusetts.

GENERAL PRINCIPLES

Part of "65 - Structural and Functional Imaging of Anxiety and Stress Disorders "

Neuroimaging research has emerged as a powerful force in shaping neurobiological models of psychiatric disorders. In this chapter, neuroimaging findings pertaining to anxiety and stress disorders are reviewed. This review necessarily extends previous ones that we have written, together with our colleagues, on this same and related topics (97 ,98 ,103).

Anxiety and Stress Disorders

Whereas anxiety disorders comprise a discrete category within the current version of the *Diagnostic and Statistical Manual* (2), the concept of stress disorders is less well defined. Although growing evidence suggests that stress may play a role across a variety of psychiatric disorders (such as major depression), posttraumatic stress disorder (PTSD) and acute stress disorder remain the only conditions for which exposure to traumatic stress is explicitly acknowledged as an etiologic factor and a criterion for diagnosis. In this chapter, we survey the imaging data pertinent to models of PTSD and other anxiety disorders, including obsessive-compulsive disorder (OCD), specific phobias (SpP), social phobia (SoP; also called social anxiety disorder) and panic disorder (PD).

In general, patients with anxiety disorders suffer either exaggerated fear responses to relatively innocuous stimuli (e.g., phobias) or spontaneous fear responses in the absence of true threat (e.g., PD). Thus, it is important to consider the mediating neuroanatomy of normal threat assessment and the fear response. In fact, contemporary models focus on these systems as candidate neural substrates for the anxiety disorders.

Relevant Neuroanatomy

The limbic system plays an important role in mediating human emotional states, including anxiety. Anterior paralimbic cortex (i.e., posterior medial orbitofrontal, anterior temporal, anterior cingulate, and insular cortex) links cortical regions subserving higher level cognition and sensory processing with deep limbic structures, such as the amygdala and the hippocampus (81).

Modern models of threat assessment and the normal fear response have focused on the role of the amygdala (1 ,75). The amygdala is positioned to receive input regarding the environment both directly and, thus, rapidly from the thalamus as well as from sensory cortex. The amygdala appears to serve several related functions, including preliminary threat assessment; facilitation of fight-or-flight responses; facilitation of additional information acquisition; and enhancement of arousal and plasticity, so the organism can learn from the current experience to guide responses optimally in future similar situations.

Conversely, several brain areas provide important feedback to the amygdala (1 ,75): medial frontal cortex (e.g., anterior cingulate and orbitofrontal cortex) may provide critical “top-down” governance over the amygdala, thus enabling attenuation of the fear response once danger has passed or when the meaning of a potentially threatening stimulus has changed; the hippocampus provides information about the context of a situation (based on information retrieved from explicit memory stores); and corticostriatothalamic circuits mediate “gating” at the level of the thalamus and thereby regulate the flow of incoming information that reaches the amygdala.

Finally, neuromodulators influence the activity within each of these various brain areas, as well as the interactions

among the nodes of the entire system outlined earlier. Ascending projections from the raphe nuclei (serotonin) and the locus ceruleus (norepinephrine), as well as widespread local γ-aminobutyric acid–ergic (GABAergic) neurons, are perhaps most relevant to the physiology of anxiety (30 ,65 ,114). These transmitter systems likely serve as the principal substrates for contemporary anxiolytic medications, including serotonergic reuptake inhibitors, monoamine oxidase inhibitors, other antidepressant medications, and benzodiazepines.

Neuroimaging Methods

Morphometric magnetic resonance imaging (mMRI) methods can be used to test hypotheses regarding abnormalities in the size or shape of particular brain structures. Functional imaging methods include positron emission tomography (PET) with tracers that measure blood flow (e.g., oxygen-15–labeled carbon dioxide) or glucose metabolism (i.e., fluorine-18–labeled fluorodeoxyglucose [FDG]), single photon emission tomography (SPECT) with tracers that measure correlates of blood flow (e.g., technetium-99– labeled hexamethylpropylene amine oxime [TcHMPAO]), and functional MRI (fMRI) to measure blood oxygenation level–dependent (BOLD) signal changes. Each of these techniques yields maps that reflect regional brain activity.

Functional imaging methods can be applied in the context of various experimental paradigms. In *neutral state paradigms*, subjects are studied during a nominal “resting” state or while they perform a nonspecific continuous task. Thus, between-group comparisons are made to test hypotheses regarding group differences in regional brain activity, without particular attention to state variables. In *treatment paradigms*, subjects are scanned in the context of a treatment protocol. In pre/posttreatment studies, subjects are scanned both before and after a trial. Then, within-group comparisons are made to test hypotheses regarding changes in brain activity profiles associated with symptomatic improvement. Alternatively, correlational analyses can be performed to identify pretreatment brain activity characteristics that predict good or poor treatment response. In *symptom provocation paradigms*, subjects are scanned during a symptomatic state (after having their symptoms intentionally induced) as well as during control conditions. Within-group comparisons can be made to test hypotheses regarding the mediating anatomy of the symptomatic state; group-by-condition interactions can be sought to distinguish responses in patient versus control groups. Behavioral and pharmacologic challenges can be used to induce symptoms. In some cases, when symptomatic states occur spontaneously, experiments are designed to capture these events without the need for provocation or induction *per se*. In *cognitive activation paradigms*, subjects are studied while they perform specially designed cognitive-behavioral tasks. This approach is intended to increase sensitivity by employing tasks that specifically activate brain systems of interest. Again, group-by-condition interactions are sought to test the functional responsivity or integrity of specific brain systems in patients versus control subjects.

Imaging studies of neurochemistry have employed PET and SPECT methods in conjunction with radiolabeled high-affinity ligands. In this way, regional receptor number and or affinity can be characterized *in vivo* (i.e., receptor-characterization studies). Other approaches include the use of magnetic resonance spectroscopy (MRS) to measure the regional relative concentration of select “MRS-visible” compounds. For instance, MRS can be used to measure the compound *N*-acetylaspartate (NAA), which is a purported marker of healthy neuronal density.

These various neuroimaging techniques should be viewed as complementary. Convergent findings across paradigms and laboratories yield the most cohesive and compelling models of pathophysiology.

RELEVANT NEUROIMAGING FINDINGS IN HEALTHY SUBJECTS

Part of "65 - Structural and Functional Imaging of Anxiety and Stress Disorders "

Anxiety and Other Negative Emotional States

Behaviorally Induced Fear and Anxiety

Fischer and colleagues used PET to study regional cerebral blood flow (rCBF) in bank officials while they viewed security camera videotape of a robbery that they had experienced previously (46). Watching the robbery video was associated with rCBF increases in orbitofrontal cortex, visual cortex, and posterior cingulate gyrus; rCBF decreases occurred in Broca's area, left angular gyrus, operculum, and secondary somatosensory cortex. Paradiso et al. studied rCBF in healthy elderly subjects who viewed emotionally evocative film clips; a fear/disgust versus neutral comparison revealed activation in orbitofrontal cortex, among other areas (93). Similarly, Kimbrell et al. studied induced emotional states in healthy adults (67); relative to a neutral condition, an anxiety condition was associated with rCBF increases in left anterior cingulate and left temporal pole, whereas rCBF decreases were found in nonparalimbic frontal cortical regions. Liotti et al. used PET and autobiographic memory scripts to examine the neural correlates of anxiety and sadness in healthy women (77); relative to a neutral condition, an anxiety condition was associated with rCBF increases in insular, orbitofrontal, and anterior temporal cortex and rCBF decreases in parahippocampal gyri.

Pharmacologically Induced Fear and Anxiety

Benkelfat et al. examined rCBF changes associated with the administration of cholecystokinin tetrapeptide versus saline in healthy persons (10). Administration of cholecystokinin

tetrapeptide was accompanied by increased heart rate and panic symptoms and rCBF increases in right cerebellar vermis, left anterior cingulate gyrus, bilateral claustrum-insula-amygdala, and bilateral temporal poles. However, further analyses suggested that the apparent temporal pole activations were attributable to extracranial artifacts of jaw muscle contraction. Ketter et al. used PET to study rCBF changes during procaine versus saline administration in healthy persons (66). In the procaine versus saline contrast, rCBF increases occurred in amygdala and anterior paralimbic structures, including anterior cingulate gyrus, insular cortex, and orbitofrontal cortex. Blood flow in left amygdala was positively correlated with fear intensity and was negatively correlated with euphoria intensity. Similar results were reported by Servan-Schreiber et al. (126).

Other Negative Emotions

Several studies examining rCBF associated with behaviorally induced sadness in healthy subjects have implicated both anterior paralimbic and nonparalimbic frontal cortical areas (4 ,54 ,55 ,77 ,94). Some studies of behaviorally induced sadness have also found rCBF changes within the amygdala (71 ,120 ,121). Of note, studies of other behaviorally induced negative emotions, including anger (39 ,67) and guilt (127), have likewise found activation of anterior paralimbic cortical territories.

Processing Unpleasant, Arousing, or Threat-Related Stimuli

In functional imaging studies of responses to unpleasant pictures (73), several studies have found amygdala activation when contrasted with a neutral (63 ,72) or pleasant picture (92) comparison condition. In a separate study, Lane et al. reported that anterior paralimbic regions were activated when study subjects attended to the emotions evoked by the pictures rather than when subjects attended to physical attributes of the depicted scenes (70). Functional imaging studies have also demonstrated a correlation between amygdala activity during encoding of emotionally arousing pictures or film clips and subjects' subsequent memory of them (29 ,59).

Several functional imaging studies have shown greater amygdala responses to overtly presented fearful human facial expressions in comparison with neutral or happy faces (14 ,86). Whalen et al. used fMRI and a technique called “backward masking” to study amygdala responses to emotional faces in the absence of explicit knowledge (142). Although subjects were unaware of seeing the “masked” emotional faces, a comparison of the masked fear and masked happy conditions yielded activation in the amygdala bilaterally.

Habituation

The term *habituation* refers to a decrement in responses over repeated presentations of a stimulus. Measures of habituation can be obtained peripherally (e.g., skin conductance) or more centrally (e.g., rCBF or fMRI BOLD signal). For example, declining fMRI BOLD signal within the amygdala has been observed in response to repeated presentations of fearful faces, regardless of whether subjects are aware the stimuli are present (14 ,142).

A few studies have directly examined the neural correlates of habituation. Fischer et al. used PET to study rCBF changes over repeated presentation of videotaped scenes in healthy women (45). In separate scanning conditions, subjects watched two repeated videotaped presentations of neutral park scenes and snakes. From the first to the second presentation of the videotapes, rCBF decreased in bilateral secondary visual cortex and right medial temporal cortex, including amygdala and hippocampus. In an fMRI study, Fischer et al. also found response decrements over repeated presentations of human face stimuli in amygdala and hippocampus, as well as thalamus, and prefrontal, inferior temporal, and posterior cingulate cortex (47). Similar results have also been reported by Wright et al. (146).

Conditioning and Extinction

Fear conditioning involves the presentation of a neutral stimulus (i.e., a conditioned stimulus [CS]), such as a tone, that is temporally paired with an aversive stimulus (i.e., the unconditioned stimulus [US]), such as a shock. After repeated presentations of the CS and US, the CS alone begins to elicit fear-related autonomic changes, such as skin conductance responses. Subsequently, over repeated presentations of the CS without the US, fear responses decline, and this process is referred to as *extinction*. Existing research suggests that the amygdala plays a critical role in fear conditioning (27 ,35 ,74 ,75 ,141), and the medial prefrontal cortex may play a critical role in the process of extinction (56 ,83 ,84).

Fredrikson and Furmark and their colleagues used PET to study healthy subjects who viewed a videotape of snakes (CS) both before and after the video was paired with shock (US) (49 ,52). The findings revealed a significant correlation between rCBF changes in right amygdala and electrodermal activity changes. Hugdahl et al. used PET to compare patterns of blood flow before and after classic conditioning in healthy male study subjects, by employing a paradigm in which a tone (CS) was paired with brief shock (US) (62). Extinction was associated with widespread activations in right prefrontal, including orbitofrontal, cortex, as well as left occipital and superior frontal cortex.

In a different conditioning paradigm, Morris et al. showed study subjects pictures of faces that were previously paired with an aversive burst of white noise (CS+) and faces that were never paired with the noise (CS-) (85). A comparison of the CS+ versus CS- conditions yielded activation in right thalamus, orbitofrontal cortex, and superior frontal gyrus. There was a positive correlation between

activation in thalamus and in right amygdala, orbitofrontal cortex, and basal forebrain. Morris and colleagues subsequently used PET and backward masking techniques to study rCBF responses to conditioned face stimuli with and without awareness in healthy male subjects (87). When all CS+ conditions were compared with all CS- conditions, bilateral activation in amygdala was observed. Right amygdala activation was found in the condition in which subjects were aware; left amygdala activation was found in the condition in which subjects were unaware of the emotionally expressive face stimuli.

In a single-trial fMRI study, LaBar et al. examined amygdala activation during both acquisition and extinction in a mixed-gender cohort (69). In the acquisition condition, a colored shape (CS+) was paired with a shock (US), whereas a different colored shape (CS-) was never paired with shock. No shocks were delivered during the extinction condition. Comparing CS+ with CS- trials revealed activation in periamygdaloid cortex and amygdala during early acquisition and early extinction trials, respectively. Activation in both these regions declined over time. Büchel and colleagues also used a single-trial fMRI to study the neural correlates of fear conditioning in healthy subjects (26). Study subjects were scanned during an acquisition phase in which two neutral faces (CS+) were presented with a loud tone (US) and two other neutral faces were presented alone (CS-). To disambiguate the effects of the CS+ and US, the US was not presented on half of the CS+ trials (i.e., $CS+_{unpaired}$). The critical comparison, $CS+_{unpaired}$ versus CS-, revealed activation in anterior cingulate, bilateral insular, parietal, supplementary motor, and premotor cortex. A time by event type interaction revealed that fMRI signal in amygdala decreased over time in the $CS+_{unpaired}$ condition relative to the CS- condition. Similar results were reported by Büchel et al. in a trace conditioning study, in which a temporal gap occurs between the offset of the CS and onset of the US (28). These researchers also found conditioning-related hippocampal activation that declined over time.

Summary

Taken together, functional imaging studies in healthy human subjects extend findings from animal research. Normal anxiety and fear reactions are associated with increased activity in limbic and paralimbic regions, whereas other territories of heteromodal association cortex exhibit decreased activity. However, similar patterns of limbic and paralimbic activation may be observed in association with other emotional states, and hence this general profile should not be taken as specific to anxiety or fear. Exposures to unpleasant, arousing, or threat-related stimuli often produce detectable amygdala responses, which can be associated with enhanced memory. Additional paralimbic recruitment may be related to a person's attention to his or her emotional state. Habituation can be observed in widely distributed brain regions, including limbic, paralimbic, and sensory areas. Consistent with animal data, human imaging results point to a role for the amygdala in fear conditioning and for the frontal cortex in extinction. The accessory role of the hippocampus in these processes remains less well defined.

POSTTRAUMATIC STRESS DISORDER

Part of "65 - Structural and Functional Imaging of Anxiety and Stress Disorders "

Amygdalocentric Neurocircuitry Model

We previously presented a neurocircuitry model of PTSD that emphasizes the central role of the amygdala and its interactions with the hippocampus, medial prefrontal cortex, and other heteromodal cortical areas purported to mediate higher cognitive functions (103). Briefly, this model hypothesizes hyperresponsivity within the amygdala to threat-related stimuli, with inadequate top-down governance over the amygdala by medial prefrontal cortex, specifically, the affective division of anterior cingulate cortex (142), and the hippocampus. Amygdala hyperresponsivity mediates symptoms of hyperarousal and explains the indelible quality of the emotional memory for the traumatic event; inadequate influence by the anterior cingulate cortex underlies deficits of habituation, and decreased hippocampal function underlies deficits in identifying safe contexts, as well as accompanying explicit memory difficulties (21). Further, we propose that in threatening situations, patients with PTSD exhibit an exaggerated reallocation of resources to regions that mediate fight-or-flight responses and away from widespread heteromodal cortical areas, as a neural substrate for dissociation.

Structural Imaging Findings

mMRI studies have reported smaller hippocampal volumes in veterans with PTSD than in comparison subjects. Bremner and colleagues (21) found that right hippocampal volumes were 8% smaller in 26 veterans with PTSD than in 22 civilians without PTSD. In addition, the PTSD group exhibited poorer performance on a standard measure of verbal memory, and their percent retention scores on this test were directly correlated with right hippocampal volume (i.e., lower scores were associated with smaller right hippocampal volumes). Gurvits and colleagues (58) used mMRI to study seven Vietnam combat veterans with PTSD, seven Vietnam combat veterans without PTSD, and eight nonveterans without PTSD. These investigators found significantly smaller hippocampal volumes bilaterally for the PTSD group in comparison with both control groups. Across the 14 veterans, hippocampal volume was inversely correlated with extent of combat exposure and PTSD symptom severity.

Similar hippocampal volumetric differences also have been reported in mMRI studies of PTSD resulting from childhood abuse. Bremner and colleagues (22) found 12%

smaller left hippocampal volumes in 17 adult survivors of childhood abuse with PTSD than in 21 nonabused comparison subjects. Stein and colleagues (134) found 5% smaller left hippocampal volumes in 21 adult survivors of childhood sexual abuse (most of whom had PTSD) than in 21 nonabused comparison subjects. Furthermore, total hippocampal volume was smaller in abused subjects with high PTSD symptom severity than in those with low PTSD symptom severity. In contrast to these results, DeBellis et al. (36) failed to find decreased hippocampal volumes in 44 maltreated children and adolescents with PTSD, compared with 61 nonabused healthy subjects. However, the PTSD group had smaller intracranial and cerebral volumes than did the comparison group.

Taken together, the results of structural neuroimaging studies of adult samples suggest that PTSD is associated with reduced hippocampal volume, which, in turn, may be associated with cognitive deficits and PTSD symptom severity. Although the extent of traumatic exposure may be correlated with hippocampal volume, it appears that differences between PTSD and control groups cannot be explained by traumatic exposure alone (58). The results of DeBellis et al. (36) suggest that hippocampal volumetric differences between groups may not be evident in samples of children and adolescents or in samples of persons with relatively recent traumatic exposures.

Stress, Glucocorticoids, and the Hippocampus

In this review about neuroimaging of anxiety and stress disorders, it is worth elaborating on the potential relationship between stress and hippocampal findings in PTSD. Animal research has provided evidence that stress is associated with damage to the hippocampus (16). For example, sustained, fatal social stress in vervet monkeys was associated with degeneration of neurons in the CA3 region of the hippocampus (139); chronic restraint stress in rats was associated with atrophy of apical dendrites of hippocampal CA3 pyramidal neurons (140); and exposure to cold water immersion stress in rats was related to structural damage to hippocampus (CA3 and CA2 fields) and deceased local CBF in hippocampus (42).

The effect of stress on hippocampus appears to be mediated by glucocorticoid hormones. Exposure to glucocorticoids is associated with hippocampal damage in both rats and primates. For example, Sapolsky et al. reported that chronic exposure to corticosterone in rats led to a loss of neurons in the CA3 region of the hippocampus (115). Woolley and colleagues found that daily corticosterone injections decreased dendritic branching and length in the CA3 region of the rat hippocampus (145). In a study of primates, Sapolsky et al. reported that chronic exposure to cortisol (through steroid-secreting pellets stereotactically implanted in hippocampus) was related to neuronal shrinkage and dendritic atrophy in the CA3 region (117). Moreover, chronic stress during development is capable of inhibiting normal cellular proliferation within the hippocampus, a process mediated by glucocorticoids and glutamatergic transmission by an *N*-methyl-D-aspartate receptor–dependent excitatory pathway (57).

Clinical research has also revealed decreased hippocampal volumes in humans with elevated cortisol levels resulting from Cushing's syndrome (130); furthermore, in these patients, a treatment-related reduction of cortisol levels is associated with increased hippocampal volumes (131). High cortisol levels and decreased hippocampal volumes have also been found in patients with major depressive disorder (25). The hippocampus is also involved in the modulation of the hypothalamic-pituitary-adrenal (HPA) axis, and lesions to hippocampus appear to increase the release of glucocorticoids during stress (43 ,60); this, in turn, may further damage the hippocampus (116).

Although these findings may have great relevance to anxiety and stress disorders, the picture is complicated by the finding that cortisol levels are characteristically reduced, rather than elevated, in PTSD (147). One parsimonious theory suggests that patients with PTSD suffer hypersensitivity to glucocorticoids, resulting in both reduced levels of cortisol (because of accentuated feedback inhibition) and reduced hippocampal volume (147).

Functional Imaging Findings

Semple and colleagues used PET to study six patients with combat-related PTSD and comorbid substance abuse versus seven normal control subjects (125). rCBF was measured in three conditions: resting state, an auditory continuous performance task, and a word generation task. Compared with the control group, the PTSD group exhibited greater rCBF during both task conditions within orbitofrontal cortex.

Rauch and colleagues studied a mixed-gender cohort of eight subjects with PTSD, using PET and a script-driven imagery method for inducing symptoms (104). In the provoked versus control condition, patients exhibited increased rCBF within anterior cingulate cortex, right orbitofrontal, insular, anterior temporal and visual cortex, and right amygdala. rCBF decreases occurred within left inferior frontal (Broca's area) and left middle temporal cortex. Interpretations of this initial study, with regard to the pathophysiology of PTSD, were limited by the absence of a comparison group.

Using a similar paradigm and a comparison group, Shin and colleagues studied eight women with childhood sexual abuse–related PTSD and eight matched trauma-exposed control subjects without PTSD (129). In the traumatic versus neutral comparison, both groups exhibited anterior paralimbic activation. However, a group-by-condition interaction revealed that the control group manifested a

significantly greater rCBF increase within anterior cingulate cortex than did the PTSD group, whereas the PTSD group showed significantly greater rCBF increases within anterior temporal and orbitofrontal cortex. Bremner and colleagues (20) also used script-driven imagery and PET to study rCBF in ten female survivors of childhood sexual abuse with PTSD and 12 without PTSD. Consistent with the findings of Shin et al. (129), Bremner and colleagues (20) reported relatively attenuated recruitment of anterior cingulate cortex in the PTSD group.

Bremner and colleagues (23) studied rCBF responses to trauma-related pictures and sounds in ten Vietnam veterans with PTSD and in ten veterans without PTSD. In the combat versus neutral comparison, the PTSD group exhibited rCBF decreases in medial prefrontal cortex (subcallosal gyrus) and anterior cingulate cortex. Liberzon and colleagues (76) used SPECT to study rCBF in 14 Vietnam veterans with PTSD, 11 veteran control subjects, and 14 healthy nonveterans. In separate scanning sessions, subjects listened to combat sounds and white noise. In the combat sounds versus white noise comparison, all three groups showed activation in anterior cingulate/medial prefrontal cortex, but only the PTSD group exhibited activation in the left amygdaloid region.

Bremner et al. (17) used PET to examine the effect of yohimbine challenge on glucose metabolic rates in ten combat veterans with PTSD and in ten nonveteran subjects without PTSD. Yohimbine administration was associated with increased anxiety and panic symptoms, as well as widespread decreases in cerebral glucose metabolism in the PTSD group.

Shin and colleagues studied seven patients with combat-related PTSD and seven matched trauma-exposed control subjects without PTSD in the context of a PET cognitive activation paradigm (128). Subjects were required to make judgments about pictures from three categories: neutral, general negative, and combat-related. Subjects performed two types of tasks: one involved responding while actually seeing the pictures (perception), and another involved responding while recalling the pictures (imagery). In the combat imagery versus control conditions, the PTSD group exhibited rCBF increases in right amygdala and ventral anterior cingulate gyrus and rCBF decreases in left inferior frontal gyrus (Broca's area).

Using another cognitive activation paradigm, Rauch et al. studied the functional integrity of the amygdala in eight combat veterans with PTSD and eight healthy combat veterans (108). During fMRI, subjects viewed fearful and happy faces temporally masked by neutral faces. Healthy persons are typically aware of seeing only the neutral faces, although they show amygdala activation to the masked fearful faces (142). Rauch and colleagues found greater amygdala activation to masked fearful faces in persons with PTSD than in control subjects (108). Furthermore, the magnitude of amygdala activation was correlated with PTSD severity. These results suggest that PTSD may be characterized by amygdala hyperresponsivity to general threat-related stimuli, consistent with our neurocircuitry model of PTSD.

Imaging Studies of Neurochemistry

Schuff and colleagues used mMRI and MRS to study seven veterans with PTSD and seven nonveteran control subjects (123). Although these investigators found a nonsignificantly smaller (6%) right hippocampus in the PTSD group by mMRI, they found an 18% reduction in right hippocampal NAA by MRS, a finding suggestive of reduced density or viability of neurons in this region.

DeBellis and colleagues used MRS to study NAA/creatine ratios in 11 maltreated children and adolescents with PTSD and 11 healthy comparison subjects without histories of maltreatment (37). The PTSD group had lower NAA/creatine ratios in pregenual anterior cingulate gyrus. This result is consistent with those of symptom provocation PET studies (20 ,23 ,129), which have reported failure to activate anterior cingulate in PTSD.

Bremner et al. (18) used SPECT and [^{123}I]iomazenil to study benzodiazepine-receptor binding in 13 veterans with PTSD and 13 nonveterans without PTSD. These investigators found decreased benzodiazepine-receptor binding in prefrontal cortex in the PTSD group, relative to the control group.

Summary

Taken together, data from neuroimaging studies are consistent with the current neurocircuitry model of PTSD that emphasizes the functional relationship among the amygdala, hippocampus, and medial prefrontal cortex. Hippocampal volumes and NAA levels appear to be decreased in persons with PTSD. These findings dovetail with animal research that points to a relationship among stress, HPA axis function, and cell viability within the hippocampus. Functionally, in comparison with control subjects, patients with PTSD exhibit the following: (a) greater activation within orbitofrontal cortex, anterior temporal poles, and the amygdala; (b) diminished activation in anterior cingulate and medial prefrontal cortex, as well as reduced NAA/creatine ratios in pregenual anterior cingulate; and (c) decreased activation within widespread areas that are associated with higher cognitive functions, such as Broca's area and dorsolateral prefrontal cortex.

OBSESSIVE-COMPULSIVE DISORDER

Part of "65 - Structural and Functional Imaging of Anxiety and Stress Disorders "

Corticostriatal Model

One current neuroanatomic model of OCD focuses on corticostriatothalamocortical circuitry (98 ,106). According to

this model, the primary disorder lies within the striatum (specifically, the caudate nucleus). This leads to inefficient gating at the level of the thalamus, which results in hyperactivity within orbitofrontal cortex (corresponding to the intrusive thoughts) and hyperactivity within anterior cingulate cortex (corresponding to anxiety, in a nonspecific manner). Compulsions are conceptualized as repetitive behaviors that are performed to recruit the inefficient striatum ultimately to achieve thalamic gating and hence to neutralize the unwanted thoughts and anxiety.

Structural Imaging Findings

The results of several mMRI investigations of OCD have suggested volumetric abnormalities involving the caudate nucleus, although the nature of the observed abnormalities has been somewhat inconsistent. Scarone et al. (119) studied a mixed-gender cohort of 20 patients with OCD versus 16 matched controls and found increased right caudate volume in the OCD group. Robinson et al. (111) studied a mixed-gender cohort of 26 patients with OCD versus 26 matched controls and found bilaterally decreased caudate volumes in the OCD group. Jenike et al. (64) studied an all-female cohort of ten patients with OCD versus matched controls and found trends toward a rightward shift in caudate volume (p = .06) as well as overall reduced caudate volume (p = .10) in the OCD group. Aylward et al. (3) studied a mixed-gender cohort of 24 patients with OCD versus 21 matched controls and found no significant differences in striatal volumes. Rosenberg et al. (112) studied 19 treatment-naive pediatric subjects with OCD and 19 case-matched psychiatrically healthy comparison subjects. These investigators found reduced striatal volumes in the OCD group and an inverse correlation between striatal volume and OCD symptom severity.

Functional Imaging Findings

Neutral state paradigms employing PET and SPECT have most consistently indicated that patients with OCD exhibit increased regional brain activity within orbitofrontal and anterior cingulate cortex, in comparison with neurologically normal control subjects (6 ,7 ,78 ,89 ,113 ,136). Observed differences in regional activity within the caudate nucleus have been less consistent (6 ,113).

Pre/posttreatment studies have reported treatment-related attenuation of abnormal regional brain activity within orbitofrontal cortex, anterior cingulate cortex, and caudate nucleus (8 ,9 ,61 ,95 ,124 ,137). In addition, both pharmacologic and behavioral therapies appear to be associated with similar brain activity changes (8 ,124). Some treatment studies have also reported that lower pretreatment glucose metabolic rates in orbitofrontal cortex predict a better response to serotonergic reuptake inhibitors (24 ,118 ,136).

Symptom provocation studies employing PET (80 ,99) as well as functional MRI (15) have also most consistently shown increased brain activity within anterior-lateral orbitofrontal cortex, anterior cingulate cortex, and caudate nucleus during the OCD symptomatic state.

Cognitive activation studies using PET and fMRI have probed the functional integrity of the striatum in OCD (102 ,105). In these studies, patients with OCD perform an implicit (i.e., nonconscious) learning paradigm shown reliably to recruit striatum in healthy individuals (101 ,107). In both studies, patients with OCD failed to recruit striatum normally and instead activated medial temporal regions typically associated with conscious information processing.

Imaging Studies of Neurochemistry

Several MRS studies have been performed to measure NAA concentrations in patients with OCD versus healthy comparison subjects. Ebert and colleagues (41) found reduced relative NAA levels in right striatum and anterior cingulate cortex in 12 patients with OCD in comparison with six healthy control subjects. Bartha and colleagues (5) found lower left striatal NAA concentrations in 13 patients with OCD than in 13 matched control subjects.

MRS has also been used to demonstrate elevated glutamatergic concentrations within the striatum of a child with OCD (82). Glutamate is the principal transmitter mediating frontostriatal communication. Interestingly, elevated striatal glutamate levels were attenuated toward normal after successful pharmacotherapy. These findings suggest that orbitofrontal hyperactivity in OCD may be mirrored by elevated glutamate at the site of orbitofrontal ramifications in striatum, and treatment-related attenuation of orbitofrontal activity may be accompanied by decreased glutamate concentration within the striatum.

Summary

Taken together, these neuroimaging findings are consistent with disorders in corticostriatothalamocortical circuitry. Consistent with the hypothesis of a primary abnormality in the striatum, MRI and MRS studies of OCD have shown reduced striatal volumes and reduced striatal NAA, respectively. PET studies have revealed hyperactivity within orbitofrontal cortex, and the magnitude of this hyperactivity predicts response to treatment. In addition, in neurologically normal persons, the performance of repetitive motor routines does facilitate striatal recruitment in the service of thalamic gating, whereas this pattern is not readily demonstrated in patients with OCD. These imaging data further support the working model of striatal pathology and striatothalamic inefficiency, together with orbitofrontal hyperactivity.

SOCIAL AND SPECIFIC PHOBIAS

Part of "65 - Structural and Functional Imaging of Anxiety and Stress Disorders "

Neuroanatomic Models

Currently, there are no cohesive neuroanatomically based models for the phobias (53 ,132). One possibility is that phobias are learned and hence reflect another example of fear conditioning to specific stimuli or situations. Alternatively, phobias may represent the product of dysregulated systems for detecting potentially threatening stimuli or situations. For instance, if humans have evolved a neural network specifically designed to assess social cues for threatening content, and another to assess threat from small animals, these may represent the neural substrates for the pathophysiology underlying phobias.

Structural Imaging Findings

Given the high prevalence of phobias and the relative ease with which medication-free phobic subjects without significant comorbidities can be recruited, it is striking how few imaging studies have been conducted in this arena. Potts et al. used mMRI to examine volumetric measures of total cerebrum, caudate, putamen, and thalamus in 22 patients with SoP and in 22 matched healthy control subjects (96). The groups did not significantly differ on any of these measures.

Functional Imaging Findings

Studies of SpP to date have principally employed PET symptom provocation paradigms and have reported somewhat inconsistent results. Mountz and colleagues found that persons with small-animal phobia exhibited increased heart rates, respiratory rates, and subjective reports of anxiety during exposure to phobic stimuli; however, no changes in rCBF measurements were observed (88).

Wik and colleagues studied six patients with snake phobias during exposure to videotapes of neutral, generally aversive, and snake-related scenes (143). During the phobic condition, they found significant rCBF increases in secondary visual cortex and rCBF decreases in prefrontal cortex, posterior cingulate cortex, anterior temporopolar cortex, and hippocampus. These findings were similar to those of two other studies of phobia from the same laboratory (50 ,51).

Using *in vivo* exposure and PET, Rauch and colleagues studied rCBF in seven persons with a variety of small animal phobias (100). In the provoked versus control condition, patients with phobias exhibited rCBF increases within multiple anterior paralimbic territories (i.e., right anterior cingulate, right anterior temporal pole, left posterior orbitofrontal cortex, and left insular cortex), left somatosensory cortex, and left thalamus.

Whereas one neutral-state SPECT study of patients with SoP and healthy control subjects found no significant between-group differences in rCBF (133), more recent cognitive activation studies performed in conjunction with fMRI have yielded more informative results. Birbaumer et al. (11) used fMRI to study seven patients with SoP and five healthy control subjects during exposure to slides of neutral human faces or aversive odors. In comparison with the control group, the SoP group exhibited hyperresponsivity within the amygdala that was specific to the human face stimuli. In a follow-up study, Schneider et al. used fMRI to study 12 patients with SoP and 12 healthy control subjects (122). The researchers used a classic conditioning paradigm in which neutral face stimuli were the conditioned stimuli and odors (negative odor and odorless air) served as the unconditioned stimuli. In response to conditioned stimuli associated with the negative odor, the SoP group displayed signal increases within amygdala and hippocampus, whereas healthy comparison subjects displayed signal decreases in these regions.

Imaging Studies of Neurochemistry

Davidson et al. (34) used MRS to study NAA in 20 patients with SoP and in 20 healthy control subjects. Relative to the control group, the SoP group exhibited decreased NAA in white matter and cortical and subcortical gray matter (e.g., caudate and thalamus).

Tiihonen et al. used SPECT and I-123–labeled ß-CIT to measure the density of dopamine reuptake sites in 11 patients with SoP and in 28 healthy comparison subjects (138). They found significantly reduced striatal dopamine reuptake binding site density in the SoP versus control group.

Summary

Although relatively few neuroimaging studies of SpP have been conducted, findings from existing research suggest activation of anterior paralimbic regions and sensory cortex corresponding to stimulus inflow associated with a symptomatic state. Although such results are consistent with a hypersensitive system for assessment of or response to specific threat-related cues, they do not provide clear anatomic substrates for the pathophysiology of SpP. Cognitive activation neuroimaging studies of SoP reveal exaggerated responsivity of medial temporal lobe structures to human face stimuli; this hyperresponsivity may reflect a neural substrate for social anxiety in SoP.

PANIC DISORDER

Part of "65 - Structural and Functional Imaging of Anxiety and Stress Disorders "

Neuroanatomic Models

Neurobiological models of PD have emphasized a wide range of disparate elements (31). Satisfactory models of PD must account for spontaneous panic attacks, which are a

defining feature of PD. It is possible that spontaneous panic corresponds to a normal physiologic anxiety response that is mediated by intact fear-anxiety circuits but, owing to homeostatic deficits, occurs in inappropriate, threat-free situations. This is consistent with theories of hypersensitivity to carbon dioxide at the level of the brainstem (i.e., the suffocation false alarm model), as well as theories regarding fundamental monoaminergic dysregulation. Another possibility is that panic attacks emerge in the context of what should be minor anxiety episodes because of failures in the systems responsible for limiting such normal responses; hippocampal deficits may underlie such a failure to inhibit anxiety responses. Finally, panic episodes described as spontaneous (i.e., without identifiable precipitants) could reflect anxiety responses to stimuli that are not processed at the conscious (i.e., explicit) level, but instead, recruit anxiety circuitry without awareness (i.e., implicitly). There is strong evidence that the amygdala can be recruited into action in the absence of awareness that a threat-related stimulus has been presented (142). By this model, PD may be characterized by fundamental amygdala hyperresponsivity to subtle environmental cues, triggering full-scale threat-related responses in the absence of conscious awareness.

Structural Imaging Findings

Fontaine et al. (48) published a qualitative MRI study involving 31 patients with PD and 20 matched healthy control subjects. The frequency of gross structural abnormalities was higher in the PD group (40%) than in the control group (10%); the most striking focal findings in the PD group involved abnormal signal or asymmetric atrophy of the right temporal lobe.

Functional Imaging Findings

In an initial PET neutral-state study, Reiman and colleagues studied 16 patients with PD and 25 normal control subjects (109). In the subset of patients who were vulnerable to lactate-induced panic (n = 8), the investigators found abnormally low left/right ratios of parahippocampal blood flow. DeCristofaro et al. used SPECT to measure rCBF at rest in seven treatment-naive patients with PD and in five age-matched healthy control subjects (38). Relative to the control group, the PD group exhibited elevated rCBF in left occipital cortex and reduced rCBF in the hippocampal area bilaterally. Nordahl et al. used PET and FDG to measure regional cerebral metabolic rate glucose (rCMRglc) in 12 patients with PD and 30 normal control subjects during an auditory continuous performance task (90). The investigators found that the PD group exhibited a lower left/right hippocampal ratio. In a follow-up experiment (91), these investigators used PET-FDG methods to study imipramine-treated subjects with PD and found a rightward shift in symmetry of rCMRglc within hippocampus and posterior inferior frontal cortex. In comparison with the untreated group, the imipramine-treated group exhibited rCMRglc decreases in posterior orbital frontal cortex. Bisaga et al. (12) used PET and FDG to study a cohort of six women with PD and six matched control subjects. In contrast to previous studies, the PD subjects displayed elevated rCMRglc in the left hippocampus and parahippocampal area.

The literature contains three symptom provocation studies of PD, all of which have employed pharmacologic challenges. Stewart et al. used SPECT and the xenon inhalation method to measure CBF during lactate infusion in ten patients with PD and in five healthy control subjects (135). The patients with PD who experienced lactate-induced panic attacks (n = 6) displayed global cortical CBF decreases. Woods et al. used SPECT and yohimbine infusions to study six patients with PD and six normal control subjects (144). In the PD group, yohimbine administration increased anxiety and decreased rCBF in bilateral frontal cortex. In a PET study, Reiman et al. measured rCBF during lactate infusions in 17 patients with PD and in 15 normal control subjects (110). The eight patients who suffered lactate-induced panic episodes exhibited rCBF increases in bilateral temporopolar cortex and bilateral insular cortex/claustrum/putamen. Healthy control subjects and patients with PD who did not experience lactate-induced panic attacks did not exhibit such rCBF changes. Of note, the temporopolar findings were subsequently questioned as possibly reflecting extracranial artifacts from muscular contractions (40 ,10). In a symptom capture case report, Fischer and colleagues (44) found that a spontaneous panic attack was associated with rCBF decreases in right orbitofrontal, prelimbic (area 25), anterior cingulate, and anterior temporal cortex.

Imaging Studies of Neurochemistry

Dager and colleagues used MRS to measure brain lactate levels during hyperventilation in seven treatment-responsive patients with PD and in seven healthy comparison subjects (32). The PD group showed a significantly greater rise in brain lactate in response to the same level of hyperventilation. Dager et al. also used MRS to measure brain lactate levels during lactate infusions in 15 patients with PD and in ten healthy comparison subjects (33). The PD group exhibited a significantly greater brain lactate level during lactate infusion, a finding consistent with the interpretation of reduced clearance, rather than higher production, of lactate in PD.

Three studies have used SPECT and [^{123}I]iomazenil to measure benzodiazepine-receptor binding in PD. Kuikka et al. (68) studied 17 subjects with PD and 17 matched healthy comparison subjects and found that the PD group exhibited a greater left/right ratio in benzodiazepine-receptor uptake that was most prominent in prefrontal cortex. Brandt et al. (13) studied 12 medication-naive patients with PD and nine

matched healthy control subjects and found that the PD group exhibited significantly elevated benzodiazepine-receptor binding within right supraorbital frontal cortex, as well as a trend toward elevated binding in the right temporal cortex. Bremner et al. (19) included 13 patients with PD and 16 healthy comparison subjects and found that the PD group showed decreased benzodiazepine-receptor binding in left hippocampus and precuneus.

Using PET and carbon-11–labeled flumazenil, Malizia et al. studied seven patients with PD and eight healthy comparison subjects (79). These investigators found that the PD group exhibited a global reduction in benzodiazepine binding that was most pronounced in right orbitofrontal and right insular cortex.

Summary

Resting state neuroimaging studies have suggested abnormal hippocampal activity in PD. Symptom provocation studies have revealed reduced activity in widespread cortical regions, including prefrontal cortex, during symptomatic states. MRS studies have reported greater brain lactate levels in response to hyperventilation and lactate infusions. Finally, receptor-binding studies of PD suggest widespread abnormalities in the GABAergic/benzodiazepine system. Consistent with prevailing neurobiological models of PD, it is possible that fundamental abnormalities in monoaminergic neurotransmitter systems, originating in the brainstem, underlie the abnormalities of metabolism, hemodynamics, and chemistry found in widespread territories of cortex. Further, regional abnormalities within the medial temporal lobes provide some support for theories regarding hippocampal or amygdala dysfunction in PD.

CONCLUSIONS AND FUTURE DIRECTIONS

Part of "65 - Structural and Functional Imaging of Anxiety and Stress Disorders "

Neuroimaging research is helping to advance neurobiological models of anxiety and stress disorders. At the current early stage of this scientific enterprise, there are hints of commonalities across anxiety disorders as well as leads regarding disorder-specific features. Beyond the need for a general expansion of the existing database, it will be critical to explore the specificity of initial findings by conducting studies with psychiatric comparison groups in addition to healthy control subjects. This is of particular relevance to the concept of stress disorders, in which common etiologic factors or vulnerability factors may have corresponding pathophysiologic profiles that are independent of our current diagnostic scheme. For instance, the relationship between early or chronic life stress and hippocampal structure and function may well span anxiety, mood, and even psychotic disorders. In this light, longitudinal and developmental studies may be of particular importance in elucidating the neural correlates and consequences of stress. Similarly, genetic studies in animals and humans will benefit from neuroimaging methods that can illuminate the bidirectional link from behavior to brain structure, function, and chemistry. For instance, research regarding the heritability of anxious temperament may be enhanced by using extended phenotypes of conditionability or distributed brain function within amygdala, hippocampus, and medial frontal cortex. In fact, the gamut of existing animal and human experimental paradigms with relevance to anxiety and stress disorders provides a promising context for advancing integrated models across scales and neuroscientific modes of inquiry. As such integrated models evolve, targets for new and improved neuropsychopharmacotherapies are destined to emerge. Indeed, neuroimaging is likely to play a role not only in conceptually motivating but also in discovering and testing such new therapies as part of the next generation of progress in this domain.

REFERENCES

1. Aggleton JP, ed. *The amygdala: neurobiological aspects of emotion, memory and mental dysfunction*. New York: Wiley–Liss, 1992.

2. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington, DC: American Psychiatric Association, 1994.

3. Aylward EH, Harris GJ, Hoehn-Saric R, et al. Normal caudate nucleus in obsessive-compulsive disorder assessed by quantitative neuroimaging. *Arch Gen Psychiatry* 1996;53:577–584.

4. Baker SC, Frith CD, Dolan RJ. The interaction between mood and cognitive function studied with PET. *Psychol Med* 1997;27:565–578.

5. Bartha R, Stein MB, Williamson PC, et al. A short echo 1H spectroscopy and volumetric MRI study of the corpus striatum in patients with obsessive-compulsive disorder and comparison subjects. *Am J Psychiatry* 1998;155:1584–1591.

6. Baxter LR, Phelps ME, Mazziotta JC, et al. Local cerebral glucose metabolic rates in obsessive compulsive disorder: a comparison with rates in unipolar depression and in normal controls. *Arch Gen Psychiatry* 1987;44:211–218.

7. Baxter L, Schwartz J, Mazziotta J, et al. Cerebral glucose metabolic rates in nondepressed patients with obsessive-compulsive disorder. *Am J Psychiatry* 1988;145:1560–1563.

8. Baxter LR Jr, Schwartz JM, Bergman KS, et al. Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:681–689.

9. Benkelfat C, Nordahl TE, Semple WE, et al. Local cerebral glucose metabolic rates in obsessive-compulsive disorder: patients treated with clomipramine. *Arch Gen Psychiatry* 1990;47:840–848.

10. Benkelfat C, Bradwejn J, Meyer E, et al. Functional neuroanatomy of CCK_4-induced anxiety in normal healthy volunteers. *Am J Psychiatry* 1995;152:1180–1184.

11. Birbaumer N, Grodd W, Diedrich O, et al. fMRI reveals amygdala activation to human faces in social phobics. *Neuroreport* 1998;9:1223–1226.

12. Bisaga A, Katz JL, Antonini A, et al. Cerebral glucose metabolism in women with panic disorder. *Am J Psychiatry* 1998;155:1178–1183.

13. Brandt CA, Meller J, Keweloh L, et al. Increased benzodiazepine receptor density in the prefrontal cortex in patients with panic disorder. *J Neural Transm* 1998;105:1325–33.

14. Breiter HC, Etcoff, NL, Whalen PJ, et al. Response and habituation of the human amygdala during visual processing of facial expression. *Neuron* 1996;17:1–20.

15. Breiter HC, Rauch SL, Kwong KK, et al. Functional magnetic resonance imaging of symptom provocation in obsessive compulsive disorder. *Arch Gen Psychiatry* 1996;53:595–606.

16. Bremner JD. Does stress damage the brain? *Biol Psychiatry* 1999;45:797–805.

17. Bremner JD, Innis RB, Ng CK, et al. Positron emission tomography measurement of cerebral metabolic correlates of yohimbine administration in combat-related posttraumatic stress disorder. *Arch Gen Psychiatry* 1997;54:246–254.

18. Bremner JD, Innis RB, Southwick SM, et al. Decreased benzodiazepine receptor binding in prefrontal cortex in combat-related posttraumatic stress disorder. *Am J Psychiatry* 2000;157:1120–1126.

19. Bremner JD, Innis RB, White T, et al. SPECT [I-123] iomazenil measurement of the benzodiazepine receptor in panic disorder. *Biol Psychiatry* 2000;47:96–106.

20. Bremner JD, Narayan M, Staib LH, et al. Neural correlates of memories of childhood sexual abuse in women with and without posttraumatic stress disorder. *Am J Psychiatry* 1999;156:1787–1795.

21. Bremner JD, Randall P, Scott TM, et al. MRI-based measurement of hippocampal volume in patients with combat-related posttraumatic stress disorder. *Am J Psychiatry* 1995;152:973–981.

22. Bremner JD, Randall P, Vermetten E, et al. Magnetic resonance imaging-based measurement of hippocampal volume in posttraumatic stress disorder related to childhood physical and sexual abuse: a preliminary report. *Biol Psychiatry* 1997;41:23–32.

23. Bremner JD, Staib LH, Kaloupek D, et al. Neural correlates of exposure to traumatic pictures and sound in vietnam combat veterans with and without posttraumatic stress disorder: a positron emission tomography study. *Biol Psychiatry* 1999;45:806–816.

24. Brody AL, Saxena S, Schwartz JM, et al. FDG-PET predictors of response to behavioral therapy versus pharmacotherapy in obsessive-compulsive disorder. *Psychiatry Res Neuroimaging* 1998;84:1–6.

25. Brown ES, Rush AJ, McEwen BS. Hippocampal remodeling and damage by corticosteroids: implications for mood disorders. *Neuropsychopharmacology* 1999;21:474–484.

26. Büchel C, Morris J, Dolan RJ, et al. Brain systems mediating aversive conditioning: an event-related fMRI study. *Neuron* 1998;20:947–957.

27. Büchel C, Dolan RJ. Classical fear conditioning in functional neuroimaging. *Curr Opin Neurobiol* 2000;10:219–223.

28. Büchel C, Dolan RJ, Armony JL, et al. Amygdala-hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *J Neurosci* 1999;19:10869–10876.

29. Cahill L, Haier RJ, Fallon J, et al. Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proc Natl Acad Sci USA* 1996;93:8016–8021.

30. Charney DS, Bremner JD, Redmond DE. Noradrenergic neural substrates for anxiety and fear: clinical associations based on pre-clinical research. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven, 1995:387–396.

31. Coplan JD, Lydiard RB. Brain circuits in panic disorder. *Biol Psychiatry* 1998;44:1264–1276.

32. Dager SR, Strauss WL, Marro KI, et al. Proton magnetic resonance spectroscopy investigation of hyperventilation in subjects with panic disorder and comparison subjects. *Am J Psychiatry* 1995;152:666–672.

33. Dager SR, Friedman SD, Heide A, et al. Two-dimensional proton echo-planar spectroscopic imaging of brain metabolic changes during lactate-induced panic. *Arch Gen Psychiatry* 1999;56:70–77.

34. Davidson JR, Krishnan KR, Charles HC, et al. Magnetic resonance spectroscopy in social phobia: preliminary findings. *J Clin Psychiatry* 1993;54[Suppl]:19–25.

35. Davis M, Falls WA, Campeau S, et al. Fear-potentiated startle: a neural and pharmacological analysis. *Behav Brain Res* 1993;58:175–198.

36. DeBellis MD, Keshavan MS, Clark DB, et al. Developmental traumatology. II. Brain development. *Biol Psychiatry* 1999;45:1271–1284.

37. De Bellis MD, Keshavan MS, Spencer S, et al. *N*-acetylaspartate concentration in the anterior cingulate of maltreated children and adolescents with PTSD. *Am J Psychiatry* 2000;157:1175–1177.

38. De Cristofaro MT, Sessarego A, Pupi A, et al. Brain perfusion abnormalities in drug-naive, lactate-sensitive panic patients: a SPECT study. *Biol Psychiatry* 1993;33:505–512.

39. Dougherty DD, Shin LM, Alpert NM, et al. Anger in healthy men: a PET study using script-driven imagery. *Biol Psychiatry* 1999;46:466–472.

40. Drevets WC, Videen TO, MacLeod AK, et al. PET images of blood flow changes during anxiety: a correction. *Science* 1992;256:1696.

41. Ebert D, Speck O, Konig A, et al. ^{1}H-magnetic resonance spectroscopy in obsessive-compulsive disorder: evidence for neuronal loss in the cingulate gyrus and the right striatum. *Psychiatry Res* 1997;74:173–176.

42. Endo Y, Nishimura J-I, Kobayashi S, et al. Chronic stress exposure influences local cerebral blood flow in the rat hippocampus. *Neuroscience* 1999;93:551–555.

43. Feldman S, Conforti N. Participation of the dorsal hippocampus in the glucocorticoid feedback effect on adrenocortical activity. *Neuroendocrinology* 1980;30:52–55.

44. Fischer H, Andersson JL, Furmark T, et al. Brain correlates of an unexpected panic attack: a human positron emission tomographic study. *Neurosci Lett* 1998;251:137–140.

45. Fischer H, Furmark T, Wik G, et al. Brain representation of habituation to repeated complex visual stimulation studied with PET. *Neuroreport* 2000;11:123–126.

46. Fischer H, Wik G, Fredrikson M. Functional neuroanatomy of robbery re-experience: affective memories studied with PET. *Neuroreport* 1996;7:2081–2086.

47. Fischer H, Wright CI, Whalen PJ, et al. Effects of repeated presentations of facial stimuli on human brain function: an fMRI study. *Neuroimage* 2000;11:S250.

48. Fontaine R, Breton G, Dery R, et al. Temporal lobe abnormalities in panic disorder: an MRI study. *Biol Psychiatry* 1990;27:304–310.

49. Fredrikson M, Wik G, Fischer H, et al. Affective and attentive neural networks in humans: a PET study of pavlovian conditioning. *Neuroreport* 1995;7:97–101.

50. Fredrikson M, Wik G, Greitz T, et al. Regional cerebral blood flow during experimental fear. *Psychophysiology* 1993;30:126–130.

51. Fredrikson M, Wik G, Annas P, et al. Functional neuroanatomy of visually elicited simple phobic fear: additional data and theoretical analysis. *Psychophysiology* 1995;32:43–48.

52. Furmark T, Fischer H, Wik G, et al. The amygdala and individual differences in fear conditioning. *Neuroreport* 1997;8:3957–3960.

53. Fyer AJ. Current approaches to etiology and pathophysiology of specific phobia. *Biol Psychiatry* 1998;44:1295–1304.

54. George MS, Ketter TA, Parekh PI, et al. Brain activity during transient sadness and happiness in healthy women. *Am J Psychiatry* 1995;152:341–351.

55. George MS, Ketter TA, Parekh PI, et al. Gender differences in regional cerebral blood flow during transient self-induced sadness or happiness. *Biol Psychiatry* 1996;40:859–871.

56. Gewirtz JC, Falls WA, Davis M. Normal conditioned inhibition and extinction of freezing and fear-potentiated startle following electrolytic lesions of medial prefrontal cortex in rats. *Behav Neurosci* 1997;111:712–726.

57. Gould E, Tanapat P. Stress and hippocampal neurogenesis. *Biol Psychiatry* 1999;46:1472–1479.

58. Gurvits TV, Shenton ME, Hokama H, et al. Magnetic resonance imaging study of hippocampal volume in chronic, combat-related posttraumatic stress disorder. *Biol Psychiatry* 1996;40:1091–1099.

59. Hamann SB, Ely TD, Grafton ST, et al. Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nat Neurosci* 1999;2:289–293.

60. Herman J, Schafer M, Young E, et al. Evidence for hippocampal regulation of neuroendocrine neurons of hypothalamo-pituitary-adrenocortical axis. *J Neurosci* 1989;9:3072–3082.

61. Hoehn-Saric R, Pearlson GD, Harris GJ, et al. Effects of fluoxetine on regional cerebral blood flow in obsessive-compulsive patients. *Am J Psychiatry* 1991;148:1243–1245.

62. Hugdahl K, Berardi A, Thompson WL, et al. Brain mechanisms in human classical conditioning: a PET blood flow study. *Neuroreport* 1995:6:1723–1728.

63. Irwin W, Davidson RJ, Lowe MJ, et al. Human amygdala activation detected with echo-planar functional magnetic resonance imaging. *Neuroreport* 1996;7:1765–1769.

64. Jenike MA, Breiter HC, Baer L, et al. Cerebral structural abnormalities in obsessive-compulsive disorder: a quantitative morphometric magnetic resonance imaging study. *Arch Gen Psychiatry* 1996;53:625–632.

65. Kent JM, Coplan JD, Gorman JM. clinical utility of the selective serotonin reuptake inhibitors in the spectrum of anxiety. *Biol Psychiatry* 1998;44:812–824.

66. Ketter TA, Andreason PJ, George MS, et al. Anterior paralimbic mediation of procaine-induced emotional and psychosensory experiences. *Arch Gen Psychiatry* 1996;53:59–69.

67. Kimbrell TA, George MS, Parekh PI, et al. Regional brain activity during transient self-induced anxiety and anger in healthy adults. *Biol Psychiatry* 1999;46:454–465.

68. Kuikka JT, Pitkanen A, Lepola U, et al. Abnormal regional benzodiazepine receptor uptake in the prefrontal cortex in patients with panic disorder. *Nucl Med Commun* 1995;16:273–280.

69. LaBar KS, Gatenby C, Gore JC, et al. Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron* 1998;20:937–945.

70. Lane RD, Fink GR, Chau PM-L, et al. Neural activation during selective attention to subjective emotional responses. *Neuroreport* 1997;8:3969–3972.

71. Lane RD, Reiman EM, Ahern GL, et al. Neuroanatomical correlates of happiness, sadness, and disgust. *Am J Psychiatry* 1997;154:926–933.

72. Lane RD, Reiman EM, Bradley MM, et al. Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia* 1997;35:1437–1444.

73. Lang PJ, Bradley MM, Cuthbert BN. *The International Affective Picture System (IAPS): photographic slides*. Gainesville, FL: University of Florida, 1995.

74. LeDoux JE: Emotion and the amygdala. In: Aggleton JP, ed. *The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction*. New York: Wiley-Liss, 1992:339–351.

75. LeDoux JE. *The emotional brain*. New York: Simon and Schuster, 1996.

76. Liberzon I, Taylor SF, Amdur R, et al. Brain activation in PTSD in response to trauma-related stimuli. *Biol Psychiatry* 1999;45:817–826.

77. Liotti M, Mayberg HS, Brannan SK, et al. Differential limbic-cortical correlates of sadness and anxiety in healthy subjects: implications for affective disorders. *Biol Psychiatry* 2000;48:30–42.

78. Machlin SR, Harris GJ, Pearlson GD, et al. Elevated medial-frontal cerebral blood flow in obsessive-compulsive patients: a SPECT study. *Am J Psychiatry* 1991;148:1240–1242.

79. Malizia AL, Cunningham VJ, Bell CJ, et al. Decreased brain GABA(A)-benzodiazepine receptor binding in panic disorder: preliminary results from a quantitative PET study. *Arch Gen Psychiatry* 1998;55:715–720.

80. McGuire PK, Bench CJ, Frith CD, et al. Functional anatomy of obsessive-compulsive phenomena. *Br J Psychiatry* 1994;164:459–468.

81. Mesulam M-M. Patterns in behavioral neuroanatomy: association areas, the limbic system, and hemispheric specialization. In: Mesulam M-M, ed. *Principles of behavioral neurology*. Philadelphia: FA Davis, 1985:1–70.

82. Moore GJ, MacMaster FP, Stewart C, et al. Case study: caudate glutamatergic changes with paroxetine therapy for pediatric obsessive-compulsive disorder. *Am Acad Child Adolesc Psychiatry* 1998;37:663–667.

83. MorganMA, LeDoux JE. Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Behav Neurosci* 1995;109:681–688.

84. Morgan MA, Romanski LM, LeDoux JE. Extinction of emotional learning: contribution of medial prefrontal cortex. *Neurosci Lett* 1993;163:109–113.

85. Morris JS, Friston KJ, Dolan RJ. Neural responses to salient visual stimuli. *Proc R Soc Lond B Biol Sci* 1997;264:769–775.

86. Morris JS, Frith CD, Perrett DI, et al. A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* 1996;383:812–815.

87. Morris JS, Öhman A, Dolan RJ. Conscious and unconscious emotional learning in the human amygdala. *Nature* 1998;393:467–470.

88. Mountz JM, Modell JG, Wilson MW, et al. Positron emission tomographic evaluation of cerebral blood flow during state anxiety in simple phobia. *Arch Gen Psychiatry* 1989;46:501–504.

89. Nordahl TE, Benkelfat C, Semple W, et al. Cerebral glucose metabolic rates in obsessive-compulsive disorder. *Neuropsychopharmacology* 1989;2:23–28.

90. Nordahl TE, Semple WE, Gross M, et al. Cerebral glucose metabolic differences in patients with panic disorder. *Neuropsychopharmacology* 1990;3:261–272.

91. Nordahl TE, Stein MB, Benkelfat C, et al. Regional cerebral metabolic asymmetries replicated in an independent group of patients with panic disorders. *Biol Psychiatry* 1998;44:998–1006.

92. Paradiso S, Johnson DL, Andreasen NC, et al. Cerebral blood flow changes associated with attribution of emotional valence to pleasant, unpleasant, and neutral visual stimuli in a PET study of normal subjects. *Am J Psychiatry* 1999;156:1618–1629.

93. Paradiso S, Robinson RG, Andreasen NC, et al. Emotional activation of limbic circuitry in elderly normal subjects in a PET study. *Am J Psychiatry* 1997;154:384–389.

94. Pardo JV, Pardo PJ, Raichle ME. Neural correlates of selfinduced dysphoria. *Am J Psychiatry* 1993;150:713–719.

95. Perani D, Colombo C, Bressi S, et al. FDG PET study in obsessive-compulsive disorder: a clinical metabolic correlation study after treatment. *Br J Psychiatry* 1995;166:244–250.

96. Potts NL, Davidson JR, Krishnan KR, et al. Magnetic resonance imaging in social phobia. *Psychiatry Res* 1994;52:35–42.

97. Rauch SL. Neuroimaging and the neurobiology of anxiety disorders. In: Davidson RJ, Scherer K, Goldsmith HH, eds. *Handbook of affective sciences*. New York: Oxford University Press, in press.

98. Rauch SL, Baxter LR. Neuroimaging of OCD and related disorders. In: Jenike MA, Baer L, Minichiello WE, eds. *Obsessive-compulsive disorders: theory and management*. Boston: CV Mosby, 1998:289–317.

99. Rauch SL, Jenike MA, Alpert NM, et al. Regional cerebral blood flow measured during symptom provocation in obsessive-compulsive disorder using ^{15}O-labeled CO_2 and positron emission tomography. *Arch Gen Psychiatry* 1994;51:62–70.

100. Rauch SL, Savage CR, Alpert, NM, et al. A positron emission tomographic study of simple phobic symptom provocation. *Arch Gen Psychiatry* 1995;52:20–28.

101. Rauch SL, Savage CR, Brown HD, et al. A PET investigation of implicit and explicit sequence learning. *Hum Brain Mapping* 1995;3:271–286.

102. Rauch SL, Savage CR, Alpert NM, et al. Probing striatal function in obsessive compulsive disorder: a PET study of implicit sequence learning. *J Neuropsychiatry* 1997;9:568–573.

103. Rauch SL, Shin LM, Whalen PJ, et al. Neuroimaging and the neuroanatomy of PTSD. *CNS Spectrums* 1998;3[Suppl 2]:30–41.

104. Rauch SL, van der Kolk BA, Fisler RE, et al. A symptom provocation study of posttraumatic stress disorder using positron emission tomography and script-driven imagery. *Arch Gen Psychiatry* 1996;53:380–387.

105. Rauch SL, Whalen PJ, Curran T, et al. Probing striato-thalamic function in OCD and TS using neuroimaging methods. In: Cohen DJ, Goetz C, Jankovic J, eds. *Tourette syndrome and associated disorders*. Philadelphia: Lippincott Williams & Wilkins.

106. Rauch SL, Whalen PJ, Dougherty DD, et al. Neurobiological models of obsessive compulsive disorders. In: Jenike MA, Baer L, Minichiello WE, eds. *Obsessive-compulsive disorders: practical management*. Boston: CV Mosby, 1998:222–253.

107. Rauch SL, Whalen PJ, Savage CR, et al. Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. *Hum Brain Mapping* 1997;5:124–132.

108. Rauch SL, Whalen PJ, Shin LM, et al. Exaggerated amygdala response to masked fearful vs. happy facial stimuli in posttraumatic stress disorder: a functional MRI study. *Biol Psychiatry* 2000;47:769–776.

109. Reiman EM, Raichle ME, Robins E, et al. The application of positron emission tomography to the study of panic disorder. *Am J Psychiatry* 1986;143:469–477.

110. Reiman EM, Raichle ME, Robins E, et al. Neuroanatomical correlates of a lactate-induced anxiety attack. *Arch Gen Psychiatry* 1989;46:493–500.

111. Robinson D, Wu H, Munne RA, et al. Reduced caudate nucleus volume in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1995;52:393–398.

112. Rosenberg DR, Keshevan MS, O'Hearn KM, et al. Frontostriatal measurement in treatment-naive children with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1997;554:824–830.

113. Rubin RT, Villaneuva-Myer J, Ananth J, et al. Regional xenon-133 cerebral blood flow and cerebral technetium 99m HMPAO uptake in unmedicated patients with obsessive-compulsive disorder and matched normal control subjects. *Arch Gen Psychiatry* 1992;49:695–702.

114. Salzman C, Miyawaki EK, le Bars P, et al. Neurobiologic basis of anxiety and its treatment. *Harv Rev Psychiatry* 1993;1:197–206.

115. Sapolsky RM, Krey LC, McEwen BS. Prolonged glucocorticoid exposure reduces hippocampal neuron number: implications for aging. *J Neurosci* 1985;5:1222–1227.

116. Sapolsky RM, Krey LC, McEwen BS. The neuroendocrinology of stress and aging: the glucocorticoid cascade hypothesis. *Endocr Rev* 1986;7:284–301.

117. Sapolsky RM, Uno H, Rebert CS, et al. Hippocampal damage associated with prolonged glucocorticoid exposure in primates. *J Neurosci* 1990;10:2897–2902.

118. Saxena S, Brody AL, Maidment KM, et al. Localized orbitofrontal and subcortical metabolic changes and predictors of response to paroxetine treatment in obsessive-compulsive disorder. *Neuropsychopharmacology* 1999;21:683–693.

119. Scarone S, Colombo C, Livian S, et al. Increased right caudate nucleus size in obsessive compulsive disorder: detection with magnetic resonance imaging. *Psychiatry Res Neuroimaging* 1992;45:115–121.

120. Schneider F, Gur RE, Mozley LH, et al. Mood effects on limbic blood flow correlate with emotional self-rating: a PET study with oxygen-15 labeled water. *Psychiatry Res Neuroimaging* 1995;61:265–283.

121. Schneider F, Grodd W, Weiss U, et al. Functional MRI reveals left amygdala activation during emotion. *Psychiatry Res Neuroimaging* 1997;76:75–82.

122. Schneider F, Weiss U, Kessler C, et al. Subcortical correlates of differential classical conditioning of aversive emotional reactions in social phobia. *Biol Psychiatry* 1999;45:863–71.

123. Schuff N, Marmar CR, Weiss DS, et al. Reduced hippocampal volume and *N*-acetyl aspartate in posttraumatic stress disorder. *Ann NY Acad Sci* 1997;821:1997:516–520.

124. Schwartz JM, Stoessel PW, Baxter LR, et al. Systematic changes in cerebral glucose metabolic rate after successful behavior modification. *Arch Gen Psychiatry* 1996;53:109–113.

125. Semple WE, Goyer P, McCormick R, et al. Preliminary report: brain blood flow using PET in patients with posttraumatic stress disorder and substance-abuse histories. *Biol Psychiatry* 1993;34:115–118.

126. Servan-Schreiber D, Perlstein WM, Cohen JD, et al. Selective pharmacological activation of limbic structures in human volunteers: a positron emission tomography study. *J Neuropsychiatry Clin Neurosci* 1998;10:148–159.

127. Shin LM, Dougherty D, Macklin ML, et al. Activation of anterior paralimbic structures during guilt-related script-driven imagery. *Biol Psychiatry* 2000;48:43–50.

128. Shin LM, Kosslyn SM, McNally RJ, et al. Visual imagery and perception in posttraumatic stress disorder: a positron emission tomographic investigation. *Arch Gen Psychiatry* 1997;54:233–241.

129. Shin LM, McNally RJ, Kosslyn SM, et al. Regional cerebral blood flow during script-driven imagery in childhood sexual abuse-related posttraumatic stress disorder: a PET investigation. *Am J Psychiatry* 1999;156:575–584.

130. Starkman MN, Gebarski SS, Berent S, et al. Hippocampal formation volume, memory dysfunction, and cortisol levels in patients with Cushing's syndrome. *Biol Psychiatry* 1992;32:756–765.

131. Starkman MN, Giordani B, Gebarski SS, et al. Decrease in cortisol reverses human hippocampal atrophy following treatment of Cushing's disease. *Biol Psychiatry* 1999;46:1595–1602.

132. Stein MB. Neurobiological perspectives on social phobia: from affiliation to zoology. *Biol Psychiatry* 1998;44:1277–1285.

133. Stein MB, Leslie WD. A brain SPECT study of generalized social phobia. *Biol Psychiatry* 1996;39:825–828.

134. Stein MB, Koverola C, Hanna C, et al. Hippocampal volume in women victimized by childhood sexual abuse. *Psychol Med* 1997;27:951–960.

135. Stewart RS, Devous MD Sr, Rush AJ, et al. Cerebral blood flow changes during sodium-lactate–induced panic attacks. *Am J Psychiatry* 1988;145:442–449.

136. Swedo SE, Shapiro MB, Grady CL, et al. Cerebral glucose metabolism in childhood-onset obsessive-compulsive disorder. *Arch Gen Psychiatry* 1989;46:518–523.

137. Swedo SE, Pietrini P, Leonard HL, et al. Cerebral glucose metabolism in childhood-onset obsessive-compulsive disorder: revisualization during pharmacotherapy. *Arch Gen Psychiatry* 1992;49:690–694.

138. Tiihonen J, Kuikka J, Bergstrom K, et al. Dopamine reuptake site densities in patients with social phobia. *Am J Psychiatry* 1997;154:239–242.

139. Uno H, Tarara R, Else J, et al. Hippocampal damage associated with prolonged and fatal stress in primates. *J Neurosci* 1989;9:1705–1711.

140. Watanabe Y, Gould E, McEwen BS. Stress induces atrophy of apical dendrites of hippocampal CA3 pyramidal neurons. *Brain Res* 1992;588:341–345.

141. Whalen PJ. Fear, vigilance, and ambiguity: initial neuroimaging studies of the human amygdala. *Curr Dir Psychol Sci* 1998;7:177–188.

142. Whalen PJ, Rauch SL, Etcoff NL, et al. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *J Neurosci* 1998;18:411–418.

143. Wik G, Fredrikson M, Ericson K, et al. A functional cerebral response to frightening visual stimulation. *Psychiatry Res Neuroimaging* 1993;50:15–24.

144. Woods SW, Koster K, Krystal JK, et al. Yohimbine alters regional cerebral blood flow in panic disorder. *Lancet* 1988;2:678.

145. Woolley CS, Gould E, McEwen BS. Exposure to excess glucocorticoids alters dendritic morphology of adult hippocampal pyramidal neurons. *Brain Res* 1990;531:225–231.

146. Wright CI, Fischer H, Whalen PJ, et al. Suppression of human brain activity by repeatedly presented emotional facial expressions. *Neuroimage* 2000;11:S252.

147. Yehuda R. Neuroendocrinology of trauma and posttraumatic stress disorder. In: Yehuda R, ed. *Psychological trauma*. Washington, DC: American Psychiatric Press, 1998:97–131.

66

Current and Emerging Therapeutics of Anxiety and Stress Disorders

Jack M. Gorman

Justine M. Kent

Jeremy D. Coplan

Jack M. Gorman, Justine M. Kent, and Jeremy D. Coplan: Columbia University, New York State Psychiatric Institute, New York, New York.

During the 1960s and 1970s the concept of “pharmacologic dissection” became popular as a putative method for differentiating among different categories of psychiatric illness. At the time, it was widely held that anxiety disorders, but not depression, respond to benzodiazepines, whereas depression, but not anxiety disorders, responds to antidepressants. Panic disorder was held to be the one exception, responding only to antidepressants. Alprazolam, selective serotonin reuptake inhibitors (SSRIs), and buspirone had not yet been tested. On the basis of these observations, it was asserted that anxiety and depression are clearly distinct categories of illness.

Thirty years later we find that the situation has changed dramatically. The first inconsistency with the notion of pharmacologic dissection was the clear finding that panic disorder does indeed respond to benzodiazepines. For a while, alprazolam and then clonazepam were regarded by some authorities and clinicians as first-line therapies for panic disorder, replacing tricyclics and monoamine oxidase inhibitors.

An even more potent challenge, however, has come from the evidence not only that anxiety disorders respond to antidepressants but also that antidepressants work better than benzodiazepines for most of them. As this chapter discusses, antidepressants are now considered the appropriate pharmacologic intervention for panic disorder, social anxiety disorder, posttraumatic stress disorder (PTSD), and generalized anxiety disorder (GAD). The latter case is particularly interesting. Once considered the sole domain of benzodiazepines, GAD was then shown to be responsive to a drug in an entirely new category, with no relationship whatsoever to the benzodiazepine receptor or γ-aminobutyric acid (GABA)—buspirone. At about the same time evidence began to accumulate from just a few studies that GAD might also respond to antidepressants. This evidence was largely ignored, and pharmaceutical companies were advised to stay away from GAD, a condition supposedly so placebo-responsive that no drug would ever be shown effective in large clinical trials. To the contrary, venlafaxine is now approved by the Food and Drug Administration (FDA) for the treatment of GAD, and there is also evidence for the efficacy of paroxetine.

At this point, rather than “dissecting” among the anxiety disorders or between anxiety disorders and depression, pharmacologic grounds might lead one to assume that these conditions are variants of each other. This also would probably be an oversimplification. Anxiety and depression are different and we can make distinctions among the anxiety disorders. Nevertheless, the finding that antidepressants are so ubiquitously effective across categories raises interesting questions and challenges. We review here the evidence for responsiveness of four anxiety disorders to medication.

PANIC DISORDER

Part of "66 - Current and Emerging Therapeutics of Anxiety and Stress Disorders "

Panic disorder (PD) has a reported lifetime prevalence of between 1.5% and 3.5% (1 ,2), is highly comorbid with major depression, and is associated in its own right with significant impairment in psychosocial functioning independent of depressive symptomatology. In the Epidemiologic Catchment Area study, subjective reports of patients with PD indicate that approximately one-third experience poor physical and emotional health, rates comparable with major depression (2).

Historical Notes

Recognized as a distinct disorder that could be distinguished from the general diagnosis of “anxious neurosis,” in part

through the pharmacologic dissection work of Klein and Fink (3 ,4) in the 1960s, PD was first categorized as a discrete diagnostic entity in the *Diagnostic and Statistical Manual of Mental Disorders*, third edition (DSM-III), in 1980. Despite some minor changes in diagnostic criteria in the third edition revised (DSM-III-R) and the fourth edition (DSM-IV), primarily involving the number and frequency of attacks required, the major criteria remain essentially the same. The key triad of symptoms are (a) the occurrence of spontaneous panic attacks; (b) the presence of anticipatory anxiety; and (c) the presence of phobic avoidance, resulting in some degree of functional impairment. The pharmacologic treatment of PD has evolved dramatically since the heterocyclic antidepressants were first established as possessing powerful antipanic properties in the early 1960s (4). Throughout the 1970s and 1980s, the heterocyclic antidepressants continued to be the mainstay of pharmacologic treatment of PD, with the monoamine oxidase inhibitors (MAOIs) used primarily in patients who failed trials of heterocyclic antidepressants. The high potency benzodiazepines were increasingly prescribed as both primary and adjunct treatments throughout this same time period. With the introduction of the SSRIs in the United States in the late 1980s and early 1990s for the treatment of depression, this class of drug began being used in the treatment of PD with promising results. In the late 1990s, several large-scale, controlled trials established the SSRIs to be effective and safe treatments for PD, thus supplanting the heterocyclic antidepressants and benzodiazepines as first-line treatment. Although the serotonin, norepinephrine, and GABA systems remain the traditional targets for the majority of antipanic medications, widely different classes of drugs targeting an array of neurochemical systems are now being explored as potential treatments for PD.

Heterocyclic Antidepressants

Numerous controlled trials have confirmed the efficacy of the heterocyclic antidepressants, since the initial observations of Klein (4), in both the acute and long-term treatment of PD. In general, heterocyclics with greatest serotoninergic reuptake inhibition effect, such as imipramine and clomipramine, appear to be most effective in the treatment of PD (5 ,6 and 7). By far the best studied of this class of antidepressants is imipramine, which due to its well-established efficacy has been generally accepted as the gold standard of PD treatment (8). In the Cross-National Collaborative Panic Study, more than 1,000 patients in 14 countries were randomized into a study comparing imipramine, alprazolam, and placebo (9). At the study's end, imipramine and alprazolam were found to have comparable efficacy, and both were significantly superior to placebo on most outcome measures. A positive dose–response relationship between imipramine levels and clinical improvement has been reported, with plasma levels of 140 mg/mL associated with the greatest improvement in panic symptoms (10). Although several other of the heterocyclic antidepressants have been used in the treatment of PD (amitryptyline, desipramine, nortriptyline, clomipramine), far less controlled data are available supporting their efficacy (11 ,12). Although effective, side effects often limit the use of this class of medication in the treatment of PD. This is particularly true in the case of clomipramine, which due to its anticholinergic and antihistaminergic effects can be difficult for patients to tolerate (13). The use of the heterocyclic antidepressants is often limited by the presence of comorbid medical conditions such as cardiac disease and glaucoma. Lethality in overdose is another concern given the reported high rates of suicide in this population when depression is comorbid (14 ,15 and 16). Although the SSRIs are often touted as offering a more tolerable side-effect profile than the heterocyclics, the side-effect burden of imipramine has recently been shown to be comparable to, although different in nature from, that of the SSRIs, with most side effects (with the exception of dry mouth, sweating, and constipation) not persisting beyond the first few weeks of treatment (17).

Monoamine Oxidase Inhibitors

Like the heterocyclic antidepressants, the MAOIs, are clearly established to be effective in the treatment of PD, yet have yielded to newer antidepressants with similar antipanic efficacies but less drug–drug and dietary interactions. Among the MAOIs, phenelzine is the best studied in PD, and its efficacy in the acute treatment of PD is supported by several studies (18 ,19 and 20). Dietary restrictions, lethality in overdose, and drug interaction concerns limit the widespread use of the traditional MAOIs. Stemming from these concerns, the reversible inhibitors of MAOI (RIMAs) were developed and have demonstrably fewer drug–drug and dietary interactions. These include moclobemide and brofaromine, neither of which is currently marketed in the United States, but are used extensively throughout Europe and other parts of the world. The efficacy of moclobemide has been shown in the treatment of PD in placebo-controlled studies (21), and moclobemide has been found to be comparable in efficacy to clomipramine (21) and fluoxetine (22). Moclobemide has also been shown to be as effective as fluoxetine in maintenance treatment of PD (23). Brofaromine was shown to be comparable to fluvoxamine (24) and clomipramine (25) in small randomized, double-blind studies lacking a placebo. In a placebo-controlled study of the efficacy of brofaromine, 30 patients with PD (DSM-III-R definition) were treated for 12 weeks. Although there was no significant reduction in the number of panic attacks for those patients treated with brofaromine, patients demonstrated clinical improvement on several other measures, including agoraphobic avoidance (26).

Benzodiazepines

The high-potency benzodiazepines, another mainstay of treatment for PD, have been shown to be effective, well tolerated, and safe. In the absence of comorbid substance abuse, concerns about abuse potential have proven largely unfounded in this population (27 ,28 ,29 and 30). Among the benzodiazepines, alprazolam and clonazepam are labeled for the treatment of panic disorder and have been shown in numerous, controlled trials to be effective treatments (31 ,32 ,33 ,34 ,35 and 36). Clonazepam, with a long half-life of 20 to 50 hours, allows fewer doses per day than the short-acting alprazolam, and may reduce the likelihood of rebound symptoms between doses. The benzodiazepines have repeatedly been shown to offer an early advantage in the treatment of anxiety by providing almost immediate relief of anxiety-related somatic symptoms such as muscle tension and insomnia (27 ,37 ,38). However, in the long term, antidepressants may offer the advantage of better targeting and relief of psychic symptoms of anxiety (37), and provide the added benefit of treating associated depressive symptoms. Discontinuing long-term pharmacotherapy with benzodiazepines can be difficult, with as many as a third of patients with PD being unable to discontinue use due to dependence/withdrawal (27). Thus, despite their efficacy and safety, many clinicians remain concerned about the risk of dependence (39).

Selective Serotonin Reuptake Inhibitors (SSRIs)

Among the antidepressants currently used in the treatment of PD, the SSRIs have become first-line treatments (40 ,41). Following in the wake of numerous promising open and controlled trials, several large, multicenter, placebo-controlled studies involving hundreds of subjects each have demonstrated the efficacy of the SSRIs in the treatment of PD (Table 66.1) (42 ,43 ,44 ,45 ,46 ,47 ,48 ,49 and 50). Both sertraline and paroxetine are labeled in the United States for the treatment of PD, and citalopram is approved in several European countries for this indication. Although fluvoxamine has not been studied in clinical trials on the scale of the other SSRIs, it has been shown to be efficacious in smaller (<100 subjects), randomized, placebo-controlled studies (25 ,44). Despite their established efficacy in the treatment of PD, there are certain problems inherent in prescribing the SSRIs in patients with PD. Of primary concern on the initiation of treatment is the commonly observed anxiogenic effect, which, despite being dose-dependent, can nonetheless make the initial several days of treatment a challenge for patients. Initiating treatment with a dose of one-fourth to one-half that of the normal starting dose may greatly reduce the patient's feelings of restlessness and increased anxiety. The delayed onset of anxiolytic action contributes to making the period of SSRI initiation difficult for patients with PD. Other effects such as sexual dysfunction and weight gain with long-term use can make these drugs as problematic as earlier classes for some patients. Overall, though, the side-effect burden associated with SSRI treatment has been shown to be more tolerable for most patients than the heterocyclics and benzodiazepines (51). Because the SSRIs have been associated with a discontinuation syndrome characterized by anxiety, tremor, dizziness, paresthesias, nausea, and other symptoms when abruptly stopped, these medications should be tapered over a few weeks, if possible, to minimize discontinuation symptoms (38).

TABLE 66.1. EFFICACY OF THE SSRIs IN THE ACUTE TREATMENT OF PANIC DISORDER BASED ON LARGE-SCALE, PLACEBO-CONTROLLED STUDIES

SSRI	Investigators	Study Design	Dose Range	Outcome
Fluoxetine	Michelson et al., 1998	Multicenter, 10-week study of 243 patients	Fixed dose: 10 or 20 mg/day	20-mg dose was most effective, demonstrating significant change versus placebo on panic attack frequency, CGI improvement scores, Hamilton Anxiety and Depression scores, phobic symptoms, and functional impairment as measured by the Sheehan Disability Scale (family life and social life); the two fluoxetine groups did not differ from placebo in the number of patients who were panic-free at endpoint
	Michelson et al., 2000	Multicenter, 12-week study of 180 patients	Flexible dose: 20–60 mg/day (mean dose = 20 mg)	A significantly greater percentage of patients on fluoxetine were panic-free at endpoint (42% versus 28% on placebo); significant change versus placebo were found for the CGI Severity, HAM-A, State Anxiety Inventory, and Sheenhan Disability Scale (work and social impairment)
Fluvoxamine	Black et al., 1993	Multicenter, 8-week study of 75 patients randomized to either fluvoxamine, cognitive therapy, or placebo	Flexible dose: up to 300 mg/day (mean dose = 230 mg)	Fluvoxamine was superior to placebo and cognitive therapy at endpoint as measured by the Clinical Anxiety Scale and CGI Severity and Improvement scales; fluvoxamine was superior to placebo as measured by a greater reduction in mean panic attack severity, and number of panic-free patients; fluvoxamine-treated patients demonstrated significant reductions versus placebo on several measures of the Sheehan Disability Scale (work, social/leisure)
Paroxetine	Oehrberg et al., 1995	Multicenter, 12-week study of 120 patients—all received cognitive therapy	Flexible dose: 20–60 mg/day	Number of patients with at least 50% reduction in panic attacks was significantly greater in the paroxetine-treated than placebo group, as was number of patients who had only one or no panic attacks during the final study week. The paroxetine group had significantly greater mean reductions in HAM-A and CGI scores versus placebo
	Ballenger et al., 1998	Multicenter, 10-week study of 278 patients	Fixed dose: 10, 20, or 40 mg/day	40-mg group demonstrated the greatest effects, with significant improvement versus placebo on measures of reduction in number of panic attacks, intensity of attacks, CGI Severity and Improvement scores, phobic fear score, HAM-A score, and MADRS
Sertraline	Pollack et al., 1998	Multicenter, 10-week study of 176 patients	Flexible dose: 50–200 mg	Significant decreases versus placebo at endpoint in number of panic attacks, CGI Improvement and Severity scales, PDSS scores; sertraline also demonstrated superiority on improvement in quality of life scores
	Sheikh et al., 2000	Pooled data from two 12-week studies with a total of 322 patients	Fixed dose: 50, 100, or 200 mg/day	All three doses of sertraline were superior to placebo on frequency of panic attacks, CGI Improvement, and change in panic burden (attack frequency X severity), with no consistent dose-response effect
Citalopram	Wade et al., 1997	Multicenter, 8-week comparative study with clomipramine and placebo of 475 patients	Fixed dose: citalopram 10 or 15 mg/day, 20 or 30 mg/day, 40 or 60 mg/day; clomipramine 60 or 90 mg/day	Patients treated with citalopram 20/30 or 40/60 mg/day were comparable to clomipramine and superior to placebo as measured by the number of patients panic-free in the final week of treatment, and by mean reduction in HAM-A total and psychic subscale, and the MADRS; only the 40/60-mg/day dose demonstrated superiority to placebo, along with clomipramine, on the HAM-A somatic subscale; on the Physician's Global Improvement Scale and Patient's Global Improvement Scale, the 20/30-mg dose exhibited greater effects; however, both the 20/30- and 40/60-mg/day doses were superior to placebo

CGI, Clinical Global Impression; HAM, Hamilton Anxiety Rating Scale; MADRS, Montgomery-Ashberg Depression Rating Scale; SSRI, selective serotonin reuptake inhibitor.

Newer Antidepressants

Among the newer antidepressants, several have demonstrated promise in PD. Venlafaxine, a serotonin-norepinephrine reuptake inhibitor, has shown efficacy (on some measures) in a small, placebo-controlled study (52). Nefazodone, a weak serotonin-norepinephrine reuptake inhibitor with serotonin receptor subtype 2C ($5\text{-}HT_{2C}$) antagonist properties, has been shown to reduce anxiety in depressed patients with comorbid PD (53). Mirtazapine enhances both noradrenergic and serotoninergic neurotransmission without reuptake inhibition. Results of an open study involving ten patients suggested that mirtazapine might be effective in the treatment of PD (54). More recently, in a double-blind randomized trial comparing mirtazapine and fluoxetine in the treatment of PD, both drugs showed comparable efficacy on the primary outcome measures and on most secondary outcome measures (55). Adverse events differed between treatments, with weight gain occurring more frequently in those patients receiving mirtazapine, whereas nausea and paresthesias occurred more often in those receiving fluoxetine.

Anticonvulsants

Among the anticonvulsants being used in the treatment of PD are valproate and carbamazepine, and the newest anticonvulsants gabapentin, lamotrigine, pregabalin, and vigabatrin. Valproate has shown promise in several open trials (56 ,57 and 58), and one small placebo-controlled study (59). It may be particularly effective when mood instability is comorbid (60). There is far less support for the use of carbamazepine in the treatment of PD, with uncontrolled studies in patients with PD with EEG abnormalities demonstrating some benefit from carbamazepine treatment (58). However, the only controlled trial of carbamazepine in a small number of PD patients (N = 14) did not report a significant difference for carbamazepine versus placebo in reducing panic attack frequency (61). Gabapentin has shown promise (62) and is recognized as having a benign side-effect profile. Lamotrigine, pregabalin, and vigabatrin are currently under investigation.

Beta-Blockers

Although the beta-blockers are more commonly used in the treatment of performance anxiety and as adjunctive treatment in PTSD, a small number of open studies suggest they may be effective in the treatment of PD (63), although they are not considered a first-line treatment.

Future Directions

Several classes of drugs, although initially viewed as promising, have shown limited efficacy in the treatment of PD. These include buspirone, bupropion, ondansetron, and the cholecystokinin (CCK) antagonists (64 ,65 ,66 and 67). A number of new classes of drugs are being studied, including the benzodiazepine partial agonists such as abecarnil and pagaclone, and the corticotropin-releasing hormone (CRH) inhibitors. Experimental agents showing promise in panic-like models in rodents include the group II metabotropic glutamate receptor agonist LY354740 (68) and drugs acting at the neuropeptide receptors, including neuropeptide-Y agonists and neurokinin substance P antagonists (69).

In summary, the expansion of the antipanic armamentarium suggests that, as with many of the psychiatric disorders, there is no single effective treatment of PD. Among the most commonly prescribed classes of drugs for the treatment of PD [benzodiazepines, SSRIs, tricyclic antidepressants (TCAs), and MAOIs], there are probably no major differences in treatment efficacy, with most reported differences in efficacy between classes probably attributable to differences in study design and samples (27). The use of antidepressants, however, and particularly the SSRIs, has supplanted the long-term use of benzodiazepines for this disorder. Antidepressant use has two main advantages over the benzodiazepines: (a) it provides antidepressant benefits in a population highly susceptible to depressive symptomatology and comorbid major depression (70), and (b) it eliminates the difficulties associated with withdrawal symptoms upon benzodiazepine discontinuation. In the case of comparable efficacy, medication choice is based on factors such as latency to onset of therapeutic action, safety, and the individual side-effect profiles of each medication. In this regard, the SSRIs are currently considered first-line treatment for PD, demonstrating comparable efficacy and superior tolerability to other treatment classes. Combination therapies are frequently used for treatment resistance, and an approach of prescribing a benzodiazepine at the initiation of treatment with an SSRI, and later tapering it, has proved to be popular with clinicians and has recently been demonstrated to be advantageous in the early stages of treatment over an SSRI alone (71).

GENERALIZED ANXIETY DISORDER

Part of "66 - Current and Emerging Therapeutics of Anxiety and Stress Disorders "

The diagnostic criteria for generalized anxiety disorder (GAD) have evolved over the past two decades, undergoing substantial revision to the current definition emphasizing excessive, unrealistic worry as the cardinal feature of this disorder. Defined in 1980 in DSM-III as a disorder of continuous or persistent worry symptoms of at least 1 month's duration, the diagnosis of GAD was reframed in DSM-III-R to require symptoms extended for 6 months or longer, and an emphasis on unrealistic worry was stressed. With DSM-IV, GAD was defined as excessive and persistent

worry, accompanied by three or more physical or psychological symptoms of anxiety, persisting 6 months or longer. Because of the shift in diagnostic criteria and required duration of symptoms over the years, comparison of pharmacologic treatment studies performed prior to the introduction of DSM-III-R is difficult.

In comparison with panic disorder, PTSD, and obsessive-compulsive disorder (OCD), there are fewer publications devoted to GAD overall, and only a limited number of published controlled clinical medication trials. Several reasons have been proposed for this deficiency, including underrepresentation in clinical settings and a view of GAD as a "mild" disorder (72). In reality, GAD is one of the most commonly diagnosed anxiety disorders, with a lifetime prevalence of 4.1% to 6.6% (73 ,74), which is often chronic (75), and associated with significant compromise in functioning (76 ,77). There remains substantial controversy surrounding the validity of GAD as a primary disorder, as opposed to a comorbid condition or a prodromal/residual phase of another disorder (78 ,79). Findings from epidemiologic studies of GAD suggest that current comorbidity with other disorders is as high as 58% to 65% (73 ,76), and lifetime comorbidity rates are between 80% and 90% (76 ,80). Non-comorbid, "pure" GAD lifetime prevalence was found to be only 0.5% in the National Comorbidity Survey (76). Overall, the sum of studies examining quality of life issues support the idea that non-comorbid GAD is relatively rare, but is associated with significant impairment in its own right (81 ,82).

Historical Notes

Prior to the introduction of the benzodiazepines, the main agents available for the treatment of anxiety were the tricyclic antidepressants (doxepin, imipramine, amitryptyline), antihistamines (diphenhydramine, hydroxyzine), barbiturates (mephobarbital), and the sedative antianxiety agent meprobamate (Milltown). The development of the benzodiazepines in the mid-1950s led to the introduction of chlordiazepoxide (Librium) in 1959, and ushered in an era of benzodiazepine use in the treatment of a wide range of anxiety symptoms related to anxiety and mood disorders, psychosis, and alcohol withdrawal. Greater tolerability and the superior safety profiles of the benzodiazepines resulted in a sharp decline in the use of barbiturates for anxiety disorders (83). The benzodiazepines have remained a common treatment choice for GAD throughout the past two decades. However, concerns about dependence and withdrawal, short-term memory impairment, interactions with alcohol, and psychomotor impairment have resulted in increased interest in alternative medications. The introduction of drugs such as buspirone (1986), SSRIs (from 1980 on), and the serotonin and norepinephrine reuptake inhibitor (SNRI) venlafaxine (1994) have broadened the available treatment armamentarium for GAD significantly.

Tricyclic Antidepressants

Tricyclic antidepressants (TCAs) have been in use in the treatment of GAD for many decades; however, data supporting their efficacy from controlled clinical trials are extremely limited. Imipramine is the only TCA shown to be effective in placebo-controlled trials of GAD patients without comorbid depression (84 ,85). In comparator trials, imipramine has been shown to have clinical efficacy comparable to the benzodiazepines (84 ,86 ,87).

Benzodiazepines

Five benzodiazepines (alprazolam, chlordiazepoxide, clorazepate, diazepam, and lorazepam) are currently labeled as treatments for anxiety (GAD) (88). Clonazepam and alprazolam are labeled specifically for the treatment of PD. There are a limited number of clinical trials demonstrating the efficacy of benzodiazepines in the treatment of GAD in its current (DSM-IV) definition (89 ,90); however, benzodiazepines have been shown effective in controlled studies using DSM-III criteria for GAD (91).

Buspirone

Buspirone is a serotonin receptor subtype 1A ($5\text{-}HT_{1A}$) partial agonist with anxiolytic properties. In a metaanalysis of placebo-controlled comparator trials with benzodiazepines, buspirone showed comparable efficacy to the benzodiazepines in eight trials in 735 patients meeting DSM-III criteria (1 month's duration of illness) for GAD (92). In a metaanalysis of eight placebo-controlled studies in over 500 GAD patients with coexisting depressive symptoms, buspirone demonstrated significant superiority to placebo (93). Prior recent treatment with a benzodiazepine (<1 month) has been shown to predict poor response to subsequent buspirone treatment in GAD (92). This may be due to a combination of factors including the presence of benzodiazepine withdrawal, exacerbation of benzodiazepine discontinuation symptoms by buspirone and its metabolite a-(2-pyridinyl)-piperazine via enhancement of noradrenergic activity (94), and psychological factors such as patient expectations.

SSRIs

The first published study of a medication with significant serotonin reuptake inhibition properties in GAD involved a small, open-label trial of clomipramine (95). The suggestion of efficacy in this study, along with the success of clomipramine in treating other anxiety disorders, raised interest in pharmacologic agents for GAD that target the serotoninergic system. Following several years later, the first comparison trial of an SSRI in the treatment of GAD was published by Rocca and colleagues (89). Treatment efficacy of paroxetine was compared with the tricyclic imipramine and the

benzodiazepine 2′-chlordesmethyldiazepam in 81 subjects with GAD. Of the 63 patients who completed the randomized, 8-week study, 68% of the paroxetine group, 72% of the imipramine group, and 55% of the 2′-chlordesmethyldiazepam group were judged to be responders as measured by a 50% or more decrease in Hamilton Anxiety Rating Scale (HAM-A) scores. The greatest improvement during the first 2 weeks occurred in the group receiving the benzodiazepine, as expected by the early relief of physical anxiety symptoms and insomnia provided by this class of medication. However, from the fourth week forward, the paroxetine and imipramine groups demonstrated superior benefits, particularly in the area of psychic symptoms of anxiety. More recently, the efficacy of paroxetine was demonstrated in a large, fixed-dose study of more than 500 patients with a DSM-IV diagnosis of GAD without major depression (96). Patients were randomized to receive paroxetine 20 mg/day, paroxetine 40 mg/day, or placebo for 8 weeks. Patients receiving both doses of paroxetine demonstrated significant differences in the primary outcome measure, reduction in HAM-A score, versus placebo, with 68% on 20 mg paroxetine and 81% on 40 mg paroxetine rated as responders based on a Clinical Global Impression (CGI-I) score of 1 or 2, versus 52% on placebo.

Venlafaxine

Venlafaxine is an inhibitor of SNRI. Venlafaxine has recently been demonstrated in humans, using peripheral measures, to have primarily 5-HT reuptake inhibition properties at low doses (75 mg/day), with increasing norepinephrine (NE) reuptake inhibition properties at higher doses (375 mg/day) (97). Shown to be effective in the treatment of anxiety symptoms associated with major depression (98 ,99), the extended release (XR) form of venlafaxine has been shown to be effective in the treatment of GAD (DSM-IV criteria) in several placebo-controlled studies (100 ,101). In a placebo-controlled multicenter comparator trial, 405 patients with GAD were randomized to receive venlafaxine XR (75 or 150 mg/day), buspirone (30 mg/day), or placebo for 8 weeks. For the 365 patients for whom efficacy measures were obtained, there was no significant difference between groups in improvement on the primary outcome measure, the HAM-A. However, both doses of venlafaxine were shown to be superior to placebo in improving HAM-A psychic anxiety and anxious mood scores at the endpoint (week 8), and venlafaxine demonstrated superiority to placebo and buspirone on the CGI-S at the same time point. More robust efficacy findings for venlafaxine were reported in a recent large, multicenter trial, involving 377 outpatients with GAD without comorbid depression (101). Patients were randomly assigned to receive either placebo or venlafaxine XR at one of three doses (75, 150, or 225 mg/day) for 8 weeks. Of the 349 patients included in the efficacy analysis, those receiving 225 mg/day demonstrated significant improvement across seven of the eight outcome measures, and the 225-mg/day group was the only group to show significant improvement in scores on both of the CGI subscales (severity, global improvement) versus placebo.

Other Agents (Trazodone, Nefazodone, Anticonvulsants, Partial GABA Agonists)

In a randomized, placebo-controlled comparator trial of trazodone, diazepam, and imipramine in the treatment of 230 patients with GAD, trazodone was found to be superior to placebo yet somewhat less effective than diazepam and imipramine at the study's endpoint (84). The antidepressant nefazodone, which antagonizes the 5-HT_{2C} receptor and inhibits the reuptake of both serotonin and NE, has shown promise in the treatment of GAD in an open trial (102). Anticonvulsants such as valproate and carbamazepine have been used in the treatment of GAD; however, evidence of their efficacy is primarily anecdotal, as there are no controlled clinical trials for either of these medications in the treatment of GAD (103). Other agents such as the partial GABA agonist abecarnil have not demonstrated significant efficacy versus placebo in GAD (101).

In summary, although the benzodiazepines have been the mainstay of pharmacotherapy for GAD since their introduction, significant concerns regarding their long-term use in GAD have fueled the search for other effective treatments. Given the chronic nature of GAD, medications such as buspirone, and the antidepressants venlafaxine and paroxetine, which have fewer effects on cognitive and psychomotor function, now represent first-line therapies for GAD.

SOCIAL PHOBIA

Part of "66 - Current and Emerging Therapeutics of Anxiety and Stress Disorders "

Social phobia (SP) (or social anxiety disorder) is reported to be the most common anxiety disorder with a 1-year prevalence of 7% to 8% and a lifetime prevalence of 13% to 14% in patients aged 15 to 54 years. Social anxiety disorder can be classified into two subtypes—discrete and generalized. Level of disability with SP can be high, and 70% to 80% of patients have comorbid psychiatric disorders, particularly depression and substance abuse (104). Given the high degree of burden of illness in SP, its treatment has become a major priority.

Historical Notes

Liebowitz et al. (105) noted that SP, like atypical depression (106), had a specific responsivity to the MAOIs, whereas TCAs, although effective for PD and “typical” major depression, were not effective for SP (107). The efficacy of the MAOIs, which block reuptake of dopamine in addition to NE and serotonin, prompted speculation about a potential

“dopaminergic” component to the neurobiology of SP. Several open clinical studies have attempted to utilize the “dopamine component” concept in phamacotherapy with some success, e.g., seligiline (108) and pergolide (109). However, as data accumulated, other systems were also implicated, and the pharmacologic dissection approach seemed less applicable (see above). Positive results with the high-potency benzodiazepine clonazepam (110) suggested a GABAergic component. However, the suitability of benzodiazepines for long-term treatment of a chronic condition such as social anxiety disorder has been questioned. In addition, the benzodiazepines are ineffective against comorbid depression.

RIMAs (Reversible Inhibitor of Monoamine Oxidase A)

Although phenelzine demonstrated efficacy, the need for dietary restrictions severely limited its use. Moclobemide is a RIMA with a much lower propensity to induce hypertensive crises and has a more favorable side-effect profile. Moclobemide had been reported to have efficacy in early studies in the treatment of social phobia (111). However, conflicting results have subsequently been reported in placebo-controlled trials. Some studies have shown moclobemide to be more effective than placebo, whereas two recent, large, randomized placebo-controlled trials conducted in the United States have reported less robust results (112 ,113). Brofaromine, another drug in the RIMA class, may still hold promise. The safety and efficacy of brofaromine were examined in a multicenter trial of 102 outpatients with SP (114). Brofaromine produced a significantly greater change from baseline in Liebowitz Social Anxiety Scale (LSAS) scores compared with placebo.

SSRIs

Based on clinical evidence, SSRIs are the first-line treatment in social anxiety disorder (115). The most extensive database for the treatment of social anxiety disorder exists for the SSRI paroxetine. Several large, multicenter, placebo-controlled trials have been completed on three different continents (116 ,117 ,118 and 119). In all cited studies, a significantly greater proportion of patients responded to paroxetine treatment compared with placebo. Paroxetine is currently the only SSRI licensed for use in this condition in the United States. The SSRIs are particularly attractive agents due to their favorable tolerance and safety profile, although typical SSRI side effects may nevertheless be problematic.

Despite promising open studies with fluvoxamine, fluoxetine, and citalopram (120 ,121) only fluvoxamine has been tested under double-blind, randomized, placebo-controlled trial conditions (122). Like paroxetine, fluvoxamine yielded efficacy data superior to placebo. A report on a multicenter sertraline trial was pending at the time of this writing.

As is the case with PD, buspirone does not appear effective for SP as monotherapy in placebo-controlled double-blind studies (123). It may, however, have a role in augmentation of the SSRIs.

Serotonin/Norepinephrine Reuptake Inhibitors

One open label study (124) aimed to evaluate the clinical response to venlafaxine in SP in 12 patients who were nonresponders to SSRIs, and to assess how the response could be influenced by Axis II comorbidity with avoidant personality disorder (APD). The duration of the study was 15 weeks using an open flexible-dosing regimen in individuals with or without concomitant APD. Venlafaxine improved SP and/or APD symptomatology, as demonstrated by decreasing LSAS total scores. Similar favorable open-labeled results have been reported for nefazodone (125 ,126). Placebo-controlled studies are warranted.

Anticonvulsants

A randomized, double-blind, placebo-controlled, parallel-group study was conducted to evaluate the efficacy and safety of gabapentin in relieving the symptoms of social phobia. A significant reduction in the symptoms of social phobia was observed among patients on gabapentin compared with those on placebo as evaluated by clinician- and patient-rated scales (127). Adverse events were consistent with the known side-effect profile of gabapentin. The efficacy of other novel anticonvulsants remains to be investigated, although encouraging results have been reported for the gabapentin-like compound, pregabalin (128).

POSTTRAUMATIC STRESS DISORDER

Part of "66 - Current and Emerging Therapeutics of Anxiety and Stress Disorders "

Despite the high prevalence, chronicity, and associated comorbidity of PTSD in the community, relatively few placebo-controlled studies have evaluated the efficacy of pharmacotherapy for this disorder. The symptom overlap between PTSD and other pharmacotherapy-responsive disorders has suggested that pharmacotherapy might be effective. Nevertheless, in those placebo-controlled trials investigating the pharmacotherapy of PTSD that have been carried out, statistically significant efficacy for the treatment being studied has traditionally been inconsistent. One of the key methodologic limitations has been the fact that most studies have been conducted with war veterans, who are likely to constitute a more treatment-refractory population.

SSRIs

More recently, a total of 187 civilian outpatients with DSM-III-R PTSD (73% were women, and 61.5% experienced

physical or sexual assault) were treated with the SSRI sertraline in a placebo-controlled design (129). Sertraline effectively diminished symptoms of PTSD of moderate to marked severity in comparison to placebo. Using a conservative last-observation-carried-forward analysis, treatment with sertraline resulted in a responder rate of 53% at the study's endpoint compared with 32% for placebo (p = .008). Sertraline is the first medication approved by the FDA for the treatment of PTSD. Similar positive results have been reported for the SSRI fluoxetine in civilian populations (130 ,131). In a study by van der Kolk et al. (132), fluoxetine was found to be most effective in the nonveteran versus veteran portion of his study sample. Although published placebo-controlled data for paroxetine are not available, Marshall et al. (133) have argued that this particular SSRI may have specific advantages because of its relatively low activating properties. Direct comparative studies are lacking.

Combat Veteran PTSD

Among combat veterans, PTSD is a highly prevalent and often chronic disorder, persisting in as many as 15% of Vietnam veterans for at least 20 years (134). Treatment response in veterans with combat-related PTSD has been disappointing. Although anxiolytics, anticonvulsants, antipsychotics, and antidepressants, including SSRIs, have been tried, none has been consistently associated with improvement in all primary symptom domains (i.e., intrusive recollections, avoidance/numbing, and hyperarousal). In an open study using nefazodone, at mean daily doses of 430 mg (range, 200 to 600 mg/day), 19 treatment-refractory PTSD patients demonstrated benefit after 12 weeks (134). Double-blind placebo controlled studies would be of interest.

The efficacy of the antidepressant drug bupropion in the treatment of male combat veterans with chronic PTSD was investigated in an open-label study of 6 weeks' duration (135). Improvement was seen in hyperarousal symptoms but was less significant than the change in depressive symptoms. Mirtazapine, a novel drug with both noradrenergic and serotoninergic properties, may be effective in individuals who demonstrate intolerance to side effects of, or a limited response to, SSRIs. Three of six severely refractory PTSD patients treated with mirtazapine were assessed as responders in a pilot study (136). Case reports have suggested benefit for refractory patients treated with venlafaxine (137) and risperidone (138). Raskind et al. (139) reported that the α_1-adrenergic antagonist prazosin ameliorated combat nightmares in a small sample of veterans with PTSD.

Monoamine Oxidase Inhibitors

Traditional MAOIs have shown efficacy in the treatment of PTSD, but their use is limited by serious drug and food interactions. Moclobemide, a RIMA, is relatively free of these limitations and is therefore potentially useful in the treatment of PTSD. Moclobemide was highly effective in an open-labeled design (140). However, in a double-blind, randomized, placebo-controlled, multicenter study, brofaromine, also a RIMA, failed to surpass efficacy levels seen with placebo. Thus, the role for RIMAs remains unclear at this time.

Anticonvulsants

Despite a long-recognized role for anticonvulsants in the treatment of PTSD (141), few placebo-controlled studies have been reported. An open study of divalproex reported favorable results (142). In a placebo-controlled study, patients treated with lamotrigine showed improvement on reexperiencing and avoidance/numbing symptoms compared to placebo-treated patients (143). The authors concluded that lamotrigine may be effective as a primary psychopharmacologic treatment in both combat and civilian PTSD and could also be considered as an adjunct to antidepressant therapy used in the treatment of PTSD. Further large-sample, double-blind, placebo-controlled trials are warranted.

Summary and Future Directions

Current management of PTSD is well summarized by Davidson et al. (144). Clearly, there are many challenges associated with the treatment of PTSD. Different patients with PTSD may not respond to pharmacotherapy in the same manner, and it is unclear whether this is related to gender, trauma type, or other factors. Antidepressants, particularly the SSRIs, are currently the form of pharmacotherapy for patients with PTSD with greatest support in the literature. Psychosocial techniques, such as cognitive-behavioral therapy or stress inoculation training, are effective and may be considered as adjunctive therapy with medication. Larger placebo-controlled studies for many different classes of medications would be desirable in moving the field forward. In addition, carefully conducted polypharmacy in which drug interactions are well understood may well be necessary for more difficult cases.

In a review by Shalev and colleagues (145), a synthesis of findings in PTSD studies is provided. Most studies explored a single treatment modality (e.g., pharmacologic, behavioral). The cumulated evidence from these studies suggests that several treatment protocols reduce PTSD symptoms and improve the patient's quality of life. The magnitude of the results, however, was often limited, and remission was rarely achieved. Given the shortcoming of unidimensional treatment of PTSD, it was suggested by the authors that combining biological, psychological, and psychosocial treatment yields the best results.

A host of novel compounds show promise for the treatment of PTSD (146). Such classes of compounds include corticotropin-releasing factor antagonists, neuropeptide-Y

enhancers, antiadrenergic compounds, drugs that down-regulate glucocorticoid receptors, more specific serotoninergic agents, agents that normalize opioid function, substance P antagonists, *N*-methyl-D-aspartate facilitators, glutamatergic antagonists, and antikindling/antisensitization anticonvulsants.

CONCLUSION

Part of "66 - Current and Emerging Therapeutics of Anxiety and Stress Disorders "

Antidepressants are the logical first choice for most patients with anxiety disorders, based on their efficacy and tolerability. Although maintaining a role, the use of benzodiazepines for first-line or long-term therapy is now less likely. Does this mean that anxiety disorders are a variant of depression? Certainly, anxiety disorders and depression are highly comorbid. Untreated, the majority of patients with anxiety disorders eventually develop depression, whereas a large fraction of depressed patients suffer from clinically significant comorbid anxiety, if not an overt syndromal anxiety disorder.

Pharmacologic dissection is clearly perilous, leaving us prone to inferences based on limited knowledge. Most of the antidepressants that successfully treat anxiety disorders manipulate the reuptake of serotonin, norepinephrine, or both. Altering the brain circuits through these modulatory neurotransmitters in turn has wide-ranging effects on many other systems in the brain. The release of CRH from extrahypothalamic sites like the amygdala is only one such system. Hence, there may be a common denominator among anxiety disorders and between anxiety disorders and depression at one or more points in these complex circuits.

The observation, then, that antidepressants work for the four anxiety disorders discussed in this chapter warrants, in our opinion, only the following inference: it is highly likely that some substrate of the serotonin and norepinephrine systems is malfunctioning in several anxiety disorders and depression. This could be the locus of a common genetic or environmental vulnerability to both categories of illness. Although it will not likely tell us that anxiety and depression are fundamentally the same thing, the search for such common substrates and vulnerabilities, suggested but not guaranteed by the psychopharmacologic findings, is likely to be very illuminating.

ACKNOWLEDGMENTS

Part of "66 - Current and Emerging Therapeutics of Anxiety and Stress Disorders "

Dr. Gorman has received research support from Pfizer, Eli Lilly, the National Institute of Mental Health, and NARSAD. In addition, he has been a consultant and received honoraria from a number of pharmaceutical companies, including Pfizer, Eli Lilly, Bristol-Myers Squibb, Wyeth-Ayerst, SmithKline Beecham, Astra Zeneca, Janssen, Organon, Forrest, Parke-Davis, Lundbeck, Solvay, Merck, Sanofi-Synthelabo, and Aventis. Dr. Kent has served as a consultant for Bristol-Myers Squibb and SmithKline Beecham. Dr. Coplan has received research support from SmithKline Beecham, Eli Lilly, and Janssen. In addition, he has served on a speakers' bureau and/or an advisory board for the following companies: SmithKline Beecham, Wyeth-Ayerst, and Bristol-Myers Squibb.

REFERENCES

1. Eaton WW, Kessler RC, Wittchen HU, et al. Panic and panic disorder in the United States. *Am J Psychiatry* 1994;151:413–420.

2. Markowitz JS, Weissman MM, Ouellette R, et al. Quality of life in panic disorder. *Arch Gen Psychiatry* 1989;46:984–992.

3. Klein DF, Fink M. Psychiatric reaction patterns to imipramine. *Am J Psychiatry* 1962;119:432–438.

4. Klein D. Delineation of two drug responses for anxiety syndromes. *Psychopharmacologia* 1964;5:397–408.

5. Modigh K, Westberg P, Eriksson E. Superiority of clomipramine over imipramine in the treatment of panic disorder: a placebo-controlled trial. *J Clin Psychopharmacol* 1992;12:251–261.

6. Fahy TJ, O'Rourke D, Brophy J, et al. The Galway study of panic disorder. I: clomipramine and lofepramine in DSM III-R panic disorder: a placebo controlled trial. *J Affective Disord* 1992;25:63–75.

7. den Boer JA, Westenberg HG, Kamerbeek WD, et al. Effect of serotonin uptake inhibitors in anxiety disorders; a double-blind comparison of clomipramine and fluvoxamine. *Int Clin Psychopharmacol* 1987;2:21–32.

8. Nair NP, Bakish D, Saxena B, et al. Comparison of fluvoxamine, imipramine, and placebo in the treatment of outpatients with panic disorder. *Anxiety* 1996;2:192–198.

9. Cross-National Collaborative Panic Study, Second Phase Investigators. Drug treatment of panic disorder. Comparative efficacy of alprazolam, imipramine, and placebo. *Br J Psychiatry* 1992;160:191–202.

10. Mavissakalian MR, Perel JM. Imipramine treatment of panic disorder with agoraphobia: Dose ranging and plasma levelresponse relationships. *Am J Psychiatry* 1995;152(5):673–682.

11. Ballenger JC. Pharmacotherapy of the panic disorder. *J Clin Psychiatry* 1986;47(suppl 6):27–32.

12. Amin MM, Ban TA, Pecknold JC, et al. Clomipramine (Anafranil) and behavior therapy in obsessive compulsive and phobic disorders. *J Int Med Res* 1997;5(suppl 5):33–37.

13. Papp AL, Schneier FR, Fyer AJ, et al. Clomipramine treatment of panic disorder: pros and cons. *J Clin Psychiatry* 1997;58:423–425.

14. Lepine J-P, Chignon JM, Teherani M. Suicide attempts in patients with panic disorder. *Arch Gen Psychiatry* 1993;50:144–149.

15. Johnson J, Weissman MM, Klerman GL. Panic disorder, comorbidity, and suicide attempts. *Arch Gen Psychiatry* 1990;47:805–808.

16. Roy-Byrne PP, Stang P, Wittchen H-U, et al. Lifetime panic-depression comorbidity in the National Comorbidity Survey. *Br J Psychiatry* 2000;176;229–235.

17. Mavissakalian MR. Burden of side effects of imipramine treatment on panic disorder. Presented at the 153rd Annual Meeting of the American Psychiatric Association, Chicago, Illinois, May 13–18, 2000.

18. Tyer P, Candy J, Kelly D. Phenelzine in phobic anxiety: a controlled trial. *Psychol Med* 1973;3:120–124.

19. Mountjoy CQ, Roth M, Garside RF, et al. A clinical trial of phenelzine in anxiety depressive and phobic neuroses. *Br J Psychiatry*, 1977;131:486–492.

20. Solyom C, Solyom L, LaPierre Y, et al. Phenelzine and exposure in the treatment of phobias. *Biol Psychiatry* 1981;16:239–247.

21. Kruger MB, Dahl AA. The efficacy and safety of moclobemide compared to clomipramine in the treatment of panic disorder. *Eur Arch Psychiatry Clin Neurosci* 1999;249(suppl 1):S19–21.

22. Tiller JW, Bouwer C, Behnke K. Moclobemide for anxiety disorders: a focus on moclobemide for panic disorder. *Int Clin Psychopharmacol* 1997;12(suppl 6):S27–S30.

23. Tiller JW, Bouwer C, Behnke K. Moclobemide and fluoxetine for panic disorder. International Panic Disorder Study Group. *Eur Arch Psychiatry Clin Neurosci* 1999;249(suppl 1):S7–10.

24. van Vliet IM, den Boer JA, Westenberg HG, et al. A double-blind comparative study of brofaromine and fluvoxamine in outpatients with panic disorder. *J Clin Psychopharmacol* 1996;16(4):299–306.

25. Bakish D, Hooper CL, Filteau MJ, et al. A double-blind, placebo-controlled trial comparing fluvoxamine and imipramine in the treatment of panic disorder with or without agoraphobia. *Psychopharmacol Bull* 1996;32:135–141.

26. van Vliet IM, Westenberg HG, den Boer JA. MAO inhibitors in panic disorder: clinical effects of treatment with brofaromine. A double-blind placebo controlled study. *Psychopharmacology (Berl)* 1993;112(4):483–489.

27. Rickels K, Schweizer E. Panic disorder: long-term pharmacotherapy and discontinuation. *J Clin Psychopharmacol* 1998;18(suppl 2):12S–18S.

28. Uhlenhuth EH, DeWitt H, Balter MB, et al. Risks and benefits of long-term benzodiazepine use. *J Clin Psychopharmacol* 1988;8:161–167.

29. Nagy LM, Krystal JH, Woods SW, et al. Clinical and medication outcome after short-term alprazolam and behavioral group treatment in panic disorder: 2.5 year naturalistic follow-up study. *Arch Gen Psychiatry* 1989;46:993–999.

30. Worthington JJ 3rd, Pollack MH, Otto MW, et al. Long-term experience with clonazepam in patients with a primary diagnosis of panic disorder. *Psychopharmacol Bull* 1998;34:199–205.

31. Charney DS, Woods SW, Goodman WK, et al. Drug treatment of panic disorder: the comparative efficacy of imipramine, alprazolam, and trazodone. *J Clin Psychiatry* 1986;47:580–586.

32. Charney DS, Woods SW. Benzodiazepine treatment of panic disorder: a comparison of alprazolam and lorazepam. *J Clin Psychiatry* 1989;50:418–423.

33. Ballenger JC, Burrows GD, Dupont RL Jr, et al. Alprazolam in panic disorder and agoraphobia: Results from a multicenter trial: I. Efficacy in short-term treatment. *Arch Gen Psychiatry* 1988;45:413–422.

34. Uhlenhuth EH, Matuzas W, Glass RM, et al. Response of panic disorder to fixed doses of alprazolam or imipramine. *J Affective Disord* 1989;17:261–270.

35. Lydiard RB, Lesser IM, Ballenger JC, et al. A fixed-dose study of alprazolam 2 mg, alprazolam 6 mg, and placebo in panic disorder. *J Clin Psychopharmacol* 1992;12:96–103.

36. Tesar GE, Rosenbaum JF, Pollack MH, et al. Double-blind, placebo-controlled comparison of clonazepam and alprazolam for panic disorder. *J Clin Psychiatry* 1991;52:69–76.

37. Rocca P, Fonzo V, Scotta M, et al. Paroxetine efficacy in the treatment of generalized anxiety disorder. *Acta Psychiatr Scand* 1997;95(5):444–450.

38. Zajecka J, Tracy KA, Mitchell S. Discontinuation symptoms after treatment with serotonin reuptake inhibitors: a literature review. *J Clin Psychiatry* 1997;58:291–297.

39. Uhlenhuth EH, Balter MB, Ban TA, et al. International study of expert judgment on therapeutic use of benzodiazepines and other psychotherapeutic medications: IV. Therapeutic dose dependence and abuse liability of benzodiazepines in the long-term treatment of anxiety disorders. *J Clin Psychopharmacol* 1999;19(suppl 2):23S–29S.

40. Ballenger JC, Davidson JR, Lecrubier Y, et al. Consensus statement on panic disorder from the international consensus group on depression and anxiety. *J Clin Psychiatry* 1998;59(8):47–54.

41. American Psychiatric Association. *Practice guideline for the treatment of patients with panic disorder*. Washington, DC: American Psychiatric Press, 1998.

42. Michelson D, Lydiard RB, Pollack MH, et al. Outcome assessment and clinical improvement in panic disorder: evidence from a randomized controlled trial of fluoxetine and placebo. The Fluoxetine Panic Disorder Study Group. *Am J Psychiatry* 1998;155:1570–1577.

43. Michelson D, Sarka N, Pemberton C. Fluoxetine in panic disorder: a randomized, placebo-controlled study. Presented at the 153rd Annual Meeting of the American Psychiatric Association, Chicago, Illinois, May 13–18, 2000.

44. Black DW, Wesner R, Bowers W, et al. A comparison of fluvoxamine, cognitive therapy, and placebo in the treatment of panic disorder. *Arch Gen Psychiatry* 1993;50:44–50.

45. Oehrberg S, Christiansen PE, Behnke K, et al. Paroxetine in the treatment of panic disorder. *Br J Psychiatry* 1995;167:374–379.

46. Ballenger JC, Wheadon DE, Steiner M, et al. Double-blind, fixed-dose, placebo-controlled study of paroxetine in the treatment of panic disorder. *Am J Psychiatry* 1998;155:36–42.

47. Pollack MH, Otto MW, Worthington JJ, et al. Sertraline in the treatment of panic disorder: a flexible-dose multicenter trial. *Arch Gen Psychiatry* 1998;55:1010–1016.

48. Sheikh JI, Londborg PD, Clary CM. The efficacy of sertraline in panic disorder: a combined fixed-dose analysis. Presented at the 153rd Annual Meeting of the American Psychiatric Association, Chicago, Illinois, May 13–18, 2000.

49. Lepola UM, Wade AG, Leinonen EV, et al. A controlled, prospective, 1-year trial of citalopram in the treatment of panic disorder. *J Clin Psychiatry* 1998;59:528–534.

50. Leinonen E, Lepola U, Koponen H, et al. Citalopram controls phobic symptoms in patients with panic disorder: randomized controlled trial. *J Psychiatry Neurosci* 2000;25:24–32.

51. Baldwin DS, Birtwistle J. The side effect burden associated with drug treatment of panic disorder. *J Clin Psychiatry* 1998;59:39–44.

52. Pollack MH, Worthington JJ 3rd, Otto MW, et al. Venlafaxine for panic disorder: results from a double-blind, placebo-controlled study. *Psychopharmacol Bull* 1996;32(4):667–670.

53. Zajecka JM. The effect of nefazodone on comorbid anxiety symptoms associated with depression: experience in family practice and psychiatric outpatient settings. *J Clin Psychiatry* 1996;57(suppl 2):10–14.

54. Carpenter LL, Leon Z, Yasmin S, et al. Clinical experience with mirtazapine in the treatment of panic disorder. *Ann Clin Psychiatry* 1999;11:81–86.

55. Kapezinski F, Ribeiro L, Busnello JV, et al. Mirtazapine versus fluoxetine in panic disorder. Presented at the 153rd Annual Meeting of the American Psychiatric Association, Chicago, Illinois, May 13–18, 2000.

56. Primeau F, Fontaine R, Beauclair L. Valproic acid and panic disorder. *Can J Psychiatry* 1990;35:248–250.

57. Woodman CL, Noyes R Jr. Panic disorder: treatment with valproate. *J Clin Psychiatry* 1994;55:134–136.

58. Keck PE, McElroy SL, Friedman LM. Valproate and carbamazepine in the treatment of panic and posttraumatic stress disorders, withdrawal states, and behavioral dyscontrol syndromes. *J Clin Psychopharmacol* 1992;12:36S–41S.

59. Lum M, Fontaine R, Elie R, et al. Divalproex sodium's antipanic effect in panic disorder: a placebo-controlled study. *Biol Psychiatry* 1990;27:164A–165A.

60. Baetz M, Bowen RC. Efficacy of divalproex sodium in patients with panic disorder and mood instability who have not responded to conventional therapy. *Can J Psychiatry* 1998;43:73–77.

61. Uhde TW, Stein MB, Post RM. Lack of efficacy of carbamazepine in the treatment of panic disorder. *Am J Psychiatry* 1988;145:1104–1109.

62. Pollack MH, Matthews J, Scott EL. Gabapentin as a potential treatment for anxiety disorders. *Am J Psychiatry* 1998;155:992–993.

63. Swartz CM. Betaxolol in anxiety disorders. *Ann Clin Psychiatry* 1998;10(1):9–14.

64. Pecknold JC. A risk-benefit assessment of buspirone in the treatment of anxiety disorder. *Drug Saf* 1997;16:118–132.

65. Sheehan DV, Davidson J, Manshreck T, et al. Lack of efficacy of a new antidepressant (bupropion) in the treatment of panic disorder with phobias. *J Clin Psychopharmacol* 1983;3:28–31.

66. Schneier FR, Garfinkel R, Kennedy B, et al. Ondansetron in the treatment of panic disorder. *Anxiety* 1996;2:199–202.

67. Pande AC, Greiner M, Adams JB, et al. Placebo-controlled trial of the CCK-B antagonist, CI-988, in panic disorder. *Biol Psychiatry* 1999;46:860–862.

68. Shekhar A, Keim SR. LY354740, a potent group II metabotropic glutamate receptor agonist prevents lactate-induced panic-like response in panic-prone rats. *Neuropharmacology* 2000;39:1139–1146.

69. Griebel G. Is there a future for neuropeptide receptor ligands in the treatment of anxiety disorders? *Pharmacol Ther* 1999;82:1–61.

70. den Boer JA. Pharmacotherapy of panic disorder: Differential efficacy from a clinical viewpoint. *J Clin Psychiatry* 1998;59(suppl 8):30–36.

71. Goddard AW, Almai AM, Jetty P, et al. SSRI and benzodiazepine treatment for panic. Presented at the 153rd Annual Meeting of the American Psychiatric Association, Chicago, Illinois, May 13–18, 2000.

72. Dugas MJ. Generalized anxiety disorder publications: So where do we stand? *J Anxiety Dis* 2000;14:31–40.

73. Blazer DG, Hughest D, George L, et al. Generalized anxiety disorder. In: Robins LN, Regier DA, eds. *Psychiatric disorders in America: The Epidemiologic Catchment Area Study*. New York: The Free Press, 1991:180–203.

74. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-II-R psychiatric disorders in the United States. *Arch Gen Psychiatry* 1994;51:8–19.

75. Angst J, Vollrath M. The natural history of anxiety disorder. *Acta Psychiatr Scand* 1991;84:446–452.

76. Wittchen HU, Zhao S, Kessler, et al. DSM-II-R generalized anxiety disorder in the National Comorbidity Survey. *Arch Gen Psychiatry* 1994;51:355–364.

77. Massion AO, Warshaw MG, Keller MB. Quality of life and psychiatric morbidity in panic disorder and generalized anxiety disorder. *Am J Psychiatry* 1993;150(4):600–607.

78. Brawman-Mintzer O, Lydiard RB. Generalized anxiety disorder: issues in epidemiology. *J Clin Psychiatry* 1996;57(suppl 7):3–8.

79. Roy-Byrne PP, Katon W. Generalized anxiety disorder in primary care: the precursor/modifier pathway to increased health care utilization. *J Clin Psychiatry* 1997;58(suppl 3):34–38.

80. Judd LL, Kessler RC, Paulus MP, et al. Comorbidity as a fundamental feature of generalized anxiety disorders: Results from the National Comorbidity Study (NCS). *Acta Psychiatr Scand Suppl* 1998;393:6–11.

81. Kessler RC, DuPont RL, Berglund P, et al. Impairment in pure and comorbid generalized anxiety disorder and major depression at 12 months in two national surveys. *Am J Psychiatry* 1999;156:1915–1923.

82. Mendlowicz MV, Stein MB. Quality of life individuals with anxiety disorder. *Am J Psychiatry* 2000:157:669–682.

83. Hollister LE. Pharmacology and clinical use of benzodiazepines. In: Usdin E, Skolnick P, Tallman JF, et al., eds. *Pharmacology of Benzodiazepines*. London: Macmillan, 1982:29–36.

84. Rickels K, Downing R, Schweizer E, et al. Antidepressants for the treatment of generalized anxiety disorder. *Arch Gen Psychiatry* 1993;50:884–895.

85. Kahn JR, McNair DM, Lipman RS, et al. Imipramine and chlordiazepoxide in depressive and anxiety disorders. II: efficacy in anxious outpatients. *Arch Gen Psychiatry* 1986;43:79–85.

86. Hoehn-Saric R, McLeod DR, Zimmerli WD. Differential effects of alprazolam and imipramine in generalized anxiety disorder: Somatic versus psychic symptoms. *J Clin Psychiatry* 1988;49:293–301.

87. Rickels K, Schweizer E. The treatment of generalized anxiety disorder in patients with depressive symptomatology. *J Clin Psychiatry* 1993;54(suppl 1):20–23.

88. *Physician's Desk Reference*, 54th ed. Montvale, NJ: Medical Economics, 2000.

89. Rocca P, Fonzo V, Scotta M, et al. Paroxetine efficacy in the treatment of generalized anxiety disorder. *Acta Psychiatr Scand* 1997;95:444–450.

90. Rickels K, DeMartinis N, Aufdembrinke B. A double-blind, placebo-controlled trial of abecarnil and diazepam in the treatment of patients with generalized anxiety disorder. *J Clin Psychopharmacol* 2000;20:12–18.

91. Laakman G, Schule C, Lorkowski G, et al. Buspirone and lorazepam in the treatment of generalized anxiety disorder in outpatients. *Psychopharmacology (Berl)* 1998;136:357–366.

92. DeMartinis N, Rynn M, Rickels K, et al. Prior benzodiazepine use and buspirone response in the treatment of generalized anxiety disorder. *J Clin Psychiatry* 2000;61:91–94.

93. Gammans RE, Stringfellow JC, Hvizdos AJ, et al. Use of buspirone in patients with generalized anxiety disorder and co-existing depressive symptoms: a meta analysis of eight randomized, controlled trials. *Neuropsychobiology* 1992;25:193–210.

94. Giral P, Soubrie P, Puech AJ. Pharmacological evidence for the involvement of 1-(2-pyridinyl)-piperazine (1-PmP) in the interaction of buspirone or gepirone with noradrenergic systems. *Eur J Pharmacol* 1987;134:113–116.

95. Wingerson D, Nguyen C, Roy-Byrne PP. Clomipramine treatment for generalized anxiety disorder [letter]. *J Clin Psychopharmacol* 1992;12(3):214–215.

96. Bellew KM, McCafferty JP, Iyengar M, et al. Paroxetine in the treatment of generalized anxiety disorder: a double blind placebo controlled trial. Presented at the Annual Meeting of the American Psychiatric Association, Chicago, Illinois, May 13–18, 2000.

97. Harvey AT, Rudolph RL, Preskorn SH. Evidence of the dual mechanisms of action of venlafaxine. *Arch Gen Psychiatry* 2000;57:503–509.

98. Feighner JP, Entsuah AR, McPherson MK. Efficacy of once-daily venlafaxine extended release (XR) for symptoms of anxiety in depressed outpatients. *J Affect Disord* 1998;47:55–62.

99. Rudolph RL, Entsuah R, Chitra R. A meta-analysis of the effects of venlafaxine on anxiety associated with depression. *J Clin Psychopharmacol* 1998;18:136–144.

100. Sheehan DV. Venlafaxine extended release (XR) in the treatment of generalized anxiety disorder. *J Clin Psychiatry* 1999;60(suppl 22):23–28.

101. Rickels K, Pollack MH, Sheehan DV, et al. Efficacy of extended-release venlafaxine in nondepressed outpatients with generalized anxiety disorder. *Am J Psychiatry* 2000;157:968–974.

102. Hedges DW, Reimherr FW, Strong RE, et al. An open trial of nefazodone in adult patients with generalized anxiety disorder. *Psychopharmacol Bull* 1996;32:671–676.

103. Roy-Byrne PP, Ward NG, Donnelly PG. Valproate in anxiety and withdrawal syndromes. *J Clin Psychiatry* 1989;50:44–48.

104. Sareen L, Stein M. A review of the epidemiology and approaches to the treatment of social anxiety disorder. *Drugs* 2000;59(3):497–509.

105. Liebowitz MR, Schneier F, Campeas R, et al. Phenelzine vs. atenolol in social phobia: a placebo-controlled comparison. *Arch Gen Psychiatry* 1992;49(4):290–300.

106. Quitkin FM, McGrath PJ, Stewart JW, et al. Phenelzine and imipramine in mood reactive depressives: further delineation of the syndrome of atypical depression. *Arch Gen Psychiatry* 1989;46(9):787–793.

107. Simpson HB, Schneier FR, Campeas RB, et al. Imipramine in the treatment of social phobia. *J Clin Psychopharmacol* 1998;18(2):132–135.

108. Simpson HB, Schneier FR, Marshall RD, et al. Low dose selegiline (L-Deprenyl) in social phobia. *Depress Anxiety* 1998;7(3):126–129.

109. Villarreal G, Johnson MR, Rubey R, et al. Treatment of social phobia with the dopamine agonist pergolide. *Depress Anxiety* 2000;11(1):45–47.

110. Davidson JR, Connor KM. Management of posttraumatic stress disorder: diagnostic and therapeutic issues. *J Clin Psychiatry* 1999;60(suppl 18):33–38.

111. Versiani M, Nardi AE, Mundim FD, et al. Pharmacotherapy of social phobia: a controlled study with moclobemide and phenelzine. *Br J Psychiatry* 1992;161:353–360.

112. Schneier FR, Goetz D, Campeas R, et al. Placebo-controlled trial of moclobemide in social phobia. *Br J Psychiatry* 1998;172:70–77.

113. Noyes R Jr, Moroz G, Davidson JR, et al. Moclobemide in social phobia: a controlled dose-response trial. *J Clin Psychopharmacol* 1997;17(4):247–254.

114. Lott M, Greist JH, Jefferson JW, et al. Brofaromine for social phobia: a multicenter, placebo-controlled, double-blind study. *J Clin Psychopharmacol* 1997;17(4):255–260.

115. Westenberg HG. Facing the challenge of social anxiety disorder. *Eur Neuropsychopharmacol* 1999;9(suppl 3):S93–99.

116. Stein MB, Liebowitz MR, Lydiard RB, et al. Paroxetine treatment of generalized social phobia (social anxiety disorder): a randomized controlled trial. *JAMA* 1998;280(8):708–713.

117. Stein DJ, Berk M, Els C, et al. A double-blind placebo-controlled trial of paroxetine in the management of social phobia (social anxiety disorder) in South Africa. *S Afr Med J* 1999;89(4):402–406.

118. Allgulander C. Paroxetine in social anxiety disorder: a randomized placebo-controlled study. *Acta Psychiatr Scand* 1999;100(3):193–198.

119. Baldwin D, Bobes J, Stein DJ, et al. Paroxetine in social phobia/social anxiety disorder. Randomized, double-blind, placebo-controlled study. Paroxetine Study Group. *Br J Psychiatry* 1999;175:120–126.

120. Bouwer C, Stein DJ. Use of the selective serotonin reuptake inhibitor citalopram in the treatment of generalized social phobia. *J Affect Disord* 1998;49(1):79–82.

121. DeVane CL, Ware MR, Emmanuel NP, et al. Evaluation of the efficacy, safety and physiological effects of fluvoxamine in social phobia. *Int Clin Psychopharmacol* 1999;14(6):345–351.

122. Stein MB, Fyer AJ, Davidson JR, et al. Fluvoxamine treatment of social phobia (social anxiety disorder): a double-blind, placebo-controlled study. *Am J Psychiatry* 1999;156(5):756–760.

123. van Vliet IM, den Boer JA, Westenberg HG, et al. Clinical effects of buspirone in social phobia: a double-blind placebo-controlled study. *J Clin Psychiatry* 1997;58(4):164–168.

124. Altamura AC, Pioli R, Vitto M, et al. Venlafaxine in social phobia: a study in selective serotonin reuptake inhibitor non-responders. *Int Clin Psychopharmacol* 1999;14(4):239–245.

125. Van Ameringen M, Mancini C, Oakman JM. Nefazodone in social phobia. *J Clin Psychiatry* 1999;60(2):96–100.

126. Worthington JJ 3rd, Zucker BG, Fones CS, et al. Nefazodone for social phobia: a clinical case series. *Depress Anxiety* 1998;8(3):131–133.

127. Pande AC, Davidson JR, Jefferson JW, et al. Treatment of social phobia with gabapentin: a placebo-controlled study. *J Clin Psychopharmacol* 1999;19(4):341–348.

128. Feltner DE, Pollack MH, Davidson JRT, et al. A placebo-controlled study of pregabalin treatment of social phobia. Anxiety Disorders Association of America Abstract, Washington, DC, 2000.

129. Brady K, Pearlstein T, Asnis GM, et al. Efficacy and safety of sertraline treatment of posttraumatic stress disorder: a randomized controlled trial. *JAMA* 2000;283(14):1837–1844.

130. Connor KM, Sutherland SM, Tupler LA, et al. Fluoxetine in post-traumatic stress disorder. Randomized, double-blind study. *Br J Psychiatry* 1999;175:17–22.

131. Malik ML, Connor KM, Sutherland SM, et al. Quality of life and posttraumatic stress disorder: a pilot study assessing changes in SF-36 scores before and after treatment in a placebo-controlled trial of fluoxetine. *J Trauma Stress* 1999;12(2):387–393.

132. van der Kolk BA, Dreyfuss D, Michaels M, et al. Fluoxetine in posttraumatic stress disorder. *J Clin Psychiatry* 1994;55(12):517–522.

133. Marshall RD, Schneier FR, Fallon BA, et al. An open trial of paroxetine in patients with noncombat-related, chronic posttraumatic stress disorder. *J Clin Psychopharmacol* 1998;18(1):10–18.

134. Zisook S, Chentsova-Dutton YE, Smith-Vaniz A, et al. Nefazodone in patients with treatment-refractory posttraumatic stress disorder. *J Clin Psychiatry* 2000;61(3):203–208.

135. Canive JM, Clark RD, Calais LA, et al. Bupropion treatment in veterans with posttraumatic stress disorder: an open study. *J Clin Psychopharmacol* 1998;18(5):379–383.

136. Connor KM, Davidson JR, Weisler RH, et al. A pilot study of mirtazapine in post-traumatic stress disorder. *Int Clin Psychopharmacol* 1999;14(1):29–31.

137. Hamner MB, Frueh BC. Response to venlafaxine in a previously antidepressant treatment-resistant combat veteran with post-traumatic stress disorder. *Int Clin Psychopharmacol* 1998;13(5):233–234.

138. Krashin D, Oates EW. Risperidone as an adjunct therapy for post-traumatic stress disorder. *Mil Med* 1999;164(8):605–606.

139. Raskind MA, Dobie DJ, Kanter ED, et al. The alpha 1-adrenergic antagonist prazosin ameliorates combat trauma nightmares in veterans with posttraumatic stress disorder: a report of 4 cases. *J Clin Psychiatry* 2000;61(2):129–133.

140. Neal LA, Shapland W, Fox C. An open trial of moclobemide in the treatment of post-traumatic stress disorder. *Int Clin Psychopharmacol* 1997;12(4):231–237.

141. Sutherland SM, Davidson JR. Pharmacotherapy for post-traumatic stress disorder. *Psychiatr Clin North Am* 1994;17(2):409–423.

142. Clark RD, Canive JM, Calais LA, et al. Divalproex in posttraumatic stress disorder: an open-label clinical trial. *J Trauma Stress* 1999;12(2):395–401.

143. Hertzberg MA, Butterfield MI, Feldman ME, et al. A preliminary study of lamotrigine for the treatment of posttraumatic stress disorder. *Biol Psychiatry* 1999;45(9):1226–1229.

144. Davidson JR, Potts N, Richichi E, et al. Treatment of social phobia with clonazepam and placebo. *J Clin Psychopharmacol* 1993;13(6):423–428.

145. Shalev AY, Bonne O, Eth S. Treatment of posttraumatic stress disorder: a review. *Psychosom Med* 1996;58(2):165–182.

146. Friedman MJ. What might the psychobiology of posttraumatic stress disorder teach us about future approaches to pharmacotherapy? *J Clin Psychiatry* 2000;61(suppl 7):44–51.

67

The Economic Burden of Anxiety and Stress Disorders

Ronald C. Kessler

Paul E. Greenberg

Ronald C. Kessler: Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts.

Paul E. Greenberg: Analysis Group, Cambridge, Massachusetts.

No society can afford to guarantee universal health insurance coverage for treatment of all illnesses for all of its citizens. The number of illnesses is simply too large and the costs of treatment too great for such a guarantee even in the most economically advantaged societies. Resource allocation rules are consequently needed (1). The most widely accepted of these rules emphasizes cost-effectiveness. According to this rule, medical interventions are appropriate only if their expected benefits clearly exceed the sum of their direct costs and their expected risks (2).

The difficulty in implementing this decision rule is that no obvious comparability exists between the single metric in which the costs of treatment are usually defined (i.e., dollars) and the many different metrics in which the benefits of treatment can be defined (e.g., physical pain, discomfort, psychological distress, and role impairment). To create transformations across these different metrics to allow for comparisons of costs and benefits on a single metric, a number of strategies have been developed, such as assessments of willingness to pay, time trade-off, standard gamble, and other utility or quasi-utility measures (3). In addition, a special interest has evolved in the indirect economic costs of illness and the benefits of treatment in terms of sickness absence and disability from work. The costs of these role impairments can be more easily assessed than the costs of other adverse effects of illness and represent the cost-benefit trade-off to purchasers of employer-sponsored health insurance plans (4).

The most ambitious effort to date to evaluate the costs of illness in terms of role impairments and disabilities is the World Health Organization (WHO) Global Burden of Disease (GBD) Study, an initiative designed to generate a rank ordering of the diseases that create the greatest societal burdens in terms of impairment and disability (5). The overarching goal of GBD is to help health policy planners prioritize disorder-specific resource allocation decisions. GBD focuses on economic costs of illness using a metric known as the disability-adjusted life year (DALY) (6), a weighted composite that combines expected years of lost life with expected years of decreased functioning due to a particular disease (or constellation of comorbid diseases).

The first generation of GBD estimates suggest that mental disorders, as a group, are the most costly diseases in the world and that major depression, in particular, is the single most costly disease among people in the middle years of life in terms of overall DALYs (5). Although the GBD rated mood disorders as considerably more costly than anxiety or stress disorders, focused cost-of-illness studies carried out subsequent to the publication of these estimates strongly suggest that the GBD underestimated the costs of anxiety and stress disorders and that the true costs of anxiety disorders are actually quite comparable to the costs of mood disorders (7 ,8).

Three reasons for the underestimation of the costs of anxiety and stress disorders in the GBD are worthy of note. The evidence to support all three of them is reviewed in this chapter. First, the epidemiologic studies used in GBD underestimated the prevalences of anxiety disorders. Second, the estimated effects of specific diseases on functioning were based on the judgments of experts rather than on objective evaluations of actual impairments in representative samples of people with the diseases. These judgments underestimated the impairments due to anxiety disorders. Third, comorbidities were ignored in making GBD cost estimates. As shown below, a consideration of comorbidities is critical in assessing the costs of anxiety disorders.

This chapter reviews available evidence on the economic burdens of anxiety and stress disorders. By focusing on eight factors that lead to the high societal costs of these disorders, we present evidence on the three sources of GBD underestimation listed above. These eight factors are as follows. First, anxiety and stress disorders are among the most commonly occurring of all chronic diseases. Second, the prevalences of

these disorders are increasing in recent cohorts in many countries. Third, these disorders have much earlier ages of onset than other commonly occurring chronic conditions. Fourth, anxiety and stress disorders are usually very chronic. Fifth, early-onset anxiety and stress disorders have a wide range of adverse effects on secondary outcomes, such as teen childbearing, marital stability, and educational attainment that have substantial economic implications. Sixth, these disorders are often associated with substantial impairments in role functioning. Seventh, anxiety and stress disorders are highly comorbid and usually temporally primary. Some of the disorders that are temporally secondary to anxiety and stress disorders, such as ulcers and substance abuse, have adverse economic effects that should be considered in part among the costs of anxiety and stress disorders. Eighth, despite the fact that effective treatments are available, only a minority of people with anxiety and stress disorders receives these treatments. Furthermore, those who receive these treatments usually do so only after many of the adverse effects of the disorders have occurred, making it very difficult to reverse the economic impacts of having had the disorders even with successful treatments. Based on all these factors, anxiety and stress disorders have to be considered among the most costly of all chronic physical and mental disorders.

PREVALENCES

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

A new generation of psychiatric epidemiologic surveys, which began with the Epidemiologic Catchment Area (ECA) Study in the early 1980s (9), has dramatically increased our knowledge about the general population prevalences and correlates of anxiety disorders. The ECA Study was the first psychiatric epidemiologic study to use a fully structured research diagnostic interview designed specifically for use by lay interviewers to operationalize the criteria of a wide range of mental disorders. This interview, known as the Diagnostic Interview Schedule (DIS) (10), was used throughout the 1980s and early 1990s to carry out parallel epidemiologic surveys in a number of countries (11 ,12). The DIS was also used as the basis for an elaborated interview developed by the WHO and known as the Composite International Diagnostic Interview (CIDI) (13). The CIDI was designed to generate diagnoses according to the definitions and criteria of both the International Classification of Diseases (ICD) and *Diagnostic and Statistical Manual of Mental Disorders* (DSM) systems. WHO auspices resulted in over a dozen large-scale, general-population CIDI surveys being carried out around the world over the past decade. Comparative analysis of these data has been facilitated by the creation of the WHO International Consortium in Psychiatric Epidemiology (ICPE) (14), which is currently coordinating national CIDI surveys in 25 countries around the world, with a combined sample size of over 150,000 respondents, as part of the WHO World Mental Health 2000 (WMH2000) Initiative (15).

The DIS and CIDI surveys show that anxiety and stress disorders are the most commonly occurring of all mental disorders. Clear illustration can be found in a recent report based on the results of six CIDI surveys carried out in Latin America, North America, and Europe (16). These surveys found that the lifetime prevalences of DSM third edition revised (III-R) anxiety disorders were as high as 25%, whereas prevalences in the year before the survey were as high as 17%. These prevalences were higher than those of any other class of mental disorders in the vast majority of the surveys. (The exceptions were a survey of adolescents in Germany and of residents of a large catchment area in Mexico City. In both of these surveys, substance use disorders were more common than anxiety disorders in the 12 months before the interview.)

It was noted above that the epidemiologic data available to the GBD researchers, which came from the DIS surveys carried out in the 1980s, underestimated the prevalence of anxiety and stress disorders. Three of the most prevalent and seriously impairing anxiety disorders were involved in this underestimation: generalized anxiety disorder (GAD), social phobia, and posttraumatic stress disorder (PTSD). The reasons for the underestimations differ from one of these disorders to the next. In the case of GAD, prevalence was underestimated in the early DIS surveys due to the fact that the excessively unrealistic criterion in the DSM-III was operationalized by requiring that respondents endorse a statement that they worried about things that were not really serious or about things that were not likely to happen. This requirement is overly restrictive in two ways. First, there is no requirement in DSM that people with GAD have insight into their worries being excessive or unrealistic. Although they must be aware that they worry more than other people do, they can perceive others as worrying too little rather than themselves as worrying too much. Second, even in the presence of a recognition that their worrying is excessive, there is no requirement in DSM that the worries of people with GAD must be exclusively focused on things that are not important or unlikely to happen. Indeed, the heterogeneous worries that are characteristic of most people with GAD (e.g., excessive concerns about job stability, how the children are going to turn out, neighborhood safety, global warming, etc.) often focus on serious matters that have nontrivial probabilities of occurring.

The restrictive assessment in the DIS led to the estimate that only about 3% of the population meet criteria for GAD at any time in their lives (17). Early CIDI surveys followed this same method of assessment and yielded similar prevalence estimates (18 ,19). Subsequent CIDI surveys expanded the assessment of excessive worry in GAD by asking respondents if there was ever a time in their lives when they were worriers or when they worried a lot more than most other people in their same situation, without requiring that

the worry be exclusively about things that are not serious or not likely to happen. Prevalence estimates were found to be considerably higher when this modification was introduced (20).

In addition, these new studies investigated the implications of the requirements in the DSM-IV and ICD-10 that the worry in GAD persists for a minimum of 6 months and found that this requirement might be too restrictive. In particular, many people with chronic excessive worry report having fairly short episodes, each of which lasts for several weeks or months, that continue in a chronic intermittent course for many years. Such individuals are currently excluded from a diagnosis of GAD and, because of their high comorbidity with depression, are classified as being depressed even though their most prominent symptoms are often associated with anxiety rather than depression. The new WHO WMH2000 Initiative is investigating this matter in some detail in an effort to evaluate whether the classification rules for GAD or mixed anxiety-depression should be modified to take these cases into consideration.

In the case of social phobia, the underestimation in the early DIS surveys was due to the fact that all phobias were assessed in a single question that presented respondents with a long checklist of feared situations and asked them if they ever had unreasonably strong fears of these situations. In addition to being mixed in with a number of specific fears, only five social phobic situations, all involving performance fears, were included in the ECA list.

This method of assessment led to the estimate that only 2.7% of the population meet criteria for social phobia at any time in their lives (21). Subsequent surveys that used the CIDI corrected this problem by screening for social phobia with a separate, longer list of social fears (both interactional and performance). These later surveys consistently found social phobia to be much higher than in the DIS surveys, with lifetime prevalences as high as 13% (18) and current prevalences as high as 8% (22).

Posttraumatic stress disorder was also wildly underestimated in the early DIS surveys. This seems to have been a result of including only a single extremely long and complex screening question for PTSD in the first version of the DIS. This question began with a statement that many people live through events that are outside the range of usual human experience, such as combat in a war or sexual assault, and that people who experience these events often have bad emotional reactions such as nightmares, flashbacks, and changes in mood. Respondents were then asked if they ever had such an event that caused such reactions and, if so, to tell the interviewer what this event was. Subsequent debriefing showed that this question was too complex for many respondents, that the absence of a detailed event list interfered with effective memory search, and that the requirement that the respondent describe the event out loud rather than give a yes or no response to event-specific questions led to underreporting of embarrassing events (23).

Assessments of PTSD in epidemiologic surveys that used the DIS led to the estimate that only about 1% of the United States population meet criteria for this disorder at any time in their life (24 ,25 and 26). Subsequent surveys that used the CIDI modified the assessment of PTSD by including a detailed traumatic event checklist and by asking respondents to give separate yes or no reports for whether each of these events ever occurred to them. In some CIDI surveys, a visual checklist was used that aimed at making it easier for respondents to report embarrassing events (e.g., “Did event number five on the list ever occur to you?” rather than “Were you ever raped?”). CIDI PTSD symptom assessment proceeded very much along the same lines as the DIS after documenting that trauma exposure had occurred. Yet the prevalence estimates obtained in the CIDI surveys were dramatically higher than in the DIS surveys, with lifetime prevalences as high as 12.2% (23 ,27).

It should also be noted that psychiatric epidemiologic surveys have not, up to now, attempted to assess either DSM acute stress disorder (a short-term disorder that occurs in reaction to traumatic stress) or adjustment disorder (a disorder that occurs in reaction to nontraumatic stress). This is important because epidemiologic surveys that include assessments of current nonspecific psychological distress typically find that a high proportion of the respondents who report clinically significant current distress in the anxiety-mood spectrum do not meet criteria for any of the anxiety or mood disorders typically assessed in these surveys (which usually include GAD, panic disorder, phobia, PTSD, obsessive-compulsive disorder, major depression, dysthymia, and mania). Given the extremely high prevalences of exposure to stressful events found in surveys of stress exposure (28), it is plausible to think that many of these people have a diagnosis of either acute stress disorder or adjustment disorder. The new WHO WMH2000 surveys mentioned earlier in this chapter are investigating this possibility by evaluating the link between stress and clinically significant nonspecific psychological distress among respondents who do not meet criteria for other anxiety or mood disorders.

Taken together, these results suggest that the combined prevalences of all anxiety and stress disorders make these among the most commonly occurring classes of seriously impairing chronic conditions. A rough comparison is provided by the recently completed Midlife Development in the U.S. (MIDUS) survey carried out by the John D. and Catherine T. MacArthur Foundation. In this survey, parallel assessments were made of commonly occurring physical and mental disorders, along with assessments of the effects of these disorders on day-to-day functioning (29). As in most other health surveys of

chronic physical conditions, of which a great many exist (e.g., 30,31), the significantly impairing physical disorders with the highest reported prevalences in the year before interview were back problems (20.3%), arthritis (19.4%), hypertension (18.2%), and seasonal allergies (15.7%). However, past health surveys of chronic physical conditions have seldom assessed emotional disorders along with these physical disorders. In doing so, the MIDUS survey found that 16.4% of respondents reported an anxiety or stress disorder exclusive of either major or minor depression, and that an additional 14.1% of respondents reported major or minor depression. These findings make anxiety-stress the fourth most commonly occurring impairing class of chronic disorders in the general population and major or minor depression the sixth most commonly occurring class of such disorders.

COHORT EFFECTS

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

In addition to anxiety and stress disorders having great importance because they are very common, they are also becoming increasingly prevalent over time. An illustration of this finding is presented in Table 67.1 , taken from ICPE surveys carried out in six countries (16). These results are based on synthetic cohort analyses using retrospective age-at-onset reports to evaluate intercohort differences in lifetime risk of anxiety disorders over a period of four decades. The data are clear in showing that the relative odds of having an anxiety disorder have steadily increased over this period in all six countries.

TABLE 67.1. THE EFFECTS (ODDS RATIOS) OF COHORT IN PREDICTING LIFETIME ANXIETY DISORDERS IN SIX COUNTRIES [a]

	Age Group				
	18–24	25–34	35–44	45–54	χ^2_3
Brazil	3.3*	3.1*	1.8*	1.0	64.4*
Canada	1.9*	1.7*	1.4*	1.0	20.7*
Mexico	2.1	2.0	2.0	1.0	2.3
Netherlands	2.2*	1.8*	1.5*	1.0	88.4*
Turkey	1.8*	1.7*	1.3	1.0	18.0*
United States	1.8*	1.4	1.1	1.0	27.9*

[a]Results are based on discrete-time survival analysis.
*Significant at the .05 level, two-sided test.
From WHO International Consortium of Psychiatric Epidemiology: cross-national comparisons of the prevalences and correlates of mental disorders: an ICPE study. *Bull WHO* 2000;78:420, with permission.

More detailed analyses of these and other data show that the increased prevalences of anxiety disorders are more pronounced than the increased prevalences of other mental disorders and that the apparent cohort effects for some other disorders, such as major depression, are largely due to increases in secondary disorders associated with primary anxiety (32). Furthermore, the increasing prevalences within the anxiety disorders have been found to be especially pronounced for GAD, generalized social phobia, and PTSD. Increases for panic, specific phobia, agoraphobia, and obsessive-compulsive disorder, in comparison, have been more modest. Although these studies have not investigated either acute stress disorder or adjustment disorder, separate evidence of secular increases in exposure to traumatic stress is consistent with the likelihood that the prevalences of these disorders have also been on the rise (33).

AGE AT ONSET

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

The discussion up to now has not clearly distinguished between lifetime and recent prevalences. This is an important distinction because the societal burden of a disorder is largely associated with its prevalence at a point in time. The latter, in turn, is a complex function of lifetime prevalence, age at onset, and chronicity. The comparatively high recent prevalence of anxiety-stress disorders found in the MIDUS survey indicates that the combined effects of these three components are strong. This is true, in part, because anxiety and stress disorders occur to a high proportion of the population at some point in the course of life. It is also true because these disorders have comparatively early ages at onset and high rates of chronicity. We focus first on age at onset.

Retrospective reports about age at onset are routinely collected in epidemiologic surveys and used to estimate synthetic onset distributions. Figure 67.1 presents Kaplan-Meier curves that show these onset distributions for any anxiety disorders in six countries the ICPE surveyed (16). The median age at onset of anxiety disorders in these surveys is less than 15 years of age. The only commonly occurring chronic physical disorder that has a similar age-at-onset distribution is hay fever. All other commonly occurring chronic physical disorders that have been shown to have an effect on role functioning have median ages at onset that occur much later, in some cases decades later, than anxiety disorders. Other mental disorders, including depression, substance use disorders, oppositional-defiant disorder, conduct disorder, and attention-deficit hyperactivity disorder, also have comparatively early ages at onset, although anxiety disorders are the temporally primary disorders in the vast majority of people with a lifetime history of any mental disorder (34). No information is available, in comparison, on age at onset of acute stress disorders or adjustment disorders.

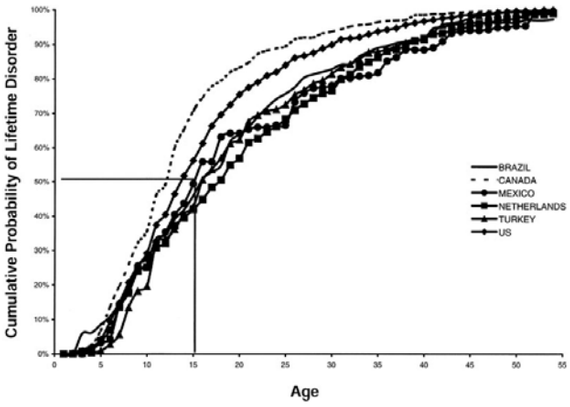


FIGURE 67.1. Age-at-onset distributions for any anxiety disorders in six countries. (Modified from WHO ICPE. Cross-national comparisons of the prevalences and correlates of mental disorders: an ICPE study. *Bull WHO* 2000;78:418, with permission.)

CHRONICITY

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

Although psychiatric epidemiologic surveys typically are cross-sectional, making it impossible to track illness course, indirect assessments of chronicity in these surveys have been carried out by comparing the ratios of current prevalence

to lifetime prevalence in subsamples of respondents with specific lifetime mental disorders. Results clearly suggest that anxiety disorders are the most chronic of all mental disorders (35). This indirect evidence is consistent with the results of longitudinal studies carried out in clinical samples, which uniformly show that anxiety disorders are typically very chronic (36 ,37 and 38). It is noteworthy that this high chronicity is not greater than that found among a number of impairing physical disorders, such as arthritis, asthma, and diabetes. However, the combined occurrence of high lifetime prevalence with early age at onset and high chronicity makes anxiety disorders unique. The one chronic physical disorder with comparable lifetime prevalence and early onset, hay fever, is active for only a few weeks each year. No systematic data exist on the chronicity of adjustment disorders, although epidemiologic data showing that PTSD is often a very persistent disorder (23 ,39) are consistent with the possibility that the same may be true for adjustment disorders.

ADVERSE EFFECTS ON SECONDARY OUTCOMES

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

Virtually all cost-of-illness studies focus on the effects of prevalent disorders on current role functioning, taking current roles as givens. The question implicitly addressed by these studies is whether it is in the financial interests of employers to invest in employee health care. Would the increased direct costs of treatment be offset by decreased indirect costs in such things as sickness absence, poor work performance, and accidents? This important question is discussed below. However, even when the focus is on narrow financial costs, the preceding is not the only question of importance in evaluating the societal costs of illness. Equally, if not more, important from a societal perspective is the question of whether the human capital potential of the individual is adversely affected by illness. Specifically, what difference does the existence of a particular chronic condition make to the individual's lifetime profile of productivity?

There is good evidence that anxiety disorders have long-term effects of this sort that are not captured in analyses of current role functioning. Both vital statistics (see Table 292A, Trend C in ref. 40) and prospective epidemiologic surveys (41) show that anxiety is associated with elevated risk of early death. Epidemiologic data also show that anxiety is associated with elevated risk of subsequent unemployment (42 ,43).

Clinical experience also suggests that anxiety is associated with more subtle decrements in role performance. It is common for patients with chronic GAD or PTSD, for example,

to work at low-paying jobs because they are unable to cope with the stresses of higher paying jobs. This would be considered a cost of illness from the societal perspective, but not from the perspective of the employer. Very little scientific evidence exists regarding opportunity costs of this sort. The most sustained examination of these costs was carried out in a series of reports from the National Comorbidity Survey (NCS) in which retrospective reports about the ages at onset of individual mental disorders were used to define time-varying predictors of subsequent transitions in educational attainment (44), teen childbearing (45), marital timing and stability (46), and earnings (42 ,43). The results clearly show that mental disorders, in general, and anxiety disorders, in particular, are associated with significantly elevated risks of several different life course events that have important adverse financial implications. In terms of standardized (for sociodemographics) odds ratios, NCS respondents with some early-onset anxiety disorders had 40% elevated odds of high school and college failure, 30% elevated odds of teenage childbearing, 60% elevated odds of marital instability, and 150% elevated odds of current unemployment at the time of interview.

It is important to recognize that this constellation of adverse individual life course consequences—especially school failure coupled with teen childbearing and marital instability—makes up the core components of welfare dependency. The costs of public assistance to single mothers with dependent children are paid by all taxpayers rather than by the welfare recipients themselves. For this reason, the component of welfare dependency costs explained by early-onset anxiety disorders should be considered a societal cost of anxiety. A number of innovative welfare-to-work programs are currently being carried out in response to welfare reform legislation in the United States (e.g., 47,48). Interestingly, early reports on these programs suggest that their success hinges on the mental health of welfare recipients (49).

EFFECTS ON CURRENT ROLE FUNCTIONING

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

As noted in the previous subsection, a number of cost-of-illness studies have evaluated the effects of chronic conditions on work role functioning. Most of these studies focus on physical disorders (e.g., 31). Most of those concerned with mental disorders focus on depression (e.g., 50). A small number of recent studies examined the effects of anxiety disorders on work functioning and found that these effects are quite substantial. These findings are an important element in the argument that anxiety disorders are among the most costly of all chronic conditions.

One of these studies, based on the NCS, examined the effects of individual mental disorders on work loss (missing a full day of work) and work cutback (either missing part of a day or working less efficiently than usual) during the month prior to the interview (51). Each of the six anxiety disorders evaluated in that study (GAD, panic disorder, specific phobia, social phobia, agoraphobia, and PTSD) had significant effects on work-cutback days, from a high of 4.9 days per month associated with PTSD to a low of 1.1 associated with social phobia. None of the six was significantly associated with work-loss days, implying that anxiety influences work largely by affecting the quality of performance on days at work rather than by reducing the amount of time spent at work.

The MIDUS survey yielded information that is even more interesting because it assessed both mental and physical disorders. Gross bivariate analyses showed that two mental disorders, both anxiety disorders, were among the top five of all chronic conditions in terms of average per capita number of past month work impairment days. These top five included GAD (6.0 work impairment days per month), thyroid disease (5.8 days), tuberculosis (5.4 days), varicose veins (5.4 days), and panic disorder (5.3 days). Furthermore, multivariate analyses controlling for age, gender, and other sociodemographic factors found that the same two anxiety disorders were among the top six in terms of unique effects on work impairment (29). Calculating the salary-equivalent magnitude of these effects, using self-reported salaries and partialing out the effects of other comorbid mental and physical disorders, led to the estimate that the excess absenteeism and lost productivity directly associated with anxiety disorders is approximately $4.1 billion per year in the United States (8).

PSYCHIATRIC COMORBIDITY

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

A number of studies in both treatment samples (52) and general population samples (35) document high rates of psychiatric comorbidity among people with anxiety disorders. Illustrative results from the NCS are reported in Table 67.2 . Shown here are odds ratios between anxiety disorders and other mental disorders both for lifetime comorbidities and for comorbidities of disorders that were active in the 6 months prior to the interview. As the latter odds ratios are generally larger than the former, there must be comorbidities between the persistence of anxiety disorders and the persistence of other disorders.

TABLE 67.2. COMORBIDITIES (ODDS RATIOS) BETWEEN DSM-III-R ANXIETY DISORDERS AND OTHER MENTAL DISORDERS ASSESSED IN THE NATIONAL COMORBIDITY SURVEY

	Panic Disorder	Phobias	GAD	PTSD
Lifetime comorbidities				
Major depression	6.6	4.1	9.4	5.2
Dysthymia	4.8	3.0	12.5	4.9
Mania	10.4	7.9	9.6	6.2
Alcohol abuse/dependence	1.6	1.7	2.0	1.7
Drug abuse/dependence	3.0	2.2	2.9	3.2
Nonaffective psychosis	20.1	4.7	15.0	9.4
Six-month comorbidities				
Major depression	14.4	6.4	17.8	7.1
Dysthymia	12.2	4.4	21.5	7.4
Mania	15.8	13.4	10.4	9.4
Alcohol abuse/dependence	1.4	2.3	2.7	2.2
Drug abuse/dependence	3.9	3.9	5.0	2.9
Nonaffective psychosis[a]	—	—	—	—

All values on table are significant at the .05 level, two-sided test.
DSM-III-R, *Diagnostic and Statistical Manual of Mental Disorders*, third edition–revised; GAD, generalized anxiety disorder; PTSD, posttraumatic stress disorder.
[a]Diagnostic hierarchy rules were suppressed in defining the disorders. Six-month nonaffective psychosis (NAP) was too rare to calculate odds-ratios with any of the anxiety disorders.
From Tsuang MT, Tohen M, Zahner GEP, eds. *Textbook in psychiatric epidemiology*. New York: Wiley, 1995:181, 183, with permission.

Several different possible explanations exist for these comorbidities. One is that prior history of other mental disorders might be associated, either as a risk factor or as a marker, with risk of the subsequent onset and persistence of anxiety disorders. The other is that anxiety disorders might be associated with the subsequent onset and persistence of other mental disorders. As briefly mentioned above, epidemiologic studies have found that the latter possibility is more consistent with the data. Comorbid anxiety disorders

are usually temporally primary to other comorbid mental disorders (34). In addition, survival analyses show that temporally primary anxiety disorders are powerful predictors of the subsequent onset and course of other mental disorders (35). In addition, panic disorder (53) and PTSD (54) are powerful predictors of suicidal behaviors.

It is not clear from these results that anxiety disorders are causal. Another possibility is that anxiety disorders are early outcomes of other causal factors, either environmental or genetic, that cause both anxiety disorders and the other mental disorders with which anxiety disorders are comorbid. To the extent that anxiety disorders are causal, the adverse effects of mental disorders that are secondary to anxiety disorders should be counted among the adverse consequences of anxiety disorders. A good case in point is secondary substance use disorders. Epidemiologic data show that early-onset anxiety disorders are significant predictors of subsequent substance use disorders, most likely mediated by self-medication (55). If these associations are causal, simulations suggest that the early intervention and successful treatment of anxiety disorders would prevent as many as one-fourth of all substance use disorders in the U.S. (56). The component of the costs of substance use disorders due to prior anxiety disorders, therefore, should be counted among the costs of anxiety in a comprehensive evaluation.

Whether anxiety disorders are causal and to what extent is an especially important issue in the case of depression, as many comparative cost-of-illness studies, including the WHO GBD study, suggest that depression is the most costly of all mental disorders (5). Yet epidemiologic data show that close to half of all cases of depression are secondary to one or more preexisting anxiety disorders (57). This priority of anxiety over depression is never taken into consideration in evaluating the costs of depression. Indeed, in those few instances where anxiety-depression comorbidity is considered, diagnostic hierarchy rules typically specify that the depression should be considered primary even though epidemiologic evidence consistently shows that anxiety is usually temporally primary.

The rationale for this hierarchy of depression over anxiety is usually that the impairments associated with such cases is thought to be due to the depression rather than to the anxiety (58), but available evidence argues against this claim. A good case in point involves comorbidity between GAD and depression. The results in Table 67.3 , taken from two U.S. national surveys, the NCS and the MIDUS, show that GAD without major depression is as important as major depression without GAD in leading to impairments in role functioning (20). Further analysis of these data showed that impairment is considerably higher among people with comorbid GAD and major depression than those with either GAD alone or major depression alone. Coupling the fact that GAD is temporally primary in the majority of these cases with the fact that GAD without major depression is associated with impairments equal to those of major depression

without GAD, argues that this interactive effect is due at least as much to GAD as to major depression.

	GAD Without Major Depression (MD)		Major Depression (MD) Without GAD		GAD Without MD vs. MD Without GAD	
	Survey 1	Survey 2	Survey 1	Survey 2	Survey 1	Survey 2
Fair/poor perceived mental health	6.0*	4.8*	3.3*	5.2*	1.6	0.8
High work impairment	3.5	3.5	3.5*	8.5*	0.9	0.5
High social impairment	2.5*	1.2	2.0*	1.6*	1.5	1.0

[a]The two surveys indicated here are the National Comorbidity Survey (Survey 1) (59) and the Midlife Development in the U.S. Survey (Survey 2) (19). Results are based on separate regression equations evaluating the effect of either GAD or MD in predicting one of the impairment measures in one of the samples controlling for sociodemographic variables (age, gender, education, race-ethnicity, employment status, marital status, and urbanicity) and other 12-month DSM-III-R disorders. Models in the first two columns evaluate the effect of 12-month GAD on the subsample of respondents who did not have 12-month major depression. Models in the middle two columns evaluate the effect of 12-month major depression on the subsample of respondents who did not have 12-month GAD. Models in the last two columns evaluate the relative impairments of GAD without MD versus MD without GAD in analyses that are confined to respondents in those two subsamples.
*Significant at the .05 level, two-sided test.
From Kessler RC, DuPont RL, Berglund P, et al. Impairment in pure and comorbid generalized anxiety disorder and major depression in two national surveys. *American Journal of Psychiatry* 1999; 156:1915–1923, with permission.

TABLE 67.3. THE EFFECTS (ODDS RATIOS) OF 12-MONTH GENERALIZED ANXIETY DISORDER (GAD) WITHOUT MAJOR DEPRESSION (MD) AND MAJOR DEPRESSION WITHOUT GAD IN PREDICTING IMPAIRMENTS IN TWO U.S. NATIONAL SURVEYS, CONTROLLING FOR SOCIODEMOGRAPHICS AND OTHER 12-MONTH DSM-III-R DISORDERS[a]

PHYSICAL COMORBIDITY

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

Although it has not been as extensively studied, evidence from clinical samples suggests that anxiety disorders have significant comorbidities with certain chronic physical disorders (60 ,61). Table 67.4 presents nationally representative general population data on some of these comorbidities. The results are odds ratios for the relationships between the 12-month prevalences of the DSM-III-R anxiety disorders assessed in the NCS and selected physical disorders assessed in the NCS chronic conditions checklist. As shown in the table, all but one of the odds ratios are greater than 1.0, indicating a positive relationship, and half are statistically significant at the .05 level.

	GAD	Panic Disorder	Simple Phobia	Social Phobia	Agoraphobia	PTSD
Arthritis	1.7	2.1*	1.4	1.2	1.5	2.0*
Asthma	1.9	2.1*	2.0*	1.4	1.2	1.7*
Hypertension	1.5	2.2*	1.5*	1.0	2.2*	1.6
Kidney/liver disease	2.0	1.4	3.5*	1.2	2.3	1.9
Ulcer	3.1*	2.7*	2.7*	2.7*	2.9*	2.0*

*Significant at the .05 level, two-sided test.

TABLE 67.4. COMORBIDITIES (ODDS RATIOS) BETWEEN 12-MONTH PREVALENCES OF DSM-III-R ANXIETY DISORDERS AND CHRONIC PHYSICAL DISORDERS IN THE NATIONAL COMORBIDITY SURVEY

The NCS did not obtain information about age at onset for these physical disorders, making it impossible to examine whether anxiety disorders are temporally primary. In some cases, such as the strong association of some anxiety disorders

with ulcers, the most plausible interpretation is that anxiety had a causal impact on the subsequent onset of the physical condition. In other cases, it is equally plausible that the physical condition helped promote the subsequent onset of anxiety. It is also possible that bidirectional causal influences were at work or that common causes led to both conditions. The eventual resolution of this uncertainty is important for an evaluation of the costs of anxiety disorders, as both the direct treatment costs and the indirect costs of physical disorders that are partly caused by anxiety should be included in evaluating the overall societal costs of anxiety.

Comorbidities of anxiety with physical disorders are also important because of evidence that anxiety disorders reduce the quality of life of patients with physical disorders (62) and complicate the expression and course of physical disease (63). The most plausible explanation for these findings is that anxiety heightens sensitivity about both physical signs and symptoms and adequacy of treatment. This possibility is consistent with the finding that adjunctive treatments for comorbid anxiety often increase adherence to physical disorder treatment regimens (64).

MENTAL HEALTH TREATMENT

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

Effective psychological (65) and pharmacologic (66) therapies exist for the treatment of most anxiety disorders. The indirect costs of anxiety disorders would consequently be expected to decline if a high proportion of people with these disorders sought treatment. However, a substantial part of the adverse effects of anxiety disorders are associated with secondary effects that occur early in life (e.g., teen childbearing, school failure), so it is important that treatment occur early in the course of the anxiety disorder. As anxiety disorders have early ages of onset, initial treatment must occur during childhood or adolescence to be maximally effective in preventing adverse effects.

We are aware of only two epidemiologic studies that investigated speed of initial treatment contact after first onset of anxiety disorders (67 ,68). These studies considered three anxiety disorders: GAD, panic disorder, and phobias. Both studies found that only a small proportion of people with childhood-onset or adolescent-onset anxiety disorders seek treatment prior to adulthood. Median delays between first onset and initial treatment contact were found to be more than a decade for some anxiety disorders. Furthermore, delays were found to be inversely related to age at onset.

Because of these delays, only a minority of people with active anxiety disorders receives treatment in a given year. This is shown clearly in Table 67.5 , which presents nationally representative U.S. data from the NCS on seeking professional help for DSM-III-R anxiety disorders during the 12 months prior to the survey. Only about one out of every four people with an anxiety disorder sought any type of treatment and only 13.3% received any type of mental health specialty treatment during this year.

	Help-Seeking in Health Care Sectors[a]						Help-Seeking in Other Sectors[a]					
	General Medical		Specialty Mental Disorders		Any		Human Services		Self-Help		Any Help-Seeking[a]	
Disorder	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)
Generalized anxiety disorder	18.6	(3.8)	19.8	(3.5)	31.8	(4.8)	10.8	(2.1)	11.0	(3.0)	38.7	(4.3)
Panic disorder	21.5	(5.1)	24.3	(4.4)	35.2	(5.6)	21.0	(4.2)	12.5	(4.1)	46.4	(6.6)
Simple phobia	8.5	(1.4)	12.5	(1.5)	16.4	(1.7)	10.6	(1.8)	8.1	(1.3)	25.7	(2.3)
Social phobia	5.9	(0.9)	11.3	(1.7)	15.3	(1.9)	8.0	(1.3)	7.0	(1.5)	23.0	(2.2)
Agoraphobia	13.6	(3.6)	15.7	(3.2)	24.9	(4.7)	12.5	(3.1)	9.2	(2.9)	33.2	(5.0)
Posttraumatic stress disorder	12.5	(2.4)	22.3	(3.4)	28.2	(3.6)	16.3	(2.3)	11.8	(2.9)	38.3	(4.2)
Any	9.0	(1.3)	13.3	(1.4)	18.7	(1.7)	9.6	(0.9)	8.2	(1.0)	26.5	(2.0)

SE, standard error.
[a]Prevalence estimates are percentaged by rows. For example, in the first row of numbers, 18.6% is the percent of people with generalized anxiety disorder who used general medical services, not the percent of people using general medical services who carried a diagnosis of generalized anxiety disorder.
American Journal of Psychiatry 1999;156:117, with permission.

TABLE 67.5. 12-MONTH PREVALENCE OF PROFESSIONAL HELP-SEEKING IN SEPARATE SERVICE SECTORS IN THE NATIONAL COMORBIDITY SURVEY, BY 12-MONTH DSM-III-R ANXIETY DISORDER

INAPPROPRIATE USE OF GENERAL MEDICAL SERVICES

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

Although anxiety disorders typically are not treated, it is a great irony that people with anxiety disorders are often high

utilizers of primary care services. Indeed, people with untreated anxiety disorders make up a large proportion of the people who overuse primary care for only vaguely defined physical complaints (69 ,70). A recent anxiety disorders cost-of-illness study estimated that unnecessary medical care costs represented the largest single component of the cost of anxiety disorders in the U.S., equal to $23 billion per year (8). There is good reason to believe that aggressive screening and outreach efforts in primary care could detect these people with untreated anxiety, channel them into appropriate treatment, and possibly have a major offset effect in reducing unnecessary primary care costs. Interventions to evaluate the magnitude of this offset effect are currently under way.

OVERALL COSTS

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

There have been two recent attempts to estimate the total annual cost of anxiety disorders in the U.S. The first, carried out by DuPont et al. (7) in 1996, estimated that the annual cost of anxiety disorders is $47 billion, whereas the second, carried out by Greenberg et al. (8) in 1999, estimated that this cost is $42 billion. These estimates are quite comparable to the annual cost of depression, which has been estimated to be between $44 billion (71) and $53 billion (72). The rough equivalence to the cost of depression is important because, as noted in the introduction, depression is generally considered the most costly of all physical or mental disorders among people in the early to middle years of life (5).

The true societal costs of anxiety disorders, however, are actually a great deal larger than these estimates suggest, as the estimates were based only on a limited set of cost components. The components included direct psychiatric treatment costs, unnecessary medical treatment costs, work performance costs in terms of sickness absence and work-cutback days, and mortality costs (evaluated as lost earnings potential). The major excluded costs were long-term opportunity costs (i.e., excess unemployment and underemployment) and costs associated with comorbidity. The first of these two excluded costs is likely to be in excess of $2,000 per year for each person with a lifetime history of anxiety disorder (42), which is equivalent to an annual cost of more than $100 billion in the total U.S. population. The second of the excluded costs is impossible to calculate with currently available data, but would have to include substantial components of the costs conventionally attributed to depression, alcohol, and drug abuse, and the many other mental and physical disorders with lifetime prevalences and courses that are influenced by the prior existence of anxiety disorders.

DISCUSSION

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

Anxiety disorders are unique among all chronic conditions, both physical and mental, in having a combination of very high prevalence, early age at onset, high chronicity, and substantial role impairment. Although our knowledge about the comparative costs of different illnesses is too primitive to make precise comparisons, this conjunction of factors arguably makes anxiety disorders one of the most costly classes of illness in existence. Increased treatment is the key to reducing these costs. Although an increase in treatment will add to direct costs, the fact that available treatments are effective and that the adverse effects of anxiety are chronic means that the costs of effective treatment can be amortized over many years.

The fact that most anxiety disorders have childhood or adolescent onsets means that early outreach and treatment could be carried out in collaboration with schools. Unfortunately, as most people with anxiety delay initial contact with the treatment system for many years and usually present for treatment only after the onset of secondary comorbid disorders, little is known about the long-term effects of early treatment of pure childhood and adolescent anxiety disorders. Demonstration projects and long-term follow-up studies are needed to evaluate these effects and to target opportunities for incremental cost-effectiveness associated with refinements in diagnosis and treatment. Although the outcomes of such studies are uncertain, it is difficult to think of another disorder where an investment in early intervention has as great a potential for long-term societal benefits.

ACKNOWLEDGMENTS

Part of "67 - The Economic Burden of Anxiety and Stress Disorders "

Preparation of this chapter was supported, in part, by U.S. Public Health Service grants K05 MH00507, R01 MH46376, R01 MH49098, and RO1 MH52861, W.T. Grant Foundation grant 90135190, and an unrestricted grant from the Anxiety Disorders Association of America (ADAA).

The authors appreciate the helpful comments of Naomi Breslau, Evelyn Bromet, Kathleen Merikangas, Bedirhan Ustun, and Uli Wittchen on an earlier version of this manuscript.

Dr. Kessler receives research support from Pfizer, SmithKline Beecham, and Wyeth-Ayerst, Inc.

REFERENCES

1. Bobadilla JL, Cowley P, Musgrove P, et al. Design, content, and financing of an essential national package of health services. *Bull WHO* 1994;72:653–662.

2. Brook RH, Chassin MR, Fink A, et al. A method for detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care* 1986;72:53–63.

3. Torrence GW, Feeny D. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care* 1989;5:559–575.

4. Kessler RC, Stang PE, eds. *Health and work productivity: emerging issues in research and policy*. Chicago: University of Chicago Press, 1999: *in press*.

5. Murray CJL, Lopez AD, eds. *The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Cambridge, MA: Harvard University Press, 1996.

6. Murray CJL. Quantifying the burden of disease: the technical basis for DALYs. *Bull WHO* 1994;72:429–445.

7. DuPont RL, Rice DP, Miller LS, et al. Economic costs of anxiety. *Anxiety* 1996;2:167–172.

8. Greenberg PE, Sisitsky T, Kessler RC, et al. The economic burden of anxiety disorders in the 1990s. *J Clin Psychiatry* 1999;60(7):427–435.

9. Robins LN, Regier DA, eds. *Psychiatric disorders in America: rhe Epidemiologic Catchment Area Study*. New York: The Free Press, 1991.

10. Robins LN, Helzer JE, Croughan JL, et al. National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. *Arch Gen Psychiatry* 1981;38:381–389.

11. Cross-National Collaborative Group. The changing rate of major depression: cross-national comparisons. *JAMA* 1992;268:3098–3105.

12. Weissman MM, Bland RC, Canino GL, et al. Cross-national epidemiology of major depression and bipolar disorder. *JAMA* 1996;276:293–299.

13. Robins LN, Wing J, Wittchen H-U, et al. The Composite International Diagnostic Interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry* 1988;45:1069–1077.

14. Kessler RC. The World Health Organization International Consortium in Psychiatric Epidemiology (ICPE): initial work and future directions—the NAPE lecture 1998. *Acta Psychiatr Scand* 1999;99:2–9.

15. Kessler RC, Ustun TB. The World Health Organization: World Mental Health 2000 Initiative. *Hosp Manag Int* 2000:195–196.

16. WHO International Consortium in Psychiatric Epidemiology. Cross-national comparisons of the prevalences and correlates of mental disorders: an ICPE survey. *Bull WHO* 2000;78:413–426.

17. Blazer DG, Hughes D, George LK, et al. Generalized anxiety disorder. In: Robins LN, Regier DA, eds. *Psychiatric disorders in America: the Epidemiologic Catchment Area Study*. New York: The Free Press, 1991:180–203.

18. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994;51:8–19.

19. Wittchen H-U, Zhao S, Kessler RC, et al. DSM-III-R generalized anxiety disorder in the National Comorbidity Survey. *Arch Gen Psychiatry* 1994;51:355–364.

20. Kessler RC, DuPont RL, Berglund P, et al. Impairment in pure and comorbid generalized anxiety disorder and major depression in two national surveys. *Am J Psychiatry* 1999;156:1915–1923.

21. Eaton WW, Dryman A, Weissman MM. Panic and phobia. In: Robins LN, Regier DA, eds. *Psychiatric disorders in America: the Epidemiologic Catchment Area Study*. New York: The Free Press, 1991:155–179.

22. Stein MB, Walker JR, Forde DR. Setting diagnostic thresholds for social phobia: considerations from a community survey of social anxiety. *Am J Psychiatry* 1994;151:408–412.

23. Kessler RC, Sonnega A, Bromet E, et al. Posttraumatic stress disorder in the National Comorbidity Survey. *Arch Gen Psychiatry* 1995;52:1048–1060.

24. Davidson JRT, Hughes D, Blazer D, et al. Posttraumatic stress disorder in the community: an epidemiological study. *Psychol Med* 1991;21:1–19.

25. Helzer JE, Robins LN, McEvoy L. Post-traumatic stress disorder in the general population. *N Engl J Med* 1987;317:1630–1634.

26. Shore JH, Vollmer WM, Tatum EI. Community patterns of posttraumatic stress disorders. *J Nerv Ment Dis* 1989;177:681–685.

27. Breslau N, Kessler RC, Chilcoat HD, et al. Trauma and Posttraumatic Stress Disorder in the community: the 1996 Detroit Area Survey of Trauma. *Arch Gen Psychiatry* 1998;55:626–632.

28. Turner RJ, Wheaton B, Lloyd DA. The epidemiology of social stress. *Am Sociol Rev* 1995;60(1):104–125.

29. Kessler RC, Mickelson KD, Barber C, et al. The chronic association between medical conditions and work impairment. In: Rossi AS, ed. *Caring and doing for others: social responsibility in the domains of family, work, and community*. Chicago: University of Chicago Press, 2001.

30. Centers for Disease Control and Prevention. Health-related quality of life and activity limitation—eight states, 1995. *MMWR* 1998;47:134–139.

31. Verbrugge LM, Patrick DL. Seven chronic conditions: their impact on US adults' activity levels and use of medical services. *Am J Public Health* 1995;85:173–182.

32. Kessler RC, McGonagle KA, Nelson CB, et al. Sex and depression in the National Comorbidity Survey. II: cohort effects. *J Affective Disord* 1994;30:15–26.

33. Kessler RC. Posttraumatic stress disorder: the burden to the individual and society. *J Clin Psychiatry* 2000;61(suppl 5):4–12.

34. Kessler RC. Epidemiology of psychiatric comorbidity. In: Tsuang MT, Tohen M, Zahner GEP, eds. *Textbook in psychiatric epidemiology*. New York: Wiley, 1995:179–197.

35. Kessler RC. The prevalence of psychiatric comorbidity. In: Wetzler S, Sanderson WC, eds. *Treatment strategies for patients with psychiatric comorbidity*. New York: Wiley, 1997:23–48.

36. Noyes R Jr, Holt CS, Woodman CL. Natural course of anxiety disorders. In: Mavissakalian MR, Prien RF, eds. *Long-term treatments of anxiety disorders*. Washington, DC: American Psychiatric Press, 1996:1–48.

37. Rogers MP, Warshaw MG, Goisman RM, et al. Comparing primary and secondary generalized anxiety disorder in a long-term naturalistic study of anxiety disorders. *Depress Anxiety* 1999;10(1):1–7.

38. Yonkers KA, Zlotnick C, Allsworth J, et al. Is the course of panic disorder the same in women and men? *Am J Psychiatry* 1998;155(5):596–602.

39. Zlotnick C, Warshaw M, Shea MT, et al. Chronicity in posttraumatic stress disorder (PTSD) and predictors of course of comorbid PTSD in patients with anxiety disorders. *J Trauma Stress* 1999;12(1):89–100.

40. Issued by funding/sponsoring agency: Office of Vital Statistics. *Death rates for 282 selected causes by 5-year age groups, color, and sex: United States: 1979–1989*. Trend C, Table 292A. Hyattsville, MD: National Center for Health Statistics; Centers for Disease Control; Public Health Service. Sponsored by the U.S. Department of Health and Human Services, 1992.

41. Bruce ML, Leaf PJ. Psychiatric disorders and 15-month mortality in a community sample of older adults. *Am J Public Health* 1989;79(6):727–730.

42. Ettner SL, Frank RG, Kessler RC. The impact of psychiatric disorders on labor market outcomes. *Ind Labor Relat Rev* 1997;51:64–81.

43. Jayakody R, Danziger S, Kessler RC. Early-onset psychiatric disorders and male socioeconomic status. *Social Sci Res* 1998;27:371–387.

44. Kessler RC, Foster CL, Saunders WB, et al. Social consequences of psychiatric disorders, I: educational attainment. *Am J Psychiatry* 1995;152:1026–1032.

45. Kessler RC, Berglund PA, Foster CL, et al. Social consequences of psychiatric disorders, II: teenage parenthood. *Am J Psychiatry* 1997;154:1405–1411.

46. Kessler RC, Walters EE, Forthofer MS. The social consequences of psychiatric disorders: III. Probability of marital stability. *Am J Psychiatry* 1998;155:1092–1096.

47. Friedlander D, Burtless G. *Five years after: the long-term effects of welfare-to-work programs*. New York: Russell Sage Foundation, 1996.

48. Gueron JM, Pauly E. *From welfare to work*. New York: Russell Sage Foundation, 1991.

49. Danzinger SK, Corcoran M, Danzinger SH, et al. Barriers to the employment of welfare recipients [discussion paper]. Madison, WI: University of Wisconsin-Madison; Institute for Research on Poverty, 1999. Available at http://www.ssc.wisc.edu/irp/dplist.htm.

50. Mintz J, Mintz LI, Arruda MJ, et al. Treatments of depression and the functional capacity to work. *Arch Gen Psychiatry* 1992;49:761–768.

51. Kessler RC, Frank RG. The impact of psychiatric disorders on work loss days. *Psychol Med* 1997;27:861–873.

52. DiNardo PA, Barlow DH. Syndrome and symptom co-occurrence in the anxiety disorders. In: Maser JD, Cloninger CT, eds. *Comorbidity of mood and anxiety disorders*. Washington, DC: American Psychiatric Press, 1990:205–230.

53. Weissman MM, Klerman GL, Markowitz JS, et al. Suicidal ideation and suicide attempts in panic disorder and attacks. *N Engl J Med* 1989;321:1209–1214.

54. Kessler RC, Borges G, Walters EE. Prevalence and risk factors of lifetime suicide attempts in the National Comorbidity Survey. *Arch Gen Psychiatry* 1999;56:617–626.

55. Kessler RC, Crum RM., Warner LA, et al. The lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Arch Gen Psychiatry* 1997;54:313–321.

56. Kessler RC, Aguilar-Gaxiola S, Andrade L, et al. [Mental-substance comorbidities in the ICPE surveys]. *Psychiatria Fennica* 2001;32(suppl. 2):62–80.

57. Kessler RC, Nelson CB, McGonagle KA, et al. Comorbidity of DSM-III-R major depressive disorder in the general population: results from the US National Comorbidity Survey. *Br J Psychiatry* 1996;168(suppl 30):17–30.

58. Roy-Byrne PP. Generalized anxiety and mixed anxiety-depression: association with disability and health care utilization. *J Clin Psychiatry* 1996;57:86–91.

59. Kessler RC. The National Comorbidity Survey of the United States. *Int Rev Psychiatry* 1994;6:365–376.

60. Sherbourne CD, Jackson CA, Meredith LS, et al. Prevalence of comorbid anxiety disorders in primary care outpatients. *Arch Fam Med* 1996;5(1):27–34.

61. Zaubler TS, Katon W. Panic disorder and medical comorbidity: a review of the medical and psychiatric literature. *Bull Menninger Clin* 1996;60:A12–A38.

62. Sherbourne CD, Wells KB, Meredith LS, et al. Comorbid anxiety disorder and the functioning and well-being of chronically ill patients of general medical providers. *Arch Gen Psychiatry* 1996;53(10):889–895.

63. Stoudemire A. Psychological factors affecting medical conditions: summary. In: Stoudemire A, ed. *Psychological factors affecting medical conditions*. Washington, DC: American Psychiatric Press, 1995:187–192.

64. Meichenbaum D, Turk DC. *Facilitating treatment adherence: a practitioner's guidebook*. New York: Plenum Press, 1987.

65. Barlow DH. *Anxiety and its disorders*. New York: Guilford, 1988.

66. Haskett RF, Desmet A, Salinas EO. Venlofaxine XR in the treatment of anxiety. *Acta Psychiatr Scand* 1999; *in press*.

67. Kessler RC, Olfson M, Berglund PA. Patterns and predictors of treatment contact after first onset of psychiatric disorders. *Am J Psychiatry* 1998;155:62–69.

68. Olfson M, Kessler RC, Berglund PA, et al. Psychiatric disorder onset and first treatment contact in the United States and Ontario. *Am J Psychiatry* 1998;155:1415–1422.

69. Katon W, Von Korff M, Lin E, et al. Distressed high utilizers of medical care: DSM-III-R diagnoses and treatment needs. *Gen Hosp Psychiatry* 1990;12(6):355–62.

70. Manning WG Jr., Wells KB. Effects of psychological distress and psychological well-being on use of medical services. *Med Care* 1992;30:541–553.

71. Greenberg PE, Stiglin LE, Finkelstein SN, et al. The economic burden of depression in 1990. *J Clin Psychiatry* 1993;54:405–418.

72. Greenberg PE, Kessler RC, Nells TL, et al. Depression in the workplace: an economic perspective. In: Feighner JP, Boyer WF, eds. *Selective serotonin re-uptake inhibitors: advances in basic research and clinical practice*. New York: Wiley, 1996:327–363.

68

Mechanism of Action of Anxiolytics

John F. Tallman

James Cassella

John Kehne

John F. Tallman, James Cassella, and John Kehne: Neurogen Corporation, Branford, Connecticut.

Drugs to reduce anxiety have been used by human beings for thousands of years. One of the first anxiolytics and one that continues to be used by humans is ethanol. A detailed description of ethanol's action may be found in Chapter 100 . A number of other drugs including the barbiturates and the carbamates (meprobamate) were used in the first half of the 20th century and some continue to be used today. This chapter focuses on current drugs that are used for the treatment of anxiety and approaches that are currently under investigation.

CORTICOTROPIN-RELEASING FACTOR (CRF)

Part of "68 - Mechanism of Action of Anxiolytics "

Corticotropin-releasing factor (CRF) is a 41 amino acid peptide that plays an important role in mediating the body's physiologic and behavioral responses to stress (1). Figure 68.1 illustrates that this role of CRF may be mediated by multiple sites of action. As a secretagogue, CRF stimulates the release of adrenocorticotropic hormone (ACTH) from the pituitary. In addition, CRF plays a neurotransmitter or neuromodulatory role through neurons and receptors distributed in diverse brain regions (2). CRF neurons, localized in the hypothalamic periventricular nucleus, are a major mediator of stress-induced activation of the hypothalamic-pituitary-adrenal (HPA) axis, whereas pathways innervating limbic and cortical areas are thought to mediate the behavioral effects of CRF.

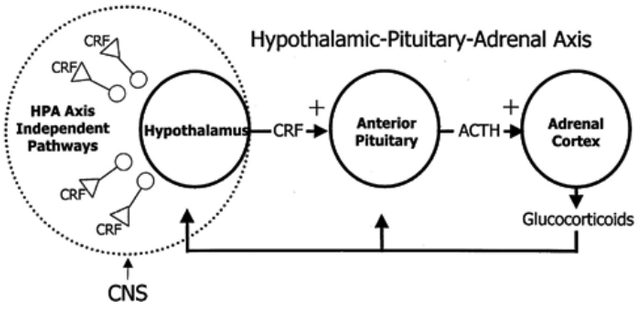


FIGURE 68.1. The role of corticotropin-releasing factor.

There is a large body of both preclinical and clinical literature implicating a key role of CRF in affective disorders such as anxiety and depression. A significant clinical literature suggests that dysfunctions of CRF in its role as a hormone in the HPA axis or as a neurotransmitter in the brain may contribute to the etiology of a variety of psychiatric conditions, including anxiety and depression (3). The link between CRF and depression is particularly strong, as numerous clinical studies have demonstrated that depressed patients show elevated cerebrospinal fluid (CSF) levels of CRF, elevated plasma cortisol, and a blunted ACTH response following intravenous CRF. Successful antidepressant treatment was shown to have a normalizing effect on CRF levels. A role of CRF in anxiety disorders has also been postulated, though the clinical evidence is not as strong as it is for depression (4).

Preclinical studies have demonstrated that CRF administered exogenously into the central nervous system (CNS) can produce behaviors indicative of anxiety and depression, for example, heightened startle responses, anxiogenic behaviors on the elevated plus maze, decreased food consumption, and altered sleep patterns. The anxiogenic effects of CRF are not blocked by adrenalectomy, suggesting that they are centrally mediated effects occurring independently of the HPA axis (5). Other studies strengthening the link between CRF and anxiety include recent work by Kalin et al. (6) demonstrating that a “fearful” phenotype in monkeys is associated with increased pituitary-adrenal activity and increased brain CRF levels. Other studies have shown that exposure to early postnatal separation stress in rat pups results in elevated levels of CRF messenger RNA (mRNA) in brain regions including the paraventricular nucleus (PVN) and the central nucleus of the amygdala (7 ,8).

Molecular Mechanism of Action

A substantial scientific effort has been directed toward characterizing the molecular biology of CRF pathways (9). Perrin and Vale (9) first isolated CRF and identified it as a secretagogue for ACTH in primary cultures of rat pituitary cells. CRF activity is shared by two nonmammalian peptides, sauvagine and urotensin I, which share a 50% homology with CRF, and by a new mammalian peptide, urocortin, which has a 45% sequence homology.

CRF acts through two G_s-protein coupled receptors, the CRF-1 and CRF-2 receptor subtypes (9, 10) . CRF-1 receptors show homology to a number of other neuropeptide receptors, including vasointestinal peptide (VIP) and calcitonin. Three splice variants of the CRF-2 receptor subtype, the CRF-2α, CRF-2β, and CRF-2γ, and two splice variants of the CRF-1 receptor, have been identified (9) . Molecular characterization studies have demonstrated that there is approximately a 70% sequence homology between CRF-1 and CRF-2 receptor subtypes. Cloning of the human *CRF-2a* gene revealed that it is 94% identical to the rat CRF-2α receptor and 70% identical to the human CRF-1 receptor. There is currently no evidence of the existence of the CRF-2β receptor in humans.

CRF-1 and CRF-2 receptors have different pharmacology and different localizations in the brain and periphery. *In situ* hybridization and receptor autoradiography techniques been used to map the relative distributions of CRF-1 and CRF-2 receptors in the rat brain (11, 12) . High expression of CRF-1 receptors was seen in the pituitary, and in a number of brain regions including the PVN of the hypothalamus, cerebral cortex, olfactory bulb, cerebellar cortex, and basolateral and medial amygdala. In contrast, high densities of CRF-2 are found in more circumscribed regions, including the lateral septum, ventromedial nucleus of the thalamus, and choroid plexus. Moderate densities of CRF-2 receptors were reported for the medial amygdala and dorsal raphe nucleus. Further characterization has indicated that the CRF-2α splice variant accounts for the brain localization of CRF-2 receptors, whereas the CRF-2β accounts for choroid plexus. Urocortin, rather than CRF, most closely maps to CRF-2 receptors, leading to the suggestion that it may be the endogenous ligand for CRF-2 receptors. Recent work has shown that the distribution of CRF-2 receptors may differ significantly in the nonhuman primate brain relative to the rodent brain such that the CRF-2 subtype may play a more significant role than previously thought (13) .

CRF receptors utilize 3′,5′-cyclic adenosine monophosphate (cAMP) as a second messenger in the pituitary and brain and can be regulated by chronic activation. Thus, desensitization following exposure to CRF has been demonstrated both *in vitro* (14) and *in vivo* (15) . Furthermore, chronic stress can down-regulate CRF receptors and decrease CRF-stimulated cAMP production in multiple brain areas (16, 17) . Down-regulation of pituitary CRF receptors following adrenalectomy presumably results from decreased ACTH mediated inhibitory feedback, which produces excess CRF stimulation.

There are a number of pharmacologic agents available for dissecting the functional significance of CRF-1 and CRF-2 receptors. Much work has been carried out using the peptide antagonists α-helical-CRF (9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 and 41) and D-Phe CRF (12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 and 41) . However, these compounds have shortcomings in that they do not penetrate the CNS and therefore have to be administered intracerebrally. Furthermore, they do not discriminate between CRF receptor subtypes and therefore do not allow a determination of their relative contributions to behavior. More recently, the development of selective, nonpeptidic antagonists of the CRF-1 receptor such as CP 154,526 (18) have provided important pharmacologic tools for the analysis of CRF-1 receptor function. Mutation studies have demonstrated that peptide and nonpeptide antagonists bind to different domains of the CRF-1 receptor (9) . To date, selective CRF-2 antagonists have not been described, though recently, nonpeptide dual antagonists of the CRF-1 and CRF-2 receptors have been described (19) .

In addition to CRF receptor subtypes, another potential target for pharmacologic manipulation of CRF is the CRF binding protein (CRF-BP), a 322 amino acid peptide. CRF-BP is a specific carrier for CRF and related peptides found in human plasma and brain. CRF-BP is thought to be a modulator of CRF activity. The CRF-BP is found in high densities in the rat amygdala, cortex, and bed nucleus of the stria terminalis.

Genetically Altered Mouse Models

Studies utilizing transgenic and knockout mouse models have provided important information with regard to the contribution of CRF and CRF receptor subtypes to processes including energy balance, emotionality, cognition, and drug dependence (20) . This chapter focuses on the evidence implicating CRF and CRF receptors in anxiety states.

Overexpression of CRF in transgenic mice produced anxiogenic effects using either the black-white box test (21) or the elevated plus maze (22) . The latter effect was reversed by central administration of the CRF receptor antagonist α-helical CRF, but not by adrenalectomy, supporting the role of central CRF pathways independent of the HPA axis (22) . Studies using antisense directed against CRF in rats have produced evidence of anxiolytic activity (23) . Finally, overexpression of CRF-BP is anxiolytic, whereas binding protein knockout mice (in which free CRF levels are elevated) display an anxiogenic phenotype in the elevated plus

maze (24). These data generally support the link between CRF and anxiety.

More recently, several studies have highlighted the importance of the CRF-1 receptor subtype in anxiety. CRF-1 knockout mice demonstrated a diminished anxiogenic response on the elevated plus maze and decreased ACTH and corticosterone responses to restraint stress (25). Similar findings were reported by Timpl et al. (26), using the black-white box anxiety paradigm. Furthermore, inactivation of the CRF-1 receptor with an antisense oligonucleotide was shown to reduce the anxiogenic effect of intraventricularly administered CRF (23). Liebsch et al. (27) provided evidence of anatomic localization by showing anxiolytic activity from CRF-1 antisense that was chronically infused into the central nucleus of the amygdala, an area of the limbic system shown by Michael Davis, Joe LeDoux, and others to be important in mediating fear and anxiety processes. Finally, CRF-2 knockout mice show anxiety-like behavior and are hypersensitive to stress (28), indicating that the CRF-2 receptor has an opposite functional role to that of the CRF-1 receptor. Thus, it could be argued that CRF-2 agonists, rather than antagonists, might be potentially useful as anxiolytic agents.

Another potential use for CRF antagonists is in the treatment of drug abuse. Several lines of evidence suggest that during the period of withdrawal from drugs of abuse such as ethanol, morphine, and cocaine, there is an activation of central CRF pathways. Anxiety is among the many physical symptoms of drug withdrawal, and given the link that has been made between CRF and anxiety, it is not surprising that CRF-1 receptor knockout mice demonstrated decreased anxiety responses during withdrawal from alcohol (26).

Current Drugs in Development

A number of nonpeptidic, small-molecule compounds that show high selectivity for the CRF-1 receptor have been proposed for the treatment of depression, anxiety, and stress disorders (29 ,30 and 31). These include CP-154,526 (Pfizer), and a methylated analogue, antalarmin (Pfizer); SC 241 (Dupont); NBI 30775 (aka R-121919; Janssen-Neurocrine); and CRA 1000 and CRA 1001 (Taisho Pharmaceuticals). An extensive preclinical literature has investigated potential anxiolytic effects of these compounds. Studies using CP-154,526 have demonstrated anxiolytic-like effects in some (32 ,33 and 34) but not all (33 ,34) preclinical anxiolytic paradigms evaluated. Griebel et al. (33) proposed that high-stress conditions may be required to demonstrate efficacy of CRF-1 receptor antagonists. CP-154,526 produces an anxiolytic effect in the separation-induced vocalization assay (35), an animal model of anxiety in which a preweaning rat pup separated from its litter emits a series of ultrasonic vocalizations that can be dose-dependently suppressed by either benzodiazepine or nonbenzodiazepine anxiolytics (36).

Of the compounds listed above, the compound discovered by Neurocrine Biosciences and licensed by Janssen Pharmaceuticals, R-121919, has proceeded the furthest in clinical evaluation. A recent report (37) describes results from a phase II, open-label, dose-escalating trial in which 20 severely depressed (Hamilton Depression Score >25) patients were administered R-121919 in one of two dose ranges: 5 to 40 mg, or 40 to 80 mg. In the low-dose group, 50% of the patients responded positively to treatment as indicated by a reduction in the Hamilton Depression Score of at least 50%, and 20% were remitters (score ≤8). In the middle-dose group, 80% responders and 60% remitters were reported. In addition, no significant untoward side effects were reported, and basal or stress-induced levels of ACTH or cortisol where unaffected, suggesting that chronic blockade of the HPA axis might not necessarily produce untoward side effects. Although these preliminary data are promising, it is important to bear in mind that they were gathered using an open label design without placebo control. Firm conclusions regarding the efficacy and safety of CRF-1 antagonists in depression and anxiety will require more rigorous double-blind, placebo-controlled trials. R-121919 is no longer under development because of reported elevations in liver enzymes; however, Neurocrine Biosciences has announced that further candidates are being pursued for clinical evaluation.

As mentioned previously, selective CRF-2 antagonists have not been described, though recently, nonpeptide dual antagonists of CRF-1 and CRF-2 receptors have been described (19). As there is contradictory evidence regarding the role of CRF-2 receptors in mediating anxiety, careful preclinical and clinical evaluation of these compounds will be needed to validate the contribution of CRF-2 receptors.

Future Drugs and Directions

As indicated above, a number of drug companies have dedicated significant efforts to identifying potent and selective CRF-1 receptor antagonists suitable for clinical development. To date, no compounds have completed phase II evaluation. Clearly, a challenge for the future will be to achieve this milestone and, in the process, validate with carefully executed clinical trials the concept that CRF-1 receptor antagonists are novel anxiolytics and/or antidepressants. Also, given increasing evidence of the importance of CRF-2 receptors in the human brain, significant efforts should be dedicated to evaluating this target for anxiety. It will be of interest to determine if agonists, rather than antagonists, of the CRF-2 receptor have anxiolytic profiles. Finally, the prospect of identifying additional CRF receptor subtypes, as well as other receptors for peptides, such as VIP, that are involved in the regulation of stress, provides fertile ground for future investigations.

GABA$_A$ RECEPTOR MODULATORS (BENZODIAZEPINES AND RELATED DRUGS)

A majority of the synapses in the mammalian CNS use the amino acids l-glutamic acid, glycine, or γ-aminobutyric acid (GABA) for signaling. GABA is formed by the decarboxylation of l-glutamate, stored in neurons, and released, and its action is terminated by reuptake; GABA's action mimics the naturally occurring inhibitory transmission in the mammalian nervous system. Because of these findings, it has been accepted for over 20 years that GABA fulfills the characteristics of a neurotransmitter (38). Along with l-glutamate, acetylcholine, and serotonin, GABA possesses two different types of receptor conserved across different species and phyla that control both excitation and inhibition. Molecular biological studies of the receptors causing these effects have indicated that GABA's effects on ionic transmission (ionotropic) and metabolism (metabotropic) are mediated by proteins in two different superfamilies. The first superfamily ($GABA_A$ receptors) is a set of ligand-gated ion channels (ligand-gated superfamily) that convey GABA's effects on fast synaptic transmission (39). When a $GABA_A$ receptor is activated, an ion channel is opened (gated) and this allows chloride to enter the cell; the usual result of chloride entry is a slowing of neuronal activity through hyperpolarization of the cell membrane potential. The second superfamily ($GABA_B$) is slower, mediating GABA's action on intracellular effectors through a seven transmembrane spanning receptor (serpentine superfamily) that modulates the action of certain guanine nucleotide binding proteins (G proteins) (40). Through their activity on other effector systems, G proteins can change second messenger levels, altering signal transduction and gene expression, or open ion channels that are dependent on the G-protein subunit activities (41). Both excitatory and inhibitory activities are possible on a time scale that is longer than $GABA_A$ receptor mediated events. There is extensive heterogeneity in the structure of the $GABA_A$ receptor members of the ligand-gated superfamily. These receptors are the targets of a number of widely used and prescribed drugs for sleep, anxiety, seizure disorders, and cognitive enhancement; they may also contribute to mediating the effects of ethanol on the body.

Structure and Molecular Pharmacology of $GABA_A$ Receptors

It is well established that the $GABA_A$ receptors possess binding sites for the neurotransmitter GABA, as well as allosteric modulatory sites for benzodiazepines, barbiturates, neurosteroids, anesthetics, and convulsants (42 ,43 and 44). The initial cloning of complementary DNAs (cDNAs) coding for the subunits of $GABA_A$ receptors indicated that the chloride channel gated by GABA is intrinsic to the structure of the receptor and that each of the binding sites also possesses specific requirements for subunit composition (45 ,46). At present, almost 20 different cDNAs have been identified and classified into six classes based upon sequence homology. Cloned from vertebrates, there are six α, four β, four γ, one δ, one ε, and two ρ subunits and some splice variants; the subunits share a basic motif where the amino acids span the membrane four times. This four transmembrane spanning motif is shared with subunits that form other receptor members of the ligand-gated superfamily (39).

Extensive mutagenesis and structural examination has been carried out with the GABA and acetylcholine family of receptors (47 ,48). Acetylcholine receptors have been shown to possess a pentameric subunit structure with a heterogeneous subunit composition; evidence for this conclusion has been obtained through the use of monoclonal antibodies and through direct electron microscopic visualization of the densely packed receptor in the Torpedo eel. Similar electron microscopic analysis of $GABA_A$ receptors has been carried out (49). It is thought that the native $GABA_A$ receptors also possess such a pentameric structure with general composition of 2 α, 2 β, and one γ subunit forming the majority of the $GABA_A$ receptors in vertebrates. Evidence of this has been more circumstantial, generated by molecular biological and pharmacologic inferences, described below, and by the behavior of solubilized recombinant complexes on sucrose gradient centrifugation in the presence and absence of different subunit specific antibodies (50). The natural receptor in endogenous tissue appears to be also pentameric (51).

Evidence from studies of acetylcholine receptors has also indicated that the second transmembrane spanning sequence forms the actual ion channel of acetylcholine receptors, and that mutations of amino acids at the inner (cellular) side of the membrane are responsible for the ability of specific cation ions to pass through the channel pore (48). The ionic selectivity can be changed by altering the charge of some of these specific amino acids, and the acetylcholine receptor can be forced to gate chloride, rather than sodium, by such changes. Thus, a relatively firm case for the involvement of this spanning region in the formation of the ion channel can be made. Because the core ion pore is highly conserved among the large number of $GABA_A$ receptor subtypes, a number of drugs that interact nonspecifically with all the members of the $GABA_A$ receptor family were identified in the past. These include anesthetic barbiturates, picrotoxin, neurosteroids, and some organic insecticides (42 ,43 and 44). More recently, because the ion channel shows little variation between $GABA_A$ receptor subtypes, it has not been as active a target for pharmaceutical discovery as the convenient allosteric modulatory site for benzodiazepines, drugs discovered by chance almost 40 years ago.

Originally, two subunits of the $GABA_A$ receptor family were cloned and, when expressed in oocytes, were capable of forming a receptor that would gate chloride in response to GABA (45). At that time, some responses were seen to barbiturates, toxins, and benzodiazepines. It is now known

that a full response to the benzodiazepines requires the incorporation of a third subunit, the γ subunit (52). One of the major forms of native $GABA_A$ receptor in vertebrates probably has the structure $\alpha_1\beta_2\gamma_2$, most likely in a 2:2:1 stoichiometry. The α_1 subunit contains the major site that is photoaffinity labeled with the benzodiazepine ^{3}H-flunitrazepam at His 101 (53). For the functional modulation of GABAergic activity by benzodiazepines, in addition to the α subunit, a γ subunit must be incorporated into the complex. Thus, there is reasonable evidence that benzodiazepines and related drugs stimulate GABA activity without opening channels directly; depending on their ability to potentiate GABA activity, they are called full or partial agonists. The binding site for these drugs incorporates a binding site composed of components of both α and γ. Other drugs (called inverse agonists) may occupy the same site to negatively modulate the action of GABA, such as β-carboline derivatives. Yet a third class of compounds exist, drugs such as flumazenil, may occupy the site as antagonists of both agonist and inverse agonists. By themselves, these antagonists have no affect on GABAergic activity and are behaviorally silent. The important allosteric modulatory effects of drugs at the benzodiazepine site were recognized early and the distribution of activities at different receptor subtypes has been an area of intense pharmacologic discovery for many years (39 ,42 ,43 and 44). The details of some of these findings are described later.

Because the expression of an α or β subunit by itself does not form a functional receptor (54), but expression of both together constitute a functional $GABA_A$ receptor, the binding sites for GABA could be associated with the combination of the two subunits. Systematic mutagenesis of α and β subunits have identified a number of amino acids on both subunits that appear to contribute to the ability of GABA to bind to the $GABA_A$ receptor and modulate chloride conductance (53). Thus, the GABA binding sites are also complex, composed of a binding pocket that is made up of amino acids from both subunits; there are two GABA binding sites per receptor [located at the two identical (or homologous in the case of heterogeneous mixtures of subunits) αβ interfaces], and electrophysiologic studies indicate that both must be occupied for the chloride channel to open.

Perhaps the most interesting aspect of these mutagenesis studies is the observation that the GABA binding site (two per αβγ complex) and the benzodiazepine binding site (one site per αβγ complex) are structurally related to one another; homologous amino acids that contribute to binding in each case are found at similar positions in each subunit. They represent in the case of the GABA binding site αβ interfaces and in the case of the benzodiazepine binding site the αγ interface (Fig. 68.2). By binding to the benzodiazepine binding site, the drugs clearly give a positive allosteric signal to the receptor and/or to the GABA binding sites; this is reflected in a change in affinity at a low-affinity GABA binding site (43). This signal is analogous to the signal that a single GABA molecule binding to one of the GABA binding sites gives to the receptor and second GABA binding site. The original finding relating GABA and benzodiazepines allosterically showed evidence of a similar signal being transmitted from the occupied GABA binding site to the benzodiazepine site, resulting in a higher affinity state at the benzodiazepine binding site in the presence of GABA at its binding site (55). This type of pseudosymmetric binding site is characteristic of allosteric protein binding interactions and is found also in acetylcholine and in glycine receptors (46 ,56 ,57). The theoretical details of such interactions allow one to rationalize how a set of drugs like benzodiazepines could stabilize a channel open state but not be able to open the channel directly. Variations in these binding sites that are dependent on different α, β, and γ subunit amino acid sequences, particularly α and γ as the components of the benzodiazepine binding site, underlie the heterogeneity of $GABA_A$ receptors and point to the possibilities for $GABA_A$ receptor subtype–specific drugs.

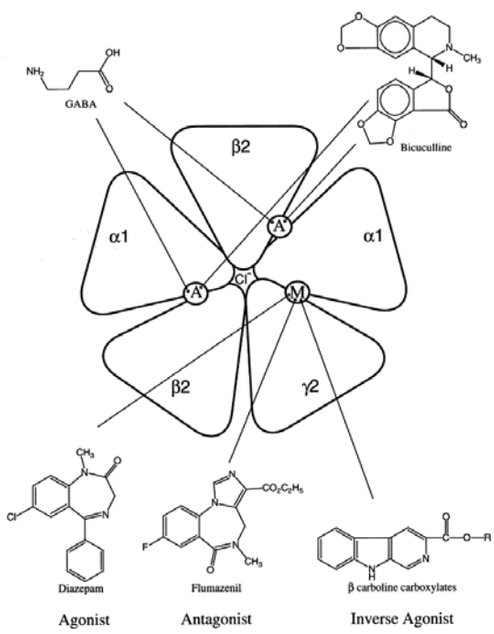


FIGURE 68.2. γ-Aminobutyric acid (GABA) binding site αβ interfaces.

Genomic Organization of $GABA_A$ Receptor Subunits

The natural association of αβγ compositions of $GABA_A$ receptors is also underscored by the genomic organization of the subunits. There is a cluster of the genes coding for α_1, β_2, and γ_2 subunits on chromosome 5 and similar clusters of other subunits on other chromosomes such as 4 and 15 (58 ,59). They point to the phylogenetic age of $GABA_A$ receptors, and similarities in organization point to the possibility that ancestral $GABA_A$ receptor gene cluster duplications spawned some of the different clusters coding for unique $GABA_A$ receptor isoforms, whereas mutations in individual subunits may have created additional diversity. Such an organization also points to probable coordinated control of subunit production and many interesting aspects of cellular and brain regional regulation of expression yet to be examined. Regional diversity also can mean functional diversity and drugs that affect certain aspects of behavior, but not others.

Functional Activities of Therapeutics at Individual Recombinant Constructs

One reason for the diversity of subtypes of $GABA_A$ receptors is that this is how neurons integrate information and change the behavioral status of the animal. Whereas $GABA_A$ receptors are found on many neurons, particularly local interneurons, some $GABA_A$ receptor subtypes are selectively localized to specific brain regions, specific cell layers within those regions, and even specific parts of cells (60 ,61 and 62). Each subtype has a unique set of electrophysiologic and pharmacologic properties. A number of factors can determine the response properties of $GABA_A$ receptor subtypes. The most

important determinant is subunit composition. Secondarily, these receptor subtypes can have unique profiles of modulation driven by membrane potential, ion gradients, biochemical conditions, and, last but not least, drugs.

Attempts have been made to categorize $GABA_A$ receptors pharmacologically in terms of responses to specific drugs that interact with the benzodiazepine binding site. Thus, in the original nomenclature, two subtypes of $GABA_A$ receptor were described (63). They were called type 1 and type II receptors and were defined in terms of differential affinity for CL 218,872 and a number of other compounds. We now know that type I $GABA_A$ receptors contain the composition $\alpha_1\beta_x\gamma_2$ and are responsive to CL 218,872, a related compound zaleplon, and zolpidem (64). Both zaleplon and zolpidem are marketed hypnotics. Type II receptors are a mixed heterogeneous class containing $\alpha_2\beta_x\gamma_2$, $\alpha_3\beta_x\gamma_2$, and $\alpha_5\beta_x\gamma_2$. α_4- and α_6-containing $\beta_x\gamma_2$ subunit combinations are generally agreed to constitute a separate category of $GABA_A$ receptors due to their lowered affinity for classic benzodiazepine site ligands (65). For example, flumazenil has a very low affinity for these α_4- and α_6-containing $GABA_A$ receptor recombinant constructs compared to α_1-, α_2-, α_3-, or α_5-containing $GABA_A$ receptors (100 nM versus 0.5 nM). The differences in amino acids between the α subunits pointed to the mutagenesis studies described above to delineate GABA and benzodiazepine binding sites.

Thus, affinity differences in the benzodiazepine binding site have been a primary method for differentiating $GABA_A$ receptor subtypes. In contrast to the affinity models of benzodiazepine binding sites, it is not clear that a simple molecular biologically based efficacy model will emerge to simplify our understanding of α_1-, α_2-, α_3-, or α_5-containing $GABA_A$ receptors.

The $GABA_A$ receptor subtypes with their benzodiazepine receptor sites are an example of a unique situation in biology. Compounds have been synthesized that can allosterically modulate GABA response over a wide range through this site. Modulation of GABA responses has spanned the range of >700% increase in response amplitude to inhibition as great as 60%. Control of efficacy by drugs with subunit specificity can be achieved. One approach of drug companies has been to develop drugs that increase GABA responses less than 100% and develop selectivity for some subunits without increasing responses at other subunit combinations. These have been called subtype-selective partial agonists and will probably represent the next generation of GABAergic modulators to enter the clinic and become drugs.

A number of studies suggest, by circumstantial evidence such as message distribution, genomic localization, and biochemical study, that the major subtype combinations in brain are $\alpha_1\beta_2\gamma_{21}$, $\alpha_2\beta_3\gamma_{21}$, $\alpha_3\beta_3\gamma_{21}$, and $\alpha_5\beta_3\gamma_{21}$. α_1 has been implicated in sedation by virtue of the fact that zolpidem, a marketed hypnotic under the trade name Ambien, is a type I (α_1) selective compound and can cause cognitive deficits (see next subsection). An ideal anxiolytic drug might have limited effects on this subtype while increasing responses at α_2- and α_3-containing subtypes, as they are located in the limbic parts of the brain directly implicated in generation and reduction of anxiety. The full examination of subtype selective drugs in humans is in the near future, and it will be interesting to see how these hypotheses fare in clinical trials.

Involvement of $GABA_A$ Receptors in Human Disease and Transgenics

The extensive investigation of the amino acids involved in the binding of GABA and benzodiazepines allows a specific and elegant approach to be made to the *in vivo* investigation of the involvement of $GABA_A$ receptor subtypes with specific neural pathways and specific behavioral activities. This approach can be made either through the examination of chromosomal deletions, the use of specific knockouts of subunit genes, or knock-ins of particular point mutations. These techniques naturally complement the development of drugs with specific activities at $GABA_A$ receptor subtypes.

Some naturally occurring chromosomal deletions of particular $GABA_A$ receptor subunits showed phenotypes of craniofacial deficits, mental retardation, and epilepsy. Deletion of large areas of human chromosome 15, containing α_5, γ_3, and β_3 subunit genes, results in this phenotype, which causes a human genetic disorder called Prater Willi/Angelman syndrome (66). In a targeted study in mice, animals with the specific deletion of the β_3 subunit shows a similar syndrome with cleft palate and neurologic abnormalities. These studies point to the importance of certain $GABA_A$ receptors in neuronal development (67 ,68). It is also clear that the deletion of an entire subunit does not result in the rescue of function by substitution of other subunits (69). Other rare genetic disorders of $GABA_A$ receptor function are likely to emerge as our knowledge of the genomic basis of neurologic disorders evolve. A very elegant approach, based on the molecular biological studies described above, has been taken to examine the significance of $GABA_A$ receptor subtypes by replacing an important amino acid for benzodiazepine binding (histidine) found in α_1, α_2, α_3, and α_5 with an arginine characteristic of α_4 and α_6. Through this conservative mutation, GABA sensitivity is retained, so the receptors function normally, but drug sensitivity is lost (70 ,71). From these studies, it appears that the α_1-containing subtypes are important in mediating the anticonvulsant, sedative, and amnestic effects of benzodiazepines, but to a smaller degree the muscle relaxant and anxiolytic effects. These animals are still susceptible to the development of tolerance to the sedative effects. In the near future, we may learn the consequence of deletion of the specific benzodiazepine modulatory sites from the other α subtypes. Thus, from a mechanistic point of view this class of drugs has a well-defined mode of action.

SEROTONIN RECEPTOR MODULATORS AND REUPTAKE INHIBITORS

Part of "68 - Mechanism of Action of Anxiolytics "

Preclinical Studies

Serotonin has long been viewed as a neurotransmitter involved in regulating emotional states. Of the 14 or so mammalian serotonin receptor subtypes that have been described in the literature, at least four have been implicated in anxiety in various animal models (72). As reported by Lucki (72) the original hypothesis implicating serotonin in anxiety surfaced from observations that reduced levels of serotonin can produce anxiolytic effects. One of the receptor subtypes implicated in anxiety is the serotonin 1A receptor subtype ($5\text{-}HT_{1A}$), which is an autoreceptor located presynaptically on serotonin neurons. When stimulated, this receptor inhibits the synthesis and secretion of serotonin. The $5\text{-}HT_{1A}$receptor agonist buspirone exhibits anxiolytic effects in

animals and was approved by the Food and Drug Administration (FDA) in 1986 for human generalized anxiety disorder. Other serotonin receptors potentially involved in anxiety include the 5-HT_{2A}, 5-HT_{2C}, and 5-HT_3 receptors. Antagonists for the 5-HT_{2A}receptor, like ritanserin, exhibit anxiolytic effects in some animal models (73 ,74). Likewise, blockade of the 5-HT_{2C} receptor produces anxiolytic effects in animals (75) and prevents the anxiogenic effects of m-CPP (76). Finally, the 5-HT_3receptor antagonist ondansetron was reported to be anxiolytic in some animal models (77).

Recent advances in molecular biology has led to the development of serotonin receptor gene knockout methodology, which generates mice lacking the 5-HT_{1A} receptor, allowing for the evaluation of this receptor subtype in a variety of measurable behaviors. Ramboz et al. (78) reported results consistent with the 5-HT_{1A} agonist's data cited above. Mice lacking this receptor displayed less exploratory activity in an open field and more anxious behavior than the wild types in the elevated plus maze. According to the serotonin hypothesis of anxiety (79), removing the negative feedback control of 5-HT with the 5-HT_{1A} receptor knockout animals should result in increased levels of 5-HT in the synaptic cleft, which would be expected to lead to the anxiogenic behavior. However, Ramboz et al. (78) reported normal levels of 5-HT, which confuses the issues related to anxiety modulation and serotonin levels. As David Julius (80) points out, the interpretation of standard gene knockout experimentation is complicated by the possibility of long-term development changes and this is true with the 5-HT_{1A}knockout animal. So despite the apparent consistency between the 5-HT_{1A} knockout animal and 5-HT_{1A} agonist studies in terms of the behavioral outcomes of each manipulation, the exact role of the 5-HT_{1A}receptor in anxiety is not absolutely clear at this time.

Clinical Studies

In 1986, the FDA approved the 5-HT_{1A} partial agonist for generalized anxiety disorder. This drug was the first to challenge the benzodiazepines for this patient group and was generally perceived as an improvement because of the lack of benzodiazepine side effects. The efficacy of buspirone, however, was not the same as that of the benzodiazepines in terms of its delayed onset of action, and it is generally accepted that when buspirone offers clinical benefit to generalized anxiety disorder (GAD) patients, it takes 3 to 4 weeks to match the efficacy of benzodiazepines such as diazepam and alprazolam (81). The 5-HT_{1A} partial agonist properties of buspirone are believed to account for its clinical effects, but it should be noted that the drug is also a D2 antagonist and is extensively metabolized. One of the major metabolites, 1-pyrimidinylpiperazine (1-PP), may contribute to the pharmacologic activity of buspirone (82). In a double-blind, placebo-controlled study of buspirone in GAD patients (83), the drug was reported to be as efficacious as lorazepam at the end of a 4-week treatment period. After the drugs were discontinued, however, the lorazepam-treated patients worsened whereas the buspirone-treated subjects maintained clinical improvement. Thus, there continues to be evidence that buspirone is effective in GAD.

The development of selective serotonin reuptake inhibitors (SSRIs) in the 1980s and 1990s widely expanded the treatment for depressive disorders, and these drugs (fluoxetine, sertraline, venlafaxin, paroxetine) have recently made inroads in treating anxiety disorders such as panic, obsessive-compulsive disorder, social phobia, and GAD. Successful treatment of GAD with a class of drugs working through the serotoninergic system will come from the SSRIs (84).

Obsessive-compulsive disorder (OCD) is a chronic, disabling anxiety disorder. In a review of the diagnosis and treatment of OCD, Goodman (85) states that the backbone of pharmacologic treatment for OCD is a 10- to 12-week trial with an SSRI in adequate doses. It is clear from a review of the role of the 5-HT_{1A} receptor (86) in OCD that partial agonists such as buspirone are generally ineffective in treating OCD. The authors also note that in studying the potential to augment efficacy of the standard OCD medication, buspirone was not different from placebo as an augmenting agent. Drugs that work through other serotonin receptor subtypes also appear to be ineffective in treating OCD. Thus, drugs modifying the 5-HT_{1A}, 5-HT_{1D}, and 5-HT_3 receptors appear ineffective in treating OCD symptoms and rule out a critical involvement of these receptor subtypes in OCD (87 ,88).

In the past, tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors, as well as high potency benzodiazepines, have been used to treat patients with panic disorder. The SSRIs have also been added to the list of effective agents for the disorder. In reviewing the pharmacotherapy of panic disorder, den Boer (89) notes that antidepressants are more effective than benzodiazepines in reducing associated depressive symptomatology and are at least as effective for improving anxiety, agoraphobia, and overall impairment. Bell and Nutt (90) remark that SSRIs improve 60% to 70% of panic patients, a similar percentage to those seen with the TCAs.

Like OCD, panic disorder is well treated by SSRIs but does not appear to be effectively treated by receptor specific compounds. Copland et al. (86) reviewed the role of 5-HT_{1A}drugs such as buspirone in panic disorder and reported that buspirone does not significantly treat panic in several well-controlled studies. Using the 5-HT_{1A} receptor agonist flesinoxan, van Vliet et al. (91) reported a worsening of symptoms in panic patients treated with high doses of the drug. It has also been reported that the 5-$HT_{2A/2C}$ antagonist ritanserin had no effects on panic attacks or phobic avoidance, and a similar negative finding has been reported with the 5-HT_3 antagonist ondansetron.

NEUROKININ RECEPTOR ANTAGONISTS

Part of "68 - Mechanism of Action of Anxiolytics "

Rationale

There is an extensive literature demonstrating that the peptide tachykinins such as substance P and their associated receptors have a widespread distribution in the brain, spinal cord, and periphery, and may play important roles in

chronic pain and inflammation processes (92 ,93 ,94 ,95 and 96). In addition, anatomic and physiologic evidence has also indicated that these peptides may affect limbic structures that are involved in the regulation of mood and affect, such as the amygdala, hypothalamus, and periaqueductal gray (97). This notion is supported by early positive clinical findings using a selective neurokinin-1 (NK-1) antagonist for the treatment of depression and anxiety (98).

Molecular Mechanism of Action

Tachykinins collectively refer to small peptides that include substance P (SP), neurokinin A (NK-A), and neurokinin B (NK-B). These peptides show preferential affinity for three receptors, designated NK-1, NK-2, and NK-3, respectively, which are members of the seven-transmembrane, G-protein–coupled family. Of these three receptors, NK-1 and NK-3 are found in the brain, whereas NK-2 is primarily localized peripherally in smooth muscle of the respiratory, urinary, and gastrointestinal tracts. Neurokinin receptors are localized in a number of different brain areas that are implicated in anxiety, including the amygdala, hypothalamus, and locus coeruleus.

Studies assessing the effects of direct administration of neurokinin agonists such as substance P into the nervous system are complicated by the findings that, depending on factors such as the site and dose, opposite effects on behavior may be achieved.

Current Drugs in Development

Numerous NK-1 antagonists have been described in the literature, including MK-869 (Merck) and an analogue, L-760,735 (Merck), SR140333 (Sanofi), CP-122,721 (Pfizer), RP67580 (Rhone-Poulenc), FK-888 (Fujisawa), SDZ NKT 343 (Novartis), and PD 154075 (Parke-Davis). NK-1 antagonists have been reported to demonstrate anxiolytic effects in animal models such as social interaction (99), though these effects are not consistently seen across all compounds (34). Researchers from Merck have reported that vocalizations elicited by maternal separation in guinea pigs are robustly blocked by NK-1 antagonists such as MK-869, an effect that is shared by a range of antidepressant and anxiolytic agents (98).

Of the compounds listed above, the primary indication has been for the treatment of conditions such as pain, chemotherapy-induced emesis, and migraine (94). MK-869, which progressed to phase III trials for emesis, has also been evaluated in a phase II depression trial in which it was reported that, in addition to showing a significant antidepressant effect, MK-869 also showed significant anxiolytic activity that emerged over the course of the 6-week study (98). The data supported the conclusion that NK-1 antagonists might be useful for the treatment of depression and anxiety. Further development of MK-869 for depression was discontinued, but these early clinical data will undoubtedly lead to further clinical evaluation of NK-1 antagonists.

NK-2 antagonists include SR48968 (Sanofi). Preclinical studies have shown that NK-2 antagonists such as GR159897 and SR48968 have also demonstrated activity in social interaction and exploration anxiolytic models, and activity has been reported in the marmoset monkey using the human “threat” test. Good therapeutic ratios were described for these agents.

NK-3 antagonists described in the literature include osnetant (Sanofi-Synthelabo), talnetant, PD-161182, and PD-157672 (Parke-Davis). The latter two have been designated for the treatment of anxiety disorders, though there have been no reports of clinical trials with any NK-3 antagonist for this indication. It should be noted that preclinical data described to date are sparse, and there is some suggestion that NK-3 agonism may produce an anxiolytic profile. Thus, intraventricular administration of the NK-3 agonist senktide produced anxiolytic effects in mice that could be blocked by administration of the NK-3 antagonist SR 142801, and SR 142801 was found to have some anxiogenic activity (100).

Future Drugs and Directions

Further depression and anxiety clinical trials with centrally active NK-1 antagonists are needed to provide further validation of the role of NK-1 receptors in treating depression and anxiety disorders. In addition, further assessment of the role of NK-2 and NK-3 subtypes is needed to determine the possible relevance, if any, of these receptor subtypes.

GLUTAMATE RECEPTOR AGONISTS AND MODULATORS

Part of "68 - Mechanism of Action of Anxiolytics "

Rationale

Glutamate is the major mediator of excitatory neurotransmission in the CNS. Despite this ubiquity, the elucidation of numerous glutamate receptor subtypes with differential localizations in the brain, and the development of selective pharmacologic agents, has led to the realization that glutamate receptors might be viable targets for a number of different neurologic and psychiatric disorders, including anxiety and depression (101 ,102 and 103).

Molecular Mechanism of Action

The molecular biology of glutamate receptors has been the subject of numerous reviews (101 ,102). Glutamate receptors are classified as either ionotropic or metabotropic. Ionotropic receptors, which mediate fast synaptic transmission, are coupled to cation-specific ion channels and bind the agonists *N*-methyl-D-aspartate (NMDA), α-amino-3-hydroxy-5-methyl-4-isoxazole

propionic acid (AMPA), and kainic acid (KA). These receptors gate both voltage-dependent and voltage-independent currents carried by Na^+, K^+, and Ca^+.

NMDA receptors, which are selectively activated by NMDA, form a receptor/channel complex that is allosterically regulated by several sites. Receptor activation results in depolarization and Ca^+ influx. Allosteric binding sites include a strychnine-insensitive glycine site, a polyamine site, a zinc site, and a channel site that binds agents such as MK-801 or phencyclidine to block channel opening. The NMDA receptor has been cloned and has two families of subunits, the NR1, with seven splice variants (1A–1G), and NR2, with four splice variants (2A–2D). NR1 receptors possess the receptor/ion channel complex, whereas the NR2 receptors lack the ion channel and appear to be modulatory. NR2 receptors, however, can form functional heteromeric channels by combining with NR1 subtypes. NMDA antagonists include competitive antagonists at the NMDA receptor such as AP5, CPPene, and CGS 19755; noncompetitive antagonists at the ion channel site such as MK-801 and phencyclidine; and noncompetitive antagonists that bind to the glycine site such as 5,7-dichlorokynurenic acid, L689,560, ACEA 1021, and MDL 105,519 (104).

“Non-NMDA” receptors refer to the class of ionotropic glutamate receptors activated by kainic acid or AMPA, agonists that activate voltage-independent channels that gate a depolarizing current mediated primarily by Na^+ ions. There are four AMPA subunits (GluR1–GluR4) and five KA subunits (GluR5–GluR7 and KA-1–KA-2). Functional channels can be produced by homomeric expression of GluR5 or GluR6 subunits, or by heteromeric expression of KA-1 or KA-2 with GluR5 or GluR6. KA sites are found in the hippocampus, cortex, and thalamus, whereas AMPA receptors are additionally localized in septal and cerebellar sites. Pharmacologic agents for blocking non-NMDA receptors include the competitive antagonists CNQX, DNQX, and NBQX, and the noncompetitive antagonist GYKI 52466.

Metabotropic glutamate receptors (mGluRs) mediate slower synaptic transmission and utilize phosphatidyl inositol (PI) and cAMP as second messengers. Opposite effects on cAMP may be mediated by either G_i or G_s stimulation. Agonists include a conformationally restricted glutamate analogue, 1-aminocyclopentane-*trans*-1,3-dicarboxylic acid (*trans*-ACPD). Metabotropic receptors have been classified into three subgroups: group I (mGluR1, mGluR5), which stimulate PI hydrolysis; and two groups that inhibit adenylate cyclase, group II (mGluR2, mGluR3) and group III (mGluR4, mGluR6). Metabotropic receptors are widely distributed throughout the brain, in areas such as the hippocampus, cerebellum, thalamus, olfactory bulb, and striatum, though the precise distribution varies considerably between groups. In addition to the agonist trans-ACPD, other agonists that are more specific for receptor subtypes include LY354740 (group II), L-AP4 (group III), and L-CCG-I (mGluR2). Pharmacologic agents for antagonizing metabotropic glutamate receptors are currently limited.

Genetically Altered Mouse Models

Site-directed mutagenesis studies have indicated that point mutations in the glycine binding site of the NR1 subunit result in mice that have reduced glycine affinity and have an anxiolytic profile as seen by decreased natural aversion to an exposed environment (105). These data supported other pharmacologic lines of evidence (see below), indicating that blockade of the glycine site can have anxiolytic actions.

Current Drugs in Development

LY354740 is an orally active group II metabotropic receptor agonist (106) currently in clinical development. Preclinically, the compound has anxiolytic activity in fear-potentiated startle (107), the elevated plus maze (107), conflict testing (108), the four-plate test (108), and in a lactate-induced panic attack model (109). LY354740 has also been shown to decrease withdrawal signs seen during naloxone-precipitated morphine withdrawal (110).

Future Drugs and Directions

Competitive and noncompetitive NMDA antagonists have primarily undergone clinical evaluation for the treatment of stroke and trauma (103). Unfortunately, clinical development for many of these compounds was halted because of severe side effects, including psychotic-like symptoms. The potential for these side effects has been a major deterrent for using NMDA antagonists for the treatment of psychiatric disorders. Although there are currently no drugs reported to be in development, preclinical studies have suggested that selective antagonists of the strychnine-sensitive glycine site can have anxiolytic properties with reduced side-effect potential relative to competitive and noncompetitive NMDA antagonists such as AP5 and MK-801 (36 ,111). Recent work has supported such a profile (112 ,113).

Antagonists of AMPA receptors have also been proposed to have anxiolytic actions in preclinical models. Thus, LY326325 was shown to have anxiolytic activity in the elevated plus maze and in a conflict test (punished drinking) in rats (114), and CNQX injected directly into the periaqueductal gray produced anxiolytic effects on the elevated plus maze (115). The antagonists CNQX and GYKI 52466 were also able to block the anxiogenic responses produced by bicuculline injected into the basolateral amygdala (116).

CCKB ANTAGONISTS

Part of "68 - Mechanism of Action of Anxiolytics "

Cholecystokinin (CCK) is a peptide found extensively both in the gut (where it was originally identified) and in the

brain (117). CCK exists in multiple forms, the most predominant of which is CCK octapeptide (CCK8) and, to a lesser extent, CCK tetrapeptide (CCK4) (118). CCK is co-localized with a number of different neurotransmitters, including serotonin, dopamine, GABA, substance P, neuropeptide Y, and VIP. CCK-like immunoreactivity has been demonstrated in anatomic regions that include the amygdala, cerebral cortex, hippocampus, striatum, hypothalamus, and spinal cord (119). There are two subtypes of CCK receptor, CCKA (sulfated CCK) and CCKB (unsulfated CCK) (120). CCKA receptors are localized in the nucleus accumbens, posterior hypothalamus, and area postrema. CCKB receptors are localized in cortex, olfactory bulb, nucleus accumbens, and other brain areas (121).

A number of selective antagonists for the CCKB receptor have been synthesized, including LY288513 (Lilly), PD 135158 (Parke-Davis), L-365,260 (Merck), and CI-988. Griebel (34) reviewed the extensive literature on the effects of CCK antagonists in models of anxiety and concluded that the results were contradictory and did not lead to a clear conclusion. L-365,260 and CI-988 have been evaluated clinically in panic disorder, but neither of these agents had significant anxiolytic effects (122 ,123 and 124).

CONCLUSION

Part of "68 - Mechanism of Action of Anxiolytics "

Many different neurotransmitter receptor systems have been shown to modulate anxiety and possess anxiolytic effects. This is not surprising because many of these transmitters control anatomic circuitry important in anxiety. The next generation of marketed anxiolytics will be determined more by their side-effect profile than by their anxiolytic activity. It will be interesting to see which of the mechanisms described in this chapter will provide the most useful anxiolytics in human populations.

ACKNOWLEDGMENT

Part of "68 - Mechanism of Action of Anxiolytics "

Drs. Tallman, Cassella, and Kehne are all full-time employees of Neurogen Corporation.

REFERENCES

1. Koob GF, Heinrichs SC, Pich EM, et al. The role of corticotropin-releasing factor in behavioral responses to stress. In: Chadwick DJ, Marsh J, Ackrill K, eds. Corticotropin-releasing factor. *Ciba Found Symp* 1993;172:277–289.

2. DeSouza EB, Grigoriadis DE. Corticotropin-releasing factor. Physiology, pharmacology, and role in central nervous system and immune disorders. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:505–517.

3. Gold PW, Licinio J, Wong ML, et al. Corticotropin releasing hormone in the pathophysiology of melancholic and atypical depression and in the mechanism of action of antidepressant drugs. *Ann NY Acad Sci* 1995;771:716–729.

4. Arborelius L, Owens MJ, Plotsky PM, et al. The role of corticotropin-releasing factor in depression and anxiety disorders. *J Endocrinol* 1999;160:1–12.

5. Berridge CW, Dunn AJ. CRF and restraint stress decrease exploratory behavior in hypophysectomized mice. *Pharmacol Biochem Behav* 1989;34:517–519.

6. Kalin NH, Helton SE, Davidson RJ. Cerebrospinal fluid corticotropin-releasing hormone levels are elevated in monkeys with patterns of brain activity associated with fearful temperament. *Biol Psychiatry* 2000;47:579–585.

7. Heim C, Owen MJ, Plotsky PM. The role of early adverse life events in the etiology of depression and posttraumatic stress disorder. Focus on corticotropin releasing factor. *Ann NY Acad Sci* 1997;821:194–207.

8. Plotsky PM, Meaney MJ. Early postnatal experience alters hypothalamic corticotropin-releasing factor (CRF) mRNA, median eminence CRF content and stress-induced release in adult rats. *Brain Res Mol Brain Res* 1993;18:195–200.

9. Perrin MH, Vale WW. Corticotropin releasing factor receptors and their ligand family. *Ann NY Acad Sci* 2000;885:312–328.

10. Dieterich KD, Lehnert H, De Souza, EB. Corticotropin-releasing factor receptors: an overview. *Exp Clin Endocrinol Diabetes* 1997;105:65–82.

11. Chalmers DT, Lovenberg TW, DeSouza EB. Localization of novel corticotropin-releasing factor receptor (CRF-2) mRNA expression to specific subcortical nuclei in rat brain: comparison with CRF-1 receptor mRNA expression. *J Neurosci* 1995;15:6340–6350.

12. Primus RJ, Yevich W, Baltazar C, et al. Autoradiographic localization of CRF-1 and CRF-2 binding sites in adult rat brain. *Neuropsychopharmacology* 1997;17:308–316.

13. Sanchez MM, Young, LJ, et al. Autoradiographic and in situ hybridization localization of corticotropin-releasing factor 1 and 2 receptors in nonhuman primate brain. *J Comp Neurol* 1999;408:365–377.

14. Dieterich, KD, Grigoriadis DE, De Souza EB. Homologous desensitization of human corticotropin-releasing factor 1 receptor in stable transfected mouse fibroblast cells. *Brain Res* 1966;710:287–292.

15. Hauger RL, Aguilera G. Regulation of pituitary corticotropin releasing hormone (CRH) receptors by CRH: interaction with vasopressin. *Endocrinology* 1993;133:1708–1714.

16. Aguilera G, Millan MA, Hauger RL, et al. Corticotropin-releasing factor receptors: distribution and regulation in brain, pituitary, and peripheral tissues. *Ann NY Acad Sci* 1987;512:48–66.

17. Anderson SM, Kant GJ, et al. Effects of chronic stress on anterior pituitary and brain corticotropin-releasing factor receptors. *Pharmacol Biochem Behav* 1993;44:755–761.

18. Schulz DW, Mansbach RS, Sprouse J, et al. CP-154,526: a potent and selective nonpeptide antagonist of corticotropin releasing factor receptors. *Proc Natl Acad Sci USA* 1996;93:10477–10482.

19. Luthin DR, Rabinovich AK, et al. Synthesis and biological activity of oxo-7H-benzo(e)perimidine-4-carboxylic acid derivatives as potent, nonpeptide corticotropin releasing factor (CRF) receptor antagonists. *Biorg Med Chem* 1999;Lett. 9:765–770.

20. Contarino A, Heinrichs SC, Gold LH. Understanding corticotropin releasing factor neurobiology: contributions from mutant mice. *Neuropeptides* 1999;33:1–12.

21. Heinrichs SC, Lapsansky J, Lovenberg TW, et al. Corticotropin releasing factor CRF-1, but not CRF-2, receptors mediate anxiogenic-like behavior. *Regul Pept* 1997;71:15–21.

22. Stenzel-Poore MP, Heinrichs SC, et al. Overproduction of corticotropin-releasing factor in transgenic mice: a genetic model of anxiogenic behavior. *J Neurosci* 1994;14:2579–2584.

23. Skutella T, Probst JC, Renner U, et al. Corticotropin-releasing hormone receptor (type I) antisense targeting reduces anxiety. *Neuroscience* 1998;85:795–805.

24. Ramesh T, Karolyi I, Nakajima M, et al. Increased anxiety in corticotropin releasing hormone-binding protein-deficient mice. *Brain Res* 1998;809:A23.

25. Smith G W, Aubry J M, Dellu F, et al. Corticotropin releasing factor receptor-1 deficient mice display decreased anxiety impaired stress response, and aberrant neuroendocrine development. *Neuron* 1999;20:1093–1102.

26. Timpl P, Spanagel R, Sillaber I, et al. Impaired stress response and reduced anxiety in mice lacking a functional corticotropin-releasing hormone receptor. *Nature Genet* 1998;19:162–166.

27. Liebsch G, Landgraf R, Gerstberger R, et al. Chronic infusion of a CRF-1 receptor antisense oligonucleotide into the central nucleus of the amygdala reduced anxiety-related behavior in socially defeated rats. *Regul Pept* 1995;59:220–39.

28. Bale, TL, Contarino A, Smith GW, et al. Mice deficient for corticotropin-releasing hormone receptor-2 display anxiety-like behavior and are hypersensitive to stress. *Nature Genet* 2000;24:410–14.

29. Gilligan PJ, Robertson DW, Zaczek R. Corticotropin releasing factor (CRF) receptor modulators: progress and opportunities. *J Med Chem* 2000;43:1–20.

30. McCarthy JR, Heinrichs SC, Grigoriadis DE. Recent advances with the CRF-1 receptor: design of small molecule inhibitors, receptor subtypes and clinical indications. *Curr Pharmaceut Design* 1999;5:289–315.

31. Christos TE, Arvanitis A. Corticotropin-releasing factor receptor antagonists. *Expert Opin Ther Patents* 1998;8:143–152.

32. Lundkvist J, Chai Z, Teheranian R, et al. A nonpeptidic corticotropin releasing factor receptor antagonist attenuates fever and exhibits anxiolytic-like activity. *Eur J Pharmacol* 1996;309:195–200.

33. Griebel G, Perrault G, Sanger DJ. Characterization of the behavioral profile of the non-peptide CRF receptor antagonist CP 154,526 in anxiety models in rodents. Comparison with diazepam and buspirone. *Psychopharmacology (Berl)* 1998;138:55–66.

34. Griebel G. Is there a future for neuropeptide receptor ligands in the treatment of anxiety disorders? *Pharmacol Ther* 1999;82:1–99.

35. Kehne JH, Coverdale S, McCloskey TC, et al. Effects of the CRF-1 receptor antagonist, CP 154,526, in the separationinduced vocalization anxiolytic test in rat pups. *Neuropharmacology* 2000;39:1357–67.

36. Kehne JH, McCloskey TC, Baron BM, et al. NMDA receptor complex antagonists have potential anxiolytic effects as measured with separation-induced ultrasonic vocalizations. *Eur J Pharmacol* 1991;193:283–92.

37. Zobel AW, Nickel T, Kunzel HE, et al. Effects of the high affinity corticotropin-releasing hormone receptor 1 antagonist R121919 in major depression: the first 20 patients treated. *J Psychiatr Res* 2000;34:171–181.

38. Paul S. GABA and glycine. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress.* New York: Raven Press, 1995:87–94.

39. Sieghart W. Structure and pharmacology of γ-aminobutyric acid A receptor subtypes. *Pharmacol Rev* 1995;47:181–234.

40. Kaupmann K, Nuggel K, Heid J, et al. Expression cloning of GABAb receptors uncovers similarity to metabotropic glutamate receptors. *Nature* 1997;386:239–246.

41. Wess J. G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition. *FASEB J* 1997;11:346–354.

42. Tallman JF, Paul SM, Skolnick P, et al. Receptors for the age of anxiety: pharmacology of the benzodiazepines. *Science* 1980;207:274–281.

43. Olsen RW, Tobin AJ. Molecular biology of $GABA_A$ receptors. *FASEB J* 1990;4:1469–1480.

44. Tallman JF, Gallager DW. The GABA-ergic system: a locus of benzodiazepine action. *Annu Rev Neurosci* 1985;8:21–44.

45. Schoefield PR, Darlison MG, Fujita N, et al. Sequence and functional expression of the $GABA_A$ receptor shows a ligand-gated receptor superfamily. *Nature* 1987;328:221–227.

46. Verdoorn TA, Draguhn A, Ymer S, et al. Functional properties of recombinant rat $GABA_A$ receptors depend upon subunit composition. *Neuron* 1990;4:919–928.

47. Galzi JL, Changeux JP. Neurotransmitter-gated ion channels as unconventional allosteric proteins. *Curr Opin Struct Biol* 1994;4:554–565.

48. Galzi JL, DeVillars-Thiery A, Hussy N, et al. Mutations in the channel domain of a neuronal nicotinic receptor convert ion selectivity from cationic to anionic. *Nature* 1992;359:500–504.

49. Nayeem N, Green TP, Martin IL, et al. Quartenary structure of the native $GABA_A$ receptor determined by electron microscopic image analysis. *J Neurochem* 1994;62:815–818.

50. Knight A, Hartnett C, Marks C, et al. Molecular size of $\alpha_1\beta_1$ and $\alpha_1\beta_1\gamma_2$ $GABA_A$ receptors expressed in SF9 cells. *Recept Channels* 1998;6:1–18.

51. Barnard EA, Skolnick P, Olsen RW, et al. IUP XV. Subtypes of $GABA_A$ receptors: classification of the basis of subunit structure and receptor function. *Pharmacol Rev* 1998;50: 291–313.

52. Prichett DB, Sonheimer H, Shivers BD, et al. Importance of a novel $GABA_A$ receptor subunit for benzodiazepine pharmacology. *Nature* 1989;338:582–584.

53. Sigel E, Buhr A. The benzodiazepine binding site of $GABA_A$ receptors. *Trends Pharmacol Sci* 1997;18:425–429.

54. Hartnett C, Brown MS, Primus, R, et al. Effect of subunit composition of $GABA_A$ receptor complex characteristics in a baculovirus expression system. *Recept Channels* 1996;4:179–195.

55. Tallman JF, Thomas JW, Gallager DW. GABAergic modulation of benzodiazepine binding site sensitivity. *Nature* 1978;274:383–385.

56. Monod J, Changeux JP, Wyman J. On the nature of allosteric transitions: a plausible model. *J Mol Biol* 1965;12:88–118.

57. Changeux JP, Edelstein SJ. Allosteric receptors after 30 years. *Neuron* 1998;21:959–980.

58. Russek SJ, Farb DH. Mapping of the β_2 subunit gene to microdissected human chromosome 5q34-q35 defines a gene cluster for the most abundant $GABA_A$ receptor isoform. *Genomics* 1994;23:528–533.

59. McLean PJ, Farb DH, Russek SJ. Mapping of the α_4 subunit to human chromosome 4 defines an α_2-α_4-β_1-γ_1 gene cluster: further evidence that modern $GABA_A$ receptor gene clusters are derived from an ancestral cluster. *Genomics* 1995;26:580–586.

60. Fritschy JM, Mohler H. $GABA_A$ receptor heterogeneity in the adult rat brain: differential regional and cellular distribution of seven major subunits. *J Comp Neurol* 1995;359:154–194.

61. Wisden W, Laurie H, Monyer H, et al. The distribution of 13 $GABA_A$ receptor subunit mRNAs in the rat brain. I. Telencephalon, diencephalon, mesencephalon. *J Neurosci* 1992;12:1040–1063.

62. Laurie DJ, Wisden W, Seeburg PH. The distribution of 13 $GABA_A$ receptor subunit mRNAs in the rat brain. II. Olfactory bulb and cerebellum. *J Neurosci* 1992;12:1063–1076.

63. Klepner CA, Lippa AS, Benson DI, et al. Resolution of two biochemically and pharmacologically distinct benzodiazepine receptors. *Pharmacol Biochem Behav* 1979;11:457–462.

64. Pritchett DB, Seeburg PH. Type I and type II $GABA_A$-benzodiazepine receptors produced in transfected cells. *Science* 1989;245:1389–1392.

65. Luddens H, Pritchett DB, Kohler M, et al. Cerebellar $GABA_A$ receptor selective for a behavioral alcohol antagonist. *Nature* 1990;346:648–651.

66. Magenis RE. Chromosomes and their disorders. In: Bennet JC, Plum F. *Cecil's textbook of medicine*. Philadelphia: WB Saunders, 1996:150–157.

67. Culiat CT, Stubbs L, Woychik J, et al. Deficiency of the $ß_3$ subunit of the type A GABA receptor causes cleft palate in mice. *Nature Genet* 1995;11:344–346.

68. Homanics, G, DeLorey TM, Firestone L, et al. Mice devoid of $GABA_A$ receptor $ß_3$ subunit have epilepsy, cleft palate and hypersensitive behavior. *Proc Natl Acad Sci USA* 1997;94:4143–4148.

69. Mohler H, Luscher B, Fritschy JM, et al. $GABA_A$-receptor assembly *in vivo*: lessons from subunit mutant mice. *Life Sci* 1998;62:1611–1615.

70. Rudolph U, Crestani F, Benke D, et al. Benzodiazepine actions mediated by specific $GABA_A$ receptor subtypes. *Nature* 1999;401:796–800.

71. McKernan R, Rosahl TW, Reynolds DS, et al. Sedative but not anxiolytic properties of benzodiazepines are mediated by the $GABA_A$ receptor α_1 subtype $GABA_A$ α_1 receptors. *Nature Neurosci* 2000;3:587–592.

72. Lucki I. Serotonin receptor specificity in anxiety disorders. *J Clin Psychiatry* 1996;57(suppl 6):5–10.

73. Critchley MAE, Handley SL. Effects in the X-maze anxiety model of agents acting at 5-HT1 and 5-HT2 receptors. *Psychopharmacology (Berl)* 1987;93:502–506.

74. Gleeson S, Ahlers ST, Mansbach RS, et al. Behavioral studies with anxiolytic drugs, VI effects on punished responding of drugs interacting with serotonin receptor subtypes. *J Pharmacol Exp Ther* 1989;250:809–817.

75. Kennett GA, Bailey F, Piper DC, et al. Effect of SB 200646A, a 5-HT $S\text{-}HT_{2b}$ receptor antagonist, in two conflict models of anxiety. *Psychopharmacology (Berl)* 1995;188:178–182.

76. Kennett GA, Whitton P, Shah K, et al. Anxiogenic-like effects of mCPP and TFMPP in animal models are opposed by $5\text{-}HT_{1c}$ receptor antagonists. *Eur J Pharmacol* 1989;164:445–454.

77. Costall B, Naylor RJ. Anxiolytic effects of $5\text{-}HT_1$ antagonists in animals. In: Rodgers RJ, Cooper SJ, eds. *$5\text{-}HT_{1A}$ agonists, $5\text{-}HT_3$ antagonists and benzodiazepines: their comparative behavioral pharmacology*. Chichester, UK: Wiley, 1991:133–157.

78. Ramboz S, Oosting R, Ait Amara D, et al. Serotonin receptor 1A knockout: an animal model of anxiety-related disorder. *Proc Natl Acad Sci USA* 1998;95(24):14476–14481.

79. Johnson Al, File SE. 5-HT and anxiety: promises and pitfalls. *Pharmacol Biochem Behav* 1986;24:1467–1470.

80. Julius D. Serotonin receptor knockouts: a moody subject. *Proc Natl Acad Sci USA* 1998;95(26):15153–15154.

81. Coplan JD, Wolk SI, Klein DF. Anxiety and the $serotonin_{1A}$ receptor. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:1301.

82. Mahmood I, Sahajwalla C. Clinical pharmacokinetics and pharmacodynamics of buspirone, an anxiolytic drug. *Clin Pharmacokinet* 1999;36(4):277–287.

83. Laakmann G, Schule C, Lorkowski G, et al. Buspirone and lorazepam in the treatment of generalized anxiety disorder in outpatients. *Psychopharmacology (Berl)* 1998;136(4):357–366.

84. Rocca P, Fonzo V, Scotta M, et al. Paroxetine efficacy in the treatment of generalized anxiety disorder. *Acta Psychiatr Scand* 1997;95:444–450.

85. Goodman WK. Obsessive-compulsive disorder: diagnosis and treatment. *J Clin Psychiatry* 1999;60(suppl 18):27–32.

86. Copland JD, Wolk SI, Klein DF. Anxiety and serotonin 1a receptor. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:1301–1311.

87. Broocks A, Pigott TA, Hill JL, et al. Acute intravenous administration of ondansetron and m-CPP, alone and in combination, in patients with obsessive-compulsive disorder (OCD): behavioral and biological results. *Psychiatry Res* 1998;79(1):11–20.

88. Pian KL, Westenberg HG, van Megen HJ, et al. Sumatriptan (5-HT1D receptor agonist) does not exacerbate symptoms in obsessive-compulsive disorder. *Psychopharmacology (Berl)* 1998;140(3):365–370.

89. den Boer JA. Pharmacotherapy of panic disorder: differential efficacy from a clinical viewpoint. *J Clin Psychiatry* 1998;59(suppl 8):30–36; discussion 37–38.

90. Bell CJ, Nutt DJ. Serotonin and panic. *Br J Psychiatry* 1998;172:465–471.

91. van Vliet IM, Westenberg HG, den Boer JA. Effects of the 5-HT1A receptor agonist flesinoxan in panic disorder. *Psychopharmacology (Berl)* 1996;127(2):174–180.

92. Otsuka M, Yoshioka K. Neurotransmitter functions of mammalian tachykinins. *Physiol Rev* 1993;73:229–308.

93. Khawaja AM, Rogers DF. Tachykinins: receptor to effector. *Int J Biochem Biol* 1996;28:721–738.

94. Longmore J, Hill RG, Hargreaves RJ. Neurokinin-receptor antagonists: pharmacological tools and therapeutic drugs. *Can J Physiol Pharmacol* 1997;75:612–621.

95. Mantyh PW, Gates T, Mantyh CR, et al. Autoradiographic localization and characterization of tachykinin receptor binding sites in the rat brain and peripheral tissues. *J Neurosci* 1989;9:258–279.

96. Tooney PA, Au GG, Chahl LA. Localization of tachykinin NK1 and NK3 receptors in the human prefrontal and visual cortex. *Neurosci Lett* 2000;283:185–188.

97. Culman J, Unger T. Central tachykinins: mediators of defence reaction and stress reactions. *Can J Physiol Pharmacol* 1995;73:885–891.

98. Kramer MS, Cutler N, Feighner J, et al. Distinct mechanism for antidepressant activity by blockade of central substance P receptors. *Science* 1998;281:1640–1645.

99. File SE. Anxiolytic action of a neurokinin 1 receptor antagonist in the social interaction test. *Pharmacol Biochem Behav* 1997;58:747–752.

100. Ribeiro SJ, Teixeira RM, Calixto JB, et al. Tachykinin NK3 receptor involvement in anxiety. *Neuropeptides* 1999;33:181–188.

101. Schoepp DD, Jane DE, et al. Pharmacological agents acting at subtypes of metabotropic glutamate receptors. *Neuropharmacology* 1999;38:1431–1476.

102. Cotman CW, Kahle JS, Miller SE, et al. Excitatory amino acid neurotransmission. In: Bloom FE, Kupfer DJ, editors. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:75–85.

103. Meldrum BS. Glutamate as a neurotransmitter in the brain: review of physiology and pathology. *J Nutr* 2000;130:1007S–10015S.

104. Baron BM, Harrison BL, Kehne JH, et al. Pharmacological characterization of MDL 105,519, an NMDA receptor glycine site antagonist. *Eur J Pharmacol* 1997;323:181–192.

105. Kew JN, Koester J, Moreau JL, et al. Functional consequences of reduction in NMDA receptor glycine affinity in mice carrying targeted point mutations in the glycine binding site. *J Neurosci* 2000;20:4037–4049.

106. Monn JA, Valli MJ, Massey SM, et al. Design, synthesis, and pharmacological characterization of (+)-2-aminobicyclo93.1.0)hexane-2,6-dicarboxylic acid (LY354740): a potent, selective, and orally active group 2 metabotropic glutamate receptor agonist possessing anticonvulsant and anxiolytic properties. *J Med Chem* 1997;40:528–537.

107. Helton DR, Tizzano JP, Monn JA, et al. Anxiolytic and side-effect profile of LY354740: a potent, highly selective, orally active agonist for group II metabotropic glutamate receptors. *J Pharmacol Exp Ther* 1998;284:651–660.

108. Klodzinska A, Chojnacka-Wojcik E, Palucha A, et al. Potential anti-anxiety, anti-addictive effects of LY 354740, a selective group II glutamate metabotropic receptors agonist in animal models. *Neuropharmacology* 1999;38:1831–1839.

109. Shekhar A, Keim SR. LY354740, a potent group II metabotropic glutamate receptor agonist prevents lactate-induced panic-like response in panic-prone rats. *Neuropharmacology* 2000;39:1139–1146.

110. Vandergriff J, Rasmussen K. The selective mGlu2/3 receptor agonist LY354740 attenuates morphine-withdrawal-induced activation of locus coeruleus neurons and behavioral signs of morphine withdrawal. *Neuropharmacology* 1999;38:217–222.

111. Kehne JH, Baron BM, Harrison BL, et al. MDL 100,458 and MDL 102,288: two potent and selective glycine receptor antagonists with different functional profiles. *Eur J Pharmacol* 1995; 284:109–118.

112. Snell LD, Claffey DJ, et al. Novel structure having antagonist actions at both the glycine site of the N-methyl-D-aspartate receptor and neuronal voltage-sensitive sodium channels: biochemical, electrophysiological, and behavioral characterization. *J Pharmacol Exp Ther* 2000;292:215–227.

113. Kotlinkska J, Liljequist S. A characterization of anxiolytic-like actions induced by the novel NMDA/glycine site antagonist, L-701,324. *Psychopharmacology* 1998;135:175–181.

114. Kotlinska J, Liljequist S. The putative AMPA receptor antagonist, LY326325, produces anxiolytic-like effects without altering locomotor activity in rats. *Pharmacol Biochem Behav* 1998;60:119–124.

115. Matheus MG, Guimaraes FS. Antagonism of non-NMDA receptors in the dorsal periaqueductal grey induces anxiolytic effect in the elevated plus maze. *Psychopharmacology* 1997;132:14–18.

116. Sajdyk TJ, Schober DA, et al. Role of corticotropin-releasing factor and urocortin within the basolateral amygdala of rats in anxiety and panic responses. *Behav Brain Res* 1999;100:207–215.

117. Mutt V, Jorpes E. Hormonal polypeptides of the upper intestine. *Biochem J* 1971;125:57–58.

118. Hokfel, T, Skirboll L, Everitt B, et al. Distribution of cholecystokinin-like immunoreactivity in the nervous system. Co-existence with classical neurotransmitters and other neuropeptides. *Ann NY Acad Sci* 1985;448:255–274.

119. Emson PC, Rehfield JF, Rossor MN. Distribution of cholecystokinin-like peptides in the human brain. *J Neurochem* 1982;38:1177–1179.

120. Moran TH, Robinson PH, et al. Two brain cholecystokinin receptors: implications for behavioral actions. *Brain Res* 1986;362:175–179.

121. Pisegna JR, de Weerth A, Huppi K, et al. Molecular cloning of the human brain and gastric cholecystokinin receptor: structure, functional expression and chromosomal localization. *Biochem Biophys Res Commun* 1992;189:296–303.

122. Adams JB, Pyke RE, Costa J, et al. A double-blind, placebo-controlled study of a CCKB receptor antagonist, CI-988, in patients with generalized anxiety disorder. *J Clin Psychopharmacol* 1995;15:428–434.

123. Kramer MS, Cutler NR, Ballenger JC, et al. A placebo-controlled trial of L-365,260, a CCKB antagonist, in panic disorder. *Biol Psychiatry* 1995;37:462–466.

124. Pande AC. Lack of efficacy of a cholecystokinin-B antagonist in anxiety disorders. *Biol Psychiatry* 1997;42:11–19.

Section VIII
Affective disorders

Charles Nemeroff

Charles Nemeroff: Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia.

Affective disorders - Introduction

> Canst thou not minister to a mind diseased
>
> Pluck from the memory a rooted sorrow,
>
> Raze out the written troubles of the brain,
>
> And with some sweet oblivious antidote,
>
> Cleanse the stuffed bosom of that perilous
>
> Stuff which weights upon the heart?
>
> --William Shakespeare, *Macbeth*

The advances in neurobiology documented comprehensively elsewhere in this volume have been applied to no area of research more so than affective disorders. This is not surprising, considering the magnitude of the public health problem these disorders represent and the long way yet to travel before an understanding of their pathophysiology is realized and optimal treatments are developed. Although the task is daunting, recent progress is heartening and informative. Major depression and bipolar disorder, the most common and severe of the mood disorders, are responsible for much of the suicide in the United States—now the eighth leading cause of death, surpassing HIV infection, in the last 2 years. The Global Burden of Disease report, a collaborative effort of the World Health Organization and the Harvard University School of Public Health, indicated that depression would be the leading cause of morbidity in the developing world in this next century. Between a fourth and a fifth of all women in the United States will experience a major depressive episode in their lifetime. It is now well established that depression is an independent risk factor for the development of coronary artery disease, stroke, and perhaps cancer. This section provides succinct yet comprehensive reviews of the most critical areas in mood disorders research by some of the leading investigators in the field. Each contribution highlights one or more major advances in the field.

Thus, there is an appropriate emphasis on the longitudinal course of affective illness and a corresponding criticism of cross-sectional studies. With the maturation of psychiatric epidemiology and the Human Genome Project comes anticipation of elucidation of both risk factors, and susceptibility and resistance genes, respectively. The pathophysiology of these disorders is being scrutinized by a myriad of techniques ranging from cellular and molecular approaches, to structural and functional brain imaging, to light and electron microscopic neuroanatomic investigations. Similar techniques are being utilized to elucidate the mechanism of action of antidepressants, mood stabilizers, and newer treatments including rapid transcranial magnetic stimulation, as well as to identify novel treatments including hormones and hormone antagonists. The data reviewing the efficacy of current and novel treatments are presented, with the realization that too many patients are improved by the current treatment regimens, but far too few achieve complete remission. Appropriately, there is also attention to the pharmacoeconomics of mood disorder treatment.

I am grateful to all of the authors for their painstaking contributions and accept all responsibility for any shortcomings in this section. This section contains much information unknown when the last edition of this book was published. We can all be heartened by the fact that the next edition will similarly replace this present effort.

69

The Course of Depression

Robert J. Boland

Martin B. Keller

Robert J. Boland: Department of Psychiatry and Human Behavior, Brown University; Department of Psychiatry, Miriam Hospital, Providence, Rhode Island.

Martin B. Keller: Department of Psychiatry and Human Behavior, Brown University; Department of Psychiatry, Butler Hospital and Brown Affiliated Hospitals, Providence, Rhode Island.

Long-term naturalistic studies have changed the way we view depression. Whereas it was often previously viewed as an episodic disease, the past two decades of research have underscored the importance of understanding depression as a lifelong disease, with a number of possible courses.

An appreciation of this longitudinal data is crucial to understanding all aspects of depression. Cross-sectional judgments of symptomatic severity provide limited prognostic information. A full understanding of a patient's prognosis or likely treatment response also requires a longitudinal perspective. Which patient is likely to recover fully, and who will suffer from a chronic mood disorder? What length of treatment will be sufficient for such patients?

Studies within the last decade have helped to shed light on these questions. This chapter examines some of these studies, and discusses their implications for our approach to depression. Limitations of the data will be discussed as well.

THE CHANGE POINTS OF DEPRESSION

Part of "69 - The Course of Depression "

Considerable confusion has resulted from the use of various terms to denote the different change points in the course of depression. Similar terms, such as “relapse” and “recurrence” have been used interchangeably and inconsistently in different studies. As a result, the MacArthur Foundations Research Network on the Psychobiology of Depression (1) recommended using the following terms:

1. *Episode*, defined as a certain number of symptoms for a certain period of time.
2. *Remission*, defined as a period of time in which an individual no longer meets criteria for the disorder. In partial remission, an individual still has more than minimal symptoms. Full remission is defined as the point at which an individual no longer meets criteria for the disorder and has no more than minimal symptoms.
3. *Recovery*, defined as a full remission that lasts for a defined period of time. Conceptually, it implies the end of an episode of an illness, not of the illness per se.
4. *Relapse*, defined as a return of symptoms sufficient to satisfy full criteria for an episode. It occurs in an interval of time before what is defined as “recovery.” Conceptually, this refers to the return of an episode, not a new episode.
5. *Recurrence*, defined as a return of full symptomatology occurring after the beginning of the recovery period. Conceptually, this represents the beginning of a new episode of an illness.

REPRESENTATIVE STUDIES

Part of "69 - The Course of Depression "

A relatively small number of studies have been particularly influential in shedding light on the course of depression.

The Collaborative Depression Study (CDS)

The CDS (2) is a prospective long-term naturalistic study of the natural course of depression. Subjects were recruited from patients with depression seeking psychiatric treatment at one of several sites (university or teaching hospitals in Boston, Chicago, Iowa City, New York, and St. Louis). This study included programs in biological and clinical studies. The data presented here are from the clinical studies program; 555 subjects in the clinical studies program had an index episode of unipolar major depression. Subjects were examined at 6-month intervals for 5 years and then annually for a minimum of 18 years. Recent National Institute of Mental Health (NIMH) funding will extend the follow-up to at least 23 years on all subjects.

The Zurich Study

Angst (3), in Zurich, has conducted the only other long-term prospective study of mood disorders. In that study, 173 hospitalized patients with unipolar depression were identified between 1959 and 1963. This group was then evaluated every 5 years for up to 21 years of follow-up.

The Medical Outcomes Study (MOS)

The MOS (4) examined the course of several diseases (myocardial infarction, congestive heart failure, hypertension, diabetes, and depression) in a variety of health care settings, including large medical group practices, small group practices, and solo practices, in three cities (Los Angeles, Boston, and Chicago). A representative sample of different medical specialties—including psychiatry—was chosen, and all patients seen from February through October 1986 were asked to participate in the study. In all, over 20,000 patients participated, and were evaluated yearly for 3 years.

THE COURSE OF DEPRESSION: CHANGE POINTS

Part of "69 - The Course of Depression "

Traditionally, depression was pictured as an acute illness, self-limited, and lasting approximately 6 to 9 months from time of onset to full recovery. A number of studies, including those mentioned above, however, show the potential for great variation from this traditional model. Recovery may take much longer, or not occur at all (i.e., chronic depression). Furthermore, the risk of relapse and recurrence of illness must be considered.

Recovery

In the CDS, approximately 70% of patients recovered from the index episode of major depression within the first year (5). However, for those patients who did not recover in the first year, most still had not recovered within 5 years. Thus by 2 years, about 20% of the original sample were still depressed—two-thirds of those still depressed at 1 year were still in their index episode of depression at 2 years. At 5 years, 12% of patients had still not recovered (6), by 10 years 7% had not recovered (7), and by 15 years, the numbers seem to have leveled off at 6%. These data are presented in Fig. 69.1 .

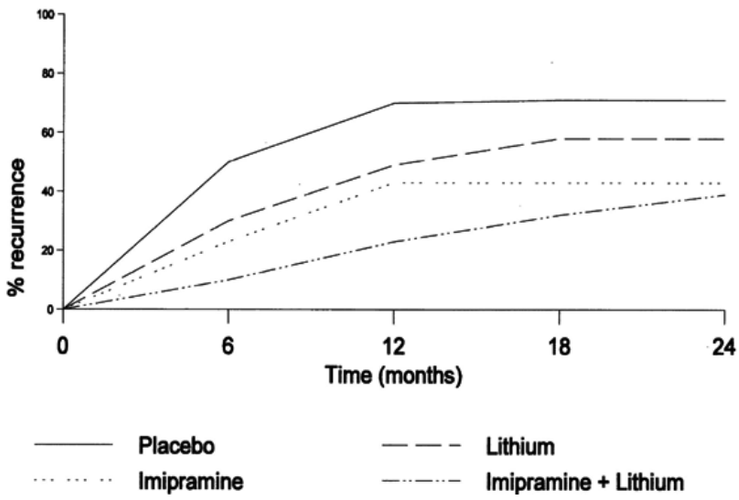


FIGURE 69.1. Outcome of maintenance therapy for depressed patients initially stabilized on imipramine plus lithium.

The long duration of the CDS allowed the investigators to observe subsequent episodes of major depression beginning during the study. This was particularly useful, as the onset of symptoms could be identified more accurately than for the retrospective determination done for an index episode. It was found that, for each new episode of depression, the rates of recovery were similar to that seen during the index episode. Thus, for the second episode (first prospectively observed episode) approximately 8% of subjects did not recover after 5 years. An analysis of subsequent episodes (second, third, and fourth prospectively observed episodes) showed similar findings. By the fifth episode, the rate decreases, but not significantly so (8). It appears that for each episode of depression, some individuals—about 10%—remain ill for at least 5 years.

The seemingly high rate of chronicity was surprising. A reasonable concern about this result was that the patient population studied may have been unusually treatment resistant. The study used a convenient sample of patients seeking inpatient or outpatient treatment at any one of five major medical centers. However, most patients studied received either no treatment or subtherapeutic doses given for very brief durations (9). Thus, the CDS cohort does not seem to be biased in the direction of treatment resistance.

Furthermore, other studies show comparable data. In the Zurich study, Angst et al. (10) reported that during the

follow-up evaluations, about 13% of patients did not recover from their episode of major depression. In the MOS, patients were divided by severity: of those with milder depression, about 65% recovered within 2 years, whereas 54% of the more severely depressed group recovered in the same period (11).

Shorter studies also give similar results. Rounsaville and colleagues (12), in a prospective follow-up of 96 patients with major depression, found that 12% of subjects had not recovered after 16 months. Kerr et al. (13), following initially hospitalized patients, found that 6% remained ill for the 4 years of the study.

Relapse

For the 141 patients in the CDS who recovered from their index episode of major depression, 22% relapsed within 1 year of follow-up (14). Factors predicting relapse included multiple episodes of major depression, older age, and a history of nonaffective psychiatric illness. The characteristics of this relapsed group were also examined, and it was found that the likelihood of remaining depressed for at least a year after relapse was 22%. Predictors of prolonged time to recovery included a longer length of the index episode, older age, and a lower family income.

Most studies look at relapse in terms of how it is affected by treatment (see below).

Recurrence

Angst (15), reporting on a 10-follow of patients in the Zurich study, found that only 25% of patients had only a single episode of depression. Thus, three-fourths of the sample had a recurrent depression, with one or more recurrences. Though Angst examined a number of sociodemographic variables, none significantly predicted the likelihood of recurrence.

Similarly high rates of recurrence have been found in other long-term studies. Weissman and Kasl (16) found that two-thirds of woman seen over 1 year had a recurrence of depression. Rao and Nammalvar (17), examining over 100 cases of depression in India for a follow-up of between 3 and 13 years, found that only about a fourth of the original group reported no recurrence of symptoms.

The rate and timing of recurrence seems most dependent on the type of recovery. Patients in the CDS who fully recovered (i.e., were asymptomatic on follow-up evaluation) had a much lower rate of recurrence (66%) than those with some residual symptoms (87%). The time to recurrence was also much longer in the asymptomatic group: mean of 180 weeks in the asymptomatic group compared with 33 weeks in the group with residual symptoms (18).

MITIGATING FACTORS

Part of "69 - The Course of Depression "

Comorbidity

Medical Illness

There are few longitudinal studies looking at the outcome of depression in medically ill patients, partly because of the difficulties inherent in recruiting such an unstable population. Studies that exist suggest that comorbid medical illness predispose individuals to a worse course of depression. The MOS, for example, found an additive effect on patient functioning when depression and other chronic medical illnesses were combined (19).

Double Depression

Double depression refers to the presence of concurrent dysthymia and major depression. In this disorder, the episodes of major depression are superimposed on a more chronic depressive disorder. It appears to be common—studies suggest that between one-fourth and two-thirds of patients with major depression will also have a comorbid dysthymia.

The comorbid presence of dysthymia can have an important effect on the course of depression. In the collaborative study, it was found that patients with double depression recovered more rapidly from episodes of major depression than those with major depression alone. However, the authors also found that the recovery tended to be not to one of “normalcy,” but to one of dysthymia. Relapse is also more frequent in patients with double depression than those with major depression alone—almost twice as likely in one study of 32 double-depressed subjects followed for 2 years (20). The MOS also found that full recovery was less likely for patients with double depression—these patients had a threefold risk of continued disease when compared with those with major depression alone (21).

Other Psychiatric Illnesses

Substance Abuse

Clearly, comorbid substance abuse has a detrimental effect on the course of depression. The CDS found that subjects who were currently alcoholic were half as likely as nonalcoholic depressed subjects to recover from their episode of major depression (22). Patients with a previous, but not current, history of alcoholism had a recovery rate comparable to those with no such history.

Anxiety Disorders

Anxiety disorders are commonly comorbid with depression. The presence of such comorbid disorders appears to exert a negative effect on the course of depression. Coryell and colleagues (23) found that depressed patients with panic disorder had a slower time to recovery than those without

comorbid panic. The CDS similarly found that patients with higher symptom ratings of anxiety had longer times to recovery from major depression (24).

Family History

A family history of depression appears to predispose an individual to depression. Two studies have looked at the relationship between parental history of depression and course of depression in the offspring. Though these studies had relatively small sample sizes, they both suggested that patients with a parental history of depression had a longer time to recovery than other patients (25 ,26).

Treatment Variables

Clearly, one of the questions of most practical interest is whether pharmacologic treatment is capable of significantly altering the course of depression for a patient. Antidepressants are generally used at all stages of depression—to hasten recovery, prevent relapse, and prevent recurrence of depression. However, as will be discussed, the further one looks down the course of depression, the less is really known about the ability of antidepressant and other pharmacologic treatments to alter the course of depression.

A wealth of data supports the efficacy of all available antidepressants in shortening the time to recovery from major depression. However, when one goes beyond the acute phase and examines pharmacotherapy during later points in the course of depression, the data become more meager.

Data support the efficacy of most of the serotonin reuptake inhibitors for continuation therapy, including fluoxetine (27), paroxetine (28), sertraline (29), citalopram (30), and mirtazapine (31). Nefazodone has also been shown in continuation treatment (32). These studies are summarized in Table 69.1 . However, when looking beyond continuation therapy to the maintenance treatment of recurrent depression, much fewer data exist.

Drug (Reference)	Weeks of Treatment	Relapse (Drug) %	Relapse (Placebo) %	*P* Value
Fluoxetine (27)	52	26	57	<.01
Paroxetine (28)	52	16	43	<.001
Sertraline (29)	44	13	46	<.001
Citalopram (30)	24	11	31	<.05
Nefazodone (32)	36	17	33	<.05
Mirtazapine (31)	20	4	23	<.0001

TABLE 69.1. RELAPSE RATES VS. PLACEBO: CONTINUATION STUDIES

Prien and colleagues (33) reported on a 2-year maintenance trial for depression. Patients who were successfully treated for acute depression were randomized to receive lithium carbonate, imipramine, both, or placebo. Treatments were continued for 2 years with doses maintained at acute treatment levels. Of 150 patients beginning maintenance treatment, 36% were successfully treated. The lowers rate of recurrence was found it the group treated with imipramine (Fig. 69.1). However, even this group had a 47% recurrent rate.

A second study, the Pittsburgh Study of Maintenance Treatment for Recurrent Depression (34), reports on up to

5 years of maintenance treatment. In this study, subjects first underwent open treatment for acute depression, using imipramine with interpersonal therapy (IPT). Patients who achieved recovery for at least 4 months were then randomized into one of five treatment conditions: (a) IPT alone, given monthly; (b) imipramine treatment alone; (c) IPT plus placebo; (d) placebo plus medication clinic; and (e) imipramine plus IPT. This portion of the study was continued for 3 years. Results from this study are summarized in Fig. 69.2 .

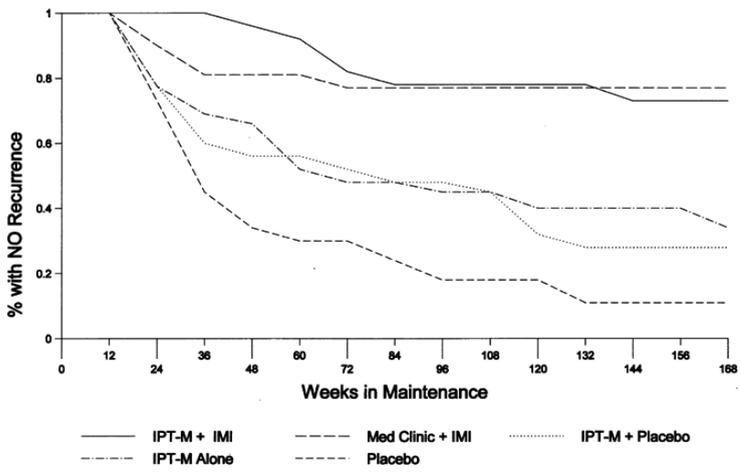


FIGURE 69.2. Pittsburgh Study of Maintenance Therapies in Recurrent Depression.

The authors intentionally chose patients who had highly recurrent depression (to maximize the likelihood of seeing a statistical difference in the groups). They found that patients who received imipramine (with or without IPT) had an approximately 20% risk of recurrence after 3 years. This compares significantly with the other conditions: IPT alone had a 60% risk of recurrence, and the placebo condition had an 80% risk.

This study was extended, with a smaller sample, for another 2 years. The subjects who had remained well during the first 3 years of the study were randomized to receive either imipramine or placebo. At that point only 20 patients remained in the study. After the two additional years, less than 10% of the patients (one patient out of 11) receiving imipramine had a recurrence, where approximately two-thirds of the placebo group had a recurrence.

Thus, the world literature of placebo-controlled studies of the treatment of recurrent depression beyond 3 years consists of only 200 subjects, who had a history of highly recurrent depression and received a medication that is rarely used by most psychiatrists today.

OTHER SUBTYPES OF DEPRESSION

Part of "69 - The Course of Depression "

The data are also limited when one considers the effect of treatment on the course of specific subtypes of depression. Two such subtypes will be considered here: chronic major depression and dysthymia.

Treatment Of Chronic Depression

Chronic depression is thought to respond more poorly to antidepressant treatment. Thus, studies of the acute and long-term treatment for this subtype are of great importance. Particularly lacking are studies of psychotherapy in this population.

Keller and colleagues (35) recently compared antidepressant treatment and cognitive-behavioral therapy, both alone and combined. For the 519 subjects completing the study, 55% of the antidepressant (nefazodone) group and 52% of the psychotherapy group responded to treatment. However, when treatment was combined, the response rate jumped to 85%. Thus, this study gives strong support to the clinical wisdom that combined treatment is preferable to either medication or psychotherapy alone.

To date, there are only two published studies on the long-term treatment of chronic depression. Kocsis and colleagues (36) report on a placebo-controlled trial of desipramine for the treatment of chronic depression. The study included patients with chronic major depression (*n* = 14).

After successful acute-phase and continuation treatment, patients were continued on treatment for a total of 2 years. During this maintenance period, patients on the placebo had four times the recurrence rate of those receiving desipramine. This rate was consistent for all diagnostic groups, including those with chronic depression.

Keller and colleagues (37) investigated the treatment of chronic depression in a larger sample of patients. Here, 161 patients who were successfully treated during an acute-phase and continuation phase were randomized to receive with sertraline or placebo for a 76-week maintenance period. The study included both chronic major depression and double depression patients in roughly equal numbers; as the results did not differ between the groups, the data was pooled.

Patients who received the placebo during the maintenance phase of treatment were four times more likely to have a recurrence of depression than those receiving sertraline. The time to recurrence was delayed for patients treated with sertraline compared to those treated with the placebo. Using a less stringent criterion of reemergence of depressive symptoms (though no necessarily meeting full criteria for depression), it was found that only 26% of patients taking sertraline experienced a reemergence of depressive symptoms, compared with 50% of those on the placebo. Similarly, only 34% of patients on sertraline maintenance therapy shored first symptoms of depression, compared with 60% of patients taking the placebo (Table 69.2).

TABLE 69.2. RECURRENCE RATES OF MAJOR DEPRESSION DURING MAINTENANCE STUDY TREATMENT

	Sertraline (*n* = 77)	Placebo (*n* = 84)	*P* Value
Suffered recurrence by strict protocol criteria (%)	6	23	.002
Suffered depression reemergence by consensus assessment (%)	26	50	.001
Showed first symptoms of reemergence of depression by consensus assessment (%)	34	60	.001

From Keller MB, Kocsis JH, Thase ME, et al. Maintenance phase efficacy of sertraline for chronic depression: a randomized controlled trial. *JAMA* 1998;280:1665–1672, with permission.

Treatment Of Dysthymia

Few studies have examined the pharmacotherapy of dysthymia, possible because of long-held beliefs that nonmajor depressions were less responsive to pharmacotherapy. What data do exist, however, do not support this belief. Most studies are of a relatively short duration of treatment, ranging from 4 to 12 weeks. For this time period there are data to support the use of most classes of antidepressants. These studies are summarized in Table 69.3 . Thus, the weight of evidence suggests that most agents that are effective for major depression are also effective for dysthymia, at least in the acute phase of treatment.

Drug	Compared to:	Study	Duration of Study (Weeks)
Imipramine	Placebo	Koesis et al., 1985 (38)	6
	Phenelzine, placebo	Stewart et al., 1993 (39)	6
Desipramine	Placebo	Stewart et al., 1983 (40)	6
Fluoxetine	Placebo	Hellerstein et al., 1993 (41); Vanelle, 1997 (42)	8; 12
Sertraline	Placebo	Ravindran et al., 1994 (43)	12
	Imipramine, placebo	Thase et al., 1996 (44)	12
	Imipramine	Keller et al., 1995 (45)	12
Ritanserin	Imipramine, placebo	Bakish et al.,1994 (46)	7
Moclobemide	Placebo	Botte et al., 1992 (47)	4
	Imipramine, placebo	Versiani et al., 1997 (48)	8
Amisulpride	Amineptine, placebo	Boyer and Lecrubier, 1996 (49)	12
	Fluoxetine	Smeraldi, 1998 (50)	12

TABLE 69.3. PHARMACOTHERAPY OF DYSTHYMIA: SELECTED AGENTS DEMONSTRATING A POSITIVE EFFECT IN RANDOMIZED-CONTROLLED TRIALS

ACKNOWLEDGMENTS

Part of "69 - The Course of Depression "

Dr. Keller has received research support and/or served as a consultant or on an advisory board for a number of different pharmaceutical companies including Pfizer, Bristol-Myers Squibb, Forrest Laboratories, Wyeth-Ayerst Laboratories, Merck, Janssen, Eli Lilly, Organon, Pharmacia/Upjohn, SmithKline Beecham, Zeneca, Mitsubishi Pharmaceuticals, Scirex, Janus Pharmaceuticals, Sepracor Pharmaceuticals, Somerset Pharmaceuticals, and Sanofi-Synthelabo.

REFERENCES

1. Frank E, Prien RF, Jarrett RB, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder: remission, recovery, relapse, and recurrence. *Arch Gen Psychiatry* 1991;48:851–855.

2. Katz M, Klerman GL. Introduction: overview of the clinical studies program. *Am J Psychiatry* 1979;136:49–51.

3. Angst J. The course of major depression, atypical bipolar disorder, and bipolar disorder. In: Hippius H, et al., eds. *New results in depression research*. Berlin/Heidelberg: Springer-Verlag, 1986:26–35.

4. Wells KB, Burnam A, Rogers B, et al. The course of depression in adult outpatients: results from the Medical Outcomes Study. *Arch Gen Psychiatry* 1992;49:788–794.

5. Keller MB, Shapiro R, Lavori PW, et al. Recovery in major depressive disorder: analysis with the life table and regression models. *Arch Gen Psychiatry* 1982;39:905–910.

6. Keller MB, Lavori PW, Mueller TI, et al. Time to recovery, chronicity, and levels of psychopathology in major depression. A 5-year prospective follow-up of 431 subjects. *Arch Gen Psychiatry* 1992;49:809–816.

7. Mueller TI, Keller MB, Leon A, et al. Recovery after five years of unremitting major depressive disorder. *Arch Gen Psychiatry* 1996;53:794–799.

8. Solomon DA, Keller MB, Leon AC. Recovery from major depression. A 10-year prospective follow-up across multiple episodes. *Arch Gen Psychiatry* 1997;54:1001–1006.

9. Keller MB, Klerman GL, Lavori PW, et al. Treatment received by depressed patients. *JAMA* 1982;248:1848–1855

10. Angst J, Baastrup P, Grof, et al. The course of monopolar depression and bipolar psychoses. *Psychiatr Neurol Neurochir* 1973;76:489–500.

11. Wells KB, Burnam A, Rogers B, et al. The course of depression in adult outpatients: results from the Medical Outcomes Study. *Arch Gen Psychiatry* 1992;49:788–794.

12. Rounsaville BJ, Prusoff BA, Padian N. Chronic mood disorders in depression outpatients: a prospective 16-month study of ambulatory patients. *J Nerv Ment Dis* 1980;168:406–411.

13. Kerr TA, Roth M, Schapira K, et al. The assessment and prediction of outcome in affective disorders. *Br J Psychiatry* 1972;121:167–174.

14. Keller MB, Lavori PW, Lewis CE, et al. Predictors of relapse in major depressive disorder. *JAMA* 1983;250:3299–3304.

15. Angst J. How recurrent and predictable is depressive illness? In: Montgomery S, Rouillon F, eds. *Long-term treatment of depression*. New York: Wiley, 1992.

16. Weissman MM, Kasl SV. Help-seeking in depressed out-patients following maintenance therapy. *Br J Psychiatry* 1976;129:252–260.

17. Rao AV, Nammalvar N. The course and outcome in depressive illness: a follow-up study of 122 cases in Madurai, India. *Br J Psychiatry* 1977;130:392–396.

18. Keller MB, Boland RJ. Implications of failing to achieve successful long-term maintenance treatment of recurrent unipolar major depression. *Biol Psychiatry* 1998;44:348–360.

19. Wells KB, Steward A, Hays RD, et al. The functioning and well-being of depressed patients—results of the Medical Outcomes Study. *JAMA* 1989;262:914–919.

20. Keller MB, Lavori PW, Rice J, et al. The persistent risk of chronicity in recurrent episodes of non-bipolar major depressive disorder: a prospective follow-up. *Am J Psychiatry* 1984;143:24–28.

21. Wells KB, Burnam A, Rogers W, et al. The course of depression in adult outpatients: results from the medical outcomes study. *Arch Gen Psychiatry* 1992;49:788–794.

22. Mueller TI, Lavori PW, Keller MB, et al. Prognostic effect of the variable course of alcoholism on the 10-year course of depression. *Am J Psychiatry* 1994;151:701–706.

23. Coryell W, Endicott J, Andreasen NC, et al. Depression and panic attacks: the significance of overlap as reflected in follow-up and family study data. *Am J Psychiatry* 1998;145:293–300.

24. Clayton PF, Grove WM, Coryell W, et al. Follow-up and family study of anxious depression. *Am J Psychiatry* 1991;148:1512–1517.

25. Warner V, Weissman MM, Fendrich M, et al. The course of major depression in the offspring of depressed parents: incidence, recurrence and recovery. *Arch Gen Psychiatry* 1992;49:795–801.

26. Hammen C, Burge D, Burney E, et al. Longitudinal study of diagnoses in children of women with unipolar and bipolar affective disorder. *Arch Gen Psychiatry* 1990;47:1112–1117.

27. Montgomery SA, Dunfour H, Brion S, et al. The prophylactic efficacy of fluoxetine in unipolar depression. *Br J Psychiatry* 1988;153(suppl 3):69–76.

28. Montgomery SA, Dunbar G. Paroxetine is better than placebo in relapse prevention and the prophylaxis of recurrent depression. *Int Clin Psychopharmacol* 1993;8:189–195.

29. Doogan DP, Caillard V. Sertraline in the prevention of depression. *Br J Psychiatry* 1992;160:271–222.

30. Montgomery SA, Rasmussen JG, Tanghol P. A 24-week study of 20 mg citalopram, 40 mg citalopram and placebo in the prevention of relapse of major depression. *Int Clin Psychopharmacol* 1993;8:181–188.

31. Montgomery SA, Reimitz PE, Zivkov M. Mirtazapine versus amitriptyline in the long-term treatment of depression: a double-blind placebo-controlled study. *Int Clin Psychopharmacol* 1998;13:63–73.

32. Feiger AD, Bielski RJ, Bremner J, et al. Double-blind placebo-substitution study of nefazodone in the prevention of relapse during continuation treatment of outpatients with major depression. *Int Clin Psychopharmacol* 1999;14:19–28.

33. Prien RF, Kupfer DJ, Mansky PA, et al. Drug therapy in the prevention of recurrences in unipolar and bipolar affective disorders. *Arch Gen Psychiatry* 1984;41:1096–1104.

34. Kupfer DJ, Frank E, Perel JM, et al. Five-year outcome for maintenance therapies in recurrent depression. *Arch Gen Psychiatry* 1992;49:769–773.

35. Keller MB, McCullough JP, Klein DN, et al. A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *N Engl J Med* 2000;342:1462–7140.

36. Kocsis JH, Friedman RA, Markowitz JC, et al. Maintenance therapy for chronic depression. A controlled clinical trial of desipramine. *Arch Gen Psychiatry* 1996;53:769–774.

37. Keller MB, Kocsis JH, Thase ME, et al. Maintenance phase efficacy of sertraline for chronic depression: a randomized controlled trial. *JAMA* 1998;280:1665–1672.

38. Kocsis JH, Frances AJ, Voss C, et al. Imipramine treatment for chronic depression. *Arch Gen Psychiatry* 1985;45:253–257.

39. Stewart JW, McGrath PJ, Quitkin FM, et al. Chronic depression: response to placebo, imipramine and phenelzine. *J Clin Psychopharmacol* 1993;13:391–396.

40. Stewart JW, Quitkin FM, Liebowitz MR, et al. Efficacy of desipramine in depressed outpatients. Response according to research diagnosis criteria diagnoses and severity of illness. *Arch Gen Psychiatry* 1983;40:202–207.

41. Hellerstein DJ, Yanowitch P, Rosenthal J, et al. A randomized double-blind study of fluoxetine versus placebo in the treatment of dysthymia. *Am J Psychiatry* 1993;150:1169–1175.

42. Vanelle JM. Controlled efficacy study of fluoxetine in dysthymia. *Br J Psychiatry* 1997;170:345–350.

43. Ravindran AV, Bialik RJ, Lapierre YD. Therapeutic efficacy of specific serotonin reuptake inhibitors (SSRIs) in dysthymia. *Can J Psychiatry* 1994;39:21–26.

44. Thase ME, Fava M, Halbreich U, et al. A placebo-controlled, randomized clinical trial comparing sertraline and imipramine for the treatment of dysthymia. *Arch Gen Psychiatry* 1996;53:777–784.

45. Keller MB, Harrison W, Fawcett JA, et al. Treatment of chronic depression with sertraline or imipramine: preliminary blinded response rates and high rates of undertreatment in the community. *Psychopharmacol Bull* 1995;31:205–212.

46. Bakish D, Ravindran A, Hooper C, et al. Psychopharmacological treatment response of patients with a DSM-III diagnosis of dysthymia disorder. *Psychopharmacol Bull* 1994;30:53–59.

47. Botte J, Evrard JL, Gilles C, et al. Controlled comparison of RO-11-1163 (moclobemide) and placebo in the treatment of depression. *Acta Psychiatry Belg* 1992;92:355–369.

48. Versiani M, Amrein R, Stabl M. Moclobemide and imipramine in chronic depression (dysthymia): an international double-blind, placebo-controlled trial. International Collaborative Study Group. *Int Clin Psychopharmacol* 1997;12:183–193.

49. Boyer P, Lecrubier Y. Atypical antipsychotic drugs in dysthymia: placebo controlled studies of amisulpride versus imipramine, versus amineptine. *Eur Psychiatry* 1996;11(suppl 3):135S–140S.

50. Smeraldi E. Amisulpride versus fluoxetine in patients with dysthymia or major depression in partial remission: a double-blind, comparative study. *J Affect Disord* 1998;48:47–56.

70

Risk Factors for Major Depression and Bipolar Disorder

Robert M. A. Hirschfeld

Myrna M. Weissman

Robert M. A. Hirschfeld: Department of Psychiatry and Behavioral Sciences, University of Texas Medical Branch, Galveston, Texas.

Myrna M. Weissman: College of Physicians and Surgeons of Columbia University, New York, New York.

Epidemiology is the study of the distribution of diseases and disorders in human populations and the variation in these distributions in different population subgroups. An observation that a disease is higher in one group or another helps to identify risk factors or correlates of these high rates whose alteration will interrupt the causal sequence that produces the disorder. Epidemiologic methods have been grouped into descriptive studies, which provide basic estimates of rates and their variation or increased risk in a population; analytic studies, which explore the variations in rates among different groups and identify risk factors; and experimental studies, which test an association between a risk factor and a disorder and seek to control or reduce the occurrence by controlling the risk factor.

Epidemiologic methods used in psychiatry are identical to those used in other branches of medicine. In psychiatry there are few clearly defined and modifiable risk factors, so studies have focused on largely unmodifiable ones such as age and gender, although even these risk factors can be useful in providing a clue to etiology.

The evidence of risk factors for depression came primarily from descriptive (large-scale epidemiologic) studies and analytic (family and high-risk offspring) studies. The former is useful as a first step because the samples include subjects regardless of treatment and thus are unbiased. The latter are useful as they include control groups and can be used to calculate relative risks.

This chapter reviews the current evidence for risk factors associated with major depression (MD) and bipolar disorder primarily based on cross-national community surveys and some family studies.

MAJOR DEPRESSION

Part of "70 - Risk Factors for Major Depression and Bipolar Disorder "

Prevalence

Data on prevalence of unipolar MD based on epidemiologic community surveys using the same diagnostic assessment, the Diagnostic Interview Schedule (DIS), are now available from different parts of the world. These population-based epidemiologic studies were conducted in the 1980s, and a cross-national collaboration was formed to analyze the data together in a standardized way. Ten countries across the world, including North America, Europe, Asia, and New Zealand, participated. These data provide the first information on cross-national rates for risk factors using the same methods. The lifetime prevalence rates of MD range from 1.5 per 100 adults in Taiwan to 19.0 per 100 in Lebanon. The results showed considerable variation in rates, but consistency in sex differences and age of onset.

In the United States, the National Institute of Mental Health (NIMH) Epidemiologic Catchment Area (ECA) study conducted in 1980 reported an overall lifetime prevalence of 5.2 per 100 and a 1-year prevalence of 3.0 per 100, which was somewhat lower than the rate reported by other studies across the world (1). In the National Comorbidity Survey (NCS) conducted a decade later in the United States, a substantially higher lifetime prevalence of MD was reported—17.1 (2). All prevalence rates have been published individually, but for the purpose of comparison between the countries they have been standardized to the U.S. age and sex distribution (Table 70.1).

Country	Lifetime Rate/100			F/M Ratio	Mean Age at Onset
	Overall	Females	Males		
United States (ECA, 1980)	5.2	7.4	2.8	2.6	25.6
United States (NCS, 1990)	17.1	21.3	12.7	1.7	23.8
Canada	9.6	12.3	6.8	1.9	24.8
Puerto Rico	4.3	5.5	3.1	1.8	29.5
France	16.4	21.9	10.5	2.1	29.2
Germany	9.2	13.5	4.4	3.1	29.7
Italy	12.4	18.1	6.1	3.0	34.8
Lebanon	19.0	23.1	14.7	1.6	25.2
Taiwan (Taipei)	1.5	1.8	1.1	1.6	29.3
Korea (Seoul)	2.9	3.8	1.9	2.0	29.3
New Zealand	11.6	15.5	7.5	2.1	27.3

CIDI, Composite International Diagnostic Interview; DIS, Diagnostic Interview Schedule; ECA, Epidemiologic Catchment Area Study; NCS, National Comorbidity Survey.
[a]Rates have been standardized to the U.S. age and sex distribution (adapted from ref. 1). All studies used the DIS with the exception of the NCS, which used the CIDI.

TABLE 70.1. CROSS-NATIONAL LIFETIME PREVALENCE OF MAJOR DEPRESSION[a]

The disparity between the rates reported in the ECA and the NCS has generated considerable controversy and discussion. Whether the differences are real (reflecting a substantial change in the prevalence rate over the decade) or artifactual (due to differences in methodology) has prompted careful examination. It is now believed that the difference is due to methodology (e.g., different diagnostic assessment, sample age, and size), and not to a true increase in the rate over the decade.

Gender

Despite the variation in rates, the most consistent finding in the cross-national studies and the two U.S. studies is the increased rate of MD in women (Table 70.1). The reasons for this disparity are not clear, but the disparity is also found in clinical studies. Interestingly, prior to puberty there are no sex differences in rates of depression. However, following puberty there is a dramatic shift in the prevalence rates, with a twofold increase in the prevalence of depression among women compared to men. A higher risk of depression in women is probably accounted for primarily by the higher risk of first onset in women. A series of analyses of the NCS data shows that there is little difference in the probability of acute recurrence in women and in men with a history of depression (3). Many theories, biological, psychosocial, and artifactual, attempt to explain this dramatic increase in the prevalence of depression among women, but none is fully satisfactory.

Age Of Onset And Secular Changes

The age of first onset of MD is fairly consistent across studies (Table 70.1). Of the ten major population-based epidemiologic studies reported by Weissman et al. (1), eight reported the age of onset between 25 and 30 years. In the NCS, the age of onset was 24 (3). Although there is consistency across studies regarding the age of onset, there is some evidence that the age of onset of depression has decreased over the last half century (4). In 1985, using the data from the NIMH Collaborative Program on the Psychobiology of Depression, the cumulative probability of a first episode of MD in female relatives of patients with MD by age 30 was less than 10% in individuals born before 1929. This rate doubled in cohorts born between 1930 and 1949, and among those born after 1950 the rate skyrocketed to 60%. The rate in males also increased in younger cohorts, but not nearly as dramatically as in women (Fig. 70.1 and Fig. 70.2).

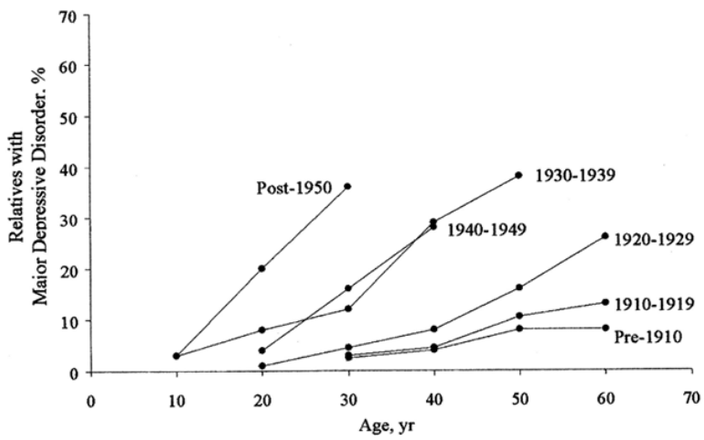


FIGURE 70.1. Cumulative probability of diagnosable major depressive disorder in male relatives by birth cohort. (Adapted from Klerman GL, Lavori PW, Rice J, et al. Birth-cohort trends in rates of major depressive disorder among relatives of patients with affective disorder. *Arch Gen Psychiatry* 1985;42(7):689–693.)

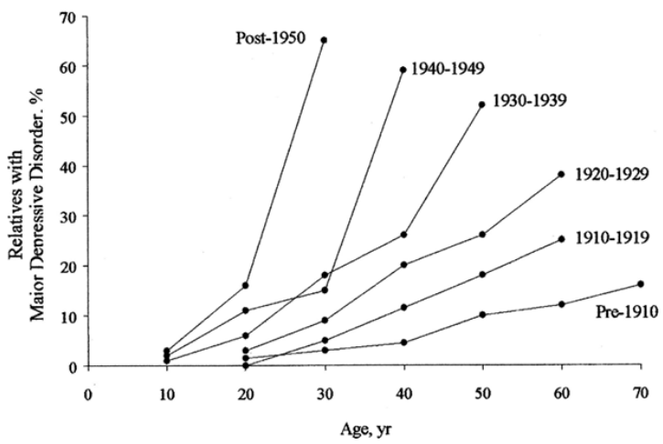


FIGURE 70.2. Cumulative probability of diagnosable major depressive disorder in female relatives by birth cohort. (Adapted from Klerman GL, Lavori PW, Rice J, et al. Birth-cohort trends in rates of major depressive disorder among relatives of patients with affective disorder. *Arch Gen Psychiatry* 1985;42(7):689–693.)

An analysis of the ECA data by Wickramaratne and colleagues (5) showed an increase in the rate of MD in the cohort born between 1935 and 1945. The rates for females stabilized after this increase. However, rates for males continued to rise in the cohort born between 1945 and 1954, and then decreased in the most recent cohort of the study,

those born between 1955 and 1964. In the NCS (2), lifetime risk of MD increased consistently through increasingly younger cohorts. In the youngest cohort (born between 1966 and 1975) there was a substantial increase in the risk of first onset in early adulthood compared to all other cohorts. The finding was true for both men and women.

With regard to prevalence across adulthood in the NCS, 30-day prevalence was fairly constant between the ages of 15 and 44, but dropped by nearly half in the decade between 45 and 54. As has already been mentioned, among women the prevalence was very high, 8.2% in the 15 to 24 cohort, and then dropped off to between 4% and 6% in subsequent decades, without a reduction as they age (6).

Marital Status

Marital status has been found to be highly associated with onset and prevalence of depression, but not with treatment outcome. In the ECA and the NCS, married and never-married persons were found to have lower rates of depression than those divorced, separated, and widowed. For example, in the ECA study, divorced and separated individuals had over a twofold increase compared to those married and never married (Table 70.2). Two countries with the lowest rate of MD in the ECA study—Korea and Taiwan—also had the lowest divorce and separation rate. Although Beirut also has a low rate of separation/divorce but a high rate of depression, the increased rate of depression may be more likely attributed to the civil war in the country that occurred around the same time as the study (1). Divorce and separation also increased the likelihood of the first depressive episode (7). The direction of the association is not clear.

	Marital Status		Lifetime Rate/100 Major Depression		
Country	Separated/ Divorced (%)	Married (%)	Separated/ Divorced (%)	Married (%)	OR (95% CI)[b]
United States	12.2	57.3	11.0	4.3	2.5 (2.0–3.2)
Canada	11.4	59.1	23.8	7.6	4.1 (2.9–5.7)
Puerto Rico	9.6	59.0	7.7	3.8	1.9 (0.9–4.3)
France	6.1	75.5	39.6	14.2	3.8 (2.4–5.4)
Germany[c]	9.5	78.1	18.4	7.8	2.2 (0.8–6.1)
Lebanon[d]	0.3	50.6	100	19.4	NA
Taiwan (Taipei)	3.7	71.9	3.8	1.2	2.9 (1.6–5.5)
Korea (Seoul)	0.7	69.6	5.3	3.0	1.6 (0.3–7.3)
New Zealand	9.4	61.4	26.5	9.8	3.5 (2.1–5.9)

NA, not available.
[a]Data standardized to U.S. age and sex distribution (adapted from ref. 1).
[b]Odds ratios (ORs) compare the rate of major depression in separated or divorced vs. married and are standardized to U.S. age and sex distribution and adjusted for age and sex within each site. CI, confidence interval.
[c]Data from former Federal Republic of Germany (West Germany) based on ages 26 to 64 years.
[d]Both cases of major depression in separated or divorced study participants (2/2).

TABLE 70.2. LIFETIME RATE PER 100 OF MAJOR DEPRESSION BY MARITAL STATUS: AGES 18 TO 64 YEARS[a]

Social Class

In the NCS, the odds ratios for MD were significantly higher for those individuals earning less than $20,000 a year, and declined as income increased (6). With regard to employment status, homemakers had a very high risk of MD (odds ratios 2.8), but it is not clear whether the employment status has been adjusted for sex. Blazer et al. (6) reported no significant association in adjusted odds ratios for

household income. In the ECA study, no association was found between socioeconomic status and MD. Although these findings are not conclusive, they suggest that socioeconomic status is not a strong risk factor for depression.

Race And Ethnicity

The prevalence of MD does not vary significantly with race and ethnicity, and in most cases can be explained by socioeconomic and educational factors. Both major epidemiologic studies (ECA and NCS) controlled for socioeconomic status and education, and both showed that race was not a significant predictor of MD.

In the NCS the lifetime prevalence of MD was lower overall among African-American subjects, except for African-Americans 35 to 44 years of age. The highest prevalence, however, was in African-American women 35 to 44 years of age. No significant association was found in adjusted odds ratios for race/ethnicity in the study, which is consistent with the ECA study. The 10-country study (1) found the lowest rate of MD in the two Asian countries included—Korea and Taiwan. Neither the ECA nor the NCS studies had sufficient Asian samples to determine whether the risk increases for Asians living in the U.S.

Familial Factors

Methods used to assess possible heritability of MD are family, twin, and adoption studies. Although these studies are suggestive of a genetic vulnerability to MD, their design does not allow identifying the genetic variables involved in the mode of transmission. More direct genetic strategies include genetic linkage and association studies.

Family Studies

Family studies assess the rate of an illness in relatives of patients with the illness as compared with the rate in a control group. The greater the difference in rates, the more likely it is that a genetic component exists. Gershon and colleagues (8) assessed rates of illness in relatives of unipolar probands and normal controls. They reported a lifetime risk of unipolar depression of 16.6% in the first-degree relatives of probands and 5.8% in those of normal controls (Table 70.3). Using these results and additional ones involving bipolar I, bipolar II, and schizoaffective disorder, the authors argued for a multifactorial genetic vulnerability model, ranging from unipolar to bipolar II, to bipolar I, and finally to schizoaffective disorder.

		Rate/100	
Form of Depression	**Study**	**Proband**	**Control**
Unipolar MDD	Gershon et al., 1982 (8)	16.6	5.8
Mild depression	Weissman et al., 1984 (9)	14.7	5.1
Severe depression		16.4–16.5	
Recurrent depression	Kupfer et al., 1989 (11)	20.7	NA
Single-episode, late onset	Bland et al., 1986 (10)	3.4	NA
Recurrent depression, late onset		8.2	
Single-episode, early onset		7.5	
Recurrent depression, early onset		17.4	
Early onset (age <20)	Weissman et al., 1984 (12)	24.2	4.8

MDD, major depressive disorder.

TABLE 70.3. FAMILY STUDIES OF MAJOR DEPRESSION

In a family study of 335 probands and 2,003 first-degree relatives, Weissman and colleagues (9) reported a lifetime rate of MD of 14.7 per 100 in the relatives of probands with mild depression, and a rate of 16.4 to 16.5 per 100 in the relatives of probands with severe depression, whereas the rate of MD in the relatives of control subjects was 5.1 per 100 (comparable to the rate in the general population). The distinction between mild and severe depression was made on the basis of hospitalization. Persons hospitalized for more than 5 days for a depressive episode met the criteria for severe depression, whereas persons with mild depression were those never hospitalized for depression.

The risk of depression is higher among the relatives of probands with early-onset recurrent MD. Bland et al. (10) reported a morbidity rate of 17.4% in relatives of probands with early-onset recurrent unipolar depression. This contrasted with 3.4% rate in relatives of probands with a single episode of depression and late age of onset. A study of 179 probands with recurrent depression found a morbid risk of 20.7% for nonbipolar depression (11). Consistent with Bland et al.'s study, relatives of probands with early-onset recurrent unipolar depression (before the age of 20) had a significantly higher risk of depression than relatives of late-onset probands (11). Similar results were obtained by Weissman et al. (12). The rates of MD were highest (24.2

per 100) in the relatives of probands with early age of onset (younger than 20) and decreased to 7.6 per 100 in relatives of probands with late age of onset (over 40).

These findings have formed the rationale for an NIMH-sponsored multisite genetic sib-pair study of early-onset recurrent MD. Nearly 1,000 sib pairs will be diagnosed and DNA made available to the scientific community for genetic study.

A variant of family studies are studies of offspring of depressed parents. These studies consistently show a threefold increased risk of MD in the offspring of depressed parents that is persistent as the offspring age and into the next generation (13 ,14 and 15). Interestingly, when parental divorce, parent-child bonding, and relationships were examined only, parental depression remained the only significant risk factor in the offspring of depressed parents. In the offspring of nondepressed parents, these factors were predictive of offspring depression. However, the rates of depression were low (16).

Twin Studies

In the early 1990s Kendler and his colleagues (17 ,18) reported a number of analyses of the data collected from female-female twin pairs identified through the population-based Virginia Twin Registry. The proband-wise concordance was 37.3% in monozygotic (MZ) twins, and 23.9% in dizygotic (DZ) twins; the estimated heritability of liability to MD was 46% (17). However, when the model was corrected for unreliability of measurement, the heritability of liability to MD was estimated at 71% (18). In a more recent study that included female-female and male-male twin pairs identified through the same registry, Kendler and Prescott (19) found proband-wise concordance of 31.1% in MZ and 25.1% in DZ male twin pairs, and 47.6% compared to 42.6% in female twin pairs. These differences were statistically significant. The estimated heritability in male twins was 39%, comparable to the only other general population twin study of lifetime MD in men (20), who reported an estimated heritability of 36%. The estimated heritability in female twin pairs was 42% as measured in the previous report by the same authors (21). These rates are approximately twice those in the general population, reflecting a genetic component to MD.

In a hospital-based twin registry study conducted by McGuffin et al. (22), the proband-wise concordance was 46% in MZ and 20% in DZ twins, a statistically significant difference. The estimates of heritability were between 48% and 75%, based on the assumed population risk.

Personality

In contrast to most of the previously discussed risk factors that are fixed and clearly antedate the onset of depression (e.g., age and sex), personality is not fixed and in fact may interact with depression. Methodologic issues have confounded the study of this issue. For example, when depressed, patients may not provide valid reports of their premorbid functioning (23). In a comparison of patients' self-report personality measures first during their depressive episode and again following complete recovery 1 year later, the depressed state significantly influenced assessment of emotional strength, interpersonal dependency, and extraversion (24).

The most methodologically clear approach is to evaluate persons before they develop a depressive disorder. Hirschfeld et al. (25) conducted a premorbid personality assessment in a large sample at risk for the development of MD. The personality features most predictive of first onset of depression among middle-aged adults (age 31 to 41) were decreased emotional strength and increased interpersonal dependency. Among younger adults (age 17 to 30) no personality features were associated with subsequent MD. This may reflect that psychosocial factors become more important in late-onset depression, whereas genetic factors are more important in early-onset depression.

Life Events

Clinicians have long described a relationship between life events (particularly adverse interpersonal events) and the onset of depressive episodes. A series of studies conducted between the middle 1960s and 1990 using standardized instruments and diagnostic procedures did find corroboration of this relationship (26). Jenaway and Paykel (26) report that events involving loss (e.g., divorce, death) and threat of separation are associated with depression. There is little specificity to this relationship, as these events precede other illnesses as well.

Kendler et al. (27) investigated how genetic liability to MD and stressful life events interact in the etiology of MD in a study of 2,164 female twins. They found that the incidence of depression increased significantly in the month of occurrence of 13 stressful events. Four of the events termed "severe"—death of a close relative, assault, serious marital problems, and divorce/breakup—predicted the incidence of MD with the odds ratios of greater than 10. Genetic liability also had a significant impact on the onset of MD. The lowest probability of onset of MD was found in individuals with the lowest genetic liability (MZ, co-twin unaffected) for those exposed and unexposed to a severe event. Probabilities of the onset of MD were substantially higher in individuals with the highest genetic liability (MZ, co-twin affected). Kendler et al. concluded that genetic factors influenced the risk of MD in part by influencing the susceptibility of individuals to the depressive effect of life events.

Early Trauma

Trauma in early life has long been considered an etiologic factor in the pathogenesis of depression by clinical theorists.

This issue has been examined in a series of studies by Brown's group (28 ,29) at the University of London. In one study of 286 working-class mothers in England, 9% reported childhood sexual abuse (28). Of these 64% suffered a depression during the period of study. In a subsequent prospective study the same investigators (29) found that childhood neglect or abuse was strongly associated with early-onset (before age 20) depression. McCauley and colleagues (30) also found that childhood physical or sexual abuse was predictive of a variety of adult afflictions, including depression. Heim and colleagues (31) found that childhood abuse causes persistent hypothalamic-pituitary-adrenocortical (HPA) hyperactivity in adulthood, which is consistent with depression. In a review of this area, Kessler and colleagues (32) concluded that adverse childhood events were associated with early-onset depression. Caution is in order in interpreting these findings because multiple-year retrospective withdrawal can be flawed, and controlling for all relevant covariates (such as family history of depression) is often impossible.

General Medical Illness

Increased rates of depression have been reported among patients with several general medical illnesses. Among these are cardiovascular disease, AIDS, respiratory disorders, cancer, and several neurologic conditions (Parkinson's disease and stroke in particular) (33 ,34). The relationship is most striking in cardiovascular disease; 13 prospective studies using structured clinical diagnostic interviews evaluating antecedent depression on subsequent cardiovascular disease were reviewed by Musselman et al. (35). Almost all the studies found a strong association between depression and subsequent cardiovascular morbidity and mortality. Several important physiologic aspects of depression may account for this association. They include HPA dysregulation, sympathomedullary hyperactivity, and diminished heart rate variability (35).

BIPOLAR DISORDER

Part of "70 - Risk Factors for Major Depression and Bipolar Disorder "

Prevalence

The lifetime prevalence rate of bipolar I disorder reported in the NCS (2) was 1.6%, and the ECA study conducted in the United States reported a lifetime rate of 0.9% (36). In contrast to MD, the prevalence rates were remarkably consistent among countries (Canada, Finland, France, Germany, Hong Kong, Italy, Korea, New Zealand, Puerto Rico, Taiwan, and the United States), and there was little variation in rates by gender (Table 70.4). The rates of bipolar spectrum disorder (bipolar I, bipolar II, cyclothymia, and others) may be considerably higher than bipolar I alone (37). The prevalence rates reported in 11 studies of bipolar II disorder were between 0.3% and 3% (38). The lifetime prevalence rates for bipolar spectrum disorder were reported between 3% and 6.5% (38).

	Lifetime Rates/100				
	Overall	Females	Males	F/M Ratio	Mean Age at Onset
United States (ECA 1980)	0.9	1.0	0.8	1.2	18.1
United States (NCS 1990)	1.6	1.7	1.6	1.1	21.0
Canada, Edmonton	0.6	0.5	0.7	0.7	17.1
Puerto Rico	0.6	0.5	0.8	0.6	27.2
Germany[b]	0.5	1.0	0.0	NA	29.0
Taiwan	0.3	0.3	0.3	1.0	22.5
Korea	0.4	0.2	0.6	0.3	23.0
New Zealand	1.5	1.2	1.7	0.7	18.2

NA, not applicable.
[a]Rates have been standardized to the U.S. age and sex distribution (adapted from ref. 1).
[b]In Germany (former Federal Republic of Germany) the one study participant in Munich with bipolar disorder was female.

TABLE 70.4. LIFETIME PREVALENCE OF BIPOLAR DISORDER

Gender and Age

In contrast to MD, epidemiologic data show no significant sex differences in rates of bipolar disorder (Table 70.4). No sex differences were reported in the ECA study (39). Similarly, the NCS reported no sex differences in either the mean number of total episodes of mania and depression combined (75.4 in men and 60.4 in women), and not a large sex difference in median number of total episodes (38 in men and 30 in women) (40). Additionally, no differences in male to female ratios of bipolar disorder were found internationally (1). Mean age of onset for bipolar depression was generally younger than for MD, ranging from 18 to 27 years of age in different countries (1). The NCS places the age of onset at 21.

Other Sociodemographic Variables

Historically bipolar disorder was thought to be more frequent among higher socioeconomic classes, and there were some pilot data to support this (e.g., 41). However, population-based studies over the last two decades have not replicated this finding. In the ECA, occupation, income, and education were not found to influence prevalence (42). In the NCS, those with bipolar I disorder were more likely to have annual incomes less than $20,000. The notion of higher rates of bipolar disorder among higher income levels was probably due to misdiagnosis of bipolar patients as schizophrenics in lower income groups.

In the ECA, bipolar disorder was much less frequent among married people, as contrasted with divorced or never-married people. Bipolar disorder was more prevalent among those with multiple divorces (42). In the NCS, bipolar disorder was more frequent among unmarried, poorly educated people.

Genetic and Familial Factors

In a review of the eight family studies of bipolar I disorder that included a control group, a metaanalysis (43) showed that bipolar I disorder was seven times more likely among relatives of bipolar I probands than of controls. These studies also demonstrate an increased risk of MD in relatives of bipolar probands, although the relative risk is lower than for bipolar I. Early age of onset and number of ill relatives increases the risk of illness in relatives, but other variables (e.g., type of relative) appear not to affect it (43).

In a pooled analysis of the six twin studies of bipolar I disorder conducted between 1962 and 1999, Craddock and Jones (43) calculated an estimate of proband-wise concordance of 50%, although the authors believe this to be an underestimate. They believe it is probably around 60%. Integrating the results of family, twin, and adoption studies, Craddock and Jones conclude that there is a substantial genetic predisposition to bipolar disorder.

There have been a number of investigations aimed at determining the actual genes involved in bipolar illness. Attempts to demonstrate linkage to the X-chromosome, to color blindness, to chromosomes 4, 11, 18, and others have not been conclusive (43). The most likely explanation for lack of success is that these strategies assume a single genetic mode of inheritance for a complex multiple-gene interaction with environmental factors.

Environmental Factors

Although biological and genetic factors have long been known to play a major role in the etiology of bipolar disorder, psychosocial factors are gaining attention. In particular, many studies have identified the association between stressful life events and social rhythm disruptions and onset of recurrence. It is unlikely that psychosocial factors play a major role in the risk of first onset of bipolar disorder, but they may have an important role in increasing the risk of recurrence.

Chrono-biological disturbances have long been associated with bipolar disorder (44). Sleep-wake and other circadian rhythm disturbances are core symptoms of bipolar episodes for both manic and depressive episodes (e.g., insomnia, and decreased need for sleep). Some have theorized that disruption in social zeitgebers (i.e., social demands or tasks that set the biological clock by environmental events) can lead to instability in circadian rhythms, which can in turn trigger bipolar episodes (45). Malkoff-Schwartz and colleagues (46) found than severe social rhythm disruptions (e.g., returning from international trips, moving, losing a job) were associated with onsets of mania, but not depression, in bipolar patients.

DISCUSSION

Part of "70 - Risk Factors for Major Depression and Bipolar Disorder "

Beliefs and hypotheses about risk factors for depression (such as undue interpersonal dependency) emerged from clinicians treating individual patients and from studies done in psychiatric settings. Both suffer from sampling bias: those presenting for treatment are evaluated, but many factors in addition to the illness itself affect treatment seeking (e.g., income and psychological mindedness).

This chapter's introduction described the value of epidemiology in testing these hypotheses. The descriptive epidemiologic studies from around the world with their unbiased population samples have supported the validity of some of the hypotheses and have not supported others.

For MD there is strong evidence that women have twofold the prevalence of men, an age of onset between 25 and 30 years, an increase following separation and divorce, and an increase in families of those with MD. There are significant differences in lifetime rates around the world. There is little support for racial, ethnic (with the possible exception of Asian), and income factors. Personality factors are nonspecific. Early life trauma, particularly sexual abuse, is associated with early-onset depression in women. Increased rates of depression are found in several general medical illnesses. The association between depression and cardiovascular illness is strong, with depression predicting increased rates of morbidity and mortality among cardiovascular patients.

Analytic studies (such as twin studies) have suggested some interaction among genetic and environmental factors. The incidence of depression was increased when several life events (such as death of a close relative) occurred in the prior month in individuals with high genetic liability. The ongoing NIMH sib-pair study of early-onset depression will shed light on the possible genetic etiology of early-onset recurrent MD.

For bipolar disorder, there are substantial differences in

risk factors compared to MD. There is little variation in lifetime rates around the world (about 1%), and nearly equal prevalence in men and women. Bipolar spectrum is much more frequent, in the range of 3% to 6%. Age of onset is in the late teens and early twenties, earlier than in MD. A range of studies strongly support a genetic predisposition to bipolar disorder, but specific replicable genetic factors have yet to be demonstrated. A complex multiple gene interaction with environmental factors is most likely, a situation that is challenging to research and to isolate. Life events, especially disruptions in social zeitgebers, increase the likelihood of manic, but not depressive, recurrences in bipolar subjects.

An intriguing new area of interest is emerging in risk factors for depression—that of differences in consumption of fish oils around the world. Hibbeln (47) documented a strong negative correlation (-0.84) between the prevalence of MD and the annual apparent fish consumption per person in nine countries worldwide. He argued that this may be related to varying amounts of long-chain polyunsaturated fatty acids in diets in different cultures (48).

An experimental epidemiologic study is under way that will test genetic risk factors for bipolar disorder. Calabrese is prophylactically treating familially high-risk adolescents with a mood stabilizer in a longitudinal double-blind placebo-controlled study.

ACKNOWLEDGMENTS

Part of "70 - Risk Factors for Major Depression and Bipolar Disorder "

The authors would like to especially thank Lana A. Vornik, MSc., for her invaluable assistance in preparing this chapter.

Dr. Hirschfeld has received research support from Abbott Laboratories, Bristol-Myers Squibb, Organon, and Pfizer. He has served as a consultant or on an advisory board for Abbott Laboratories, Bristol-Myers Squibb, Glaxo Wellcome, Forest Laboratories, Eli Lilly & Company, Pfizer, SmithKline Beecham, Organon, Pharmacia & Upjohn, and Janssen Pharmaceutica. In addition, he has served on speakers' bureaus for the following companies: Abbott Laboratories, Bristol-Myers Squibb, Forest Laboratories, Eli Lilly & Company, Organon, SmithKline Beecham, and Parke-Davis.

REFERENCES

1. Weissman MM, Bland RC, Canino GJ, et al. Cross-national epidemiology of major depression and bipolar disorder. *JAMA* 1996;276(4):293–299.

2. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994;51(1):8–19.

3. Kessler RC, McGonagle KA, Swartz M, et al. Sex and depression in the National Comorbidity Survey. I: Lifetime prevalence, chronicity and recurrence. *J Affective Disord* 1993;29(2 and 3):85–96.

4. Klerman GL, Lavori PW, Rice J, et al. Birth-cohort trends in rates of major depressive disorder among relatives of patients with affective disorder. *Arch Gen Psychiatry* 1985;42(7):689–693.

5. Wickramaratne PJ, Weissman MM, Leaf PJ, et al. Age, period and cohort effects on the risk of major depression: results from five United States communities. *J Clin Epidemiol* 1989;42(4):333–343.

6. Blazer DG, Kessler RC, McGonagle KA, et al. The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. *Am J Psychiatry* 1994;151(7):979–986.

7. Coryell W, Endicott J, Keller M. Major depression in a nonclinical sample. Demographic and clinical risk factors for first onset. *Arch Gen Psychiatry* 1992;49(2):117–125.

8. Gershon ES, Hamovit J, Guroff JJ, et al. A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. *Arch Gen Psychiatry* 1982;39(10):1157–1167.

9. Weissman MM, Gershon ES, Kidd KK, et al. Psychiatric disorders in the relatives of probands with affective disorders. The Yale University—National Institute of Mental Health Collaborative Study. *Arch Gen Psychiatry* 1984;41(1):13–21.

10. Bland RC, Newman SC, Orn H. Recurrent and nonrecurrent depression. A family study. *Arch Gen Psychiatry* 1986;43(11):1085–1089.

11. Kupfer DJ, Frank E, Carpenter LL, et al. Family history in recurrent depression. *J Affective Disord* 1989;17(2):113–119.

12. Weissman MM, Wickramaratne P, Merikangas KR, et al. Onset of major depression in early adulthood. Increased familial loading and specificity. *Arch Gen Psychiatry* 1984;41(12):1136–1143.

13. Hammen C, Burge D, Burney E, et al. Longitudinal study of diagnoses in children of women with unipolar and bipolar affective disorder. *Arch Gen Psychiatry* 1990;47(12):1112–1117.

14. Warner V, Weissman MM, Mufson L, et al. Grandparents, parents, and grandchildren at high risk for depression: a three-generation study. *J Am Acad Child Adolesc Psychiatry* 1999;38(3):289–296.

15. Weissman MM, Warner V, Wickramaratne P, et al. Offspring of depressed parents. 10 Years later. *Arch Gen Psychiatry* 1997;54(10):932–940.

16. Fendrich M, Warner V, Weissman MM. Family risk factors, parental depression and childhood psychopathology. *Dev Psychol* 1990;26:40–50.

17. Kendler KS, Neale MC, Kessler RC, et al. The lifetime history of major depression in women. Reliability of diagnosis and heritability. *Arch Gen Psychiatry* 1993;50(11):863–870.

18. Kendler KS, Neale MC, Kessler RC, et al. A longitudinal twin study of 1-year prevalence of major depression in women. *Arch Gen Psychiatry* 1993;50(11):843–852.

19. Kendler KS, Prescott CA. A population-based twin study of lifetime major depression in men and women. *Arch Gen Psychiatry* 1999;56(1):39–44.

20. Lyons MJ, Eisen SA, Goldberg J, et al. A registry-based twin study of depression in men. *Arch Gen Psychiatry* 1998;55(5):468–472.

21. Kendler KS, Neale MC, Kessler RC, et al. A population-based twin study of major depression in women. The impact of varying definitions of illness. *Arch Gen Psychiatry* 1992;49(4):257–266.

22. McGuffin P, Katz R, Watkins S, et al. A hospital-based twin registry study of the heritability of DSM-IV unipolar depression. *Arch Gen Psychiatry* 1996;53:129–136.

23. Hirschfeld RM, Shea MT, Holzer CE III. Personality dysfunction and depression. In: Honig A, Van Praag HM, eds. *Depression: neurobiological, psychopathological and therapeutic advances*. New York: Wiley, 1997:327–341.

24. Hirschfeld RM, Klerman GL, Clayton PJ, et al. Personality and depression. Empirical findings. *Arch Gen Psychiatry* 1983;40(9):993–998.

25. Hirschfeld RM, Klerman GL, Lavori P, et al. Premorbid personality assessments of first onset of major depression. *Arch Gen Psychiatry* 1989;46(4):345–350.

26. Jenaway A, Paykel ES. Life events and depression. In: Honig A, Van Praag HM, eds. *Depression: neurobiological, psychopathological and therapeutic advances*. New York: Wiley, 1997:279–295.

27. Kendler KS, Kessler RC, Walters EE, et al. Stressful life events, genetic liability, and onset of an episode of major depression in women. *Am J Psychiatry* 1995;152(6):833–842.

28. Bifulco A, Brown GW, Adler Z. Early sexual abuse and clinical depression in adult life. *Br J Psychiatry* 1991;159:115–122.

29. Bifulco A, Brown GW, Moran P, et al. Predicting depression in women: the role of past and present vulnerability. *Psychol Med* 1998;28(1):39–50.

30. McCauley J, Kern DE, Kolodner K, et al. Clinical characteristics of women with a history of childhood abuse: unhealed wounds. *JAMA* 1997;277(17):1362–1368.

31. Heim C, Newport DJ, Heit S, et al. Pituitary-adrenal and autonomic responses to stress in women after sexual and physical abuse in childhood. *JAMA* 2000;284(5):592–597.

32. Kessler RC, Gillis-Light J, Magee WJ, et al. Childhood adversity and adult psychopathology. In: Gotlib IH, Weaton B, eds. *Trajectories and turning points: stress and adversity over the life course*. New York: Cambridge University Press, 1996.

33. Evans DL, Staab JP, Petitto JM, et al. Depression in the medical setting: biopsychological interactions and treatment considerations. *J Clin Psychiatry* 1999;60(suppl 4):40–55.

34. Stoner SC, Marken PA, Sommi RW. Psychiatric comorbidity and medical illness. *Med Update Psychiatrists* 1998;3(3):64–70.

35. Musselman DL, Evans DL, Nemeroff CB. The relationship of depression to cardiovascular disease: epidemiology, biology, and treatment. *Arch Gen Psychiatry* 1998;55(7):580–592.

36. Weissman MM, Bruce ML, Leaf PJ, et al. Affective disorders. In: Robins L, Regier D, eds. *Affective disorders*. New York: Free Press, 1990:53–80.

37. Hirschfeld RMA, Williams JBW, Spitzer RL, et al. Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. *Am J Psychiatry* 2000;157:1873–1875.

38. Angst J. The emerging epidemiology of hypomania and bipolar II disorder. *J Affective Dis* 1998;50(2 and 3):143–151.

39. Weissman MM, Leaf PJ, Tischler GL, et al. Affective disorders in five United States communities. *Psychol Med* 1988;18(1):141–153.

40. Kessler RC, Rubinow DR, Holmes C, et al. The epidemiology of DSM-III-R bipolar I disorder in a general population survey. *Psychol Med* 1997;27(5):1079–1089.

41. Weissman MM, Myers JK. Affective disorders in a US urban community: the use of research diagnostic criteria in an epidemiological survey. *Arch Gen Psychiatry* 1978;35(11):1304–1311.

42. Weissman MM, Kidd KK, Prusoff BA. Variability in rates of affective disorders in relatives of depressed and normal probands. *Arch Gen Psychiatry* 1982;39(12):1397–1403.

43. Craddock N, Jones I. Genetics of bipolar disorder. *J Med Genet* 1999;36(8):585–594.

44. Wehr TA, Wirz-Justice A, Goodwin FK, et al. Phase advance of the circadian sleep-wake cycle as an antidepressant. *Science* 1979;206(4419):710–713.

45. Ehlers CL, Frank E, Kupfer DJ. Social zeitgebers and biological rhythms. A unified approach to understanding the etiology of depression. *Arch Gen Psychiatry* 1988;45(10):948–952.

46. Malkoff-Schwartz S, Frank E, Anderson B, et al. Stressful life events and social rhythm disruption in the onset of manic and depressive bipolar episodes: a preliminary investigation. *Arch Gen Psychiatry* 1998;55(8):702–707.

47. Hibbeln JR. Fish consumption and major depression. *Lancet* 1998;351:1213.

48. Hibbeln JR, Salem N. Dietary polyunsaturated fats and depression: when cholesterol does not satisfy. *Am J Clin Nutr* 1995;62:1–9.

71

Bipolar Disorders: Review of Molecular Genetic Linkage Studies

Wade Berrettini

Wade Berrettini: University of Pennsylvania School of Medicine, Center for Neurobiology and Behavior, Philadelphia, Pennsylvania.

This chapter reviews evidence of the heritability of bipolar (BP) and recurrent unipolar (RUP) disorders. This evidence has accumulated through family, twin, adoption, linkage, and candidate gene studies. Both the twin and family studies suggest that unipolar and bipolar disorders share some fraction of genetic susceptibility. Data from the twin and family studies can be used for genetic counseling, and this will be discussed briefly from a clinical perspective.

Efforts to find susceptibility genes through linkage studies have yielded several confirmed regions of the genome where such genes will be found. These linkage studies will be discussed and summarized from methodologic perspectives.

Candidate gene approaches to BP and RUP disorders will be reviewed, with some suggestions for improving methods. A few of the most promising candidate genes will be noted.

Mechanisms on nonmendelian inheritance may be involved in the complex genetics of BP and RUP disorders. Imprinting, triplet repeat expansion, and mitochondrial inheritance are reviewed briefly as examples of nonmendelian mechanisms possibly involved in these disorders.

Finally, the implications of the Human Genome Project for future progress will be discussed.

GENETIC EPIDEMIOLOGY

Part of "71 - Bipolar Disorders: Review of Molecular Genetic Linkage Studies "

Family Studies

Family studies can answer three critical questions concerning the inheritance of a human phenomenon:

1. Is the phenomenon found more frequently among the biological relatives of an affected individual compared to biological relatives of unaffected persons? Alternatively, are relatives of an affected subject at increased risk for the disorder compared to relatives of control subjects?
2. What other phenomena (possibly genetically related) are also found more frequently among relatives of an affected individual? Alternatively, what other disorders (or clinical characteristics) may share a common genetic vulnerability with the phenomenon in question?
3. Can a specific mode of inheritance be discerned?

Family studies are executed as follows. A proband that (most likely) has the phenomenon in question is examined to determine its presence. This person's biologic relatives are similarly examined for its presence. Simultaneously, relatives of unaffected probands are examined in the same fashion for its presence. The importance of simultaneous examination of control families cannot be overestimated, as the risk of psychiatric disorders in the general population varies according to the diagnostic criteria and according to the judgment of specific interviewers and diagnosers. Thus, it is rarely acceptable to rely on data collected by others to estimate risk for a control population. The risk of a particular disorder can then be calculated for the relatives of both affected probands and control probands. If the disorder in question is heritable, a family study should show that relatives of a proband with the disorder are at significantly higher risk to show the disorder themselves, compared to relatives of control probands. Often, the risk for a certain class of relatives of affected probands is expressed as a ratio. Shown below is an example for siblings:

$$\text{Sibling relative risk} = \frac{\text{[Risk for siblings of affected probands]}}{\text{[Risk for siblings of control probands]}}$$

A family study is not the same as a “family history” study, in which the relatives are not directly examined, but information from the proband or other persons is used to establish the affections status of relatives. The reliability of the family history method is not as high as the family study

method. The discrepancy in reliability is a function of the phenomenon under study, but for psychiatric diseases the discrepancy is often great enough to render the family history method undesirable.

As with any disorder with a variable age of onset, the prevalence of illness among relatives must be corrected for the fraction of the age of risk that each relative has yet to live through. There are several ways of performing this correction. Commonly, using age-at-onset data from the relevant population, the number of relatives in a particular age decade is multiplied by the fraction of affected people who became ill by that decade of life. This product is known as the *bezugziffer*. A bezugziffer is calculated for each decade of life and the sum of them represents the total number of relatives at risk. When the total number of ill relatives is divided by this sum of bezugziffers, a morbid risk value (risk of developing the illness at some point in life) is determined for the relatives.

Family studies of bipolar disorder (BPD) show that a spectrum of mood disorders is found among the first-degree relatives of BPD probands: BP I disorder, BP II disorder with major depression (hypomania and RUP illness in the same person), and schizoaffective (SA) disorders and RUP depression, as described by multiple investigators (1 ,2 ,3 ,4 ,5 ,6 ,7 ,8 ,9 and 10). Nearly all family studies of BPD probands reveal that their biologic relatives are at increased risk for BPD, SA, and RUP diagnoses. These results suggest that RUP and SA diagnoses, within a familial context of BPD disorders, may be alternative phenotypic expressions of the same genetic susceptibility. Thus, this spectrum of mood disorders represents a reasonable BPD affection status model, in that it can be expected that some genetic susceptibility to BPDs may partially explain genetic susceptibility to SA and RUP disorders.

The first-degree relatives of RUP probands are at increased risk for BPD and RUP disorder (1 ,2). If the general population risk for BPDs is ˜1%, then the risk to first-degree relatives of BPD probands is ˜10%. Similarly the first-degree relatives of BPD probands are at increased risk of RUP disorders: if the general population risk if ˜8% (1 ,2), the RUP risk to first-degree relatives of BPD probands is ˜15% (1). Historically, SA disorder has been defined as a nosologic category that may describe psychotic affective disorders in some diagnostic systems. This makes estimates of risk of SA disorder among relatives of BPD probands difficult to compare across family studies.

Several family studies have also reported a higher risk of RUP illness among the first-degree relatives of RUP probands (2 ,4 ,11 ,12 ,13 and 14). The first-degree relatives of RUP probands also show significant increases in risk for BPD diagnoses (1 ,2). Thus, first-degree relatives of RUP and first-degree relatives of BPD probands are at increased risk for both RUP and BPD diagnoses. These data suggest that the two nosologic entities may share some genetic susceptibility and/or environmental risk. This tentative conclusion, which must await molecular studies, is supported by twin studies (see below).

Thus, family studies of BPD suggest that the first-degree relatives are at increased risk of BPD, SA disorder, and RUP disorder. Similarly, a review of SZ family studies reveals that the first-degree relatives of schizophrenia (SZ) probands are at increased risk of SZ, SA, and RUP disorders (10 ,15). Kendler et al. (16) describe an increased risk of psychotic affective illness among relatives of SZ probands. Despite numerous carefully conducted investigations, no family study of SZ reports increased risk for BPD among first-degree relatives of SZ probands. However, the first-degree relatives of SZ probands *and* the first-degree relatives of BP probands are at increased risk of SA and RUP disorders. The overlap in elevated risk of SA and RUP diagnoses is evident. Therefore, these family studies are consistent with *partial* overlap in familial susceptibility for BPD and SZ disorders. It will be instructive to determine whether putative BPD susceptibility loci are mapped to regions of the genome at which SZ susceptibility loci are thought to exist (see below).

Twin Studies

A phenomenon that is under genetic control should be more “concordant” (similar) in monozygoric (MZ) twins compared to dizygotic (DZ) twins. By comparing the concordance rate (how often the second member of a twin pair demonstrates the phenomenon in question when the first member has it) for MZ and DZ twin pairs, evidence of the genetic determination of a phenomenon can be obtained. There are two methods for calculating concordance rates in twin studies. “Pairwise concordance” involves a simple calculation of the percentage of twin pairs in which both members demonstrate the phenomenon in question. In “proband-wise concordance” every affected person is considered a proband (within a family, a proband is usually the first person who comes into contact with the investigators), independent of status within a twin pair, meaning that twin pairs in which both are affected are counted twice. Studies in which twin subjects are selected at random from a population most appropriately apply the pairwise method. However, in most twin studies of psychiatric diseases, the twins are not selected at random, but are selected because one member has the disease in question. Therefore, these studies most appropriately use the proband-wise concordance, although both types of analyses can yield similar results. The results of both types of analyses are often reported. Differences between the MZ and DZ concordances are usually larger when the proband-wise calculation is used.

If the variable being measured is genetically determined, then the concordance rate may be significantly higher in MZ twins compared to that for DZ twins. This suggests, but is by no means conclusive, that the variable in question may be heritable. The magnitude of the difference for the

MZ versus DZ concordance rates may provide an estimate of the heritability of the variable in question. Although there are several methods for calculating this heritability, one of the most simple is Holzinger's index:

$$\text{Heritability} = \frac{[\text{MZ Concordance (\%)} - \text{DZ Concordance (\%)}]}{[100 - \text{DZ Concordance (\%)}]}$$

When the MZ concordance rate is 100% and the DZ rate is 50%, the variable is a purely genetically determined phenomenon. More complex path modeling of heritability estimates from twin data is commonly used today (e.g., 16 ,17). MZ concordance rates lower than 100% suggest reduced penetrance (the gene is present but the disease is not expressed because of protective environmental events) or the presence of phenocopies (individuals appear to exhibit the variable in question but do not have the genetic diathesis for it).

Twin studies conducted over the past 70 years have indicated greater MZ twin concordance compared to DZ twin concordance for BPD and RUP disorders (for review see ref. 18). More recent twin studies (16 ,17 ,19), conducted with operationalized diagnostic criteria, validated semistructured interviews, and blinded assessments, confirm the earlier research, showing significantly greater MZ twin concordance. The MZ twin concordance rate (˜65%) indicates decreased penetrance of inherited susceptibility or the presence of phenocopies (nongenetic cases). Among MZ twin pairs concordant for mood disorder, when one twin has a BPD diagnosis, RUP illness is present among 20% of the ill co-twins (20 ,21). This suggests that BPD and RUP syndromes share some common genetic susceptibility factors. This result is nicely concordant with the family study results reviewed above, which demonstrate that the first-degree relatives of BPD and the first-degree relatives of RUP probands are at increased risk for both BPD and RUP disorders. The twin and family study data suggest that BPD and RUP disorders share genetic susceptibility.

This has some clinical implications. For example, these genetic epidemiologic data suggest that RUP individuals who have lithium-responsive BPD relatives should be treated with lithium for prophylaxis of RUP, if maintenance treatment with an antidepressant is not successful.

Adoption Studies

Most adoption studies proceed through identification of affected probands that have been adopted early in life. Similarly, a control group of unaffected, adopted probands is identified. Risk for the disorder is compared in four groups of relatives: the adoptive and biological relatives of affected adoptees and the adoptive and biological relatives of unaffected adoptees. For partially genetic phenomena, there will be an increased risk among the biological relatives of affected probands, compared to the other three groups of relatives. Alternatively, risk for illness in adopted-away children of ill parents can be compared with risk for illness in adopted-away children of well parents.

In the “cross-fostering” design, researchers ascertain two groups of adopted-away children—those of ill parents and those of unaffected parents—both of whom are adopted away early in life and are raised by well parents. Researchers also ascertain two additional groups of adopted-away children—those of ill parents and those of unaffected parents—both of whom are similarly adopted away early in life and raised by affected adults. If the presence of the illness in the family environment increases risk for development of the disease, then the risk for children raised by affected parents will be greater than the risk for children raised by unaffected parents. If only genetic factors are important in the pathogenesis, then children with ill biologic parents will be at increased risk for illness, independent of the presence of illness in the adoptive parents. These adoption studies cannot exclude intrauterine or perinatal events, which may yield results similar to those of genetic diseases.

Mendlewicz and Rainer (22) reported a controlled adoption study of BPD probands, including a control group of probands with poliomyelitis. The biological relatives of the BPD probands had a 31% risk for BPD or RUP disorders, compared to 2% risk in the relatives of the control probands. The risk of affective disorder in biological relatives of adopted BPD patients was similar to the risk in relatives of BPD patients who were not adopted away (26%). Adoptive relatives do not show increased risk compared to relatives of control probands.

Wender et al. (23) and Cadoret (24) studied RUP and BPD probands. Although evidence of genetic susceptibility was found, *adoptive* relatives of affective probands had a tendency to excess affective illness themselves, compared with the adoptive relatives of controls. Von Knorring et al. (25) did not find concordance in psychopathology between adoptees and biological relatives when examining the records of 56 adoptees with unipolar (UP) disorders. Sample size may have limited the conclusions of this study. It is also possible that heritable factors may be more prominent for BPD disorders, compared to RUP disorders.

Genetic Counseling

Frequently, patients with BPD or RUP disorder are aware of the genetic component to these illnesses, and, quite naturally, they are concerned about risks of illness to other members in the family, most often children. Clinician responses to these requests for genetic counseling properly use risk estimates derived from controlled family studies (e.g., 1,2). Family studies of BP and RUP illness have established the risk for offspring of affected parents. When one parent is BP (and the other parent is unaffected), the risk for a child of developing BP illness is ˜9%, whereas the risk of RUP

disorder is ˜18%. When one parent is UP (and the other parent is unaffected), there is ˜16% risk of RUP illness and a ˜3% risk of BPD. These risks are elevated compared to the general population risk of ˜1% for BP and ˜8% for RUP disorders. When both parents are BP or one parent is BP and the other RUP, risk of either BPD or RUP disorder is ˜75%.

In the near future, when BP and RUP susceptibility genes are identified, and risks of the common variants (alleles) at those genes are estimated (through large-scale population studies), it may be possible to assay patient DNA samples to determine which of the common susceptibility gene alleles are present in an individual at risk. With this information, it may be possible to provide improved estimates of risk.

METHODOLOGIC CONSIDERATIONS IN LINKAGE

Part of "71 - Bipolar Disorders: Review of Molecular Genetic Linkage Studies "

Linkage and association analyses are commonly employed methods to locate and define susceptibility genes for diseases. In the family shown in Fig. 71.1 , a BPD mother has alleles X,Y at some anonymous DNA marker, while unaffected father has alleles U,V. Mother transmits allele Y to affected children and allele X to the unaffected children. The probability that a parent will transmit a specific allele to each child within a family is 50%. A logarithm of odds (LOD) score statistic assesses the probability that, *within a family*, co-segregation of illness and a marker allele has occurred randomly, versus the probability that the co-segregation of illness and a marker allele has occurred because the marker allele is located near a disease gene on the same chromosome, such that the two are transmitted together more often than expected by chance (= 50%). In a single family, as shown, segregation of BPD illness with an allele at this marker locus could be a random event. However, if such a segregation was observed in >25% of 50 such BPD families, the probability that this is a random event would be remote. LOD scores for individual families can be summed to provide evidence that a region contains a susceptibility gene.

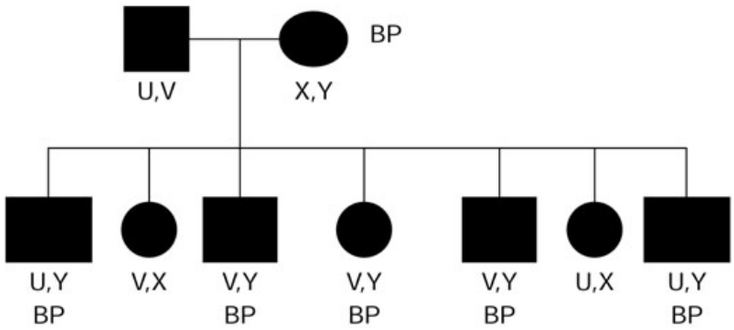


FIGURE 71.1. A family in which the bipolar disorder (BPD) mother has alleles X,Y at some anonymous DNA marker, whereas the unaffected father has alleles U,V.

LOD score calculations require specification of the disease allele frequency in the population, the mode of inheritance (dominant or recessive or some intermediate model), and the penetrance. If the mode of inheritance is misspecified, then the LOD score may not detect linkage when it is present (26). For psychiatric diseases, none of these parameters is known. In practice, investigators usually calculate LOD scores under dominant and recessive models of inheritance with reduced penetrance.

A commonly employed analysis in complex trait studies is the affected sibling pair (ASP) statistic. Pairs of siblings will share 50% of their alleles randomly, and the expected distribution of this allele sharing is as follows:

No. of alleles shared:	0	1	2
Percent of all sibling pairs:	25%	50%	25%

Pairs of affected siblings will tend to share alleles to a greater extent when the DNA marker alleles are located near a disease gene that contributes to the illness in the affected siblings pairs. Consider the ten possible affected siblings in the pedigree diagram above. Whereas four affected sibling pairs share two alleles, six pairs share one allele, but none share no alleles. This skewing of the expected random distribution of allele sharing toward greater sharing is consistent with the hypothesis that the DNA marker is located near a BPD susceptibility gene (i.e., linkage is present). This statistical method has been extended to all types of affected relative pairs (27 ,28). Because this approach does not require specification of several *parameters*, including mode of inheritance, penetrance, and disease allele frequency (as is necessary for the LOD score method), these statistical methods are often described as *nonparametric* methods.

The validity of a linkage study is demonstrated by delineation of the functional DNA sequence variants that explain the linkage statistics, or through independent confirmation in another set of families. Statistical guidelines for judging validity of linkage reports in complex disorders have been described (29). These guidelines suggest thresholds for an initial report of “significant” linkage (LOD score = 3.6 or nominal p = .00002) and for confirmation (LOD score = 1.2 or p = .01). These guidelines should limit false positives to less than 5%. It should be remembered that these guidelines refer to analysis of a single phenotypic definition (e.g., BP I and BP II disorders). If multiple phenotypes are analyzed, some statistical adjustments for multiple hypothesis testing may be necessary.

In genetic linkage analysis of common complex disorders, failure of subsequent studies to confirm previously nominated susceptibility loci has become commonplace. This is as true for diabetes mellitus (30 ,31) as it is for SZ and bipolar disorders. This failure to confirm has multiple origins most probably the manifestations of multiple susceptibility loci, within an affected individual, which interact

with each other and the environment to produce these well-known syndromes. For these disorders, no single susceptibility locus has a major effect on risk for illness in a majority of the ill population. Loci that increase risk by factors greater than 2 are unusual for common, complex disorders. Despite Herculean efforts in numerous disorders, only two loci that increase risk by a factor >2 in a large fraction of ill people have been detected: one is human leukocyte antigen (HLA) for insulin-dependent diabetes mellitus [increased risk = ˜3 (32)], and the other is apolipoprotein E in late-onset Alzheimer's disease (33 ,34 and 35 ,79 ,80). Substantial sample sizes are required to detect such loci that increase risk by factors ≤ 2. As Hauser et al. (36) have shown, ˜400 affected sibling pairs are needed to have >90% power to detect (p <.0001 or LOD >3) loci, which increases risk by a factor of 2. No single linkage study of BPD or SZ disorders published in the 1990s has exceeded this sample size, although metaanalyses of multiple independent data sets have larger sample sizes.

A second major reason for lack of confirmation in linkage studies of common complex disorders has been delineated by Suarez et al. (37), who conducted simulation studies to evaluate the power to replicate linkage. They simulated linkage data for a complex disease caused in part by six equally frequent independent (unlinked) disease loci. They found that a larger sample size was required to confirm linkage of a previously detected locus, because independent pedigree samples might (through sampling variation) contain an overrepresentation of different susceptibility loci, rather than the locus initially detected. Given that investigators often draw their pedigrees from different ethnic backgrounds (in which prevalence of a particular susceptibility locus might vary), sampling variation is an important origin of confirmation failure. Thus, expectations of universal agreement (even when sample size is adequate) regarding susceptibility loci for common complex traits are unrealistic.

If three loci of equal effect size are used in an interactive, multiplicative model to explain the increased relative risk in BPD disorder (each locus increases relative risk by ˜2), then these three hypothetical interactive loci explain most of the relative risk (2 × 2 × 2 = 8). Thus, loci that increase risk of BPD disorder will have minor to moderate effects. Substantial sample sizes are required to detect such loci of minor effect. As Hauser et al. (36) have shown, ˜400 affected sibling pairs are needed to have >95% power to detect initially (LOD ≥3.6, or p ≤.00002) loci that increase risk by a factor of 2, whereas 200 pairs are needed to have >95% power to provide confirmation (p ≤.01) of a previously detected locus.

Linkage Studies of Bipolar Disorder

To focus attention on the most promising linkage reports in BPD, this chapter limits consideration to those BPD linkages that meet criteria for validity (29), as noted above. Table 71.1 describes those BPD linkages that have at least one principal report with p = ˜.00002 and at least one independent confirmation at p = ˜.001. It is undoubtedly true that each of these confirmed linkages has been the subject of multiple negative reports. This is unavoidable when detecting loci of modest or minor effect, where the locus-specific relative risk is less than 2. Nearly all the negative reports are perhaps secondary to inadequate power to detect the initially described evidence of linkage. These negative reports will not be reviewed here.

Genomic Location	Principle Report	Independent Confirmations	Comments
18p11.2	Berrettini et al., 1994 (38) and 1997 (39)	Stine et al., 1995 (40); Nothen et al., 1999 (41); Turecki et al., 1999 (42)	Paternal parent-of-origin effect; see Schwab et al., 1998 (43)
21q22	Straub et al., 1994 (44)	Detera-Wadleigh et al., 1996 (45); Smyth et al., 1996 (46); Kwok et al., 1999 (47); Morissette et al., 1999 (48)	
22q11–13	Kelsoe et al., 2001 (49)	Detera-Wadleigh et al., 1997 (50) and 1999; (51)	Velocardiofacial syndrome region; possible overlap with a schizophrenia locus
18q22	Stine et al., 1995 (40)	McInnes et al., 1996 (52); McMahon et al., 1997 (53); De Bruyn et al., 1996 (54)	See Freimer et al., 1996 (55)
12q24	Morissette et al., 1999 (48)	Ewald et al., 1998 (56); Detera-Wadleigh et al., 1999 (51)	Principal report in a Canadian isolate
4p15	Blackwood et al., 1996 (57)	Ewald et al., 1998 (58); Nothen et al., 1997 (59); Detera-Wadleigh et al., 1999 (51)	See Ginns et al., 1998 (60)

TABLE 71.1. CONFIRMED LINKAGES IN BIPOLAR DISORDER

Berrettini et al. (38 ,39) reported evidence of a BPD susceptibility locus on 18p11 using ASP and affected pedigree member (APM) methods (p = 10^{-4} to 10^{-6}). Independent evidence of confirmation of this finding was reported

by Stine et al. (40), Nothen et al. (41), and Turecki et al. (42). Evidence of linkage was found most often among those families with paternally transmitted illness (40 ,41 ,61). As part of Genetic Analysis Workshop no. 10, independent BPD chromosome 18 linkage data sets, including ~1,200 samples, were assembled for metaanalyses (62). An affected sibling pair (*N* = 382 sibling pairs) metaanalysis yielded $p = 2.8 \times 10^{-8}$ at marker D18S37 (63).

In light of the family studies suggesting partial overlap in susceptibility for BPD and SZ (see above), it is of interest to determine whether any of these confirmed BPD loci might overlap with confirmed SZ susceptibility loci. Schwab et al. (43) employed ~20 chromosome 18 markers in a linkage study of 59 multiplex German and Israeli SZ pedigrees, in which there were 24 affective disorder cases (two were BP). When these data were analyzed in two-point parametric methods, the maximum LOD score was 3.1 at D18S53. A multipoint nonparametric analysis using Genehunter (28) revealed *p* = .002 at D18S53.

One possible explanation for the results of Schwab et al. (43) is that their kindreds were misdiagnosed or unusual in some undetected characteristics. If the SZ kindreds of Schwab et al. (64) were nosologically unique (perhaps misclassified affective disorder kindreds), then one would not expect to find confirmations of other SZ loci in those kindreds. For example, these kindreds show linkage to chromosome 6p (65), as reported in other series of multiplex SZ kindreds (66 ,67). Similarly, Faraone et al. (68) and Straub et al. (69) report SZ linkage to 10p14, as did Schwab et al. (64). Nosologic misclassification does not explain the chromosome 18p11.2 linkage to SZ detected by Schwab et al. (43). Thus, one region of partial overlap in genetic susceptibility to BPD and SZ may be 18p11.2.

Straub et al. (44) initially described linkage of BPD disorder to 21q21 markers, in a study of American and Israeli BPD kindreds. One BPD pedigree with a LOD score of 3.41 was reported from a series of 57 BPD kindreds; further nonparametric analysis provided evidence of linkage (*p* <.0003 for the phosphofructokinase locus). An emendation of this original work has been published by Aita et al. (70). A confirmation has been described in a two-locus analysis of genotypic data from 21q21 and 11p15.5 (46). This 21q21 BPD susceptibility locus has been confirmed by Detera-Wadleigh et al. (45), who employed multipoint nonparametric analyses (*p* <.001). Kwok et al. (47) described confirmatory evidence for linkage to this region in nonparametric analyses. Kelsoe et al. (49) report a LOD >2 in this region. Morisette et al. (48) also report a confirmation for this locus in a population isolate of French ancestry.

Lachman et al. (71), Edenberg et al. (72), Kelsoe et al. (49) described evidence of a BPD susceptibility locus on chromosome 22q11-13, near the velocardiofacial syndrome (VCFS) locus. Kelsoe et al. (49) report a LOD score of 3.8 at D22S278. Detera-Wadleigh et al. (51) report *p* = .008 for markers in this region. This VCFS has been associated with microdeletions of the 22q region. These individuals have a psychosis in ~30% of cases. The syndromal form of the psychosis has been termed schizophrenia-like (73), whereas others have described it in terms of bipolar disorder (74 ,75).

Another region of the genome that harbors a BPD susceptibility locus is 18q22. McMahon et al. (53) initially reported linkage to this region in 28 American BPD kindreds (LOD is 3.51 for D18S41) and the ASP method (*p* = .00002 at D18S41). In an extension of this work, McMahon et al. (53) provided additional evidence for linkage to 18q21-2 in 30 new BPD kindreds. This locus may have been detected by Freimer et al. (55) and McInnes et al. (52) who studied Costa Rican BPD kindreds. McInnes et al. described evidence of increased allele sharing at some of the same markers identified by McMahon et al. For example, at D18S55, McMahon et al. (53) reported a nonparametric LOD score of 2.2, whereas McInnes et al. (52) at this same marker reports a maximum likelihood estimate of the LOD score as 1.67. Although the genetic map position of greatest significance for these two studies are not identical, there is sufficient map location overlap so that the simplest conclusion is that the two studies have detected the same locus.

Morissette et al. (48) reported evidence of a chromosome 12q24 BPD susceptibility locus, detected through the study of a population isolate (French ancestry) from the Saguenay River region of Quebec province. Detera-Wadleigh et al. (51) observed modest support for this locus in a study of 22 American kindreds of European origin.

An extended Scottish kindred showed linkage (LOD 4.1 at D4S394) to 4p16 DNA markers (57). Confirmation of the 4p locus has been reported in a paper by Nothen et al. (59), in which increased allele-sharing was noted at D4S394 (*p* = .0009). Another confirmation was described by Ewald et al. (58), who noted a LOD of 2.0 at D4S394. Ginns et al. (60) reported linkage to this region for a *mental health locus*, meaning absence of any psychiatric disorder. This requires additional investigation. Thus, the 4p16 region has a confirmed BPD susceptibility locus.

CANDIDATE GENE APPROACHES

Part of "71 - Bipolar Disorders: Review of Molecular Genetic Linkage Studies "

The confusing array of disputed claims for association of candidate genes with psychiatric disorders becomes more comprehensible (and expected) if we recall two key issues. First, these candidate genes confer a small risk (if any), so that adequate power to confirm the originally described effect size is frequently absent in subsequent reports. Second, because the population genetics history of our species is unknown, associations detected in one ethnic group may not be detected so easily in another ethnic group. For example, the protective effect of aldehyde dehydrogenase deficiency on risk of alcoholism is easily demonstrated in

Chinese, Korean, and Japanese populations, because the deficiency allele has a frequency of ~30% (76 ,77). Much larger sample sizes are required to detect this influence in European populations, because the protective allele frequency is lower by an order of magnitude.

Candidate gene influences on risk of disease can be detected by demonstrating that certain candidate gene alleles are found more frequently among affected individuals compared to unaffected individuals. These studies are often termed “case-control association” investigations. This process is quite reliable when the effect size is robust. Candidate gene effect sizes can be considered as genotype relative risk (GRR), in which the risk associated with a particular genotype is compared to the general population risk. In general, there are four possible models of GRR: dominant, recessive, additive, and multiplicative. Let us consider each of these models for the general population risk, R, for a given disease, caused partially by a disease allele D, which triples the general population risk (the normal allele being d):

Model/genotype	DD	Dd	dd
Dominant	3R	3R	R
Recessive	3R	R	R
Additive	6R	3R	R
Multiplicative	9R	3R	R

An often-cited example of a multiplicative effect is apolipoprotein E4 in Alzheimer's disease (33 ,35 ,50 ,51 ,52 ,53 ,54 ,55 ,56 ,57 ,58 ,59 ,60 ,61 ,62 ,63 ,64 ,65 ,66 ,67 ,68 ,69 ,70 ,71 ,72 ,73 ,74 ,75 ,76 ,77 and 78), where one copy of the E4 allele increases risk for Alzheimer's disease by a factor of ~4 (at age 75), whereas homozygosity for E4 increases risk by a factor of ~14. Thus the influence of E4 on Alzheimer's disease risk may follow a multiplicative model.

Thus, one can genotype cases and unaffected individuals, comparing risk for disease across the three possible genotypes. However, in complex diseases such as BPD and RUP illness, the same disease allele may act in dominant or recessive mechanisms, depending on genetic background and environmental influence. Thus, a straightforward comparison of disease allele frequency in cases and controls can be recommended.

The major difficulty in comparing genotypes (or allele frequencies) in cases and controls is the risk of false positives because of subtle genetic differences between the case and control populations, differences that are independent of disease risk (81). Sometimes this is termed “population stratification.” This danger can be illustrated by the following example. Suppose we are interested in testing the hypothesis that glucose-6-phosphate dehydrogenase (G6PD) is a disease gene in diabetes mellitus, using the case-control method. Let us also suppose that we are unaware that G6PD deficiency protects against malaria, and is found at increased frequency among individuals of Mediterranean origins. We select cases from a population enriched with individuals of Mediterranean origin, where G6PD deficiency is fairly common. Our controls also come from a population of individuals of European ancestry, but mostly northern Europe, where G6PD deficiency is relatively uncommon. We test our cases and controls, and find that the diabetics have increased frequencies of alleles that result in marked enzyme deficiency. We conclude falsely that these G6PD alleles are risk factors for diabetes mellitus.

One method to protect against such errors is known as a family-based association test (82 ,83 and 84). Such methods generally employ DNA samples from an affected individual and his/her parents. In one form, the transmission disequilibrium test (TDT) (84), the putative susceptibility allele is examined for excess transmission from heterozygous parents to affected children. Consider this nuclear family (Fig. 71.2), consisting of an affected child and two parents, whose affection status is unknown. Genotypes at a putative candidate gene are listed. Randomly, the affected child has a 50% probability of inheriting allele 1 from her heterozygous father. Let us hypothesize that allele 1 increases risk for the disorder present in the child. If this hypothesis is true, allele 1 should be inherited from heterozygous parents by affected offspring greater than 50% of the time. If DNA samples are collected from 500 parent-affected child trios, and those samples are genotyped at the candidate gene, the hypothesis can be tested. Note that this method does not require any diagnostic information from parents, only their DNA samples. Clearly, this method is not applicable to disease with onset late in life, such as Parkinson's disease or Alzheimer's disease. However, derivatives of this approach that use discordant siblings have been described (85 ,86 and 87).

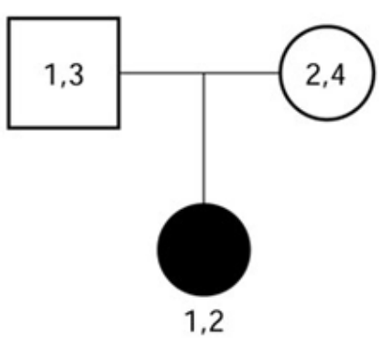


FIGURE 71.2. A nuclear family consisting of an affected child and two parents, whose affection status is unknown. Genotypes at a putative candidate gene are listed.

The disadvantages to the family-based association methods include the greater difficulty in collecting parent-child trios, compared to unrelated cases and controls. Also, only those parents who are heterozygous are informative. Given that most variants in the human genome have only two alleles, parental homozygosity can be a significant problem. One compromise paradigm is to conduct a large case-control association study of candidate gene polymorphisms. Where one sees a positive result, it can be confirmed in a smaller family-based confirmation. If the family-based sample provides confirmation, one can have greater confidence that the original case-control positive result was not due to population stratification.

Although a multitude of candidate genes have been examined in populations of BPD and RUP patients, there are no candidate genes that have been unequivocally established. It is instructive to review one candidate gene intensively studied, monoamine oxidase A (*MAOA*), located on the short arm of the X chromosome, Xp11.23. These studies exemplify the difficulties outlined above, including possible population stratification, limited power, and different ethnic groups. Although no linkage studies have suggested Xp11 as a genomic location of a susceptibility gene for affective disorder, the role of *MAOA* in the deamination of serotonin and norepinephrine and the therapeutic efficacy of MAO inhibitors suggest that this gene should be evaluated as a potential risk factor for affective disorders.

There have been numerous independent association studies of BPD and RUP and an *MAOA* (CA)n repeat polymorphism in European (88 ,89 ,90 ,91 ,92 ,93 and 94) and Asian (95 ,96) studies. Those studies reporting a positive association (89 ,92 ,94 ,95) generally detect an overrepresentation of allele 5 or 6 of the *MAOA* (CA)n repeat among BPD patients, compared to controls, an observation that may be particularly evident among women. The effect size is small, the odds ratio being 1.49 (94), and the sample size required for adequate power to detect is larger than most of the negative studies (88 ,90 ,91 ,93 ,96). There is also an *MAOA* promoter polymorphism (97). These studies involve multiple ethnic groups, case-control methods, and family-based designs, with some studies having limited power to detect a small effect size. Thus, it is understandable that conflicting studies are reported.

PARENT-OF-ORIGIN EFFECTS

Part of "71 - Bipolar Disorders: Review of Molecular Genetic Linkage Studies "

Parent-of-origin effects refer to unequal rates of transmission of a disorder from fathers, compared to mothers. In BPD, McInnis et al. (98) first observed an excess of maternal transmission in multiplex kindreds selected for a linkage study. This observation was confirmed by Gershon et al. (15) in an independent series of multiplex BPD kindreds. Although this observation has not been confirmed by other investigators (99), it raises the possibility that the complex genetics of BPD may involve mitochondrial inheritance and/or imprinting.

Mitochondrial DNA is a nonnuclear circular 16,500 base pair molecule that is solely maternal in origin. It contains genes for oxidative phosphorylation and genes for transfer RNA (tRNA) molecules, among others. Defects in mitochondrial DNA sequence can contribute to genetic susceptibility for complex disorders, such a diabetes mellitus (100) and some forms of nonsyndromic deafness (101). If a fraction of all BPD included a mitochondrial susceptibility gene, then this would be consistent with the excess maternal transmission observed in BPD (15 ,98). However, variations in mitochondrial DNA have not been associated with BPD.

Another mechanism consistent with excess maternal transmission is that of imprinting (see ref. 102 for review). Imprinting results in the transcriptional silencing of the allele of a sex-specific parent. For some genes, imprinting is *paternal*, meaning that the paternal allele is not transcribed into messenger RNA (mRNA). For other genes, *maternal* imprinting is present, and the maternal allele is transcriptionally silent. This results in a "functional hemizygosity," so that defects in the single active allele may have a greater impact on the phenotype. How might this mechanism give rise to an apparent excess of maternal transmission, as has been observed by McInnis et al. (98) and Gershon et al. (15)? If the putative BPD susceptibility gene is *paternally* imprinted, then the paternal allele is transcriptionally silent, and defects in the expressed maternal allele may be more often detected in the phenotypes of the offspring. In the embryonic gonads of each generation, the imprinting mark is reset, so that the alleles can be properly regulated in the next generation. Consider an example of *paternal* imprinting (Fig. 71.3). Note that individuals are heterozygous at the DNA level, but they are hemizygous at the mRNA level, expressing only the maternal allele. Note also that the imprinting mark is reset with each generation, so the woman who inherits (and expresses) allele 3 from her mother transmits allele 2 to her son, and this allele is transcriptionally active in the son, who does not express allele 5, which he inherits paternally. Thus if allele 2 is a BPD susceptibility gene, it will be expressed in the third generation, and influence the risk of disease, due to maternal imprinting. However, allele 2 is not expressed in the second generation, and will not influence risk of BPD.

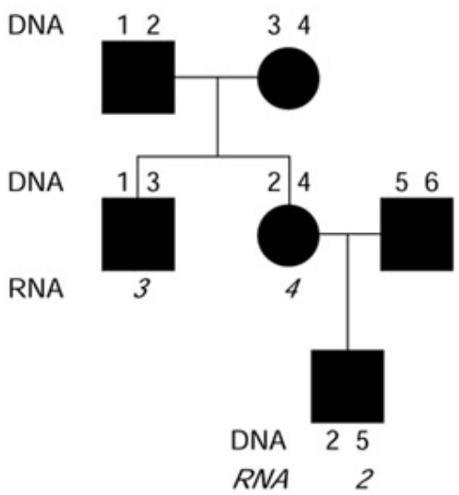


FIGURE 71.3. An example of paternal imprinting.

Molecular mechanisms of imprinting are complex, but involve methylation of the promoter regions of the target genes (102), resulting in nontranscription. Imprinting defects give rise to human diseases, the classic examples of which are Prader-Willi and Angelman's syndromes. Imprinting can be tissue-specific, such as the serotonin receptor subtype 2A (5-HT_{2A}) gene (103), which is imprinted in specific brain regions. Imprinting has been described for

˜50 human genes. It is likely that imprinting will explain some of the complex inheritance of BPD.

ANTICIPATION AND TRIPLET REPEATS

Part of "71 - Bipolar Disorders: Review of Molecular Genetic Linkage Studies "

Anticipation is the term used to define an observation that a familial disorder occurs with earlier age at onset and/or increasing severity among younger generations, compared to older generations. Anticipation occurs in several neurodegenerative diseases, including Huntington's disease, fragile X, myotonic dystrophy, spinocerebellar ataxias, and others (see ref. 104 for review). The molecular explanation for anticipation in these disorders involves unstable intragenic trinucleotide repeats, which expand in subsequent generations, giving rise to increasing levels of gene disruption and thus to earlier age at onset and increasingly severe phenotype in younger generations.

Evidence of anticipation has been reported in several family studies of BPD illness (98 ,105 ,106 and 107), but some authorities suggest that there is intractable ascertainment bias (108 ,109). Individuals with earlier age-at-onset BPD disorder may have reduced capacity to reproduce, so parents with such early-onset disorders may be underrepresented in the general population. Individuals with familial BPD disorder may come to treatment earlier than those with sporadic disease, such that less severe mood disorder episodes are detected medically, and an earlier age at onset is defined. Such individuals (by virtue of their familiarity with mood disorder symptoms) may be more likely to report minor mood disturbance in terms of “diagnosable syndromes.” Some evidence of anticipation in BPD comes from extensive studies of multiplex BPD families for linkage studies. These linkage studies select for earlier age-at-onset cases, because preference is given to densely affected kindreds. Among broader cultural factors possibly underlying the evidence of anticipation, if stigma concerning mood disorders is less among younger affected persons (compared to older individuals), then younger cohorts might describe their experiences more easily in terms of a diagnosable mood disorder, because denial (due to stigma) is less prevalent among the younger cohorts. These potential confounding factors make detection of anticipation in BPD disorder difficult.

The hypothesis that anticipation in BPD disorder reflects causative expanding trinucleotide CTG repeat sequences has generated genomic searches for such sequences (110 ,111 ,112 and 113), using the repeat expansion detection method (114). These three groups have noted increased lengths of CTG repeats in BPD disorders, especially among those with familial disease. However, not all studies have reported this difference (115), and no report shows transmission of an expanding repeat within BPD families, the definitive evidence. Furthermore, greater than 90% of the expanded CTG repeats detected by the method of Schalling et al. (114) are from two apparently nonpathogenic unstable CTG repeats on 17q and 18q21 (116). The hypothesis that unstable trinucleotide repeats represent BPD susceptibility factors warrants continued study.

IMPLICATIONS OF THE HUMAN GENOME PROJECT

Part of "71 - Bipolar Disorders: Review of Molecular Genetic Linkage Studies "

The Human Genome Project is nearing completion of one of its primary goals, the sequence of the human genome. However, the additional goal of defining all expressed sequences (genes) may require several additional years of work. Once this goal is achieved, the most important task will be the definition of function for each expressed sequence (functional genomics). Implied (but not included) in this goal is the function of each protein (functional proteomics), involving interactions between proteins. A third goal of the Human Genome Project is the definition of ˜300,000 common single nucleotide polymorphisms (SNPs), including several within each gene.

The cloning of individual susceptibility genes for BPD and RUP disorder will be facilitated remarkably by the completion of these Human Genome Project milestones noted above. At present our knowledge base regarding central nervous system (CNS) function and the biochemical etiology of BPD and RUP disorder is so poor that too many brain-expressed genes may be considered candidates. This limitation is made more severe by the fact that linkage studies of all complex traits result in genomic regions of interest that typically span 20,000,000 base pairs of DNA. Because ˜20 to 50 genes (most of which are unknown today) are usually found in every 1,000,000 base pairs of DNA, the task of discovering a single disease gene within such a region, implicated by linkage, is the proverbial equivalent of finding a needle in a haystack, with currently available DNA technologies. However, once all expressed sequences are known, and their functions are understood, it is possible to focus on the few best candidates. This reduces an intractable problem (from a DNA technology perspective) to a manageable size. Thus, finding susceptibility genes implicated by linkage results will become progressively less difficult as the Human Genome Project goals are approached.

SUMMARY

Part of "71 - Bipolar Disorders: Review of Molecular Genetic Linkage Studies "

Despite the extensive data (from twin, family, and adoption studies) for genetic factors in BPD, gene identification through linkage studies has been elusive. There are multiple confirmed BPD linkage regions across the human genome, but the effect sizes are uniformly small at each locus. Cloning genes from these small effect size regions is a challenge for current molecular techniques. Part of the complexity of BPD genetics may be due to imprinting, mitochondrial

inheritance, and trinucleotide repeat expansion. These nonmendelian influences require additional research.

ACKNOWLEDGMENTS

Part of "71 - Bipolar Disorders: Review of Molecular Genetic Linkage Studies "

This paper was prepared with the support of National Institutes of Health (NIH) grant MH59553 and a National Alliance for Research on Schizophrenia and Depression Distinguished Investigator Award.

REFERENCES

1. Gershon ES, Hamovit J, Guroff JJ, et al. A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. *Arch Gen Psychiatry* 1982;39:1157–1167.

2. Weissman MM, Gershon ES, Kidd KK, et al. Psychiatric disorders in the relatives of probands with affective disorder. *Arch Gen Psychiatry* 1984;41:13–21.

3. Baron M, Gruen R, Anis L, et al. Schizoaffective illness, schizophrenia and affective disorders: morbidity risk and genetic transmission. *Acta Psychiatr Scand* 1983;65:253–262.

4. Winokur G, Tsuang MT, Crowe RR. The Iowa 500: affective disorder in relatives of manic and depressed patients. *Am J Psychiatry* 1982;139:209–212.

5. Winokur G, Coryell W, Keller M, et al. A family study of manic-depressive (bipolar I) disease. Is it a distinct illness separable from primary unipolar depression? *Arch Gen Psychiatry* 1995;52(5):367–373.

6. Helzer JE, Winokur G. A family interview study of male manic-depressives. *Arch Gen Psychiatry* 1974;31:73–77.

7. James NM, Chapman CJ. A genetic study of bipolar affective disorder. *Br J Psychiatry* 1975;126:449–456.

8. Johnson GFS, Leeman MM. Analysis of familial factors in bipolar affective illness. *Arch Gen Psychiatry* 1977;34:1074–1083.

9. Angst J, Frey R, Lohmeyer R, et al. Bipolar manic depressive psychoses: results of a genetic investigation. *Hum Genet* 1980;55:237–254.

10. Maier W, Lichtermann D, Minges J, et al. Continuity and discontinuity of affective disorders and schizophrenia. Results of a controlled family study. *Arch Gen Psychiatry* 1993;50:871–883.

11. Slater E. The inheritance of manic-depressive insanity. *Proc R Soc Med* 1936;29:981–990.

12. Mendlewicz J, Baron M. Morbidity risks in subtypes of unipolar depressive illness: differences between early and late onset forms. *Br J Psychiatry* 1981;139:463–466.

13. Bland RC, Newman SC, Orn H. Recurrent and nonrecurrent depression. *Arch Gen Psychiatry* 1986;43:1085–1089.

14. Kupfer DJ, Frank E, Carpenter LL, et al. Family history in recurrent depression. *J Affective Disord* 1989;17:113–119.

15. Gershon ES, DeLisi LE, Hamovit J, et al. A controlled family study of chronic psychoses. *Arch Gen Psychiatry* 1988;45:328–336.

16. Kendler KS, McGuire M, Gruenberg AM, et al. The Roscommon family study. *Arch Gen Psychiatry* 1993;50:527–540.

17. Kendler KS, Neale MC, Kessler RC, et al. A population-based twin study of major depression in women: the impact of varying definitions of illness. *Arch Gen Psychiatry* 1992;49:257–266.

18. Nurnberger JL Jr, Berrettini WH. *Psychiatric genetics*. London: Chapman and Hall, 1998.

19. McGuffin P, Katz R, Watkins S, et al. A hospital-based twin register of the heritability of DSM-IV unipolar depression. *Arch Gen Psychiatry* 1996;53:129–136.

20. Bertelsen A, Harvald B, Hauge M. A Danish twin study of manic-depressive disorders. *Br J Psychiatry* 1977;130:330–351.

21. Allen MG, Cohen S, Pollin W, et al. Affective illness in veteran twins. A diagnostic review. *Am J Psychiatry* 1974;131:1234–1239.

22. Mendlewicz J, Rainer JD. Adoption study supporting genetic transmission in manic-depressive illness. *Nature* 1977;368(5618):327–329.

23. Wender H, Kety SS, Rosenthal D, et al. Psychiatric disorders in the biological and adoptive families of adopted individuals with affective disorders. *Arch Gen Psychiatry* 1986;43:923–929.

24. Cadoret RJ. Evidence for genetic inheritance of primary affective disorder in adoptees. *Am J Psychiatry* 1978;135(4):463–466.

25. Von Knorring AL, Cloninger CR, Bohman M, et al. An adoption study of depressive disorders and substance abuse. *Arch Gen Psychiatry* 1983;40:943–950.

26. Clerget-Darpoux F, Bonaiti-Pellie C, Hochez J. Effects of mis-specifying genetic parameters in LOD score analysis. *Biometrics* 1986;42:393–399.

27. Weeks DE, Lange K. The affected-pedigree-member method of linkage analysis. *Am J Hum Genet* 1988;42:315–326.

28. Kruglyak L, Daly MJ, Reeve-Daly MP, et al. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996;58:1347–1363.

29. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genet* 1995;11:241–247.

30. Hanis CL, Boerwinkle E, Chakraborty R, et al. A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nature Genet* 1996;13(2):161–166.

31. Concannon P, Gogolin-Ewens KJ, Hinds DA, et al. A second-generation screen of the human genome for susceptibility to insulin-dependent diabetes mellitus. *Nature Genet* 1998;19:292–296.

32. Davies JL, Kawaguchi Y, Bennett ST, et al. A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 1994;371:130–136.

33. Corder EH, Saunders AM, Strittmatter WJ, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 1993;261:921–923.

34. Mayeux R, Stern Y, Ottman R, et al. The apolipoprotein E4 allele in patients with Alzheimer's disease. *Ann Neurol* 1993;34:752–754.

35. Tsai M-S, Tangalos EG, Petersen RC, et al. Apolipoprotein E: risk factor for Alzheimer disease. *Am J Hum Genet* 1994;54:643–649.

36. Hauser ER, Boehnke M, Guo S-W, et al. Affected sib pair interval mapping and exclusion for complex genetic traits. *Genet Epidemiol* 1996;13:117–137.

37. Suarez B, Harpe CL, Van Eerdewegh P. Problems of replicating linkage claims in psychiatry. In: Gershon ES, Cloninger CR, eds. *Genetic approaches to mental disorders*. Proceedings of the 82nd Annual Meeting of the American Psychopathological Association. Washington, DC: American Psychiatric Press, 1994:23–46.

38. Berrettini W, Ferraro T, Goldin L, et al. Chromosome 18 DNA markers and manic depressive illness: evidence for a susceptibility gene. *Proc Natl Acad Sci USA* 1994;91:5918–5921.

39. Berrettini W, Ferraro T, Choi H, et al. Linkage studies of bipolar illness. *Arch Gen Psychiatry* 1997;54:32–39.

40. Stine OC, Xu J, Koskela R, et al. Evidence for linkage of bipolar disorder to chromosome 18 with a parent-of-origin effect. *Am J Hum Genet* 1995;57:1384–1394.

41. Nothen MM, Cichon S, Rohleder H, et al. Evaluation of linkage of bipolar affective disorder to chromosome 18 in a sample of 57 German families. *Mol Psychiatry* 1999;4:76.

42. Turecki G, Grof P, Cavazzoni P, et al. Lithium responsive bipolar disorder, unilineality and chromosome 18: a linkage study. *Am J Med Genet* 1999;88:411–415.

43. Schwab SG, Hallmayer J, Lerer B, et al. Support for a chromosome 18p locus conferring susceptibility to functional psychoses in families with schizophrenia, by association and linkage analysis. *Am J Hum Genet* 1998;63:1139.

44. Straub RE, Lehner T, Luo Y, et al. A possible vulnerability locus for bipolar affective disorder on chromosome 21q22.3. *Nature Genet* 1994;8:291–296.

45. Detera-Wadleigh SD, Badner JA, Goldin LR, et al. Affected sib-pair analyses reveal support of prior evidence for a susceptibility locus for bipolar disorder on 21q. *Am J Hum Genet* 1996;58:1279–1285.

46. Smyth C, Kalsi G, Brynjolfsson J, et al. Further tests for linkage of bipolar affective disorder to the tyrosine hydroxylase gene locus on chromosome 11p15 in a new series of multiplex British affective disorder pedigrees. *Am J Psychiatry* 1996;153:271–274.

47. Kwok JB, Adams LJ, Salmon JA, et al. Non-parametric simulation-based statistical analyses for bipolar affective disorder locus on chromosome 21q22.3. *Am J Med Genet* 1999;88:99–102.

48. Morissette J, Villeneuve A, Bordeleau L, et al. Genome-wide search for linkage of bipolar affective disorders in a very large pedigree derived from a homogeneous population in Quebec points to a locus of major effect on chromosome 12q23-q24. *Am J Med Genet* 1999;88:567–587.

49. Kelsoe JR, Spence MA, Loetscher E, et al. A genome survey indicates a susceptibility locus for bipolar disorder on chromosome 22. *Proc Natl Acad Sci USA* 2001;98:585–590.

50. Detera-Wadleigh SD, Hsieh W-T, Berrettini WH, et al. Genetic linkage mapping for a susceptibility locus to bipolar illness: Chromosomes 2, 3, 4, 7, 9, 10p, 11p, 22, and Xpter. *Neuropsychiatr Genet* 1997;74:206–218.

51. Detera-Wadleigh SD, Badner JA, Berrettini WH, et al. Evidence for a bipolar susceptibility locus on 13q32 and other potential loci on 1q32 and 18p11.2. *Proc Natl Acad Sci USA* 1999;96:5604–5609.

52. McInnes LA, Escamilla MA, Service SK, et al. A complete genome screen for genes predisposing to severe bipolar disorder in two Costa Rican pedigrees. *Proc Natl Acad Sci USA* 1996;93:13060–13065.

53. McMahon FJ, Hopkins PJ, Xu J, et al. Linkage of bipolar affective disorder to chromosome 18 markers in a new pedigree series. *Am J Hum Genet* 1997;61:1397–1404.

54. De Bruyn A, Souery D, Mendelbaum K, et al. Linkage analysis of families with bipolar illness and chromosome 18 markers. *Biol Psychiatry* 1996;39:679–688.

55. Freimer NB, Reus VI, Escamilla MA, et al. Genetic mapping using haplotype, association and linkage methods suggests a locus for severe bipolar disorder (BPI) at 18q22-q23. *Nature Genet* 1996;12:436–441.

56. Ewald H, Degn B, Mors B, et al. Significant linkage between bipolar affective disorder and chromosome 12q24. *Psychiatr Genet* 1998;8:131–140.

57. Blackwood DHR, He L, Morris SW, et al. A locus for bipolar affective disorder on chromosome 4p. *Nature Genet* 1996;12:427–430.

58. Ewald H, Degn B, Mors B, et al. Support for the possible locus on chromosome 4p15 for bipolar affective disorder. *Mol Psychiatry* 1998;3:442–448.

59. Nothen MM, Cichon S, Franzek E, et al. Systematic search for susceptibility genes in bipolar affective disorder—evidence for disease loci at 18p and 4 p. *Am J Hum Genet* 1997;61(S):A288.

60. Ginns EI, St Jean P, Philibert RA, et al. A genome search for chromosomal loci linked to mental health wellness in relatives at high risk for bipolar affective disorder among the Old Order Amish. *Proc Natl Acad Sci USA* 1998;95:15531.

61. Gershon ES, Badner JA, Ferraro TN, et al. Maternal inheritance and chromosome 18 allele sharing in unilineal bipolar illness pedigrees. *Neuropsychiatr Genet* 1996;67:1–8.

62. Goldin LR, Gershon ES, Berrettini WH, et al. Description of the Genetic Analysis Workshop 10 bipolar disorder linkage data sets. *Genet Epidemiol* 1997;14:563–568.

63. Lin JP, Bale SJ. Parental transmission and D18S37 allele sharing in bipolar affective disorder. *Genet Epidemiol* 1997;14:665–668.

64. Schwab SG, Hallmayer J, Albus M, et al. Further evidence for a susceptibility locus on chromosome 10p14-p11 in 72 families with schizophrenia by non-parametric linkage analysis. *Am J Med Genet* 1999;81:302–307.

65. Schwab SG, Albus M, Hallmayer J, et al. Evaluation of a susceptibility gene for schizophrenia on chromosome 6p by multipoint affected sib-pair linkage analysis. *Nature Genet* 1995;11:325–327.

66. Straub RE, MacLean CJ, O'Neill FA, et al. A potential vulnerability locus for schizophrenia on chromosome 6p22-24: evidence for genetic heterogeneity. *Nature Genet* 1995;11:287–293.

67. Moises HW, Yang L, Kristbjarnarson H, et al. An international two-stage genome-wide search for schizophrenia susceptibility genes. *Nature Genet* 1995;11:321–324.

68. Faraone SV, Matise T, Svrakic D, et al. Genome scan of European-American schizophrenia kindreds: results of the NIMH Genetics Initiative and Millennium consortium. *Am J Med Genet* 1998;81:290–295.

69. Straub RE, MacLean CJ, Martin RB, et al. A schizophrenia locus may be located in region 10p15-p11. *Am J Med Genet* 1998;81:296–301.

70. Aita VM, Liu J, Knowles JA, et al. A comprehensive linkage analysis of chromosome 21q22 supports prior evidence for a putative bipolar affective disorder locus. *Am J Hum Genet* 1999;64:210–217.

71. Lachman HM, Kelsoe JR, Remick RA, et al. Linkage studies support a possible locus for bipolar disorder in the velocardiofacial syndrome region on chromosome 22. *Am J Med Genet* 1996;74:121–128.

72. Edenberg H, Foroud T, Conneally M, et al. Initial genomic scan of the NIMH genetics initiative bipolar pedigrees: chromosomes 3,5,15,16,17, and 22. *Neuropsychiatr Genet* 1997;74:238–246.

73. Pulver AE, Nestadt G, Goldberg R, et al. Psychotic illness in patients diagnosed with velo-cardio-facial syndrome and their relatives. *J Nerv Ment Dis* 1994;182(8):476–478.

74. Papolos DF, Faedda GL, Veit S, et al. Bipolar spectrum disorders in patients diagnoses with VCFS: does a hemizygous deletion of chromosome 22q11 result in bipolar affective disorder? *Am J Psychiatry* 1996;153:1541–1547.

75. Carlson C, Papolos D, Pandita RK, et al. Molecular analysis of VCFS patients with psychiatric disorders. *Am J Med Genet* 1997;60:851–859.

76. Harada S, Agarwal DP, Goedde HW. Aldehyde dehydrogenase deficiency as cause of the flushing reaction to alcohol in Japanese. *Lancet* 1982;2(8253):982.

77. Thomasson HR, Edenberg HJ, Crabb DW, et al. Alcohol and aldehyde dehydrogenase genotypes and alcoholism in Chinese men. *Am J Hum Genet* 1991;48:677.

78. Mayeux R, Stern Y, Ottman R, et al. The apolipoprotein e4 allele in patients with Alzheimer's disease. *Ann Neurol* 1993;34:752–754.

79. Saunders AM, Schmader KE, Breitner JC, et al. Apolipoprotein E4 allele distributions in late-onset Alzheimer's disease and in other amyloid-forming diseases. *Lancet* 1993;342:710–711.

80. Martin ER, Lai EH, Gilbert JR, et al. SNPing away at complex diseases: analysis of SNPs around APOE in Alzheimer disease. *Am J Hum Genet* 2000;67:383–394.

81. Berrettini WH. On the interpretation of association studies in behavioral disorders. *Mol Psychiatry* 1997;2:274–275.

82. Falk CT, Rubenstein P. Hablotype relative risks: an eary reliable way to construct proper control samples for risk calculations. *Ann Hum Genet* 1987;51:227–233.

83. Ott J. Statistical properties of the haplotype relative risk. *Genet Epidemiol* 1989;6:127–130.

84. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993;52:506–516.

85. Spielman RS, Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am J Hum Genet* 1998;62:450–458.

86. Lazzeroni LC, Lange K. A conditional inference framework for extending the transmission/disequilibrium test. *Hum Hered* 1998;48:67–81.

87. Monks SA, Kaplan NL, Weir BS. A comparative study of sibship tests of linkage and/or association. *Am J Hum Genet* 1998;63:1507–1516.

88. Craddock N, Daniels J, Roberts E, et al. No evidence for allelic association between bipolar disorder and MAOA gene polymorphisms. *Am J Med Genet* 1995;60:322–324.

89. Lim LC, Powell J, Sham P, et al. Evidence for a genetic association between alleles of MAOA gene and bipolar affective disorder. *Am J Med Genet* 1995;60:325–331.

90. Nothen MM, Eggermann K, Albus M, et al. Association analyses of the MAOA gene in bipolar affective disorder by using family based controls. *Am J Hum Genet* 1995;57:975–977.

91. Parsian A, Todd RD. Genetic association between monoamine oxidase and manic-depressive illness: comparison of relative risk and haplotype relative risk data. *Am J Med Genet* 1997;74(5):475–479.

92. Furlong RA, Ho L, Rubinszstein JS, et al. Analysis of the MAOA gene in bipolar affective disorder by association studies, meta-analyses and sequencing of the promoter. *Am J Med Genet* 1999;88:398–406.

93. Turecki G, Grof P, Cavazzoni P, et al. MAOA: association and linkage studies with lithium responsive bipolar disorder. *Psychiatr Genet* 1999;9:13–16.

94. Preisig M, Bellivier F, Fenton BT, et al. Association between bipolar disorder and MAOA gene polymorphisms: results of a multicenter study. *Am J Psychiatry* 2000;157:948–959.

95. Kawada Y, Hattori M, Dai XY. Possible association between MAOA gene and bipolar affective disorder. *Am J Hum Genet* 1995;56:335–336.

96. Muramatsu T, Matsushita S, Kanba S, et al. MAO genes polymorphisms and mood disorder. *Am J Med Genet* 1997;19:494–496.

97. Kunugi H, Ishida S, Kato T, et al. A functional polymorphism in the promoter region of MAOA gene and mood disorders. *Mol Psychiatry* 1999;4:393–395.

98. McInnis MG, McMahon FJ, Chase GA, et al. Anticipation in bipolar affective disorder. *Am J Hum Genet* 1993;53:385–390.

99. Grigoroiu-Serbanescu M, Martinez M, Nothen MM, et al. Patterns of parental transmission and familial aggregation models in bipolar affective disorder. *Am J Med Genet* 1998;81:397–404.

100. Kadowaki T, Kadowaki H, Mori Y, et al. A subtype of diabetes mellitus associated with a mutation of mitochrondrial DNA. *N Engl J Med* 1994;330:962–967.

101. Prezant TR, Agapian JV, Bohlman MC, et al. Mitochondrial ribosomal RNA mutation associated with both antibioticinduced and non-syndromic deafness. *Nature Genet* 1994;4:289–294.

102. Tilghman SM. The sins of the fathers and mothers: Genomic imprinting in mammalian development. *Cell* 1999;96:185–193.

103. Bunzel R, Blumke I, Cichon S, et al. Polymorphic imprinting of the serotonin-2A (5-HT2A) receptor gene in human adult brain. *Brain Res Mol Brain Res* 1998;59(1):90–92.

104. Margolis RL, McInnis MG, Rosenblatt A, et al. Trinucleotide repeat expansion and neuropsychiatric disease. *Arch Gen Psychiatry* 1999;56:1019–1031.

105. Lipp O, Souery D, Mahieu B, et al. Anticipation in affective disorders. *Psychiatr Genet* 1995;5:S8.

106. Nylander PO, Engstrom C, Shotai J, et al. Anticipation in Swedish families with bipolar affective disorder. *J Med Genet* 1994;31:686–689.

107. Gershon ES, Hamovit JH, Guroff JJ, et al. Birth-cohort changes in manic and depressive disorders in relatives of bipolar and schizoaffective patients. *Arch Gen Psychiatry* 1987;440:314–319.

108. Penrose LS. The problem of anticipation in pedigrees of dystrophia myotonica. *Ann Eugenics* 1948;14:125.

109. Hodge SE, Wickramaratne P. Statistical pitfalls in detecting age-at-onset anticipation: the role of correlation in studying anticipation and detecting ascertainment bias. *Psychiatr Genet* 1995;5:43–47.

110. O'Donovan MC, Guy C, Craddock N, et al. Expanded CAG repeats in schizophrenia and bipolar disorder. *Nature Genet* 1995;10:380.

111. O'Donovan MC, Guy C, Craddock N, et al. Confirmation of association between expanded CAG/CTG repeats and both schizophrenia and bipolar disorder. *Psychol Med* 1996;26:1145–53.

112. Lindblad K, Nylander P-O, De Bruyn A, et al. Detection of expanded CAG repeats in bipolar affective disorder using the repeat expansion detection (RED) method. *Neurobiol Dis* 1995;2:55–62.

113. Oruc L, Lindblad K, Verheyen GR, et al.CAG repeat expansions in bipolar and unipolar disorders. *Am J Hum Genet* 1997;60:730–732.

114. Schalling M, Hudson TJ, Buetow KH, et al. Direct detection of novel expanded trinucleotide repeat in the human genome. *Nature Genet* 1993;4:135–139.

115. Vincent JB, Klempan T, Parikh SS, et al.Frequency analysis of large CAG/CTG trinucleotide repeats in schizophrenia and bipolar affective disorder. *Mol Psychiatry* 1996;1:141–148.

116. Sidransky E, Burgess C, Ikeuchi T, et al. A triplet repeat on 17q accounts for most expansions detected by the repeat-expansion detection technique. *Am J Hum Genet* 1998;62:1548.

72

Molecular and Cellular Mechanisms in Depression

Alan F. Schatzberg

Stephen J. Garlow

Charles B. Nemeroff

Stephen J. Garlow and Charles B. Nemeroff: Departments of Psychiatry and Behavioral Sciences, Emory University, Atlanta, Georgia.

Alan F. Schatzberg: Departments of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, California.

Over the past three decades, considerable progress has been made in our understanding of the biology of depressive disorders. Still, there are a great number of unanswered questions regarding the relative roles specific biological systems may play in pathogenesis. This debate in part reflects a number of methodologic factors: a possibly over broad definition of the clinical syndrome of major depression; limitations inherent in studies using indirect measurement of brain neuronal activity; problems inherent in postmortem studies; and an overemphasis on cross-sectional rather than longitudinal studies. In this chapter, we review the current status of the neurochemical and cellular features of depressive disorders.

BACKGROUND

Part of "72 - Molecular and Cellular Mechanisms in Depression "

Although Freud put forth a hypothesis for understanding the psychological causes of depression in his classic paper, “Mourning and Melancholia,” he noted that some depressions were clearly biological in etiology. Research over the past 40 years has done much to point to likely “culprits” that are involved in the etiology of the disorder as well as in the mediation of treatment response; these have been reviewed several times recently (1 ,2).

Early research revolved around monoaminergic theories with particular emphasis first on norepinephrine and later serotonin. The basis for invoking these systems rested largely on a number of pharmacologic observations that have been termed “the psychopharmacologic bridge.” These observations included: reserpine, an early antihypertensive, caused depression in some patients and depleted monoamine stores in rat brain; iproniazid, a drug studied as an antitubercular agent, elevated depressed mood and inhibited monoamine degradation by the enzyme, monoamine oxidase; imipramine, a tricyclic compound originally studied as an antipsychotic, had potent antidepressant effects and blocked the reuptake of norepinephrine (and to some extent serotonin) into presynaptic neurons.

These observations led two groups of investigators (3 ,4) to argue that norepinephrine (NE) activity was decreased in depressive disorders and elevated in manic or excited states. Although a low norepinephrine state was the cornerstone of Schildkraut's catecholamine hypothesis (30), he also argued for other types of dysregulation, including altered receptor functioning. Indeed, more recent data have pointed to biological heterogeneity of norepinephrine activity in depression with some patients demonstrating low and others apparently elevated activity (5). Serotonin (5-HT) theories, in contrast, have emphasized decreased production or reuptake in depression.

As research has continued, investigators have noted a number of other alterations in depressed patients, including among others: elevated corticotropin-releasing hormone (CRH); elevated acetylcholine activity; increased γaminobutyric acid (GABA) levels; excessive glucocorticoid activity in psychotic major depression; hippocampal volume loss, perhaps reflecting the effects of excessive glucocorticoids on neurogenesis, and so on. These have in turn led to or been associated with a number of new biological hypotheses regarding why some individuals become depressed or develop specific symptoms. In the following sections we review the current status of these approaches.

NOREPINEPHRINE

Part of "72 - Molecular and Cellular Mechanisms in Depression "

Norepinephrine is a catecholamine that is found in various tissues, including brain, plasma, sympathetic nervous system, heart, and so on. It is synthesized from the amino acid tyrosine, which forms L-Dopa via the enzyme tyrosine

hydroxylase. L-Dopa is converted to dopamine via dopa decarboxylase and then in turn is converted to norepinephrine via dopamine ß-hydroxylase. In the adrenal and other tissues, norepinephrine is converted to epinephrine via phenyl-*N*-methyltransferase (PNMT). NE is degraded by the enzymes catechol-o-methyltransferase and monoamine oxidase.

Norepinephrine measured in urine or plasma is largely derived from non-central nervous system (CNS) sources. In contrast, much early work in this area emphasized 3-methoxy-4-hydroxyphenylglycol (MHPG), 20% to 30% of which is derived from brain. The earliest studies on urinary MHPG reported significantly lower levels in depressed patients than healthy controls (6). Further research revealed low urinary MHPG levels were seen, particularly in bipolar depressives and a subgroup of unipolar patients. As diagnostic nomenclature differentiated bipolar I from II patients, investigators reported low MHPG levels were characteristic of bipolar I and not II patients (5 ,7 ,8).

Unipolar depressed patients are heterogeneous in their MHPG levels. As indicated previously, a subgroup of unipolar patients demonstrate low MHPG levels, similar to those seen in bipolar I patients. In contrast, some unipolar patients demonstrate elevated MHPG levels (9). In these patients, urinary free cortisol is similarly elevated (10).

Catecholamine levels have been reported to parallel the state of the disorder in bipolar patients. Bipolar patients demonstrate significantly lower plasma NE and E levels when depressed than when euthymic or manic. Manic bipolar patients demonstrate elevated CSF, urine, or plasma MHPG levels than depressives or healthy controls (11 ,12 and 13). These data provide a rationale for measuring catecholamine output in mood disorder and invoking NE as playing an etiologic role; however, critics argue that some of the changes in levels may be secondary to such features as activity or agitation.

Urinary MHPG levels have been explored as possible tests for predicting antidepressant response. The earliest studies (14 ,15) pointed to low MHPG levels predicting positive responses to imipramine but not amitriptyline. High-MHPG patients responded to amitriptyline (14 ,15). These findings led Maas (14) to hypothesize that there were two forms of depression—one a low MHPG state reflected a norepinephrine depression; another characterized by high MHPG levels signified a serotonin depression. This hypothesis, although heuristic, has not been borne out. Subsequent studies failed: (a) to demonstrate that high MHPG levels predicted amitriptyline response (16); and (b) in the light of the development of selective serotonin reuptake inhibitors (SSRIs) the serotonergic potency of amitriptyline was also thrown into question. In contrast, several studies have reported that low urinary MHPG levels do indeed predict response to noradrenergic agents—nortriptyline, desipramine, and maprotiline (17 ,18). Application of urinary MHPG levels has been limited in part because of: difficulty collecting 24-hr urine samples; the need for patients being drug free when studied; and the lack of surety of the optimal treatment for high-MHPG patients.

Tyrosine Hydroxylase/Locus Ceruleus

The locus ceruleus (LC) is the nucleus of the NE system in brain. Neurons project from the LC to various parts of the brain, particularly the frontal cortex. The LC has been the focus of several postmortem studies of depressed patients or suicide victims. Tyrosine hydroxylase activity has been reported to be up-regulated in brains of suicide victims, perhaps reflecting the effect of chronic stress (19). In another study, NE neurons were reported to be modestly decreased in suicide victims relative to controls (20). A third study reported that NE transporter sites were decreased in depressed subjects who committed suicide but NE neurons were not (21). These studies are somewhat contradictory in direction of NE changes in suicide but suggest the system is altered in suicide. A possible unifying hypothesis revolves around up-regulation of TH in some neurons in an attempt to compensate for loss of neurons or transporter sites.

Receptors

Receptors for NE are grouped into α_1, α_2, $ß_1$, and $ß_2$ subtypes. α_2 Receptor numbers and activity can be studied using platelets; ß receptors, using leukocytes. Both have also been explored in postmortem brain. α_2 Receptors are found both presynaptically and postsynaptically. Presynaptic α_2 receptors act as thermostats to regulate NE production and release. α_2 Receptors are universally connected to adenylate cyclase second messenger systems such that agonists inhibit cAMP formation. In contrast, ß receptors, which are entirely postsynaptic, stimulate adenylate cyclase and cAMP formation.

α_2 Receptor numbers and activity have been reported in multiple studies to be increased in the platelets of depressed patients (22 ,23), although there is also at least one negative study (24). α_2 Receptor activity can be explored by measuring cAMP responses to challenges with agonists. Mooney and associates (25) reported that epinephrine suppression of prostaglandin-E induced cAMP formation is decreased in the platelets of depressed patients. Siever and colleagues (26) reported norepinephrine stimulation results in blunted adenylate cyclase responses in an E_1–α_2 prostaglandin coupled model. Platelet aggregation that results from α_2 stimulation has also been reported to be altered in depressed patients (27). Mooney and colleagues (25) using stimulation of α_2 receptors with a variety of agents, including NaF, which directly affects G_1 coupled proteins have hypothesized that this down-regulation is not agonist specific and have argued that a fundamental uncoupling of the receptor–G-protein–AC complex occurs in some depressed patients.

Growth hormone (GH) responses to challenge with clonidine, an α_2 agonist, has also been employed as a functional test of α_1 activity. Consistent blunting of GH responses in depressed patients has been reported (28 ,29). Blunted GH response appears to be a trait marker; it is found in remitted patients (30). The significance of blunted GH responses to clonidine is not entirely clear, however. Clonidine could be affecting presynaptic or postsynaptic receptors (31). Also, somatostatin, an inhibitor of GH release, may play a role in the GH response to clonidine challenge.

α_2 Receptors have also been explored in postmortem brain of suicide victims. Increased binding sites have been reported in several studies (32 ,33 and 34), although findings regarding the specific isoform and location of the increased binding sites have not been consistent.

ß Receptors have been studied in both leukocytes and postmortem brain. Results have been less consistent than with the α_2 receptor. Decreased binding in leukocytes of depressives has been reported inconsistently (35 ,36). Similarly, in postmortem brain tissue, increased ß-adrenergic receptor density has been reported by Mann and colleagues (37); however, decreased Bmax was reported by Crow and co-workers (38) in the hippocampus of depressives. Effects of previous medication may enter into these discrepant findings, as could biological heterogeneity of catecholamine secretion in depressed patients.

Depletion Strategies

α-Methylparatyrosine (AMPT) inhibits TH and ultimately synthesis of norepinephrine. When given to healthy controls or depressives it does little to lower mood (39 ,40); however, remitted depressed patients on noradrenergic antidepressants show a worsening of symptoms when challenged with AMPT, suggesting that norepinephrine availability or tone is needed for maintaining response to NE agents (41). In contrast, patients on SSRIs do not relapse when challenged with AMPT challenge. AMPT in previously depressed patients who are not currently on medication causes a recurrence of depressive symptoms (42). Taken together, these data suggest the test could be a possible trait marker for depressive vulnerability and that maintaining NE tone is important for sustaining responses to noradrenergic drugs.

SEROTONIN

Part of "72 - Molecular and Cellular Mechanisms in Depression "

Serotonin (5-HT) is a monoamine neurotransmitter involved in mood and appetite regulation. In brain, it is synthesized within the raphe from l-tryptophan. Serotonin itself does not cross the blood–brain barrier. Synthesis includes an initial conversion to 5-hydroxytryptophan (5 HTP) via the enzyme, tryptophan hydroxylase. 5-HTP is decarboxylated by L-aromatic-amino acid decarboxylase to form 5-HT. The principal metabolite of 5-HT is 5-hydroxyindole acetic acid (5-HIAA), which is easily measurable in cerebrospinal fluid (CSF) and urine. MAO mediates part of the metabolism of serotonin.

Metabolite Studies

Much of the early interest in serotonin was generated by observations that low CSF 5-HIAA levels in hospitalized depressives were associated with an increased risk for suicide (43). Further studies revealed a relationship, particularly with violent methods of suicide (e.g., hanging) (44) and subsequently with difficulties with impulse control in subjects with antisocial personality (45). Current theories emphasize a more general relationship between low serotonin metabolite concentrations and impulse control problems; the latter may predispose to suicide in subjects who become depressed (46).

Transporter

The serotonin transporter (SERT), a 12-transmembrane molecule, actively pumps 5HT into the presynaptic neuron. Originally, the transporter was studied in platelets using tritiated (^{3}H) imipramine and more recently with the higher affinity (^{3}H)-paroxetine. Numerous studies have reported decreased binding (Bmax) in the platelets of depressed patients as compared to healthy controls. A metaanalysis by Ellis and Salmond (47) of 70 studies demonstrated an overall significant difference between patients and controls, although not all studies concur. Medication did not appear to account for these differences. Although mean values appear to differ between patients and controls, there is considerable overlap in values among patients and controls such that there are numbers of patients who do not appear to have decreased binding; this overlap limits the use of the test as a diagnostic measure.

Decreased ^{3}H imipramine binding was once thought to be a trait marker, that is, did not normalize with treatment. Further study, however, has revealed that decreased ^{3}H-imipramine binding does normalize with treatment but one must wait for sufficient periods to allow for protein regeneration.

The transporter has also been the subject of examination in postmortem brain tissue. Early studies in this area pointed to decreased binding in suicide brains (48); however, more recent studies have failed to confirm these findings (49). These data raise questions regarding the significance of abnormalities in the activity SERT in the pathophysiology of depression and the relationship of peripheral and central forms of the transporter. There are a number of methodologic problems inherent in postmortem tissue that may account for differences among studies, including accuracy of diagnosis, time from death to collection of brain tissue preparation of tissue, and so on.

One approach to studying the activity of the transporter has been to apply functional imaging (e.g., SPECT to determine relative activity). The development of ligands (e.g., ^{123}I-ß-CIT), that bind selectively to the transporter has allowed *in vivo* study in humans. In one study, a significant difference in binding using SPECT was observed between depressed patients and controls (50). In this study, significant differences were not observed in platelet binding to ^{3}H-paroxetine, raising questions regarding whether the transporter is regulated differently in the two tissues.

Receptors

Presynaptic and postsynaptic 5HT receptors have also been studied in depressed patients. Over a dozen serotonin receptors have been identified, although the possible roles for many have not. Two that have attracted most study for the longest periods are the $5HT_{1A}$ and $5HT_{2a}$ types.

$5HT_{2a}$ receptors are located postsynaptically in the CNS and can also be found in platelets as well as in other non-CNS tissue. As with the transporter, multiple studies have investigated $5HT_{2a}$ binding sites in the platelets of depressed patients. An increase in binding sites (B-max) has been reported in depressed and suicidal patients with some suggestion that increased binding in suicidal patients may be independent of a diagnosis of major depression (51 ,52 and 53). Generally, $5HT_{2a}$ binding has been thought to be a state marker, although one recent study has suggested binding may not normalize with SSRI treatment (54).

$5HT_{2a}$ binding has also been studied in postmortem brain tissue. As with the serotonin transporter, results here have been mixed with some studies demonstrating increased prefrontal cortical binding but others not (37 ,55 ,56 and 57). $5HT_{2a}$ receptors are found in frontal cortex suggesting a role in the cognitive aspects of depression.

PET ligands have been developed to study $5HT_{2a}$ activity in brain. One study employed [^{18}F]- altanserin and reported a reduction in activity in right posterolateral frontal, orbitofrontal, and anterior cingulate regions in depressives (58). In another study, no differences were found between nonsuicidal depressives and controls using [^{18}F-] setoperone (59). The exclusion of patients with recent suicidal ideation may have played a role in not finding differences between patients and controls. Studies on effects of antidepressants on $5HT_{2a}$ binding using PET have also yielded mix results. One group reported that SSRIs appear to increase ^{18}F-setoperone binding (60), whereas another recently reported that 3 to 4 weeks of desipramine treatment resulted in a significant decrease in $5HT_{2}$ activity in multiple areas, particularly in frontal cortex (61). This group was unable to conclude whether the ligand was binding to $5HT_{2a}$ or $5HT_{2c}$ receptors.

$5HT_{2A}$ receptors are coupled to the phosphoinositide second messenger system. When $5HT_{2a}$ receptors are activated by agonists, phosphatidyl inositol 4,5 bisphosphate is hydrolyzed by phospholipase C to form two second messengers, diacylglycerol and inositol 1,4,5-triphosphate. Protein kinase C is activated by diacylglycerol. This system has been studied in the brains of suicide victims. Pandey and associates reported [^{3}H] phorbol dibutyrate binding to protein kinase C in prefrontal cortex was lower in teenage suicide victims (62). More recently they observed that both phospholipase C activity and the $ß_{1}$ isozyme protein level were decreased of teenage suicide victims (63). Depression per se did not appear to affect the differences between suicide victims and controls. In contrast Hrdina and associates (64) reported unaltered protein kinase C activity in antidepressant free depressives who suicided, and Coull and colleagues (65) reported that phorbol dibutyrate binding sites were increased in the prefrontal cortex of adults with similar histories. Age, diagnosis of depression, antidepressant use, and time to collection of brain may play a role in these disparate findings.

The $5HT_{1a}$ autoreceptor controls release of serotonin from the presynaptic neuron. Over the past few years, multiple groups have explored the potential use of pindolol, a $5HT_{1a}$ antagonist, to hasten response to antidepressants or bring out responses in refractory patients. These studies have yielded mixed results suggesting that pindolol may hasten response to antidepressants in milder or first-episode patients seen in primary care settings. $5HT_{1a}$ receptor number and activity have been studied in postmortem brain. Increased $5HT_{1a}$ Bmax has been reported in suicide victims using nonviolent means compared to violent completers or controls but others have failed to find alterations in $5HT_{1a}$ activity in suicide victims (66 ,67 and 68).

Genetic Studies

A number of studies have explored the possible role of genetics may play, particularly vis-á-vis transporter activity. Long and short forms of the transporter gene appear to be relatively common in the general population. An early study indicated a relationship of the short form with an increased frequency of a variety of neurotic or behavioral traits (69). Allelic variation has also been applied to predicting drug response. In three studies in Europe and the United States, homozygotes or heterozygotes for the S-form were reported to show sluggish responses to SSRIs (70 ,71 and 72). The opposite was found in a Korean study (73). Clearly, further work is needed to understand the importance of genetic forms of the transporter in major depression. More recently, Mann and colleagues (69) reported that the short form genotype was associated with a diagnosis of major depression but not with suicide or 5HT-transporter binding in postmortem tissue.

Depletion Studies

Brain concentrations of serotonin are highly dependent on circulating levels of tryptophan, which competes with other

amino acids for transport into the brain. Charney and Delgado have pioneered in the use of an amino acid cocktail that is relatively devoid of L-tryptophan to rapidly decrease plasma tryptophan and ultimately brain serotonin. In these studies, the drink was first administered to subjects who had responded to various antidepressants and who were being maintained on medication. Diphenhydramine has been commonly used as the comparison cocktail. Euthymic patients on SSRIs but not TCAs rapidly experienced depressive symptoms when depleted of L-tryptophan, suggesting the need for maintaining adequate serotonin levels to ensure continued drug response (74 ,75). Parallel decreases in glucose utilization in frontal and thalamic regions using PET have also been reported in depressives who experience a relapse in symptoms (76). In contrast, there are multiple reports of depletion not causing a clear recurrence of symptoms in patients treated with bupropion or electroconvulsive therapy (75 ,77 ,78 and 79). Studies have used a variety of different methods (e.g., patients' being on or off medication, inclusion or exclusion of suicidal patients, etc.), and these differences may account for the discrepant findings. The degree and duration of response observed before the depletion challenge is administered may be of particular importance (79). Patients who are in remission or have shown a prolonged response are unlikely to demonstrate significant worsening of moods (79). These data suggest recent responders are those who are susceptible to experiencing relapse with depletion strategies. Depletion of unmedicated euthymic depressives does not appear to induce recurrence, indicating maintaining serotonin tone is important primarily for continuance of response in recently remitted patients (79).

Of interest is a recent report that women controls show much lower rates of 5HT synthesis and a greater decrease in response to depletion than men do (80). This gender-based difference is consistent with a recent observation that chronically depressed women are more responsive to an SSRI than men are (81).

Fenfluramine Challenge

Fenfluramine, previously marketed as an appetite suppressant, causes a release of serotonin from presynaptic neurons and results in an elevation of prolactin. Prolactin responses to fenfluramine challenge are blunted in depressed patients (82 ,83) and there are some data to suggest this may be a trait marker (84). However, bipolar manic and axis II patients may also demonstrate blunted prolactin responses, raising questions regarding the specificity of the test. (See refs. 1 and 2 for further review.)

DOPAMINE

Part of "72 - Molecular and Cellular Mechanisms in Depression "

As indicated, dopamine (DA) is a precursor for norepinephrine. Although NE has played a central role in etiologic theories of depression, DA has been emphasized far less in depression in spite of its being widely distributed in brain.

CSF levels of homovanillic acid (HVA), a major metabolite, are decreased in depressives (2 ,85 ,86), although some studies have reported elevated CSF DA, but not HVA levels (87). Urinary DOPAC levels are decreased in depressives compared with controls (88); in one study, DOPAC levels appeared associated with suicidal behavior (85). Dopaminergic agents such as psychostimulants, nomifensine, and the dopamine agonist pramipexole all have antidepressant effects in nondelusional patients.

In contrast, elevated mesolimbic DA activity has been hypothesized to play a role in delusional depression (89). Elevated CSF HVA levels have been associated with psychotic symptoms and agitation in major depression (89), and increased plasma DA and HVA levels have also been reported in delusional depression (90 ,91). Increased mesolimbic DA activity has been postulated to occur secondary to elevated hypothalamic–pituitary–adrenal (HPA) axis activity (89). Recent studies in rats, nonhuman primates, and psychotic depressives suggest elevated glucocorticoid activity could also result in altered or decreased prefrontal cortical dopamine metabolism and to alterations in attention and response inhibition (92 ,93). These data suggest increased HPA axis activity could affect DA turnover differently in specific brain regions—alterations that have been suggested in schizophrenia. Antipsychotic drugs appear to play a key role in treatment of delusional depression and glucocorticoid receptor antagonists are being actively studied in the disorder.

GABA

Part of "72 - Molecular and Cellular Mechanisms in Depression "

GABA has become a focus of greater study over the past several years with the increasing use of anticonvulsants in mood disorders. GABA is a major inhibitory neurotransmitter in brain and regulates seizure threshold as well as norepinephrine and dopamine turnover. There are two types of GABA receptors. $GABA_A$ receptors have been studied in anxiolysis because of their location within a benzodiazepine–GABA receptor complex that is coupled to chloride channels. $GABA_B$ receptors are coupled to Ca^{+2} channels. In rats, antidepressants and mood stabilizers appear to up-regulate frontal-cortical $GABA_B$, but not $GABA_A$, receptors (94 ,95). $GABA_B$ agonists may enhance cAMP responses to norepinephrine and ß-adrenergic down-regulation in response to tricyclic antidepressants, suggesting a facilitative role for $GABA_B$.

GABA is also enhanced by anticonvulsants such as valproic acid, which act as mood stabilizers. GABA levels have been reported to be decreased in the CSF of depressed patients in some but not all studies (96 ,97). Plasma GABA levels have also been reported to be lower in unipolar depressives (98 ,99), and this may not normalize with treatment (100).

Alcoholism can also be associated with low plasma GABA levels (101). In refractory depressed patients undergoing cingulotomy, GABA levels are inversely related to degree of depression (102). A number of groups are actively exploring using fMRI to image GABA in the brains of patients with mood disorders, both before and after treatment.

NEUROENDOCRINE SYSTEMS

Part of "72 - Molecular and Cellular Mechanisms in Depression "

Neuroendocrine systems were originally studied as gateways to the exploration of neurotransmitter activity, such as norepinephrine and serotonin, in depression. Over time, emphasis has shifted to exploring the roles components of several of these axes may play in the pathogenesis of specific symptoms or disease states. Three axes, hypothalamic–pituitary–adrenal (HPA), hypothalamic–pituitary–thyroid (HPT), human growth hormone (HGH), in particular have been examined in major depression.

HPA Axis

The hypothalamic–pituitary–adrenal (HPA) axis is frequently activated during periods of stress and depression. The axis consists of three major components: (a) corticotropin releasing hormone (CRH), which is located in the paraventricular nucleus of the hypothalamus and stimulates the pituitary to release adrenocorticotropin hormone (ACTH); (b) ACTH which stimulates the adrenal cortex to release cortisol; and (c) cortisol, which feeds back to the pituitary, hypothalamus, and hippocampus to decrease release of CRH and ACTH.

Multiple lines of evidence point to abnormalities of the axis in depression. Initial studies focused on excretion of cortisol and its precursors in patients with depression. Activation of the axis was also associated with suicidal ideation. Sachar in a classic study reported elevated serum cortisol levels over a 24-hour period in severely depressed patients (103). Elevations were particularly seen in the evening and overnight, times when the axis should be quiescent. These data generally were interpreted as indicating that the depressed patient was highly stressed.

One method for challenging the axis is to administer the synthetic steroid dexamethasone (DEX) (104). The expected response is to suppress the axis because the pituitary and perhaps the hypothalamus read the DEX signal as indicating sufficient glucocorticoid activity, and shuts off production or release of cortisol. Depressed patients demonstrate a significantly higher nonsuppression rate than do controls, although the rates of nonsuppression are relatively low in many studies (105). Patients with severe or psychotic depression demonstrate relatively high rates of nonsuppression or high postdexamethasone cortisol levels (106). Indeed, psychosis appears to be the greatest symptom or syndrome contributor to DEX nonsuppression, greater than the effect of severity or melancholia (107). Outpatients with milder and nonpsychotic disorders show much lower rates of nonsuppression. This difference in types of patients studied may help explain the variability in DEX nonsuppression rates across studies. DEX responses have also been used to assess adequacy of treatment with patients who are nonsuppressors after treatment showing a significantly increased risk for relapse (108).

Glucocorticoid overactivity has been hypothesized to play a direct role in the development of cognitive impairment and delusions in patients with psychotic major depression (89). Trials are currently underway exploring the efficacy of glucocorticoid antagonists in psychotic depression (109). Moreover, glucocorticoids have been hypothesized to cause increases in glutamate activity, decrease nerve growth factor activity, and hippocampal volume loss on MRI in patients with a history of severe depression, but there are no studies that have simultaneously explored these various dimensions in depressed patients (110). Recently, Rojkowska and colleagues did report that neuronal size and density and glial densities were reduced in prefrontal cortical regions in postmortem tissue from subjects with major depression as compared to controls (111). Overall, there has been a shift from viewing excessive glucocorticoid activity in major depression as an epiphenomenon to its having specific effects on cognition or symptom formation.

Study of more proximal components of the axis have also pointed to marked abnormalities in major depression. In most of the relevant studies, CRH levels have been reported to be elevated in the CSF or plasma of depressed patients (1 ,112). Challenge with ovine or human forms of CRH results in blunted ACTH responses in depressives suggesting increased central CRH release (113). Remission of episodes appears to be associated with normalization of CRH studies. Postmortem studies have reported that CRH mRNA expression was increased (114) and CRH Bmax was decreased in the frontal cortex of suicide victims (115). These data suggest CRH release from the hypothalamus may be associated with a down-regulation of CRH in other brain regions (2).

Imaging studies have reported increased pituitary and adrenal size during depression, which appear to normalize with recovery (116 ,117). Increased pituitary size and elevated CSF CRH levels are associated with DEX nonsuppression (118). Elevated plasma ACTH levels have been reported in psychotic depression (119).

ACTH release is not only stimulated by CRH. For example, arginine vasopressin (AVP) may enhance CRH's stimulation of ACTH. AVP neurons are increased in the PVN of suicide victims (120) and serum AVP has been reported in one study to be elevated in hospitalized depressives (121).

CRH is also found in extrahypothalamic brain regions. In the amygdala, CRH appears to play a key role in fear responses and over-activation of these systems may lead to panic and depression (2). Amygdala CRH has been reported

to be under positive (stimulatory) feedback by cortisol and this observation has spurred on much research to develop specific CRH antagonists to treat anxiety and depressive disorders. A recent report on an open label trial suggested that a CRH antagonist might be effective in hospitalized depressives (122).

Although the literature has emphasized elevated CRH and cortisol activity in major depression (in part because of the emphasis on DST nonsuppression), there is emerging evidence that CRH and cortisol activity may only be elevated in some subtypes of major depression and that some depressed patients may actually have low HPA activity. Recent data suggest that depressed patients with a history of early abuse (as well as those with psychosis) may be most consistently at risk for demonstrating elevated ACTH levels in response to social stress (123). Depressives who were not abused as children did not show similar responses. In a recent study, we reported *decreased* levels of ACTH or cortisol activity over 24 hours in nonpsychotic depressives as compared to controls (119). Similarly, low values have been reported in several other types of patients, including atypical depression, posttraumatic stress disorder, so-called burn out syndromes, and so on. Thus, both decreased and elevated HPA axis activity may be found in specific depressive subtypes. In many ways this parallels the findings in catecholamine activity in depressed patients.

This seeming contradiction in findings or emphasis over time may have several explanations. First, the DST as we use it may not measure cortisol overactivity as much as it does central CRH overdrive in response to suppressing the pituitary because DEX poorly penetrates brain at the doses used in the test. Second, previous studies have often not explored the role of psychosis or early abuse. Third, relatively few studies on the HPA axis in depression have explored cortisol activity over the full 24-hour period.

HPT Axis

The overlap in symptoms between patients with hypothyroidism and those with major depression has led to number of studies on the hypothalamic–pituitary–thyroid (HPT) axis in patients with mood disorders. These studies have yielded intriguing, although at times, conflicting results.

Thyrotropin-releasing hormone (TRH) is released from the hypothalamus and stimulates TRH receptors in the pituitary to release thyroid-stimulating hormone (TSH), which in turn stimulates specific receptors in the pituitary to release triiodothyronine (T_3) and thyroxine (T_4) hormones. Thyroid hormones provide feedback to both the hypothalamus and pituitary to regulate the axis.

Activity of the axis can be measured in several ways: circulating levels of T_3 and T_4—both bound and unbound; TRH levels in the CSF; TSH responses to TRH administration (TRH-stimulation test); and circulating TSH levels. In addition some patients demonstrate antibodies to thyroid tissue suggestive of an autoimmune thyroiditis, often in the face of normal T_4, T_3, or TSH levels.

TRH is found in extra-hypothalamic regions in brain. CSF TRH was increased in two small studies of depressed patients as compared to controls (124 ,125), although not all studies agree (126). Elevated TRH levels should be accompanied by a blunted TSH response to TRH because TRH levels in the pituitary would be expected to be down-regulated in the face of elevated TRH. Indeed, multiple studies have reported such blunting in a relatively high percentage (approximately 25%) of patients with major depression. A recent review concluded that 41 of 45 studies reported blunted TSH responses to TRH in major depression (127). Blunting of TSH responses to TRH in these patients is not owing to clinical or subclinical hypothyroidism because thyroid parameters were generally within normal limits in these patients.

Type I hypothyroidism is characterized by decreased levels of T_3 and/or T_4, increased TSH, and increased TSH responses to TRH (1). Antithyroid antibodies may be present. Type II hypothyroidism is characterized by normal T_3 or T_4 levels but otherwise similar abnormalities as in Type I disease (1). Rates of Type III or IV subclinical hypothyroidism have been reported to be elevated in depressed patients. These syndromes are both characterized by normal circulating levels of T_3, T_4, and TSH but have other abnormalities such as elevated TSH responses to TRH or the presence of antithyroid antibodies. In one study, depressed patients with high normal thyroid levels were also reported to demonstrate exaggerated TSH responses to TRH (128). These data have been interpreted as indicating some patients with major depression may have subclinical hypothyroidism. Indeed, asymptomatic autoimmune thyroiditis with positive antibodies has been reported to be relatively high (9% to 25%) in several surveys of depressed patients (127 ,129). Taken together, TSH stimulation test data suggest elevated or decreased TRH activity could be involved in major depression, depending on whether patients met criteria for subclinical hypothyroidism.

T_3 has been reported to be an effective augmentor of responses to antidepressants and appears to exert greater effects than does T_4 (130). Patients with a history of thyroid disease (e.g., adenoma) who were taking suppressing or replacement doses of thyroxine or T_4 demonstrated an improvement in mood and cognitive function when T_3—but not placebo—was added (131). One possible explanation for the differential effects of T_4 and T_3 on mood rests with local tissue conversion of T_4 to T_3 that in brain is mediated by Type II 5′ deiodinase and may be dysregulated in some patients. Depressed patients have been reported to demonstrate increased reversed T_3 levels in CSF (130), which suggests inhibition of the Type II 5′ deiodinase and subsequent increased activity of the Type III of the enzyme. Cortisol can inhibit Type II of the enzyme and may play a role in the increased rT_3 levels. Of interest is a recent report (133)

that in depressed patients low T_3 levels predicted earlier relapse, pointing further to an important role for T_3 in mood relation.

Transthyretin is a protein that transports thyroid hormone from the periphery to the brain via the choroid plexus. Transthyretin levels have been reported to be low in refractory depressed patients (134). This may also help explain possible CNS enhancing effects of T_3 in the face of normal circulating thyroid hormone levels.

Overall, research on the HPT axis has produced some intriguing leads for understanding the pathophysiology of depression and improving its treatment. However, there are still a number of seeming contradictions regarding the direction and specific nature of HPT alterations in depression. Data point to both elevations in central TRH activity and subtle forms of hypothyroidism (suggestive of low T_3 and TRH activity) as playing potential roles in major depression.

Human Growth Hormone

Growth hormone (GH) is synthesized in the anterior pituitary. Two hypothalamic hormones, growth hormone releasing factor (GRF) and somatostatin modulate its release from the pituitary. GRF is stimulating; in contrast, somatostatin inhibits release. Somatostatin is also found in extra-hypothalamic regions, and appears to act as a neurotransmitter. The major neurotransmitters involved in mood regulation (e.g., norepinephrine, serotonin, and dopamine) all affect GH release, and these systems can be challenged by specific compounds (e.g., apomorphine, clonidine, etc.).

Diurnal rhythms of GH, as measured in plasma, are disrupted in depression. Nocturnal GH is elevated in depression (135), but daylight-stimulated GH levels are increased in both unipolar and bipolar depressives (136).

As indicated, GH responses to clonidine are blunted in depression (28). GH responses to GRF have also been explored in patients with major depression with several, but not all, groups reporting blunted GH responses (137 ,138 and 139). CSF levels of somatostatin, which inhibits GH, CRH, and ACTH release, are also reduced in depression (140 ,141), such that somatostatin does not appear to provide an explanation for the blunted GH responses to GRF in depression. Low somatostatin levels in depression may reflect increased cortisol activity (1 ,142) and appear to normalize with treatment (2). Low CSF somatostatin has also been observed in various neurological disorders (e.g., Alzheimer's disease).

CONCLUSION

Part of "72 - Molecular and Cellular Mechanisms in Depression "

Proliferation of research into the biology of depression has resulted in a number of intriguing leads for understanding the pathophysiology of major depression. Most studies have focused on single biological systems such that there is a dearth of studies that simultaneously explore multiple systems and their complex interactions in depression. Also, research has tended to emphasize cross-sectional rather than longitudinal designs such that we have little understanding of the biological underpinnings of initiation, maintenance, and termination of depressive episodes. Future research that combines genetic risk factors with longitudinal study of multiple systems will likely lead to breakthroughs in our understanding of the biology of the disorder. Also, greater emphasis on the biology of specific depressive subtypes (e.g., delusional depression) or of symptom dimensions may provide greater insights.

REFERENCES

1. Musselman DL, DeBattista C, Nathan KI, et al. Biology of mood disorders. In: Schatzberg AF, Nemeroff CB, eds. *Textbook of psychopharmacology*, second ed. Washington, DC: The American Psychiatric Press, 1998:549–588.

2. Garlow SJ, Nemeroff CB. Neurobiology of depressive disorders. In: Davidson RJ, Post RM, eds. *Handbook of affective sciences*. New York: Oxford University Press, in press.

3. Schildkraut JJ. The catecholamine hypothesis of affective disorders: a review of supporting evidence. *Am J Psychiatry* 1965;122:509–522.

4. Bunney WE, Davis M. Norepinephrine in depressive reactions. *Arch Gen Psychiatry* 1965;13:137–152.

5. Schatzberg AF, Samson JA, Bloomingdale KL, et al. Toward a biochemical classification of depressive disorders. X: Urinary catecholamines, their metabolites, and D-type scores in subgroups of depressive disorders. *Arch Gen Psychiatry* 1989;46(3):260–268.

6. Maas JW, Fawcett JA, Dekirmenjian H. Catecholamine metabolism, depressive illness and drug response. *Arch Gen Psychiatry* 1972;26:252–262.

7. Schildkraut JJ, Orsulak PJ, Schatzberg AF, et al. Toward a biochemical classification of depressive disorders. I: Differences in urinary excretion of MHPG and other catecholamine metabolites in clinically defined subtypes of depression. *Arch Gen Psychiatry* 1978;35(12):1427–1433.

8. Muscettola G, Potter WZ, Pickar D, et al. Urinary 3-methoxy-4-hydroxyphenylglycol and major affective disorders: a replication and new findings. *Arch Gen Psychiatry* 1984;41:337–342.

9. Schatzberg AF, Orsulak PJ, Rosenbaum AH, et al. Toward a biochemical classification of depressive disorders. V: Heterogeneity of unipolar depressions. *Am J Psychiatry* 1982;139(4):471–475.

10. Rosenbaum AH, Maruta T, Schatzberg AF, et al. Toward a biochemical classification of depressive disorders. VII: Urinary free cortisol and urinary MHPG in depressions. *Am J Psychiatry* 1983;140:314–318.

11. Maj M, Ariano MG, Arena F, et al. Plasma cortisol, catecholamine and cyclic AMP levels, response to dexamethasone suppression test and platelet MAO activity in manic-depressive patients. A longitudinal study. *Neuropsychobiology* 1984;11(3):168–173.

12. Swann AC, Koslow SH, Katz MM, et al. Lithium carbonate treatment of mania. *Arch Gen Psychiatry* 1987;44(4):345–354.

13. Halaris A. 3-methoxy-4-hydroxyphenyl-glycol in manic psychosis. *Am J Psychiatry* 1978;135:493–494.

14. Maas JW, Fawcett JA, Dekirmenjian H. Catecholamine metabolism, depressive illness and drug response. *Arch Gen Psychiatry* 1972;26:252–262.

15. Schildkraut JJ. Norepinephrine metabolites as biochemical criteria for classifying depressive disorders and predicting response to treatment. Preliminary findings. *Am J Psychiatry* 1973;130:695–699.

16. Janicak PG, Davis JM, Chan C, et al. Failure of urinary MHPG levels to predict treatment response in patients with unipolar depression. *Am J Psychiatry* 1986;143:1398–1402.

17. Hollister LE, Davis KL, Berger PA. Subtypes of depression based on excretion of MHPG and response to nortriptyline. *Arch Gen Psychiatry* 1980;37:1107–1110.

18. Schatzberg AF, Samson JA, Schildkraut JJ, et al. Relationships between catecholamine and cortisol measures in depressed patients. Presented at the American College of Neuropsychopharmacology Annual Meeting. Maui, Hawaii, 1989.

19. Ordway GA. Pathophysiology of the locus ceruleus in suicide. *Ann NY Acad Sci* 1997;836:233–252.

20. Arango V, Underwood MD, Mann JJ. Fewer pigmented locus ceruleus neurons in suicide victims: preliminary results. *Biol Psychiatry* 1996;39(2):112–120.

21. Klimek V, Stockmeier C, Overholser J, et al. Reduced levels of norepinephrine transporters in the locus ceruleus in major depression. *J Neurosci* 1997;17(21):8451–8458.

22. Garcia-Sevilla JA, Udina C, Fuster MJ, et al. Enhanced binding of [^{3}H]-adrenaline to platelets of depressed patients with melancholia: effect of long-term clomipramine treatment. *Acta Psychiat Scand* 1987;75(2):150–157.

23. Halaris A, Piletz J. Platelet adrenoreceptor binding as a marker in neuropsychiatric disorders. *Abstr 17th CINP Congr* 1990;28.

24. Maes M, Van Gastel A, Delmeire L, et al. Decreased platelet alpha-2 adrenoceptor density in major depression: effects of tricyclic antidepressants and fluoxetine. *Biol Psychiatry* 1999;45(3):278–284.

25. Mooney JJ, Schatzberg AF, Cole JO, et al. Rapid antidepressant response to alprazolam in depressed patients with high catecholamine output and heterologous desensitization of platelet adenylate cyclase. *Biol Psychiatry* 1988;23:543–559.

26. Siever LJ, Kafka MS, Targum SM, et al. Platelet alpha adrenergic binding and biochemical responsiveness in depressed patients and controls. *Psychiatry Res* 1984;11:287–302.

27. Garcilla-Sevilla JA, Pardo D, Giralt MT, et al. Alpha-2-adrenoreceptor-mediated inhibition of platelet adenylate cyclase and induction of aggregation in major depression: effect of long-term antidepressant drug treatment. *Arch Gen Psychiatry* 1990;47:125–132.

28. Siever LJ, Uhde TW, Jimerson DC, et al. Differential inhibitory noradrenergic responses to clonidine in 25 depressed patients and 25 normal control subjects. *Am J Psychiatry* 1984;141(6):733–741.

29. Amsterdam JD, Maislin G, Skolnick B, et al. Multiple hormone responses to clonidine administration in depressed patients and healthy volunteers. *Biol Psychiatry* 1989;26(3):265–278.

30. Siever LJ, Terestmen RL, Coccaro E, et al. The growth hormone response to clonidine in acute and remitted depressed male patients. *Neuropsychopharmacology* 1992;6(3):165–177.

31. Schatzberg AF, Schildkraut JJ. Recent studies on norepinephrine systems in mood disorders. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:911–920.

32. Meana JJ, Barturen F, Garcia-Sevilla JA. Alpha 2-adrenoceptors in the brain of suicide victims: increased receptor density associated with major depression. *Biol Psychiatry* 1992;31(5):471–490.

33. De Paermentier F, Mauger JM, Lowther S, et al. Brain alpha-adrenoceptors in depressed suicides. *Brain Res* 1997;757(1):60–68.

34. Ordway GA. Pathophysiology of the locus ceruleus in suicide. *Ann NY Acad Sci* 1997;836:233–252.

35. Extein I, Tallman J, Smith CC, et al. Changes in lymphocyte beta-adrenergic receptors in depression and mania. *Psychiatry Res* 1979;1(2):191–197.

36. Healy D, Carney PA, O'Halloran A, et al. Peripheral adrenoceptors and serotonin receptors in depression. Changes associated with response to treatment with trazodone or amitriptyline. *J Affect Dis* 1985;9(3):285–296.

37. Mann JJ, Stanley M, McBride PA, et al. Increased serotonin 2 and beta-adrenergic receptor binding in the frontal cortices of suicide victims. *Arch Gen Psychiatry* 1986;43:954–959.

38. Crow TJ, Cross AJ, Cooper SJ, et al. Neurotransmitter receptors and monoamine metabolites in the brains of patients with Alzheimer-type dementia and depression, and suicides. *Neuropharmacology* 1984;23:1561–1569.

39. Salomon RM, Miller HL, Krystal JH, et al. Lack of behavioral effects of monoamine depletion in healthy subjects. *Biol Psychiatry* 1997;41(1):58–64.

40. Miller HL, Delgado PL, Salomon RM, et al. Effects of alpha-methyl-para tyrosine (AMPT) in drug-free depressed patients. *Neuropsychopharmacology* 1996;14(3):151–157.

41. Miller HL, Delgado PL, Salomon RM, et al. Clinical and biochemical effects of catecholamine depletion on antidepressant-induced remission of depression. *Arch Gen Psychiatry* 1996;53(2):117–128.

42. Berman RM, Narasimhan M, Miller HL, et al. Transient depressive relapse induced by catecholamine depletion: potential phenotypic vulnerability marker? *Arch Gen Psychiatry* 1999;56(5):395–403.

43. Asberg M, Traskman L, Thoren P. 5-HIAA in the cerebrospinal fluid: a biochemical suicide predictor? *Arch Gen Psychiatry* 1976; 33:1193–1197.

44. Traskman L, Asberg M, Bertilsson L, et al. Monoamine metabolites in CSF and suicidal behavior. *Arch Gen Psychiatry* 1981;38(6):631–636.

45. Linnoila VM, Virkkunen M, Scheinin M, et al. Low cerebrospinal fluid 5-hydroxyindoleacetic acid concentration differentiates impulsive from nonimpulsive violent behavior. *Life Sci* 1983;33:2609–2614.

46. Linnoila VM, Virkkunen M. Aggression, suicidality and serotonin. *J Clin Psychiatry* 1992;54(Suppl):46–51.

47. Ellis PM, Salmond C. Is platelet imipramine binding reduced in depression? A meta-analysis. *Biol Psychiatry* 1994;36:292–299.

48. Perry EK, Marshall BG, Tomlinson BE, et al. Decreased imipramine binding in the brains of patients with depressive illness. *Br J Psychiatry* 1983;142:188–192.

49. Little KY, McLaughlin DP, Ranc J, et al. Serotonin transporter binding sites and mRNA levels in depressed persons committing suicide. *Biol Psychiatry* 1997;41(12):1156–1164.

50. Malison RT, Price LH, Berman R, et al. Reduced brain serotonin transporter availability in major depression as measured by [^{123}I]-2-beta-carbomethoxy-3-beta-(4-iodophenyl)tropane and single photon emission computed tomography. *Biol Psychiatry* 1998;44(11):1090–1098.

51. Arora R, Meltzer H. Increased serotonin1 (5-HT2) receptor binding as measured by 3H-lysergic diethylamide (^{3}H-LSD) in the blood platelets of depressed patients. *Life Sci* 1989;44:725–734.

52. Biegon A, Essar N, Israeli M, et al. Serotonin 5-HT2 receptor binding on blood platelets as a state dependant marker in major affective disorder. *Psychopharmacology* 1990;102:73–75.

53. Pandey GN, Pandey SC, Janicak, et al. Platelet serotonin-2 receptor binding sites in depression and suicide. *Biol Psychiatry* 1990;28:215–222.

54. Bakish D, Cavazzoni P, Chudzik J, et al. Effects of selective serotonin reuptake inhibitors on platelet serotonin parameters in major depressive disorders. *Biol Psychiatry* 1997;41(2):184–190.

55. Arango V, Ernsberger P, Marzuk P, et al. Autoradiographic demonstration of increase serotonin 5-HT2 and beta-adrenergic receptor binding sites in the brain of suicide victims. *Arch Gen Psychiatry* 1990;47:1038–1047.

56. Stanley M, Mann JJ. Increased serotonin-2 binding sites in frontal cortex of suicide victims. *Lancet* 1983;1:214–216.

57. Hrdina P, Emeter E, Vu T, et al. 5-HT uptake sites and 5-HT_2 receptors in brain of antidepressant-free suicide victims/depressives. Increase in 5-HT_2 sites in cortex and amygdala. *Brain Res* 1993;614:37–44.

58. Biver F, Wikler D, Lotstra F, et al. Serotonin 5-HT2 receptor imaging in major depression: Focal changes in orbito-insular cortex. *Br J Psychiatry* 1997;171:444–448.

59. Meyer JH, Kapur S, Houle S, et al. Pefrontal cortex 5-HT2 receptors in depression: An [18F] setoperone PET imaging study. *Am J Psychiatry* 1999;156(7):1029–1034.

60. Massou JM, Trichard C, Attar-Levy D, et al. Frontal 5-HT2A receptors studied in depressive patients during chronic treatment by selective serotonin reuptake inhibitors. *Psychopharmacology* 1997;133(1):99–101.

61. Yaham LN, Liddle PF, Dennie J, et al. Decrease in brain serotonin 2 receptor binding in patients with major depression following desipramine treatment. *Arch Gen Psychiatry* 1999;56:705–711.

62. Pandey GN, Dwivedi Y, Pandey SC, et al. Protein kinase C in postmortem brain of teenage suicide subjects. *Neurosci Lett* 1997;228:111–114.

63. Pandey GN, Dwivedi Y, Pandey SC, et al. Low phosphoinositide-specific phospholipase C activity and expression of phospholipase C B_1 protein in the prefrontal cortex of teenage suicide subjects. *Am J Psychiatry* 1999;156:1895–1901.

64. Hrdina P, Faludi G, Li Q, et al. Growth-associated protein (GAP43), its mRNA, and protein Kinase C (PKc) izoenzymes in brain regions of depressed suicides. *Mol Psychiatry* 1988;3:411–418.

65. Coull MA, Lowther S, Katona CLE, et al. Altered brain protein kinase C in depression: a post-mortem study. *Eur Neuropsychopharmacol* 2000;10:283–288.

66. Matsubara S, Arora RC, Meltzer HY. Serotonergic measures in suicide brain: 5-HT1A binding sites in frontal cortex of suicide victims. *J Neural Transm* 1991;85:181–194.

67. Lowther S, Cheetham SC, Crompton MR, et al. 5-HT1A receptor binding sites in post-mortem brain samples from depressed suicides and controls. *J Affect Dis* 1997;42(2 and 3):199–207.

68. Stockmeier CA, Dilley GE, Shapiro LA, et al. Serotonin receptors in suicide victims with major depression. *Neuropsychopharmacology* 1997;16:162–173.

69. Mann JJ, Huang X-Y, Underwood MD, et al. A serotonin transporter gene promoter polymorphism (5-HT TLPR) and prefrontal cortical binding in major depression and suicide. *Arch Gen Psychiatry* 2000;57:729–738.

70. Smeraldi E, Zanardi R, Benedetti F, et al. Polymorphism within the promoter of the serotonin transporter gene and antidepressant efficacy of fluvoxamine. *Mol Psychiatry* 1998;3(6):508–511.

71. Zanardi R, Benedetti F, Di Bella D, et al. Efficacy of paroxetine in depression is influenced by a functional polymorphism within the promoter of the serotonin transporter gene. *J Clin Psychopharmacol* 2000;20(1):105–107.

72. Pollock BG, Ferrell RE, Mulsant BH, et al. Allelic variation in the serotonin transporter promoter affects onset of paroxetine treatment response in late-life depression. *Neuropsychopharmacology* 2000;23(5):587–590.

73. Kim DK, Lim SW, Lee S, et al. Serotonin transporter gene polymorphism and antidepressant response. *Neuroreport* 2000;11(1):215–219.

74. Delgado PL, Price LH, Miller HL, et al. Serotonin and the neurobiology of depression. Effects of tryptophan depletion in drug-free depressed patients. *Arch Gen Psychiatry* 1994;51(11):865–874.

75. Delgado PL, Miller HL, Salomon RM, et al. Tryptophan-depletion challenge in depressed patients treated with desipramine or fluoxetine: implications for the role of serotonin in the mechanism of antidepressant action. *Biol Psychiatry* 1999;46(2):212–220.

76. Bremner JD, Innis RB, Salomon RM, et al. Positron emission tomography measurement of cerebral metabolic correlates of tryptophan depletion-induced depressive relapse. *Arch Gen Psychiatry* 1997;54(4):364–374.

77. Delgado PL, Price LH, Miller HL, et al. Rapid serotonin depletion as a provocative challenge test for patients with major depression. *Psychopharmacol Bull* 1991;27:320–330.

78. Cassidy F, Murray E, Weiner RD, et al. Lack of relapse with tryptophan depletion following successful treatment with ACT. *Am J Psychiatry* 1997;154:1151–1152.

79. Moore P, Landolt HP, Seifritz E, et al. Clinical and physiological consequences of rapid tryptophan depletion. *Neuropsychopharmacology* 2000;23:601–622.

80. Nishizawa S, Benkelfat C, Young SN, et al. Differences between males and females in rates of serotonin synthesis in human brain. *Proc Natl Acad Sci USA* 1997;94:5308–5313.

81. Kornstein SG, Schatzberg AF, Thase ME, et al. Gender differences in treatment response to sertraline versus imipramine in chronic depression. *Am J Psychiatry* 2000;157:1445–1452.

82. Mitchell P, Smythe G. Hormonal responses to fenfluramine in depressed and control subjects. *J Affect Dis* 1990;19(1):43–51.

83. O'Keane V, Dinan TG. Prolactin and cortisol responses to d-fenfluramine in major depression: evidence of diminished responsivity of central serotonergic function. *Am J Psychiatry* 1991;148(8):1009–1015.

84. Flory JD, Mann JJ, Manuck SB, et al. Recovery from major depression is not associated with normalization of serotonergic function. *Biol Psychiatry* 1998;43(5):320–326.

85. Roy A, Karoum F, Pollock S. Marked reduction in indexes of dopamine transmission among patients with depression who attempted suicide. *Arch Gen Psychiatry* 1991;49:447–450.

86. Reddy PL, Khanna S, Subash MN, et al. CSF amine metabolites in depression. *Biol Psychiatry* 1992;31:112–118.

87. Gjerris A, Werdelin L, Rafaelson OJ, et al. CSF dopamine increased in depression: CSF dopamine, noradrenaline and their metabolites in depressed patients and controls. *J Affect Dis* 1987;13(3):279–286.

88. Roy A, Pickar D, Douillet P, et al. Urinary monoamines and monoamine metabolites in subtypes of unipolar depressive disorder and normal controls. *Psychol Med* 1986;16(3):541–546.

89. Schatzberg AF, Rothschild AJ, Langlais PJ, et al. A corticosteroid/dopamine hypothesis for psychotic depression and related states. *J Psychiatr Res* 1985;19(1):57–64.

90. Devanand DP, Bowers MB, Hoffman FJ, et al. Elevated plasma homovanillic acid in depressed females with melancholia and psychosis. *Psychiatry Res* 1985;15(1):1–4.

91. Rothschild AF, Schatzberg AF, Langlais PJ, et al. Psychotic and non psychotic depressions. I: Comparison of plasma catecholamines and cortisol measures. *Psychiatry Res* 1987;20:143–153.

92. Lindley SE, Bengoechea TG, Schatzberg AF, et al. Glucocorticoid effects on mesotelencephalic dopamine neurotransmission. *Neuropsychopharmacology* 1999;21:399–407.

93. Lyons DM, Lopez JM, Yang C, et al. Stress-level cortisol treatment impairs inhibitory control of behavior in monkeys. *J Neurosci* 2000;20:7816–7821.

94. Lloyd KG, Thuret F, Pilc A. Upregulation of gamma-aminobutyric acid (GABA) B binding sites in rat frontal cortex. *J Pharmacol Exp Ther* 1985;235(1):191–199.

95. Kimber JR, Cross JA, Horton RW. Benzodiazepine and GABA receptors in rat brain following chronic antidepressant drug administration. *Biochem Pharmacol* 1987;36(23):4173–4175.

96. Gold BI, Bowers MB Jr, Ruth RH, et al. GABA levels in CSF of patients with psychiatric disorders. *Am J Psychiatry* 1980;137(3):362–364.

97. Roy A, Dejong J, Ferraro T. CSF GABA in depressed patients and normal controls. *Psychol Med* 1991;21(3):613–618.

98. Petty F, Schlesser MA. Plasma GABA in affective illness, a preliminary investigation. *J Affect Dis* 1981;3:339–343.

99. Petty F, Kramer GL, Gullion CM, et al. Low plasma gamma-aminobutyric acid levels in male patients with depression. *Biol Psychiatry* 1992;32:354–363.

100. Berretini WH, Nurnberger JI Jr, Hare TA, et al. Reduced plasma and CSF gamma-aminobutyric acid in affective illness: effect of lithium carbonate. *Biol Psychiatry* 1983;18:185–194.

101. Petty F, Sherman AD. Plasma GABA levels in psychiatric illness. *J Affect Disord* 1984;6:131–138.

102. Honig A, Bartlett JR, Bouras N, et al. Amino acid levels in depression: a preliminary investigation. *J Psychiatric Res* 1989;22:159–164.

103. Sacher E, Hellman L, Fukushima D, et al. Cortisol production in depressive illness. *Arch Gen Psychiatry* 1970;23:289–298.

104. Carroll BJ. Use of the dexamethasone test in depression. *J Clin Psychiatry* 1982;43:44–50.

105. Arana GW, Baldessarini RJ, Ornsteen M. The dexamethasone suppression test for diagnosis and prognosis in psychiatry. *Arch Gen Psychiatry* 1985;42:1193–1204.

106. Schatzberg AF, Rothschild AJ, Stahl JB, et al. The dexamethasone suppression test: Identification of subtypes of depression. *Am J Psychiatry* 1983;140:88–91.

107. Nelson JC, Davis JM. DST studies in psychotic depression: a meta-analysis. *Am J Psychiatry* 1997;154:1497–1503.

108. Ribeiro SC, Tandon R, Grunhaus L. the DST as a prediction of outcome in depression: a meta-analysis. *Am J Psychiatry* 1993;150:1618–1629.

109. Belanoff J, Flores B, Kalezhan M et al. Rapid reversal of psychotic major depression using mifepristone (RU 486). *J Clin Psychopharmacol* 2001;21:516–521.

110. Sheline Y. 3D MRI studies of nueroanatomic changes in unipolar major depression: the role of stress and medical comorbidity. *Biol Psychiatry* 2000;48:791–800.

111. Rajkowska G. Postmortem studies in mood disorders indicate altered numbers of neurons and glial cells. *Biol Psychiatry* 2000;48:766–777.

112. Nemeroff CB, Widerlov E, Bissette G, et al. Elevated concentrations of CSF corticotropin-releasing factor-like immunoreactivity in depressed patients. *Science* 1984;226:1342–1344.

113. Gold PW, Loriaux DL, Roy A, et al. Responses to corticotropin releasing hormone in the hypercortisolism of depression and Cushing's disease: pathophysiologic and diagnostic implications. *N Engl J Med* 1986;314:1329–1335.

114. Raadsheer FC, van Heerikhuize JJ, Lucassen PJ, et al. Corticotropin-releasing hormone mRNA levels in paraventricular nucleus of patients with Alzheimer's disease and depression. *Am J Psychiatry* 1995;152(9):1372–1379.

115. Nemeroff CB, Owens MJ, Bissette AC, et al. Reduced corticotropin releasing factor binding sites in the frontal cortex of suicide victims. *Arch Gen Psychiatry* 1988;45(6):577–579.

116. Krishnan KRR, Doraiswamy PM, Lurie SN, et al. Pituitary size in depression. *J Clin Endocrinol Metab* 1991;72:256–259.

117. Rubin RT, Phillips JJ, Sadow TF. Adrenal gland volume in major depression: increase during the depressive episode and decrease with successful treatment. *Arch Gen Psychiatry* 1995;52:213–218.

118. Axelson DA, Doraiswamy PM, Boyko OB, et al. In vivo assessment of pituitary volume using MRI and systemic stereology: relationship to dexamethasone suppresion test results in patients with affective disorder. *Psychiatry Res* 1992;46:63–70.

119. Posener JA, DeBattista C, Williams GH, et al. Twenty-four hour monitoring of cortisol and corticotropin secretion in psychotic and nonpsychotic major depression. *Arch Gen Psychiatry* 2000;57:755–760.

120. Purba JS, Hoogendijk WJG, Hofman MA, et al. Increased number of vasopressin- and oxytocin-expressing neurons in the paraventricular nucleus of the hypothalamus in depression. *Arch Gen Psychiatry* 1996;53:137–143.

121. Van Londen L, Goekoop JG, Van Kempen GMJ. Plasma levels of arginine vasopressin elevated in patients with major depression. *Neuropsychopharmacology* 1997;17:284–292.

122. Zobel AW, Nickel T, Kunzel HE, et al. Effects of the high-affinity corticotrophin-releasing hormone receptor 1 antagonist R121919 in major depression: the first 20 patients treated. *J Psychiatric Res* 2000;34:171–181.

123. Heim C, Newport DJ, Heit S, et al. Pituitary-adrenal and autonomic responses to stress in women after sexual and physical abuse in childhood. *JAMA* 2000;284:592–597.

124. Banki CM, Bissette G, Arato M, et al. Elevation of immunoreactive CSF TRH in depressed patients. *Am J Psychiatry* 1988;145:1526–1531.

125. Kirkegaard C, Faber J, Hummer L, et al. Increased levels of TRH in cerebrospinal fluid from patients with endogenous depression. *Psychonueroendocrinology* 1979;4:227–235.

126. Roy A, Wolkowitz OM, Bissette G, et al. Differences in CSF concentrations of thyrotropin-releasing hormone in depressed patients and normal subjects: negative findings. *Am J Psychiatry* 1994;151(4):600–602.

127. Esposito S, Prange AJ, Golden RN. The thyroid axis and mood disorders: overview and future prospects. *Psychopharmacol Bull* 1997;33:205–217.

128. Kraus RP, Phoenix E, Edmonds MW, et al. Exaggerated TSH responses to TRH in depressed patients with “normal” baseline. *J Clin Psychiatry* 1997;58(6):266–270.

129. Nemeroff CB, Simon JS, Haggerty J Jr, et al. Antithyroid antibodies in depressed patients. *Am J Psychiatry* 1985;142:840–843.

130. Joffe RT, Singer W. A comparison of triiodothyronine and thyroxine in the potentiation of antidepressants. *Psychiatry Res* 1990;32:241–252.

131. Bunevicius R, Kazanavicius G, Zalinkevicius R, et al. Effects of thyroxine as compared with thyroxine plus triiodothyronine in patients with hypothyroidism. *N Engl J Med* 1999;340:424–429.

132. Linnoila VM, Cowdry R, Lanberg BA, et al. CSF triiodothyronine (rT_3) levels in patients with affective disorders. *Biol Psychiatry* 1983;18:1489–1492.

133. Joffe RT, Marriott M. Thyroid hormone levels and recurrence of major depression. *Am J Psychiatry* 2000;157:1689–1691.

134. Hatterer JA, Herbert J, Hidaka C, et al. CSF transthyretin in patients with depression. *Am J Psychiatry* 1993;150:813–815.

135. Schilkrut R, Chandra O, Osswald M, et al. Growth hormone during sleep and with thermal stimulation in depressed patients. *Neuropsychobiology* 1975;1:70–79.

136. Mendlewicz J, Linkowski P, Kerkhofs M, et al. Diurnal hypersecretion of growth hormone in depression. *J Clin Endocrinol Metab* 1985;60:505–512.

137. Lesch KP, Laux G, Erb A, et al. Attenuated growth hormone response to growth hormone RH in major depressive disorder. *Biol Psychiatry* 1987;22:1495–1499.

138. Contreras F, Navarro MA, Menchon JM, et al. Growth hormone response to growth hormone releasing hormone in non-delusional and delusional depression and healthy controls. *Psychol Med* 1996;26:301–307.

139. Krishnan KR, Manepalli AN, Ritchie JC, et al. Growth hormone-releasing factor stimulation test in depression. *Am J Psychiatry* 1988;145:190–192.

140. Agren H, Lundquist G. Low levels of somatostatin in human CSF mark depressive episodes. *Psychoneuroendocrinology* 1984;9:233–248.

141. Rubinow DR, Gold PW, Post RM, et al. CSF somatostatin in affective illness. *Arch Gen Psychiatry* 1983;40:409–412.

142. Rubinow DR. Cerebrospinal fluid somatostatin and psychiatric illness. *Biol Psychiatry* 1986;21:341–365.

73

Neurocircuitry of Mood Disorders

Gregory A. Ordway

Violetta Klimek

J. John Mann

Gregory A. Ordway and Violetta Klimek: Department of Psychiatry and Human Behavior, University of Mississippi Medical Center, Jackson, Mississippi.

J. John Mann: Department of Neuroscience, New York State Psychiatric Institute, New York, New York.

One of the first neurochemical theories of depression was the monoamine deficiency hypothesis (139 ,143 ,153). Over the past 30 years, this hypothesis has been the most scrutinized of any theories regarding the biology of depression. Unfortunately, the biology of depression remains an elusive issue, despite intense biological research. It is widely held that most, if not all, antidepressant drug treatments produce their therapeutic antidepressant effects, at least in part, by modulating monoamine systems (noradrenergic, serotonergic, and dopaminergic); however, less is known about the neurochemical pathology of these monoamine systems in depression. Early attempts to evaluate monoamine systems in depressive disorders led to diverse and not clearly integrated findings. As a result, many other neurochemical theories have been generated in efforts to explain the biological basis of depression. These theories include HPA axis hyperactivity (111), the GABA hypothesis (132), the galanin hypothesis (186), the substance P hypothesis (82), the glutamate hypothesis (162), the neurotrophin hypothesis (39), and many others. A substantial portion of the evidence supporting these “other neurotransmitter” theories derives from studies of the pharmacologic and behavioral effects of antidepressant drugs in laboratory animals. Of course, these antidepressant drugs have prominent actions on norepinephrine (NE), serotonin (5HT), and to a lesser extent, dopamine (DA). Hence, originators of new hypotheses are continuously forced to place new theories in the context of the old monoamines. The principal reason for this is that, despite years of pharmaceutical development, drugs with primary actions on monoamine systems remain the mainstay of treatment for depressive disorders. In fact, evidence that there has been an improvement of medication over the past 20 years is highly debated, in terms of greater efficacy or even faster onset of action. If improvements are evident in antidepressant medications, then they are as a result of a reduction of adverse side effects with newer antidepressant compounds, rather than novel pharmacologic mechanisms with enhanced activity. For this reason, as research on depression biology progresses into this new century, the monoamine hypotheses continue to be among the most popular biological theories and continue to be heavily investigated and debated.

Much of the past 10 years of research in the biology of mood disorders has led to advancements in our understanding of the role that monoamines play in these disorders. New modern approaches have been applied, including the use of *in vivo* imaging techniques in live patients, morphologic and neurochemical investigations with high levels of anatomic resolution, use of postmortem brain tissues from psychiatrically characterized subjects, and genetic studies. Considerable evidence has accumulated implicating multiple system pathology in mood disorders, including abnormalities of monoamine as well as other neurotransmitter systems. These approaches and findings have led researchers to propose broader theories regarding depression biology (e.g., depression as a spreading neuronal adjustment disorder or limbic–cortical dysregulation disorder) (67 ,102). In this chapter, the authors reconcile new findings of multiple system pathologies specifically with regard to monoaminergic systems. Emphasis is placed on the cellular sources of monoamine systems and their circuitry, the communication between these monoamine nuclei, and the influence of other neurotransmitter systems that are putatively disrupted in depression on monoaminergic neuronal activity.

NORADRENERGIC CIRCUITRY AND DEPRESSION

Part of "73 - Neurocircuitry of Mood Disorders "

The original speculation that NE is deficient in depression hinged partly on clinical observations of depression in some

individuals receiving reserpine for hypertension. Reserpine depletes brain monoamines by blocking vesicular monoamine storage; however, only a fraction of individuals administered reserpine develops depression. In fact, short-term depletion of NE by administration of alpha-methyl-*p*-tyrosine to normal control subjects does not result in a significant change in mood (152). These findings, along with a history of inconsistent findings regarding levels of NE and its metabolites in depressive disorders (19 ,138 ,147 ,180), casted doubt about NE's role in depression for many years. Recently, however, Charney and co-workers (28) demonstrated that rapid pharmacologic depletion of NE in patients taking noradrenergic antidepressants causes a rapid relapse of depression. This latter finding demonstrates that NE is critical to the therapeutic action of noradrenergic antidepressant drugs. Most recently, and of considerable significance, is the demonstration that depletion of catecholamines (NE and DA) in unmedicated fully remitted subjects with histories of major depression resulted in relapse into depression (21). Together, these studies demonstrate that acute NE depletion is insufficient to induce depressive symptoms by itself, but depletion in a susceptible individual induces depressive symptoms.

The neurobiological relationship between NE and mood is poorly understood. Most of the NE in the brain arises from cell bodies in the locus ceruleus (LC). The projections of these neurons are diffuse and overlapping with respect to the brain regions innervated. Some of the brain regions that are most densely innervated by the LC are limbic brain regions, including the amygdala and hippocampus. The LC is part of the reticular activating system and neurons of the LC have tonic pacemaker activity. LC activity is elevated during states of arousal. In contrast, LC activity slows during sleep and is inactive during random eye movement (REM) sleep. The LC is robustly activated by stress, contributing to the alerting of the organism to stimuli relevant to survival. In addition to innervation of forebrain regions, the LC densely innervates other monoaminergic nuclei, including the serotonergic raphe nuclei and the dopaminergic ventral tegmental area (VTA) (10 ,116). Given the large number of brain regions innervated by the LC, noradrenergic transmission is in an ideal position to globally modulate brain function, and modulate the activity of other monoamines. Foote and colleagues (51) and Aston-Jones and associates (10) have extensively reviewed the physiologic consequences of LC activation.

The hypothesis that NE plays a role in the neurochemical pathology of depression raises the possibility that the activity and biochemistry of the noradrenergic LC is abnormal in this illness. Recent studies using postmortem brain tissues from psychiatrically characterized subjects reveal a complex pathology of the noradrenergic LC in major depression. These studies are unique with respect to many previous postmortem research studies on the noradrenergic system in depression because of the focus on neurons that are the source of NE and because brain tissues utilized were from subjects whose psychiatric status was rigorously characterized (80 ,195). Prominent among exclusion criteria for subjects is the absence of any antidepressant (or antipsychotic) drug use, determined both by next-of-kin interview and from a toxicology examination. Elevated amounts of tyrosine hydroxylase (TH) (195) elevated binding to α_2-adrenoceptors (119 ,122) and reduced amounts of NE transporter binding (80) have been reported in the LC of major depressive subjects as compared to psychiatrically normal control subjects. In contrast, other proteins measured in the LC of major depressives appear to occur in normal amounts (e.g., monoamine oxidase A, MAO-A) (118). Interestingly, a lower number of noradrenergic neurons have been observed in the rostral LC from victims of suicide relative to normal control subjects (5). In contrast, Klimek and co-workers (80 ,195) report no differences in noradrenergic neuron cell counts in the middle to caudal portion of the LC between depressed suicide victims and psychiatrically normal control subjects, or in psychiatrically uncharacterized suicide victims compared to control subjects (120). Elevation of radioligand binding to some (3), but not all, noradrenergic receptors (4 ,79) has also been identified in some projection areas of the LC, comparing suicide victims to control subjects.

To interpret these postmortem findings, it is very useful, if not necessary, to utilize information from studies of NE in laboratory animals. In animals, stress activates the LC (130) and exposure to repeated stress elevates the demand for NE (110 ,124) revealed by increases in LC TH (105 ,175 ,183). Uncontrollable shock, a stress-based animal model of depression, increases the release of NE (187) and reduces NE stores (187). Because TH is the enzyme catalyzing the rate-limiting step in the synthesis of NE, increased tyrosine hydroxylase in the LC may adjust the set point of basal activity in the NE system, to keep pace with increased demand. Similar examples of up-regulation of TH expression in the LC occur after administration of reserpine (31) and intraventricular infusion of 6-hydroxydopamine (106), both of which cause loss of brain NE. Pharmacologic depletion of NE also up-regulates binding to α_2-adrenoceptors (60 ,172) and down-regulates binding to NE transporter (86). Hence, postmortem findings of LC biochemistry (e.g., elevated tyrosine hydroxylase and α_2-adrenoceptor binding) and reduced NE transporter binding are predictive of premortem increases in LC activity and decreases of NE availability. Decreased NE availability could also contribute to compensatory up-regulation of ß-adrenoceptors in LC projection areas, such as has been observed in the frontal cortex from suicide victims (3). As mentioned, the link between reduced brain NE and depressive symptoms, at least in susceptible individuals, has been made (21). Moreover, the relevance of stress-induced biochemical abnormalities in the LC is underscored by studies demonstrating a relationship between life stress and development of depression (26), and

that stress plays a role in the etiology of depression (131). A shortcoming of the cited postmortem studies, however, is that most of the depressive subjects who have been studied died as a result of suicide, and the relationship between the biological abnormalities found in the central noradrenergic system and behaviors related to suicide that are distinct from those related to depression has not been investigated.

If one accepts that biological abnormalities in the noradrenergic LC are relevant to the symptoms of depression, then it follows that treatment with antidepressant drugs might reverse these abnormalities. Again, it is presently necessary to look to laboratory animal studies to examine this issue. Repeated treatment of rats with antidepressants from many different pharmacologic classes (including 5HT uptake inhibitors), but not with non-antidepressant drugs, down-regulates LC tyrosine hydroxylase (114). Antidepressant drug treatment blocks stress-induced elevation of tyrosine hydroxylase mRNA in the rat LC (154). Repeated antidepressant drug treatment also down-regulates α_2-adrenoceptors in the rat LC (154). These findings demonstrate that repeated antidepressant treatment down-regulates tyrosine hydroxylase and α_2-adrenoceptors, proteins that are apparently up-regulated in the LC of human major depressives. Returning to the suggestion that LC activity may be elevated in depressives, recent studies demonstrate that repeated treatment of rats with antidepressant drugs of many different pharmacologic classes (including 5HT uptake inhibitors) reduces LC activity (58 ,68). Hence, animal data strongly support the contention that drugs produce antidepressant effects, at least in part, by reducing demand for NE, that is, reducing biochemical measures of demand (reduced biosynthetic enzyme for NE) and reducing LC activity.

In summary, evidence of: 1) norepinephrine depletion-induced depression in susceptible human subjects, 2) abnormal levels of noradrenergic proteins in the LC of human major depressives, 3) the ability of antidepressant medication to produce effects that would be expected to reverse noradrenergic pathology in depression provide strong support for the venerable theory that norepinephrine plays a role in the pathobiology of depression.

SEROTONERGIC CIRCUITRY AND DEPRESSION

Part of "73 - Neurocircuitry of Mood Disorders "

The biological basis for the indoleamine hypothesis was similar to that for the catecholamine hypothesis (178). That is, reserpine depletes not only catecholamines, but also 5HT. Since that time, several (but not all) studies have found reduced levels of CSF 5-hydroxyindole acetic acid (5-HIAA) in depressed patients (101 ,179); however, the degree of reduction of CSF 5-HIAA does not correlate with severity of depression. Oddly, many antidepressant medications, particularly 5HT reuptake inhibitors and monoamine oxidase inhibitors (MAOIs), reduce CSF 5-HIAA, possibly because of feedback inhibition resulting from increased synaptic concentrations of 5HT. Levels of CSF 5-HIAA are lower in depressed patients with a history of serious suicidal behavior, as compared to depressed patients with no history of suicide attempts. (See ref. 99 for review.) CSF 5-HIAA levels appear to exhibit a bimodal distribution in depressed patients. CSF 5-HIAA is not distinguished by more severe depression, but by a history of serious suicide attempts (55). Rapid tryptophan depletion causes transient, mild, nonclinical increases in negative mood in healthy young men (36). In depressed patients who had recent therapeutic responses to antidepressant medications, tryptophan depletion causes a transient depressive relapse (35). Rapid tryptophan depletion of many, but not all, patients with a history of depression and that are antidepressant drug-free causes a depressive relapse (35 ,164). In symptomatic, medication-free patients with depression, tryptophan depletion causes no significant behavioral effects (36), perhaps because of a floor effect (36). Together, these findings suggest that depression is often, but not always, associated with a serotonergic deficit.

A number of neuroendocrine challenge tests have demonstrated impaired serotonergic activity in depressed patients (49 ,56 ,69), although conflicting findings have also been reported. (See ref. 98 for review.) Numerous researchers have utilized postmortem brain tissues to study the serotonergic system in depression. Measurement of both $5HT_{1A}$ and $5HT_{2A}$ receptors in the prefrontal cortex from suicide victims has yielded no clear conclusions. Most of the studies cited in a recent review of these data by Stockmeier and colleagues (165) utilized tissues primarily from victims of suicide. Because there is an association of serotonergic abnormalities with suicidal risk, it is difficult to determine what effects are attributable to the presence of major depression from those associated with suicidal behavior (98). Significant increases in $5HT_{2A}$ receptor binding and a decrease in 5-HIAA in major depressives dying of causes other than suicide have been reported (47 ,103).

Given substantial evidence that 5HT plays a role in the pathology of major depression, it is expected that neurons supplying affected areas of the brain would display neuronal or neurochemical pathology. Two major nuclei in the brain, from which the majority of brain serotonergic innervation originates, are the dorsal raphe and median raphe nuclei. These nuclei provide an extensive innervation of cortical and subcortical target areas. The dorsal and median raphe nuclei give rise to separate axonal pathways to different brain regions. For example, the septum and hippocampus are innervated predominantly by the median raphe nuclei. In contrast, the striatum and substantia nigra are innervated by the dorsal raphe nuclei. Serotonergic terminals densely innervate various components of the limbic system. The widespread innervation of the brain by serotonergic neurons is the anatomic basis for the influence of 5HT on many diverse brain functions.

Several recent studies have investigated the serotonergic raphe in depression. Using single photon emission computed tomography (SPECT), Malison and associates (97) reported a decrease in 5HT transporter availability in the brainstem of living subjects with major depression. Because of issues related to spatial resolution, it is difficult to conclude from this study that the reduction in 5HT transporter occurred in raphe nuclei and/or other brainstem nuclei where the 5HT transporter occurs (e.g., substantia nigra or VTA). Little and co-workers (90) found no significant change in mRNA for the 5HT transporter in the dorsal raphe and median raphe nuclei from depressed persons who had committed suicide. Consistent with these findings, Bligh-Glover and colleagues (25) found no significant differences between depressed suicide victims and normal control subjects in [^{3}H]paroxetine binding to the 5HT transporter in the entire dorsal raphe or in its constituent subnuclei, as determined using postmortem tissues from psychiatrically characterized subjects. An increase in radioligand binding to $5HT_{1A}$ autoreceptors in dorsal raphe nuclei from depressed suicide victims has been observed (166). In apparent contrast, a decrease in the binding potential to $5HT_{1A}$ receptors in the midbrain raphe nuclei has been observed using positron emission tomography (PET) in patients with familial mood disorder (38). Recent studies also provide evidence of morphologic abnormalities of brainstem serotonergic nuclei. Underwood and associates (173) have demonstrated elevated numbers and densities of 5HT neurons in the dorsal raphe of suicide victims, most of whom had major depression. In addition, Becker and colleagues (18) have demonstrated significantly low echogenicity of the dorsal raphe nucleus in patients with major depression using a novel transcranial ultrasound technique. Together, these findings are strongly suggestive of a neuropathologic involvement of brainstem serotonergic nuclei in depression, but the study by Underwood and associates ruled out a loss of serotonergic neurons in depressed suicides, suggesting that the postulated hypofunction of the serotonergic system is not owing to fewer serotonergic neurons, but dysfunction of serotonergic neurons. As is the case with noradrenergic pathology, the specificity of serotonergic pathology for major depression versus suicidal behavior is yet to be clarified.

Evidence of a serotonergic deficit in depression predicts that drugs that are effective antidepressants should enhance serotonergic transmission. In fact, repeated treatment of rats with antidepressant drugs results in a net enhancement of serotonergic transmission (24). This effect is regardless of the primary pharmacologic site of action of the drug and includes selective 5HT transporter inhibitors, MAOIs, tricyclic antidepressants, and electroconvulsive shock. Selective 5HT transporter inhibitors and MAOIs enhance serotonergic transmission by desensitizing the somatodendritic $5HT_{1A}$ autoreceptors (23 ,24) and enhancing responsiveness of postsynaptic $5HT_{1A}$ receptors (63). Chronic administration of some tricyclic antidepressants or a course of electroconvulsive shock to rats does not appear to desensitize somatodendritic autoreceptors, although these treatments enhance the responsiveness of postsynaptic 5HT receptors (24 ,107). Hence, the mechanism by which different antidepressant drugs regulate serotonergic activity appears to differ, but the net effect of enhancing serotonergic transmission is similar. These preclinical findings are consistent with the hypothesis that there is a deficit in serotonergic transmission in depressive disorders that is normalized or corrected by antidepressant drug administration.

DOPAMINERGIC CIRCUITRY AND DEPRESSION

Part of "73 - Neurocircuitry of Mood Disorders "

Since the discovery that tricyclic antidepressant drugs can block DA reuptake *in vitro* (66), and that elevation of the functional activity of DA has antidepressant efficacy (142 ,143), there has been interest in the potential role of DA in the pathophysiology of depression. The contribution of DA to emotion-laden behaviors such as reward seeking, motivation, and environmental responsiveness also raises speculation that DA plays a role in the pathobiology of depression (48 ,167). In fact, clinical, pharmacologic, and laboratory animal evidence suggests that dopaminergic neurotransmission is decreased in depression. Lower concentrations of homovanillic acid (HVA), a DA metabolite, have been observed in CSF of patients with depression, and depression-inducing effects of DA-depleting agents or DA antagonists have been reported (143 ,144 ,189). In contrast, agents that enhance DA transmission, at least in part, such as bupropion, nomifensine, and amineptine, exert antidepressant effects in humans. Given that DA is intimately involved in motivational process and affect (73 ,167), these findings suggest that a deficiency of mesolimbic and/or mesocortical DA is a leading candidate for the etiology of core symptoms of depression, such as difficulty in the experience of pleasure (anhedonia), social isolation, loss of motivation (lack of interest), and psychomotor retardation (190).

At least three DA systems putatively involved in neurologic and psychiatric disorders have been extensively characterized in the brain: the nigrostriatal, mesolimbic, and mesocortical systems. A loss of nigrostriatal DA neurons causes the motor impairment of Parkinson's disease (PD), whereas dysfunction or activation of mesolimbic and/or mesocortical DA systems are implicated in psychiatric disorders, including depression, schizophrenia, and psychostimulant drug abuse disorders. However, some overlap in the pathology of PD and psychiatric disorders apparently occurs because cell loss in the VTA (in addition to substantia nigra) has been observed in patients with PD who have complications of co-morbid mood and cognitive disorders (171).

Anatomic, electrophysiologic, and neurochemical studies have delineated reciprocal pathways linking various limbic

and cortical regions with dopaminergic brainstem nuclei. Kalivas and Nakamura described the neuronal circuit that mediates the integration of reward perception and adaptive behavioral responses (75). This circuit includes the nucleus accumbens, amygdala, prefrontal cortex, mediodorsal thalamus, ventral pallidum, and midbrain neurons located in the VTA. Of brain limbic structures, the nucleus accumbens (ventral striatum) has been considered an important anatomic substrate of psychiatric illness, because of its established role in motivation and affect (167 ,75). Neurons in the nucleus accumbens receive a highly compressed input from the amygdala, hippocampus, cingulate gyrus, and prefrontal cortex (68 ,194). Ascending to synapse onto the same neurons in the nucleus accumbens are DA-containing fibers from the VTA (68), suggesting that the nucleus accumbens may integrate information coming from the prefrontal cortex and limbic regions with those originating from the VTA. Besides projecting to the nucleus accumbens, DA neurons ascending from the VTA project to other limbic structures, including discrete regions of amygdala, to cortical areas, and to the septum (116). Prefrontal, orbitofrontal, and cingulate cortices receive robust innervation from the VTA. Interestingly, most of the areas receiving DA projections from the VTA project back to the VTA.

If DA neurotransmission were disrupted in depression, then antidepressant drug treatment would be expected to produce effects on the brain dopaminergic system. Numerous studies demonstrate that antidepressant drugs enhance mesolimbic DA activity. Repeated treatment of rats with antidepressant drugs (tricyclics, mianserin, or citalopram) enhances DA agonist-induced locomotor hyperactivity, an effect observed when DA agonists are administered either systemically or injected directly into the nucleus accumbens (94 ,95 and 96). It is noteworthy that stereotypy (a behavioral effect reflecting the activity of nigrostriatal system) induced by D-amphetamine or apomorphine, is not increased by repeated treatment with antidepressant drugs (156); therefore, it has been assumed that the mesolimbic DA system mediates the increased behavioral responses to DA agonists following antidepressant treatments. Consistent with this effect, antidepressant drug treatment increases the affinity of D_2 receptors for their agonist in the limbic forebrain, but not in the striatum (78) and chronic treatment with antidepressant drugs results in postsynaptic DA receptor supersensitivity in the nucleus accumbens (40). Recent autoradiography studies confirm these findings by showing that when [^{3}H]raclopride, an antagonist at $D_{2/3}$ receptors, is used as a radioligand, no significant differences in the density of $D_{2/3}$ receptors are observed after chronic antidepressant drug administration. In marked contrast, when [^{3}H]quinpirole, an agonist at $D_{2/3}$ receptors is used as a radioligand, a significant increase in its binding is observed in the caudate and NAC of antidepressant-treated rats (146).

The level of mRNA encoding the D_1 receptor and [^{3}H]SCH 23390 binding to D_1 receptors are decreased in the limbic regions following these antidepressant drug treatments (42 ,37). A lower density of D_1 receptors induced by chronic antidepressant medication might contribute to enhancement of D_2 receptor functions as a result of a reduction in the inhibitory interactions between these two receptors at the level of the βγ subunit of G proteins (182).

Behavioral models in laboratory animals also point to a role for DA in antidepressant drug action. Following exposure to uncontrollable foot shock, an animal model of depression, rats display a pronounced reduction of responding for electrical brain stimulation of the nucleus accumbens. This response is attenuated by repeated treatment with the antidepressant drug desipramine (193). Rats exposed to chronic mild stress, another animal model of depression, experience decreased responsiveness to rewards (anhedonia), which is antidepressant-reversible (191). These behavioral changes are accompanied by lower $D_{2/3}$ receptor binding in the limbic forebrain that is reversed by 5 weeks of imipramine treatment (127). Overall, preclinical findings imply that a putatively important pharmacologic effect of antidepressant treatment is the augmentation of mesolimbic DA activity.

A few recent studies have measured the DA receptors in depressed patients *in vivo* using brain imaging techniques. Two studies demonstrate an increase of D_2 receptor density in the striatum in depression, possibly reflecting reduced DA function and a consequential up-regulation of these receptors (33 ,157). On the other hand, Ebert and associates, 1996 (43) found striatal D_2 receptor binding unchanged in major depression.

INTERACTIONS BETWEEN THE MONOAMINE NUCLEI AND MONOAMINES AND OTHER NEUROTRANSMITTERS

Part of "73 - Neurocircuitry of Mood Disorders "

Abnormalities of the biochemistry of one or more monoamine systems may cause depressive disorders. Alternatively, disrupted monoamine biochemistry may be secondary to other root biological, environmental, and/or psychological causes. A multitude of experimental approaches will be required to determine the core cause(s) of depressive disorders, even as considerable evidence of monoamine dysfunction in depression accumulates. Nevertheless, it is interesting to consider the relationship of the monoamines with other neurotransmitters that modulate monoaminergic chemistry. The LC, raphe nuclei and VTA receive a variety of neuronal inputs, including the monoamines themselves, that regulate their activity (Fig. 73.1). Several neurotransmitter inputs to monoamine nuclei are of particular relevance to major depression because of the accumulation of evidence that these systems are also disrupted in depression. For example, abnormalities in GABA, substance P, corticotropin releasing factor (CRF) and glutamate neurochemistry have been implicated

in depression. Theoretically, disruption of the activity of monoamine systems could result from disease-related abnormalities in neurotransmitter input to the monoamine nuclei. In addition, disruption of serotonergic activity in depression could be secondary to deficient noradrenergic input to the serotonergic raphe nuclei. That is, disruption of one monoamine would be expected to result in general monoamine imbalance because of the inter-connectivity between these nuclei. Hence, the relationship between the monoamine systems, and between monoamines and other neurotransmitter systems suspected of playing a role in depression biology, is worthy of discussion.

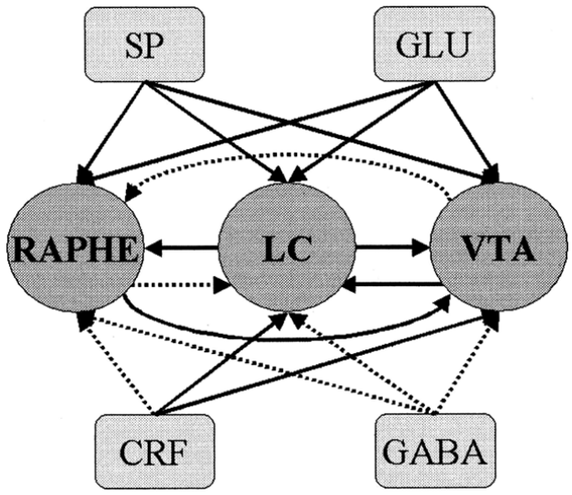


FIGURE 73.1. Neurotransmitter interactions at the level of monoaminergic cell bodies. Solid lines represent excitatory inputs to the raphe nuclei, LC, and VTA; dashed lines represent inhibitory inputs. In some cases, neurotransmitter inputs may be both direct and indirect via synapses with other neurons projecting to the nuclei. Neurotransmitters inputs shown are only those that are discussed in this chapter.

Monoamine Interactions at the Level of the Monoaminergic Nuclei

The monoaminergic nuclei are highly interconnected and physiologically integrated. For example, noradrenergic neurons innervate the serotonergic raphe nuclei and the dopaminergic VTA. Both the dorsal and median raphe nuclei receive noradrenergic innervation (15 ,121). In fact, the rat dorsal raphe receives one of the richest noradrenergic innervations in the brain (15 ,89 ,150). Overall, NE appears to be excitatory at serotonergic raphe neurons. For example, interruption of noradrenergic transmission by systemic administration of an α-adrenoceptor antagonist or iontophoretic application of an α-adrenoceptor antagonist in the vicinity of serotonergic neurons completely suppresses their spontaneous firing (14). Iontophoretic application of NE during the suppression of serotonergic cell activity produced by phentolamine or WB-4104, antagonists of α-adrenoceptors, can rapidly restore firing of these neurons to their normal activity (14). Most of the noradrenergic innervation of the VTA arises from LC neurons and noradrenergic input to the VTA is excitatory, mediated by excitatory α_1-adrenoceptors (59). In rats, chemical denervation of noradrenergic projections by DSP_4 treatment suppresses mesolimbic DA release (83) and reduces the effectiveness of positive reinforcers (109). Hence, if noradrenergic transmission is reduced, as has been hypothesized to occur in major depression, then reduced noradrenergic input to the raphe nuclei and VTA would be expected to contribute to reductions in serotonergic and dopaminergic transmission (109).

The LC and the VTA receive serotonergic terminals originating in the raphe nuclei. Serotonergic innervation to the LC originates from several sources including the dorsal and median raphe nuclei (93 ,108 ,181). The effects of 5HT on the activity of the LC are complex, and depend on whether drugs used to manipulate the serotonergic system are administered directly into the LC or whether they are administered systemically. Systemic administration of $5HT_1$ and $5HT_2$ receptor agonists or antagonists modulates the activity of the LC. These effects appear to be mediated indirectly, at least in part, rather than by actions at these receptors within the LC (30 ,57 ,64). 5HT affects LC activity, at least in part, by attenuating glutamatergic activation of the LC (8 ,159). Messenger RNAs encoding $5HT_{1A}$, $5HT_{1C}$, and $5HT_{2C}$ receptors are found in rat LC neurons (137 ,192). Interestingly, neurotoxic destruction of serotonergic terminals results in an increase in firing of the LC (71) and increases in LC mRNA and tyrosine hydroxylase activity (71). Based on these findings, the overall effect of 5HT release in the LC appears to be inhibitory.

The VTA receives afferent projections from 5HT-containing axon terminals originating in the dorsal and median raphe nuclei (70). Moreover, 5HT neurons innervate both dopaminergic and nondopaminergic (e.g., GABA) neurons in the VTA and may influence mesocortical and mesolimbic efferent systems through synaptic as well as nonsynaptic mechanisms (70). 5HT-induced release of [^{3}H]DA from rat VTA slice preparations is blocked by methysergide, but not cyproheptadine, suggesting an involvement of the $5HT_1$ receptor (17). Local application (in the VTA) of agonists at 5-HT_{1A} receptors increases the firing of DA neurons in the VTA (6 ,87) and administration of a $5HT_{1A}$ receptor agonist systemically increases DA release in the medial PFC (7). Microinfusion of 5HT into the VTA in rats results in an increased release of DA in the NAC (61). It is tempting to speculate that the firing mode of VTA DA neurons is dependent, among other factors, on the activity of serotonergic terminals originating in the raphe. Because 5HT increases extracellular DA (128), a serotonergic deficit, which has been suggested as one primary abnormality in depression, would also lead to a DA deficit.

High levels of DA are found in the dorsal raphe and LC (89). Moreover, D_2 receptors and D_2 receptor mRNA are expressed in both regions (185). Lesions of the VTA cause LC DA levels to fall by about 50%. DA neurotransmission is important for the rewarding effects of LC stimulation, without which such stimulation appears to be aversive (41). In the dorsal raphe nuclei, a moderate number of DAimmunoreactive fibers cover rather homogenously all subdivisions of the region (136). It has been postulated that dopaminergic neurotransmission to the dorsal raphe inhibits the activity of dorsal raphe neurons by increasing extracellular concentrations of 5HT in the dorsal raphe and, consequently, by increasing somatodendritic 5HT autoreceptor stimulation in this nucleus (46).

Monoamine Systems and Other Neurotransmitters

CRF

Much evidence has accumulated implicating a state of CRF hypersecretion in major depression (112 ,113 ,125). Interestingly, the LC, raphe nuclei and VTA receive moderate to dense innervation by CRF neurons. The LC receives excitatory CRF input from several sources, and these afferents appear to be topographically organized with respect to the type of information conveyed (177). The nucleus paragigantocellularis and Barrington's nucleus sends afferents directly to the nuclear elements of the LC. The LC also receives CRF input from limbic brain regions, including the central nucleus of the amygdala, as well as the bed nucleus of the stria terminalis and hypothalamic subregions (177). These limbic CRF neurons project to the peri-cerulea area, and in particular to the rostrolateral peri-LC. CRF terminals form direct contacts with noradrenergic dendrites (176). CRF, injected intracerebroventrically or directly into the LC, activates LC neurons and enhances release of NE in projection areas (163). Internal and external stressors are known to activate the LC via CRF, including colonic distension, hypotensive challenge, and foot shock. The ability of these stressors to activate the LC is blocked by CRF antagonists (32 ,85 ,104 ,174). In general, the pontine-medullary CRF projections to the LC are thought to coordinate cognitive and autonomic responses to internal physiologic challenges, whereas the limbic CRF projections mediate LC activation by external stressors that have emotional content (81). Interestingly, administration of a CRF antagonist blocks stress-induced increases in LC tyrosine hydroxylase (104), an effect shared by antidepressant drugs (105).

The serotonergic raphe nuclei also receive CRF innervation. CRF terminals in raphe nuclei originate from local and distant cell bodies (148 ,151). The effects of CRF on raphe firing are complex (77). At low doses, CRF produces primarily inhibitory effects on raphe discharge. In contrast, higher doses of CRF excite raphe neurons. Likewise, the effects of intracerebroventricularly administered CRF on striatal 5HT release are biphasic (140). Low doses of CRF decrease 5HT release in the striatum, whereas high doses increase striatal 5HT. Price and associates (140) suggest that CRF has predominantly inhibitory actions at the level of the raphe. Hence, a putative hypersecretion of CRF in major depression could contribute to the deficit in serotonergic transmission at the level of the raphe nuclei.

The VTA is densely innervated by CRF-positive fibers, whereas the substantia nigra receives only scattered CRF innervation (11). Intracerebroventricular administration of CRF to mice produces behavioral activation and a “stress-like” increase in DA metabolism in several brain regions. Direct injections of CRF into the VTA produces dose-dependent increase in locomotor activity, an affect that is not antagonized by the DA receptor blocker, haloperidol (74). Intracerebroventricular or intraperitoneal administration of low doses of CRF increases DA and DA metabolite levels in the rat medial prefrontal cortex (84). Together, these findings suggest that CRF exerts an excitatory action in the VTA. The long-term effects of CRF administration on DA metabolism have not been studied.

Substance P

Recent studies suggest that substance P antagonists may have antidepressant properties (82), although there have been questions regarding their efficacy (45). Interestingly, there is a relatively dense network of substance P immunoreactive fibers in the human LC and surrounding regions (50). Many of these fibers may originate from the nucleus of the solitary tract (50 ,100 ,145). In addition, there is a high density of binding of radiolabeled substance P to neurokinin-1 receptors in the LC (34). Substance P potently stimulates the firing of LC neurons (62). There is considerable evidence that substance P plays a role in the central response to stress (13 ,65). Interestingly, substance P antagonists (in particular, selective neurokinin-1 receptor antagonists), when administered intracerebroventricularly, attenuate restraint stress-induced biochemical indices of LC activation (65). Repeated administration of rats with antidepressant drugs (perhaps not all types) down-regulates substance P in several brain regions (27 ,158).

Substance P is co-localized with 5HT in 25% to 50% of the neurons in the human median and dorsal raphe nuclei, respectively (12 ,155). Substance P-containing serotonergic neurons are not randomly located within the raphe nuclei, but are localized to specific subregions, suggesting that substance P co-releases with 5HT in specific brain regions. The dorsal raphe nuclei also receive innervation from substance P-containing neurons with cell bodies occurring outside the region of the raphe (92). There is a high density of substance P receptors in the region of the dorsal raphe nuclei (91). Substance P appears to activate raphe neurons and microinjection

of substance P into the dorsal raphe increases hippocampal levels of 5HT.

Substance P receptor mRNAs (NK1 and NK3) are found in DA neurons of the human and rat midbrain (188); substance P-immunoreactive terminals, making synaptic contacts with TH-positive neurons in the VTA, also have been demonstrated (169). Infusion of a substance P receptor agonist into the VTA stimulates locomotor activity and increases DA turnover in the nucleus accumbens (149), indicating an excitatory action of substance P on DA neurotransmission.

Glutamate

NMDA receptor antagonists have antidepressant actions in animal models of depression (129) and demonstrated antidepressant effects in humans (20). High levels of serum glutamate levels in depressed subject have been reported (2 ,76) with exception (1). In addition, alterations in the allosterism of NMDA receptor binding in the frontal cortex of suicide victims (115), and elevated levels of CSF glutamine (glutamate metabolite/precursor) in depressed patients have been reported (88). Such findings have led to speculation that there may be excessive glutamate neurotransmission in depressive disorders. Glutamatergic neurons provide the major excitatory neurotransmitter input to the LC. Glutamatergic innervation of the LC derives largely from the nucleus paragigantocellularis (9). Glutamate activates the LC through activation of both NDMA and non-NMDA (aspartate) receptors (117). Handling and immobilization stress increases glutamate measured in the rat LC by microdialysis (161 ,170). Interestingly, noise stress-induced enhancement of glutamate release in the LC is abolished by superfusion of the LC with a CRF antagonist (160), demonstrating an important interaction between CRF and glutamate systems at the level of the LC (88). It is tempting to speculate that a deficit in noradrenergic transmission in major depression is secondary to a chronic elevation in glutamatergic input into the LC and a resulting depletion of central NE.

The raphe nuclei also receive glutamatergic input. At least part of the glutamatergic input to the dorsal raphe nuclei originates in the habenula (72). As is the case for the LC, glutamate is excitatory in the raphe nuclei. The activity of DA neurons in the mesolimbic and mesocortical circuitry can also be modulated by excitatory amino acids (73). DA neurons in the VTA receive direct glutamatergic innervation from the prefrontal cortex (73). Glutamate excites DA cell activity via inotropic and metabotropic receptors (167).

GABA

There is considerable preclinical and clinical evidence that depression is associated with reduced GABA function. Petty has reviewed this topic (135). To summarize, plasma GABA is low in patients with major depression (133 ,134 and 135). GABA agonists have activity in animal models useful for identifying antidepressants (16 ,196). Finally, GABA agonists appear to have some antidepressant activity in humans (135).

GABA provides a major inhibitory input to the LC. GABAergic neurons arriving in the LC originate largely from the nucleus prepositus, stimulation of which inhibits the firing of LC neurons (44). There are apparently no GABA cell bodies intrinsic to the LC, but glutamic acid decarboxylase immunoreactive nerve terminals are present, closely juxtaposed to noradrenergic cell bodies and dendrites (22). GABA inhibits the firing of LC neurons primarily by activation of $GABA_A$ receptors (123), and these receptors have been autoradiographically identified in the LC (29 ,126). The dorsal raphe nuclei receive GABAergic innervation from local interneurons and from multiple distant sources (54 ,184) and dorsal raphe neurons express $GABA_A$ receptors (53). Iontophoretic application of GABA strongly inhibits the firing of dorsal raphe nuclei neurons (52). DA neurons in the VTA are innervated by GABAergic afferents projecting mainly from the forebrain. GABA terminals also synapse on GABA interneurons that themselves synapse onto DA neurons (73). GABA inhibits the activity of DA neurons by acting through GABA receptors ($GABA_B$) on DA neurons (167).

INTEGRATION OF MONOAMINE AND OTHER NEUROTRANSMITTER THEORIES

Part of "73 - Neurocircuitry of Mood Disorders "

Investigations of the neurochemical pathology of depressive disorders reveal abnormalities in monoamine systems as well as other neurotransmitter systems. Nevertheless, it is conceivable that a root cause of depression is a failure or deficit in a single neurotransmitter system. Because of the interconnectivity of the monoamine systems, it is likely that failure in one system to adequately respond to demand would quickly lead to compensations, or possibly failure, of the other monoamine systems, as well as changes and/or biochemical compensations of numerous systems that are directly regulated by the pathogenic neurotransmitter system. Hence, evidence suggesting that there is dysfunction of monoaminergic, as well as nonmonoaminergic, neurotransmitter systems in depression compels us to integrate neurotransmitter interactions into theoretical models of the neurochemical circuitry of depression. This is a difficult undertaking and requires translation and integration of clinical, preclinical, and basic research findings. The postulate that depression is associated with a deficiency of NE or an increase in the demand for NE (as discussed) provides a good example of concatenation of clinical/postmortem findings, experimental/laboratory animal findings, and basic research findings regarding transmitter interactions. That is, elevated tyrosine hydroxylase in the LC, as observed in major depressive suicide victims, can be experimentally produced

by pharmacologically depleting NE or chronically stressing rats. Depletion of 5HT can also up-regulate tyrosine hydroxylase in the LC, as can chronic administration of CRF. Interestingly, CRF is reported to be elevated in depression, is released by stress, and CRF excites LC neurons. Together, these data suggest that elevated CRF in depression increases demand for NE, probably leading to elevated tyrosine hydroxylase expression. Elevated CRF may also contribute to reduce serotonergic transmission in depression, given the CRF can inhibit dorsal raphe neurons. Furthermore, it is conceivable that other excitatory inputs to the LC, such as substance P, might also exhibit elevated activity in depression. If so, substance P antagonists with antidepressant actions may elicit their effects on mood, at least in part, through actions at the LC. In contrast to excitatory transmitters, elevated demand for NE could also result from reduced inhibitory input to the LC. Here, of interest is the putative association of low levels of GABA with major depression (discussed in the preceding) and the fact that GABA provides an inhibitory input to the LC.

Another interesting possibility is that disruption of a single neurotransmitter system may be common to depressive disorders, whereas different mood disorders or subtypes of major depression itself may be a result of the type of altered neuronal input to that particular system. Using the hypothesis of increased demand for NE in depression as an example, elevated activity of the LC may be a result of elevated CRF input to the LC in some depressives, whereas others may experience elevated LC activity as a result of overactive substance P or glutamate input or decreased GABA input. Presently, there is little evidence to support the idea of serotonergic or noradrenergic depressives that respond selectively to serotonergic or noradrenergic antidepressant drugs, respectively. However, NE and 5HT containing neurons may be downstream of disrupted input systems that may actually differentiate subtypes of depressive disorders at a neurochemical/neuroanatomic level. Future research on neuropathology of psychiatric disorders should benefit greatly if designed to simultaneously measure multiple neurotransmitter systems. Ultimately, a thorough understanding of neurotransmitter interactions and integrated neuronal systems as they relate to the neurochemical pathology of depressive disorders will likely yield novel therapeutic interventions.

ACKNOWLEDGMENTS

Part of "73 - Neurocircuitry of Mood Disorders "

Dr. Ordway has received research support from Eli Lilly, Pharmacia-Upjohn, and Merck. In addition, he served as a consultant for both Eli Lilly and Pharmacia-Upjohn.

REFERENCES

1. Altamura C, Maes M, Dai J, et al. Plasma concentrations of excitatory amino acids, serine, glycine, taurine and histidine in major depression. *Eur Neuropsychopharmacol* 1995;5(Suppl):71–75.

2. Altamura CA, Mauri MC, Ferrara A, et al. Plasma and platelet excitatory amino acids in psychiatric disorders. *Am J Psychiatry* 1993;150:1731–1733.

3. Arango V, Ernsberger P, Marzuk PM, et al. Autoradiographic demonstration of increased serotonin 5-HT2 and B-adrenergic receptor binding sites in the brain of suicide victims. *Arch Gen Psychiatry* 1990;47:1038–1047.

4. Arango V, Ernsberger P, Sved AF, et al. Quantitative autoradiography of α_1 and α_2 adrenergic receptors in the cerebral cortex of controls and suicide victims. *Brain Res* 1993;630:271–282.

5. Arango V, Underwood MD, Mann JJ. Fewer pigmented locus ceruleus neurons in suicide victims: preliminary results. *Biol Psychiatry* 1996;39:112–120.

6. Arborelius L, Chergui K, Murase S, et al. The 5-HT1A receptor selective ligands, (R)-8-OH-DPAT and (S)-UH-301, differentially affect the activity of midbrain dopamine neurons. *Naunyn Schmiedebergs Arch Pharmacol* 1993;347:353–362.

7. Arborelius L, Nomikos GG, Hacksell U, et al. (R)-8-OH-DPAT preferentially increases dopamine release in rat medial prefrontal cortex. *Acta Physiol Scand* 1993;148:465–466.

8. Aston-Jones G, Akaoka H, Charlety P, et al. Serotonin selectively attenuates glutamate-evoked activation of noradrenergic locus ceruleus neurons. *J Neurosci* 1991;11:760–769.

9. Aston-Jones G, Ennis M, Pieribone VA, et al. The brain nucleus locus coeruleus: restricted afferent control of a broad efferent network. *Science* 1986;234:734–737.

10. Aston-Jones G, Shipley MT, Chouvet G, et al. Afferent regulation of locus ceruleus neurons: anatomy, physiology and pharmacology. *Prog Brain Res* 1991;88:47–75.

11. Austin MC, Rhodes JL, Lewis DA. Differential distribution of corticotropin-releasing hormone immunoreactive axons in monoaminergic nuclei of the human brainstem. *Neuropsychopharmacology* 1997;17:326–341.

12. Baker KG, Halliday GM, Hornung JP, et al. Distribution, morphology and number of monoamine-synthesizing and substance P-containing neurons in the human dorsal raphe nucleus. *Neuroscience* 1991;42:757–775.

13. Bannon MJ, Elliott PJ, Alpert JE, et al. Role of endogenous substance P in stress-induced activation of mesocortical dopamine neurones. *Nature* 1983;306:791–792.

14. Baraban JM, Aghajanian GK. Suppression of serotonergic neuronal firing by α-adrenoceptor antagonists: evidence against GABA mediation. *Eur J Pharmacol* 1980;66:287–294.

15. Baraban JM, Aghajanian GK. Noradrenergic innervation of serotonergic neurons in the dorsal raphe: demonstration by electron microscopic autoradiography. *Brain Res* 1981;204:1–11.

16. Bartholini G. Experimental basis for the antidepressant action of the GABA receptor agonist progabide. *Neurosci Lett* 1984;47:351–355.

17. Beart PM, McDonald D. 5-Hydroxytryptamine and 5-hydroxytryptaminergic-dopaminergic interactions in the ventral tegmental area of rat brain. *J Pharmacol* 1982;34:591–593.

18. Becker G, Struck M, Bogdahn U, et al. Echogenicity of the brainstem raphe in patients with major depression. *Psychiatry Res* 1994;55:75–84.

19. Beckmann H, Goodwin FK. Urinary MHPG in subgroups of depressed patients and normal controls. *Neuropsychobiology* 1980;6:91–100.

20. Berman RM, Cappiello A, Anand A, et al. Antidepressant effects of ketamine in depressed patients. *Biol Psychiatry* 2000;47:351–354.

21. Berman RM, Narasimhan M, Miller H, et al. Transient depressive relapse induced by catecholamine depletion. *Arch Gen Psychiatry* 1999;56:395–403.

22. Berod A, Chat M, Paut L, et al. Catecholaminergic and GABAergic anatomic relationship in the rat substantia nigra, locus coeruleus, and hypothalamic median eminence: immunocytochemical visualization of biosynthetic enzymes on serial semithin plastic-embedded sections. *J Histochem Cytochem* 1984;32:1331–1338.

23. Blier P, de Montigny C. Current advances and trends in the treatment of depression [see comments]. *Trends Pharmacol Sci* 1994;15:220–226.

24. Blier P, de Montigny C, Chaput Y. A role for the serotonin system in the mechanism of action of antidepressant treatments: peclinical evidence. *J Clin Psychiatry* 1990;5:14–20.

25. Bligh-Glover W, Kolli TN, Shapiro-Kulnane L, et al. The serotonin transporter in the midbrain of suicide victims with major depression. *Biol Psychiatry* 2000;47:1015–1024.

26. Brady LS. Stress, antidepressant drugs, and the locus coeruleus. *Brain Res Bull* 1994;35:545–556.

27. Brodin E, Ogren SO, Theodorsson-Norheim E. Effects of subchronic treatment with imipramine, zimelidine and alaproclate on regional tissue levels of substance P- and neurokinin A/neurokinin B-like immunoreactivity in the brain and spinal cord of the rat. *Neuropharmacology* 1987;26:581–590.

28. Charney DS. Monoamine dysfunction and the pathophysiology and treatment of depression. [Review] [22 refs]. *J Clin Psychiatry* 1998;59(Suppl 14):11–14.

29. Cheetham SC, Crompton MR, Katona CL, et al. Brain GABAA/benzodiazepine binding sites and glutamic acid decarboxylase activity in depressed suicide victims. *Brain Res* 1988;460:114–123.

30. Chiang C, Aston-Jones G. A 5-hydroxytryptamine2 agonist augments gamma-aminobutyric acid and excitatory amino acid inputs to noradrenergic locus ceruleus neurons. *Neuroscience* 1993;54:409–420.

31. Cubells JF, Kim KS, Baker H, et al. Differential in vivo regulation of mRNA encoding the norepinephrine transporter and tyrosine hydroxylase in rat adrenal medulla and locus ceruleus. *J Neurochem* 1995;65:502–509.

32. Curtis AL, Grigoriadis DE, Page ME, et al. Pharmacologic comparison of two corticotropin-releasing factor antagonists: in vivo and in vitro studies. *J Pharmacol Exp Ther* 1994;268:359–365.

33. D'haenen HA, Bossuyt A. Dopamine D2 receptors in depression measured with single photon emission computed tomography. *Biol Psychiatry* 1994;35:128–132.

34. Dam TV, Martinelli B, Quirion R. Autoradiographic distribution of brain neurokinin-1/substance P receptors using a highly selective ligand [^{3}H]-[Sar9,Met(O2)11]-substance P. *Brain Res* 1990;531:333–337.

35. Delgado PL. Depression: the case for a monoamine deficiency. *J Clin Psychiatry* 2000;61(Suppl 6):7–11.

36. Delgado PL, Price LH, Miller HL, et al. Serotonin and the neurobiology of depression. Effects of tryptophan depletion in drug-free depressed patients. *Arch Gen Psychiatry* 1994;51:865–874.

37. Demontis MG, Fadda P, Devoto P, et al. Sleep deprivation increases dopamine D1 receptor antagonist [3H]SCH 23390 binding and dopamine-stimulated adenylate cyclase in the rat limbic system. *Neurosci Lett* 1990;117:224–227.

38. Drevets WC, Frank E, Price JC, et al. PET imaging of serotonin 1A receptor binding in depression. *Biol Psychiatry* 1999;46:1375–1387.

39. Duman RS, Heninger GR, Nestler EJ. A molecular and cellular theory of depression [see comments]. *Arch Gen Psychiatry* 1997;54:597–606.

40. Durlach-Misteli C, Van Ree JM. Dopamine and melatonin in the nucleus accumbens may be implicated in the mode of action of antidepressant drugs. *Eur J Pharmacol* 1992;217:15–21.

41. Duvauchelle CL, MacConell LA, Eremia AD, et al. Pimozide prevents the development of conditioned place preferences induced by rewarding locus ceruleus stimulation. *Behav Brain Res* 1992;50:85–92.

42. Dziedzicka-Wasylewska M, Rogoz R, Klimek V, et al. Repeated administration of antidepressant drugs affects the levels of mRNA coding for D1 and D2 dopamine receptors in the rat brain. *J Neural Transm* 1997;104:515–524.

43. Ebert D, Feistel H, Loew T, et al. Dopamine and depression—striatal dopamine D2 receptor SPECT before and after antidepressant therapy. *Psychopharmacology (Berl)* 1996;126:91–94.

44. Ennis M, Aston-Jones G. GABA-mediated inhibition of locus ceruleus from the dorsomedial rostral medulla. *J Neurosci* 1989;9:2973–2981.

45. Enserink M. Can the placebo be the cure? [news] [see comments]. *Science* 1999;284:238–240.

46. Ferre'S, Cor'tes R, Artigas F. Dopaminergic regulation of the serotonergic raphe-striatal pathway: microdialysis studies in freely moving rats. *J Neurosci* 1994;14(8):4839–4846.

47. Ferrier IN, McKeith IG, Cross AJ, et al. Postmortem neurochemical studies in depression. *Ann N Y Acad Sci* 1986;487:128–142.

48. Fibiger HC, Phillips AG. Mesocorticolimbic dopamine systems and reward. *Ann N Y Acad Sci* 1988;537:206–215.

49. Flory JD, Mann JJ, Manuck SB, et al. Recovery from major depression is not associated with normalization of serotonergic function. *Biol Psychiatry* 1998;43:320–326.

50. Fodor M, Gorcs TJ, Palkovits M. Immunohistochemical study on the distribution of neuropeptides within the pontine tegmentum—particularly the parabrachial nuclei and the locus ceruleus of the human brain. *Neuroscience* 1992;46:891–908. [Published erratum appears in *Neuroscience* 1992;48(3):753.]

51. Foote SL, Bloom FE, Aston-Jones G. Nucleus locus coeruleus: new evidence of anatomic and physiological specificity. *Physiol Rev* 1983;63:844–914.

52. Gallager DW. Benzodiazepines: potentiation of a GABA inhibitory response in the dorsal raphe nucleus. *Eur J Pharmacol* 1978;49:133–143.

53. Gao B, Fritschy JM, Benke D, et al. Neuron-specific expression of GABAA-receptor subtypes: differential association of the alpha 1- and alpha 3-subunits with serotonergic and GABAergic neurons. *Neuroscience* 1993;54:881–892.

54. Gervasoni D, Peyron C, Rampon C, et al. Role and origin of the GABAergic innervation of dorsal raphe serotonergic neurons. *J Neurosci* 2000;20:4217–4225.

55. Gibbons RD, Davis JM. Consistent evidence for a biological subtype of depression characterized by low CSF monoamine levels. *Acta Psychiatr Scand* 1986;74:8–12.

56. Golden RN, Ekstrom D, Brown TM, et al. Neuroendocrine effects of intravenous clomipramine in depressed patients and healthy subjects. *Am J Psychiatry* 1992;149:1168–1175.

57. Gorea E, Davenne D, Lanfumey L, et al. Regulation of noradrenergic coerulean neuronal firing mediated by 5-HT2 receptors: involvement of the prepositus hypoglossal nucleus. *Neuropharmacology* 1991;30:1309–1318.

58. Grant MM, Weiss JM. Effects of chronic antidepressant drug administration and electroconvulsive shock and locus ceruleus electrophysiological activity. *Biol Psychiatry* 2001;49:117–129.

59. Grenhoff J, Svensson TH. Prazosin modulates the firing pattern of dopamine neurons in rat ventral tegmental area. *Eur J Pharmacol* 1993;233:79–84.

60. Gross G, Gothert M, Glapa U, et al. Lesioning of serotoninergic and noradrenergic nerve fibres of the rat brain does not decrease binding of ^{3}H-clonidine and ^{3}H-rauwolscine to cortical membranes. *Naunyn-Schmiedeberg's Arch Pharmacol* 1985;328:229–235.

61. Guan XM, McBride WJ. Serotonin microinfusion into the ventral tegmental area increases accumbens dopamine release. *Brain Res Bull* 1989;23:541–547.

62. Guyenet PG, Aghajanian GK. Excitation of neurons in the nucleus locus ceruleus by substance P and related peptides. *Brain Res* 1977;136:178–184.

63. Haddjeri N, Blier P, de Montigny C. Long-term antidepressant treatments result in a tonic activation of forebrain 5-HT1A receptors. *J Neurosci* 1998;18:10150–10156.

64. Haddjeri N, de Montigny C, Blier P. Modulation of the firing activity of noradrenergic neurones in the rat locus ceruleus by the 5-hydroxytryptamine system. *Br J Pharmacol* 1997;120:865–875.

65. Hahn MK, Bannon MJ. Stress-induced C-fos expression in the rat locus ceruleus is dependent on neurokinin 1 receptor activation. *Neuroscience* 1999;94:1183–1188.

66. Halaris AE, Belendiuk KT, Freedman DX. Antidepressant drugs affect dopamine uptake. *Biochem Pharmacol* 1975;24:1896–1897.

67. Harro J, Oreland L. Depression as a spreading neuronal adjustment disorder. *Eur Neuropsychopharmacol* 1996;6:207–223.

68. Heimer L, Alheid GF, De Olmos JS, et al. The accumbens: beyond the core-shell dichotomy. *J Neuropsychiatry Clin Neurosci* 1997;9:354–381.

69. Heninger GR, Charney DS, Sternberg DE. Serotonergic function in depression. Prolactin response to intravenous tryptophan in depressed patients and healthy subjects. *Arch Gen Psychiatry* 1984;41:398–402.

70. Herve D, Pickel VM, Joh TH, et al. Serotonin axon terminals in the ventral tegmental area of the rat: fine structure and synaptic input to dopaminergic neurons. *Brain Res* 1987;435:71–83.

71. Kaehler ST, Singewald N, Philippu A. Dependence of serotonin release in the locus ceruleus on dorsal raphe neuronal activity. *Naunyn Schmiedebergs Arch Pharmacol* 1999;359:386–393.

72. Kalen P, Pritzel M, Nieoullon A, et al. Further evidence for excitatory amino acid transmission in the lateral habenular projection to the rostral raphe nuclei: lesion-induced decrease of high affinity glutamate uptake. *Neurosci Lett* 1986;68:35–40.

73. Kalivas PW, Churchill L, Klitenick MA. The circuitry mediating the translation of motivational stimuli into adaptive motor responses. In: Kalivas PW, Barnes CD, eds. *Limbic motor circuits and neuropsychiatry*. Boca Raton, FL: CRC Press, 1993:237–287.

74. Kalivas PW, Duffy P, Latimer LG. Neurochemical and behavioral effects of corticotropin-releasing factor in the ventral tegmental area of the rat. *J Pharmacol Exp Ther* 1987;242:757–763.

75. Kalivas PW, Nakamura M. Neural systems for behavioral activation and reward. *Curr Opin Neurobiol* 1999;9:223–227.

76. Kim JS, Schmid-Burgk W, Claus D, et al. Effects of amitriptyline on serum glutamate and free tryptophan in rats. *Arch Psychiatr Nervenkr* 1982;232:391–394.

77. Kirby LG, Rice KC, Valentino RJ. Effects of corticotropin-releasing factor on neuronal activity in the serotonergic dorsal raphe nucleus. *Neuropsychopharmacology* 2000;22:148–162.

78. Klimek V, Maj J. Repeated administration of antidepressants enhances agonist affinity for mesolimbic D2-receptors. *J Pharm Pharmacol* 1989;41:555–558.

79. Klimek V, Rajkowska G, Luker SN, et al. Brain noradrenergic receptors in major depression and schizophrenia. *Neuropsychopharmacology* 1999;21:69–81.

80. Klimek V, Stockmeier CA, Overholser JC, et al. Reduced levels of norepinephrine transporters in the locus ceruleus in major depression. *J Neurosci* 1997;17:8451–8458.

81. Koob GF. Corticotropin-releasing factor, norepinephrine, and stress. *Biol Psychiatry* 1999;46:1167–1180.

82. Kramer MS, Cutler N, Feighner J, et al. Distinct mechanism for antidepressant activity by blockade of central substance P receptors [see comments]. *Science* 1998;281:1640–1645.

83. Lategan AJ, Marien MR, Colpaert FC. Suppression of nigrostriatal and mesolimbic dopamine release in vivo following noradrenaline depletion by DSP-4: a microdialysis study. *Life Sci* 1992;50:995–999.

84. Lavicky J, Dunn AJ. Corticotropin-releasing factor stimulates catecholamine release in hypothalamus and prefrontal cortex in freely moving rats as assessed by microdialysis. *J Neurochem* 1993;60:602–612.

85. Lechner SM, Curtis AL, Brons R, et al. Locus ceruleus activation by colon distention: role of corticotropin-releasing factor and excitatory amino acids. *Brain Res* 1997;756:114–124.

86. Lee C-M, Javitch JA, Snyder SH. Recognition sites for norepinephrine uptake: regulation by neurotransmitter. *Science* 1983;220:626–629.

87. Lejeune F, Millan MJ. Induction of burst firing in ventral tegmental area dopaminergic neurons by activation of serotonin (5-HT)1A receptors: WAY 100,635-reversible actions of the highly selective ligands, flesinoxan and S 15535. *Synapse* 1998;30:172–180.

88. Levine J, Panchalingam K, Rapoport A, et al. Increased cerebrospinal fluid glutamine levels in depressed patients. *Biol Psychiatry* 2000;47:586–593.

89. Levitt P, Moore RY. Origin and organization of brain stem catecholamine innervation in the rat. *J Comp Neurol* 1979;186:505–528.

90. Little KY, McLauglin DP, Ranc J, et al. Serotonin transporter binding sites and mRNA levels in depressed persons committing suicide. *Biol Psychiatry* 1997;41:1156–1164.

91. Liu RP, Swenberg ML. Autoradiographic localization of substance P ligand binding sites and distribution of immunoreactive neurons in the periaqueductal gray of the rat. *Brain Res* 1988;475:73–79.

92. Lorens SA, Paris JM, Brodin E. Neurokinin innervation of the rat median raphe nucleus does not originate in the brain stem. *Ann NY Acad Sci* 1991;632:431–434.

93. Maeda T, Kojima Y, Arai R, et al. Monoaminergic interaction in the central nervous system: a morphological analysis in the locus ceruleus of the rat. *Comp Biochem Physiol C* 1991;98:193–202.

94. Maj J, Rogoz Z, Skuza G, et al. Repeated treatment with antidepressant drugs increases the behavioural response to apomorphine. *J Neural Transm* 1984;60:273–282.

95. Maj J, Rogoz Z, Skuza G, et al. Repeated treatment with antidepressant drugs potentiates the locomotor response to (+)-amphetamine. *J Pharm Pharmacol* 1984;36:127–130.

96. Maj J, Wedzony K, Klimek V. Desipramine given repeatedly enhances behavioural effects of dopamine and *d*-amphetamine injected into the nucleus accumbens. *Eur J Pharmacol* 1987;140:179–185.

97. Malison RT, Price LH, Berman R, et al. Reduced brain serotonin transporter availability in major depression as measured by [^{123}I]-2 beta-carbomethoxy-3 beta-(4-iodophenyl)tropane and single photon emission computed tomography [see comments]. *Biol Psychiatry* 1998;44:1090–1098.

98. Mann JJ. Role of the serotonergic system in the pathogenesis of major depression and suicidal behavior. *Neuropsychopharmacology* 1999;21:99S–105S.

99. Mann JJ, Malone KM, Psych MR, et al. Attempted suicide characteristics and cerebrospinal fluid amine metabolites in depressed inpatients. *Neuropsychopharmacology* 1996;15:576–586.

100. Mantyh PW, Hunt SP. Neuropeptides are present in projection neurones at all levels in visceral and taste pathways: from periphery to sensory cortex. *Brain Res* 1984;299:297–312.

101. Martensson B, Nyberg S, Toresson G, et al. Fluoxetine treatment of depression. Clinical effects, drug concentrations and monoamine metabolites and N-terminally extended substance P in cerebrospinal fluid. *Acta Psychiat Scand* 1989;79:586–596.

102. Mayberg HS. Limbic-cortical dysregulation: a proposed model of depression. *J Neuropsychiatry Clin Neurosci* 1997;9:471–481.

103. McKeith IG, Marshall EF, Ferrier IN, et al. 5-HT receptor binding in postmortem brain from patients with affective disorder. *J Affect Disord* 1987;13:67–74.

104. Melia KR, Duman RS. Involvement of corticotropin-releasing factor in chronic stress regulation of the brain noradrenergic system. *Proc Natl Acad Sci USA* 1991;88:8382–8386.

105. Melia KR, Nestler EJ, Duman RS. Chronic imipramine treatment normalizes levels of tyrosine hydroxylase in the locus ceruleus of chronically stressed rats. *Psychopharmacology* 1992;108:23–26.

106. Melia KR, Rasmussen K, Terwilliger RZ, et al. Coordinate regulation of the cyclic AMP system with firing rate and expression of tyrosine hydroxylase in the rat locus coeruleus: effects of chronic stress and drug treatments. *J Neurochem* 1992;58:494–502.

107. Mongeau R, de Montigny C, Blier P. Electrophysiologic evidence for desensitization of alpha 2-adrenoceptors on serotonin terminals following long-term treatment with drugs increasing norepinephrine synaptic concentration. *Neuropsychopharmacology* 1994;10:41–51.

108. Morgane PJ, Jacobs MS. Raphe projections to the locus ceruleus in the rat. *Brain Res Bull* 1979;4:519–534.

109. Morley MJ, Shah K, Bradshaw CM, et al. DSP4 and Herrnstein's equation: further evidence for a role of noradrenaline in the maintenance of operant behaviour by positive reinforcement. *Psychopharmacology (Berl)* 1988;96:551–556.

110. Nankova B, Kvetnansky R, Hiremagalur B, et al. Immobilization stress elevates gene expression for catecholamine biosynthetic enzymes and some neuropeptides in rat sympathetic ganglia: effects of adrenocorticotropin and glucocorticoids. *Endocrinology* 1996;137:5597–5604.

111. Nemeroff CB. The role of corticotropin-releasing factor in the pathogenesis of major depression. [Review]. *Pharmacopsychiatry* 1988;21:76–82.

112. Nemeroff CB, Owens MJ, Bissette G, et al. Reduced corticotropin releasing factor binding sites in the frontal cortex of suicide victims. *Arch Gen Psychiatry* 1988;45:577–579.

113. Nemeroff CB, Widerlov E, Bissette G, et al. Elevated concentrations of CSF corticotropin-releasing factor-like immunoreactivity in depressed patients. *Science* 1984;226:1342–1344.

114. Nestler EJ, McMahon A, Sabban EL, et al. Chronic antidepressant administration decreases the expression of tyrosine hydroxylase in the rat locus coeruleus. *Proc Natl Acad Sci USA* 1990;87:7522–7526.

115. Nowak G, Ordway GA, Paul IA. Alterations in the *N*-methyl-D-aspartate (NMDA) receptor complex in the frontal cortex of suicide victims. *Brain Res* 1995;675:157–164.

116. Oades RD, Halliday GM. Ventral tegmental (A10) system: neurobiology. 1. Anatomy and connectivity. *Brain Res* 1987;434:117–165.

117. Olpe H, Steinmann MW, Brugger F, et al. Excitatory amino acid receptors in rat locus coeruleus. *Naunyn-Schmiedeberg's Arch Pharmacol* 1989;339:312–314.

118. Ordway GA, Farley IJ, Dilley GE, et al. Quantitative distribution of monoamine oxidase A in brainstem monoamine nuclei is normal in major depression. *Brain Res* 1999;847:71–79.

119. Ordway GA, Schenck JE, Dilley GE, et al. Increased p-[^{125}I]iodoclonidine binding to α_2-adrenoceptors in the locus ceruleus in major depression. *Soc Neurosci Abs* 1999;2139.

120. Ordway GA, Smith KS, Haycock JW. Elevated tyrosine hydroxylase in the locus ceruleus of suicide victims. *J Neurochem* 1994;62:680–685.

121. Ordway GA, Stockmeier CA, Cason GW, et al. Pharmacology and distribution of norepinephrine transporters in the human locus ceruleus and raphe nuclei. *J Neurosci* 1997;17:1710–1719.

122. Ordway GA, Widdowson PS, Smith KS, et al. Agonist binding to α_2-adrenoceptors is elevated in the locus ceruleus from victims of suicide. *J Neurochem* 1994;63:617–624.

123. Osmanovic SS, Shefner SA. γ-Aminobutyric acid responses in rat locus ceruleus neurones in vitro: a current-clamp and voltage-clamp study. *J Physiol* 1990;421:151–170.

124. Osterhout CA, Chikaraishi DM, Tank AW. Induction of tyrosine hydroxylase protein and a transgene containing tyrosine hydroxylase 5′ flanking sequences by stress in mouse adrenal gland. *J Neurochem* 1997;68:1071–1077.

125. Owens MJ, Nemeroff CB. Preclinical and clinical studies with corticotropin-releasing factor: implications for affective disorders. [Review]. *Psychopharmacology Bull* 1988;24:355–359.

126. Palacios, J. M., Wamsley, J. K., and Kuhar, M. J. High affinity GABA receptors—autoradiographic localization. *Brain Res* 1981;222:285–307.

127. Papp M, Klimek V, Willner P. Parallel changes in dopamine D2 receptor binding in limbic forebrain associated with chronic mild stress-induced anhedonia and its reversal by imipramine. *Psychopharmacology (Berl)* 1994;115:441–446.

128. Parsons LH, Justice JBJ. Perfusate serotonin increases extracellular dopamine in the nucleus accumbens as measured by in vivo microdialysis. *Brain Res* 1993;606:195–199.

129. Paul IA. NMDA receptors and affective disorders. In: Skolnick P, ed. *Antidepressants: new pharmacologic strategies*. Totowa, NJ: Humana Press, 1997:145–148.

130. Pavcovich LA, Cancela LM, Volosin M, et al. Chronic stress-induced changes in locus ceruleus neuronal activity. *Brain Res Bull* 1990;24:293–296.

131. Paykel ES, Myers JK, Dienelt MN, et al. Life events and depression. A controlled study. *Arch Gen Psychiatry* 1969;21:753–760.

132. Petty F. GABA and mood disorders: a brief review and hypothesis. *J Affect Disord* 1995;34:275–281.

133. Petty F, Kramer GL, Gullion CM, et al. Low plasma gamma-aminobutyric acid levels in male patients with depression. *Biol Psychiatry* 1992;32:354–363.

134. Petty F, Schlesser MA. Plasma GABA in affective illness. A preliminary investigation. *J Affect Disord* 1981;3:339–343.

135. Petty F, Sherman AD. Plasma GABA levels in psychiatric illness. *J Affect Disord* 1984;6:131–138.

136. Peyron C, Luppi PH, Kitahama K, et al. Origin of the dopaminergic innervation of the rat dorsal raphe nucleus. *Neuroreport* 1995;6:2527–2531.

137. Pompeiano M, Palacios JM, Mengod G. Distribution of the serotonin 5-HT2 receptor family mRNAs: comparison between 5-HT2A and 5-HT2C receptors. *Mol Brain Res* 1994;23:163–178.

138. Potter WZ, Grossman G, Rudorfer MV. *Noradrenergic function in depressive disorders. Part A: A systems perspective*. New York: Plenum, 1993:1–27.

139. Prange AJ Jr. The pharmacology and biochemistry of depression. *Dis Nerv Syst* 1964;25:217–221.

140. Price ML, Curtis AL, Kirby LG, et al. Effects of corticotropin-releasing factor on brain serotonergic activity. *Neuropsychopharmacology* 1998;18:492–502.

141. Prisco S, Pagannone S, Esposito E. Serotonin-dopamine interaction in the rat ventral tegmental area: an electrophysiological study in vivo. *J Pharmacol Exp Ther* 1994;271:83–90.

142. Qian Y, Melikian HE, Rye DB, et al. Identification and characterization of antidepressant-sensitive serotonin transporter proteins using site-specific antibodies. *J Neurosci* 1995;15:1261–1274.

143. Randrup A, Munkvad I, Fog R, et al. Mania, depression and brain dopamine. In: Essman WB, Valzelli L, eds. *Current developments in psychopharmacology*, volume 2. New York: Spectrum, 1975:206–248.

144. Reddy PL, Khanna S, Subhash MN, et al. CSF amine metabolites in depression. *Biol Psychiatry* 1992;31:112–118.

145. Riche D, De Pommery J, Menetrey D. Neuropeptides and catecholamines in efferent projections of the nuclei of the solitary tract in the rat. *J Comp Neurol* 1990;293:399–424.

146. Rogoz R, Dziedzicka-Wasylewska M. Effects of antidepressant drugs on the dopamine D2/D3 receptors in the rat brain differentiated by agonist and antagonist binding—an autoradiographic analysis. *Naunyn Schmiedebergs Arch Pharmacol* 1999;359:178–186.

147. Roy A, Pickar D, Linnoila M, et al. Plasma norepinephrine level in affective disorders: relationship to melancholia. *Arch Gen Psychiatry* 1985;42:1181–1185.

148. Ruggiero DA, Underwood MD, Rice PM, et al. Corticotropic-releasing hormone and serotonin interact in the human brainstem: behavioral implications. *Neuroscience* 1999;91:1343–1354.

149. Rupniak NM, Kramer MS. Discovery of the antidepressant and anti-emetic efficacy of substance P receptor (NK1) antagonists. *Trends Pharmacol Sci* 1999;20:485–490.

150. Saavedra JM, Grobecker H, Zivin J. Catecholamines in the raphe nuclei of the rat. *Brain Res*. 1976;114:339–345.

151. Sakai K, Salvert D, Touret M, et al. Afferent connections of the nucleus raphe dorsalis in the cat as visualized by the horseradish peroxidase technique. *Brain Res* 1977;137:11–35.

152. Salomon R, Miller HL, Krystal JH, et al. Lack of behavioral effects of monoamine depletion in healthy subjects. *Biol Psychiatry* 1997;41:58–64.

153. Schildkraut JJ. The catecholamine hypothesis of affective disorders: a review of supporting evidence. *Am J Psychiatry* 1965;122:509–522.

154. Schultzberg M, Austin MC, Crawley JN, et al. Repeated administration of desmethylimipramine blocks reserpine-induced increase in tyrosine hydroxylase mRNA in locus ceruleus neurons of the rat. *Mol Brain Res* 1991;10:307–314.

155. Sergeyev V, Hokfelt T, Hurd Y. Serotonin and substance P co-exist in dorsal raphe neurons of the human brain. *Neuroreport* 1999;10:3967–3970.

156. Serra G, Collu M, D'Aquila PS, et al. Possible role of dopamine D1 receptor in the behavioural supersensitivity to dopamine agonists induced by chronic treatment with antidepressants. *Brain Res* 1990;527:234–243.

157. Shah PJ, Ogilvie AD, Goodwin GM, et al. Clinical and psychometric correlates of dopamine D2 binding in depression. *Psychol Med* 1997;27:1247–1256.

158. Shirayama Y, Mitsushio H, Takashima M, et al. Reduction of substance P after chronic antidepressants treatment in the striatum, substantia nigra and amygdala of the rat. *Brain Res* 1996;739:70–78.

159. Singewald N, Kaehler ST, Hemeida R, et al. Influence of excitatory amino acids on basal and sensory stimuli-induced release of 5-HT in the locus coeruleus. *Br J Pharmacol* 1998;123:746–752.

160. Singewald N, Zhou GY, Chen F, et al. Corticotropin-releasing factor modulates basal and stress-induced excitatory amino acid release in the locus ceruleus of conscious rats. *Neurosci Lett* 1996;204:45–48.

161. Singewald N, Zhou GY, Schneider C. Release of excitatory and inhibitory amino acids from the locus ceruleus of conscious rats by cardiovascular stimuli and various forms of acute stress. *Brain Res* 1995;704:42–50.

162. Skolnick P, Layer RT, Popik P, et al. Adaptation of N-methyl-D-aspartate (NMDA) receptors following antidepressant treatment: implications for the pharmacotherapy of depression. *Pharmacopsychiatry* 1996;29:23–26.

163. Smagin GN, Swiergiel AH, Dunn AJ. Corticotropin-releasing factor administered into the locus coeruleus, but not the parabrachial nucleus, stimulates norepinephrine release in the prefrontal cortex. *Brain Res Bull* 1995;36:71–76.

164. Smith KA, Fairburn CG, Cowen PJ. Relapse of depression after rapid depletion of tryptophan [see comments]. *Lancet* 1997;349:915–919.

165. Stockmeier CA, Dilley GE, Shapiro LA, et al. Serotonin receptors in suicide victims with major depression. *Neuropsychopharmacology* 1997;16:162–173.

166. Stockmeier CA, Shapiro LA, Dilley GE, et al. Increase in serotonin-1A autoreceptors in the midbrain of suicide victims with major depression—postmortem evidence for decreased serotonin activity. *J Neurosci* 1998;18:7394–7401.

167. Swerdlow NR, Koob GF. Dopamine, schizophrenia, mania, and depression: toward a unified hypothesis of cortico-striato-pallido-thalamic function. *Behav Brain Sci* 1987;10:197–245.

168. Szabo ST, de Montigny C, Blier P. Progressive attenuation of the firing activity of locus ceruleus noradrenergic neurons by sustained administration of selective serotonin reuptake inhibitors. *Int J Neuropsychopharmacology* 2000;3:1–11.

169. Tamiya R, Hanada M, Kawai Y, et al. Substance P afferents have synaptic contacts with dopaminergic neurons in the ventral tegmental area of the rat. *Neurosci Lett* 1990;110:11–15.

170. Timmerman W, Cisci G, Nap A, et al. Effects of handling on extracellular levels of glutamate and other amino acids in various areas of the brain measured by microdialysis. *Brain Res* 1999;833:150–160.

171. Torack RM, Morris JC. The association of ventral tegmental area histopathology with adult dementia. *Arch Neurol* 1988;45:497–501.

172. U'Prichard DC, Bechtel WD, Rouot BM, et al. Multiple apparent alpha-noradrenergic receptor binding sites in rat brain: effect of 6-hydroxydopamine. *Mol Pharmacol* 1979;16:47–60.

173. Underwood MD, Khaibulina AA, Ellis SP, et al. Morphometry of the dorsal raphe nucleus serotonergic neurons in suicide victims. *Biol Psychiatry* 1999;46:473–483.

174. Valentino RJ, Page ME, Curtis AL. Activation of noradrenergic locus ceruleus neurons by hemodynamic stress is due to local release of corticotropin-releasing factor. *Brain Res* 1991;555:25–34.

175. Van Bockstaele EJ, Aston-Jones G. Collateralized projections from neurons in the rostral medulla to the nucleus locus coeruleus, the nucleus of the solitary tract and the periaqueductal gray. *Neuroscience* 1992;49:653–668.

176. Van Bockstaele EJ, Colago EE, Valentino RJ. Corticotropin-releasing factor-containing axon terminals synapse onto catecholamine dendrites and may presynaptically modulate other afferents in the rostral pole of the nucleus locus ceruleus in the rat brain. *J Comp Neurol* 1996;364:523–534.

177. Van Bockstaele EJ, Colago EE, Valentino RJ. Amygdaloid corticotropin-releasing factor targets locus ceruleus dendrites: substrate for the co-ordination of emotional and cognitive limbs of the stress response. *J Neuroendocrinol* 1998;10:743–757.

178. Van Praag HM. Indoleamines in depression. In: Usdin E, et al, eds. *Neuroregulators and psychiatric disorders*. New York: Oxford University Press, 1977:163–176.

179. Van Praag HM. New evidence of serotonin-deficient depressions. *Neuropsychobiology* 1977;3:56–63.

180. Veith R, Lewis N, Linares O, et al. Sympathetic nervous system activity in major depression: basal and desipramine-induced alterations in plasma norepinephrine kinetics. *Arch Gen Psychiatry* 1994;51:411–422.

181. Vertes RP, Kocsis B. Projections of the dorsal raphe nucleus to the brainstem: PHA-L analysis in the rat. *J Comp Neurol* 1994;340:11–26.

182. Waddington JL, Deveney AM. Dopamine receptor multiplicity: "D1-like"- "D2-like" interactions and "D1-like" receptors not linked to adenylate cyclase. *Biochem Soc Trans* 1996;24:177–182.

183. Wang P, Kitayama I, Nomura J. Tyrosine hydroxylase gene expression in the locus ceruleus of depression-model rats and rats exposed to short- and long-term forced walking stress. *Life Sci* 1998;62:2083–2092.

184. Wang QP, Ochiai H, Nakai Y. GABAergic innervation of serotonergic neurons in the dorsal raphe nucleus of the rat studied by electron microscopy double immunostaining. *Brain Res Bull* 1992;29:943–948.

185. Weiner DM, Levey AI, Sunahara RK, et al. D1 and D2 dopamine receptor mRNA in rat brain. *Proc Natl Acad Sci USA* 1991;88:1859–1863.

186. Weiss JM, Bonsall RW, Demetrikopoulos MK, et al. Galanin: a significant role in depression? *Ann NY Acad Sci* 1998;863:364–382.

187. Weiss JM, Glazer HI, Pohorecky LA. Coping behavior and neurochemical changes: An alternative explanaion for the original "learned helplessness" experiments. In: Kling SG, ed. *Animal models in human psychobiology*. New York: Plenum, 1976:133–145.

188. Whitty CJ, Paul MA, Bannon MJ. Neurokinin receptor mRNA localization in human midbrain dopamine neurons. *J Comp Neurol* 1997;382:394–400.

189. Willner P. Dopamine and depression: a review of recent evidence. I. Empirical studies. *Brain Res* 1983a;287:211–224.

190. Willner P. Dopamine and depression: a review of recent evidence. II. Theoretical approaches. *Brain Res* 1983b;287:225–236.

191. Willner P. Animal models as simulations of depression. *TiPS* 1991;12:131–136.

192. Wright DE, Seroogy KB, Lundgren KH, et al. Comparative localization of serotonin 1A, 1C, and 2 receptor subtype mRNAs in rat brain. *J Comp Neurol* 1995;351:357–373.

193. Zacharko RM, Bowers WJ, Anisman H. Responding for brain stimulation: stress and desmethylimipramine. *Prog Neuropsychopharmacol Biol Psychiatry* 1984;8:601–606.

194. Zahm DS. Is the caudomedial shell of the nucleus accumbens part of the extended amygdala? A consideration of connections. *Crit Rev Neurobiol* 1998;12:245–265.

195. Zhu M–Y, Klimek V, Dilley GE, et al. Elevated levels of tyrosine hydroxylase in the locus ceruleus in major depression. *Biol Psychiatry* 1999;46:1275–1286.

196. Zivkovic B, Scatton B, Worms P, et al. Pharmacologic and therapeutic actions of GABA receptor agonists. *J Neural Transm Suppl* 1983;18:319–326.

74

Structural and Functional Imaging of Affective Disorders

Yvette I. Sheline

Mark A. Mintun

Yvette I. Sheline: Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri.

Mark A. Mintun: Department of Radiology, Washington University School of Medicine, St. Louis, Missouri.

With advances in imaging technology it is now possible to examine subtle changes in both structure and regional function that are associated with the pathophysiology of affective illness. Understanding how these changes fit together with findings from clinical studies, postmortem findings, and animal studies will yield insight into the neuroanatomic pathways involved in affective illness. Combining anatomic MRI studies with functional studies has improved the localization of abnormalities in blood flow, metabolism, and neurotransmitter receptor function and has the potential to provide a better-integrated model of depression. Functional and structural mapping also provide a bridge between the hypotheses stemming from rapidly increasing knowledge of molecular biology, psychopharmacology, and clinical and treatment applications. In this chapter we review studies of structural brain changes associated with early-onset recurrent depression (EORD), late-onset depression (LOD), bipolar disorder, and potential etiologic mechanisms. We also review functional studies in affective illness, including positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and single-photon emission computed tomography (SPECT) studies.

STRUCTURAL STUDIES

Part of "74 - Structural and Functional Imaging of Affective Disorders "

Historically, the major psychiatric illnesses, including affective disorders, were not thought to be associated with structural brain pathology. With the development of new imaging tools in the last two decades, increasing evidence has accumulated that challenges this assumption. Studies using high-resolution three-dimensional (3D) magnetic resonance imaging (MRI) now have a resolution of 1.0 mm or better and are available to examine smaller brain structures with precision. Initially primarily focused on older subjects, structural studies have found both generalized and localized structural brain changes in major depression and bipolar disorder across the age spectrum.

Recurrent Unipolar Major Depression

Studies of neuroanatomic structure in early-onset recurrent depression (EORD) have recently found evidence for depression-associated structural change (Table 74.1). Brain changes associated with early-onset major depression have been reported in the hippocampus, amygdala, caudate, putamen, and frontal cortex, structures that are extensively interconnected (20). They comprise a neuroanatomic circuit, which has been termed the limbic-cortical-striatal-pallidal-thalamic (LCSPT) circuit (21). As discussed herein, functional aspects of this circuit are also altered in depression as measured by blood flow and metabolism.

Author	Brain Region	Sample (Number and Diagnosis)	Age (Mean ± SD)	Methods and Resolution	Findings
Krishnan et al., 1992 (1)	Frontal lobe	50 NC 50 MDD	49.3 ± 18 48.3 ± 17	1.5 T 5 mm	Bifrontal distances smaller Bifrontal brain widths smaller
Coffey et al., 1993 (2)	Frontal lobe	76 NC 48 MDD (44 unipolar; 4 bipolar) ECT referred	62.4 ± 16.4 61.6 ± 15.9	1.5 T 5 mm 5.2 interval	Smaller frontal lobe volumes in major depression
Drevets et al., 1997 (3)	Subgenual prefrontal cortex	33 NC 13 MDD 10 MDD, remitted	36.2 ± 8.9 33.6 ± 10.0 30.1 ± 7.8	1.5 T 1 mm	Smaller subgenual prefrontal cortex volumes in major depression
Krishnan et al., 1992 (1)	Caudate	50 NC 50 MDD	49.3 ± 18 48.3 ± 17	1.5 T 5 mm	Decreased caudate volumes in major depression
Greenwald et al., 1997 (4)	Caudate	30 NC 36 MDD	72.8 ± 6.6 75.9 ± 6.7	1.0 T 3 mm	Decreased left caudate volume in major depression
Husain et al., 1991 (5)	Putamen	44 NC 41 MDD	56.4 ± 19.2 55.3 ± 18.8	1.5 T 5 mm (2 patients with 7 mm)	Decreased putamen volume in major depression
Dupont et al., 1995 (6)	Caudate and lenticular nucleus	26 NC 36 MDD-bipolar 30 MDD-unipolar	39.1 ± 9.4 36.6 ± 10.8 38.6 ± 10.6	1.5 T 5 mm–2.5 mm gap	No significant difference
Lenze, Sheline, 1999 (7)	Caudate and putamen	24 NC 24 MDD-remitted	52.8 ± 17.8 52.8 ± 18.4	1.5T 0.5 mm	No significant difference
Axelson et al., 1992 (8)	Pituitary	21 MDD (1 bipolar; 1 adjustment disorder; 1 multi-infarct dementia; 1 schizo affective)	47.9 ± 18.4 39.6 ± 13.2	1.0 T 3 mm	Increase in pituitary volume in major depression
Sheline et al., 1996 (9)	Hippocampus	10 NC 10 MDD, remitted	68.0 ± 9.5 68.5 ± 10.4	1.5 T 0.5 mm	Decreased hippocampal gray matter volume in major depression
Shah et al., 1998 (10)	Hippocampus	20 NC 20 MDD 20 TRD	49.3 ± 11.8 47.7 ± 9.9 48.9 ± 9.8	1.0 T 2 mm	Decreased hippocampal volume in treatment resistant depression
Sheline et al., 1999 (11)	Hippocampus	24 NC 24 MDD-remitted	52.8 ± 17.8 52.8 ± 18.4	1.5 T 0.5 mm	Decreased hippocampal volume in major depression
Bremner et al., 2000 (12)	Hippocampus	16 NC 16 MDD (1 panic disorder)	45.0 ± 10.0 43.0 ± 8.0	1.5 T 3 mm	Decreased hippocampal volume in major depression
Mervaala et al., 2000 (13)	Hippocampus	17 NC 34 MDD (6 bipolar, 28 monopolar)	42.1 ± 14.6 42.2 ± 12.2	1.5 T 3 mm	No significant difference
Vakili et al., 2000 (14)	Hippocampus	20 NC 38 MDD	40.3 ± 10.4 38.5 ± 10.0	1.5 T 3 mm	No significant difference
Sheline, Gado, Price, 1998 (15)	Amygdala	20 NC 20 MDD, remitted	53.8 ± 17.7 54.1 ± 18.1	1.5 T 0.5 mm	Decreased amygdala core nuclei volume in major depression
Bremner et al., 2000 (12)	Amygdala	16 NC 16 MDD (1 panic disorder)	45.0 ± 10.0 43.0 ± 8.0	1.5 T 3 mm	Increased right amygdala volume in major depression
Mervaala et al., 2000 (13)	Amygdala	17 NC 34 MDD (6 bipolar, 28 unipolar)	42.1 ± 14.6 42.2 ± 12.2	1.5 T 3 mm	Significant asymmetry in amygdalar volume (right smaller than left)
Swayze et al., 1992 (16)	Amygdala/ hippocampus complex	55 Schizophrenic 48 MDD-bipolar 47 NC	32.3 ± 35.4 33.4 ± 34.6	0.5 T 1-cm thick slices 8 cuts	No significant difference
Axelson et al., 1993 (17)	Amygdala/ hippocampus complex	30 NC 19 MDD	46.7 ± 20.4 56.6 ± 19.1	1.5 T 5 mm	No significant difference
Pantel et al., 1997 (18)	Amygdala/ hippocampus complex	13 NC 19 MDD 27 AD	68.2 ± 5.3 72.4 ± 8.8 71.9 ± 8.0	1.5 T 1.25 mm	No significant difference
Ashtari et al., 1999 (19)	Amygdala/ hippocampus complex	46 NC 40 MDD	71.4 ± 0.3 74.3 ± 6.0	1.0 T 3.1 mm	No significant difference

[a]AD, Alzheimer's disease; MDD, major depressive disorder; NC, normal control; T, tesla; TRD, treatment resistant depression.

TABLE 74.1. BRAIN STRUCTURAL CHANGES REPORTED IN MAJOR DEPRESSIVE DISORDER

Several studies have examined hippocampal volume in depression. In some (9 ,10 ,11 and 12) but not all (13 ,14 ,16 ,17 ,19) significant reductions in hippocampal volumes were found with depression. In some studies the volume loss appears to have functional significance with an association between acute depression and abnormalities of declarative memory (22) as well as an association between severe depression in remission and lower scores on tests of verbal memory (11). One study (10) found hippocampal atrophy in patients with chronic depression but not in patients with remitted depression. Vakili and colleagues (14) also observed correlations between depression severity and hippocampal volumes, although no group differences between depressed and control subjects.

Methodologic differences may account for some of the discrepancies. The studies reporting negative findings typically had lower resolution, ranging from 3 to 10 mm (6 ,14 ,16 ,17 ,19), compared with 0.5 mm to 3 mm (9 ,10 ,11 and 12) for studies reporting significant differences. Some studies have reported negative findings for the amygdala–hippocampus complex (16 ,17 and 18), using methodology that does not separate

the two structures. Clinical selection of subjects may also contribute to different findings. Most of these studies used a mixed group of early-onset and late-onset depression, which may therefore have different contributing etiologies. In some studies (9 ,11 ,12) subjects were case control-matched for age and education, whereas in other studies subjects were group-matched or the results were corrected statistically for significant covariates. Some but not all studies excluded subjects with other physical illness or any current or past drug or alcohol abuse. In summary, in most studies that assessed depression severity in unipolar subjects and used high-resolution MRI techniques (Fig. 74.1), depression was associated with hippocampal atrophy, ranging from 8% to 19%. The significance and source of the volume loss has not been demonstrated. It is intriguing that a recent postmortem study (23) has found glial cell loss in the dentate gyrus of the hippocampus as well as in the amygdala in major depression.

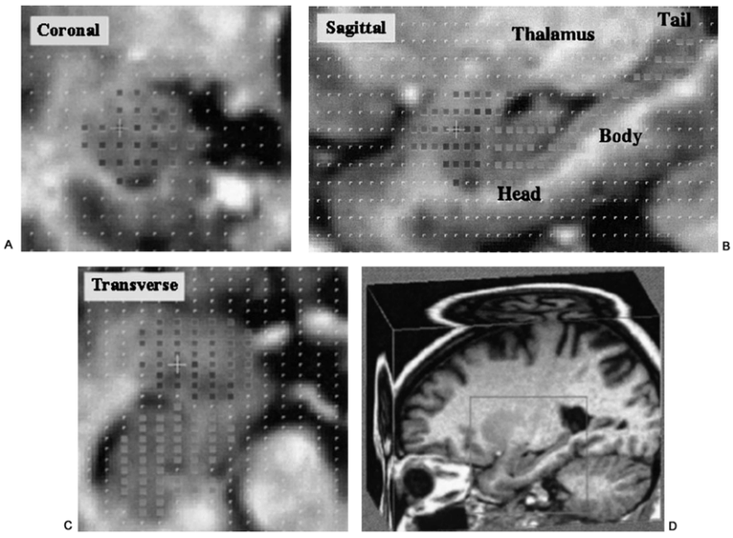


FIGURE 74.1. Orthogonal views (coronal, sagittal, and transverse) represent hippocampal and amygdala volumes as they are shown using stereological techniques. A 5 × 5 mm^2 grid with a 0.5-mm slice thickness is used to determine volumes for each structure. The total amygdala is made up of green plus blue. Labels for orientation of the hippocampus are included on the sagittal view. These measurements are derived from a cubic subvolume depicted in the last panel. See color version of figure.

Study of the amygdala has yielded inconsistent results with depressed subjects exhibiting an increase in volume in the right amygdala (12), loss of normal asymmetry (13), or reduction in the bilateral core nuclei (15). The amygdala is a difficult structure to measure, because in many areas the “cortical amygdala” merges with surrounding cortex. Particular boundaries selected can vary widely.

A number of studies have found decreased volume of basal ganglia structures in major depression, particularly late-onset depression (1 ,4 ,5 ,24), as further discussed. One study has reported negative findings in caudate and putamen in depressed subjects who were otherwise physically healthy (7), a criterion not clearly present in other studies.

Volume reductions in frontal cortex have been reported, ranging from 7% overall reduction in frontal lobe volume in major depression (25) to 48% in the subgenual prefrontal cortex (3). A postmortem study of prefrontal cortex in major depression (26) showed that depressed subjects differed significantly from controls in several prefrontal cortical areas. These included rostral orbitofrontal cortex decreases in cortical thickness, neuronal size decrease, and loss of glial cells in layers II to IV. In caudal orbitofrontal cortex there were reductions in glial cells in layers V to VI and decreases in neuronal sizes. Reductions in glial and neuronal cells throughout all layers as well as reduction in cell size were reported in dorsolateral prefrontal cortex. Glial cell loss in the subgenual region of prefrontal cortex has also been reported in major depression (27). Neuropathologic changes such as these could account for some of the MRI volumetric findings in frontal cortex. The prefrontal cortex is particularly important as a target of monoamine projections. Abnormalities in monoamine receptors, transporters, and second messenger systems (28 ,29 ,30 ,31 and 32) are reported to occur in major depression. It is also possible that overactivation in one part of the interconnected LCSPT neuroanatomic circuit may lead to over excitation in the other components, resulting in excitotoxic damage. The orbitomedial prefrontal cortex has high concentrations of glucocorticoid receptors, potentially rendering it vulnerable to stress-mediated damage (see the following).

Potential Mechanisms for Volume Loss in Early-Onset Recurrent Depression

It has been proposed that hypothalamic–pituitary–adrenal (HPA) axis dysfunction can produce repeated episodes of hypercortisolemia. Currently volume studies do not routinely include measures of cortisol, nor can they ascertain past episodes of hypercortisolemia. However, several different mechanisms could explain volume loss including neuronal loss through exposure to repeated episodes of hypercortisolemia, glial cell loss, resulting in increased vulnerability to glutamate neurotoxicity, stress-induced reduction in neurotrophic factors, and stress-induced reduction in neurogenesis. A mechanism that could account for hippocampal, amygdala, and prefrontal cortex volume loss, areas that have high concentrations of GC receptors is GC-mediated neurotoxicity (33), with repeated hypercortisolemic episodes of depressions giving rise to atrophy. Early life stress may produce a permanent hypersensitivity to stress (34), with the production of ongoing HPA axis dysregulation, particularly in subjects who develop depression. In the case of hippocampal volume loss, the inverse correlations between the total amount of time patients have been depressed and hippocampal volume (9 ,11) found in some studies but not all (12) supports recurrent depressive episodes having a causal relationship. Further, a study by Lupien and colleagues (35) demonstrated a correlation between higher cortisol levels measured longitudinally and greater hippocampal volume loss in normal human aging.

Glial cell loss either directly or indirectly is another potential mechanism for producing volume loss. Gray matter atrophy in the prefrontal cortex in an area ventral to the genu of the corpus callosum (3), an area associated in postmortem studies with glial cell loss (27) has been reported. Glial cell loss was also found in another postmortem study of depressed subjects in two different areas of prefrontal cortex (26). In addition, glial cell loss has been reported in postmortem studies of major depression in the amygdala and hippocampus (23).

Through excitatory connections between the amygdala and hippocampus (36) it is possible that damage in one structure could produce damage in the connected structure. Similarly, interconnections between prefrontal cortex and hippocampus (37) could produce excitotoxic damage. Glial

cells sequester glutamate, maintain metabolic and ionic homeostasis, and produce trophic factors, including brain derived neurotrophic factor (BDNF) (38 ,39). Therefore, loss of glial cells could increase vulnerability to neurotoxic damage, supporting the idea that glutamate neurotoxicity may be involved in the volume loss in the limbic–cortical–striatal–pallidal–thalamic (LCSPT) circuit.

A recently developed concept—stress-induced inhibition of neurogenesis (40)—may also explain depression-related volume loss, although high rates of baseline neurogenesis would be needed to produce atrophy of the scale required by volume loss in depression. Psychosocial stress was shown to suppress neurogenesis in the tree shrew (40). Likewise, corticosterone treatment in adult rats produced suppression of neurogenesis, which was reversed by removal of the adrenal gland (41). More recent findings (42) indicate that neurogenesis may occur in the frontal cortex as well as the hippocampus and subventricular zone. Thus, if depression inhibits neurogenesis, this could potentially explain both cortical and hippocampal volume loss.

Late-Onset Major Depression

Partly because of the increased prevalence of comorbid illness with age (including cerebrovascular disease, Parkinson's disease, etc.), patients with late life depression have an increased prevalence of structural brain changes relative to EORD. Depression onset in late age frequently occurs in patients with medical and neurologic disorders and compared with EORD it is characterized by greater medical morbidity and mortality (43), higher rates of neuroradiologic abnormalities, particularly white-matter hyperintensities (44 ,45), and lower familial frequency of affective disorders (46). In some studies, it is associated with higher rates of neuropsychological impairment and treatment refractoriness (47 ,48).

Both computed tomography (CT) and MRI studies have shown diffuse cortical and subcortical atrophy and ventricular enlargement in late life depression (18 ,49 ,50 and 51). Specific illnesses that have been associated with brain atrophy include hypertension (52), diabetes (53), Cushing's disease (54), and alcohol abuse (55). Any condition that produces neuronal ischemia or neurotoxicity is a potential candidate for producing brain atrophy.

Neurologic illnesses associated with both cortical and subcortical atrophy are associated with unusually high rates of depression, including Huntington's disease (56), post-stroke syndromes (57), dementia of the Alzheimer's type (58), and Parkinson's disease (59). Each of these illnesses produces damage to brain structures critical in emotional functioning. Importantly, these same brain structures are involved in more classical or early-onset major depression, namely frontal cortex, hippocampus, thalamus, and basal ganglia. Some studies do not find evidence for generalized atrophy in addition to volume loss in structures of the LCSPT circuit. Kumar and colleagues (60), for example, have found loss in prefrontal lobe volume in late-onset depression in the absence of generalized atrophy, suggesting that as in early-onset depression some subjects with late-onset depression may also have focal volume loss. It is not known whether this focal volume loss involves the same etiologic mechanisms.

A well replicated finding in elderly subject groups with depression is the increased numbers of hyperintensities seen on T2-weighted scans (T2H) (51 ,61 ,62 ,63 ,64 and 65). Some studies that included younger subjects with depression have also found increased T2H (2 ,66), although negative findings have also been reported with younger groups (5 ,67). The underlying causes of T2H are unknown, and indeed, it is important to note that T2H also occur at rates of up to 60% in healthy elderly (68), in whom their significance is unknown (69). Fujikawa and associates (70 ,71) found a higher rate of “silent” cerebral infarctions (T2H) in late- compared to early-onset MDD. Clinical correlates of MRI-defined T2H in late-life depression have included older age, vascular risk factors, neuropsychological impairment, and late age of onset (2 ,48 ,72). A subtype of “vascular depression” with increased CVD risk factors and increased T2H has been proposed (47 ,73).

Bipolar Disorder

Structural abnormalities reported in bipolar disorder have been intermittently reported. These include diffuse gray matter tissue loss, enlarged ventricles, increased numbers of T2-signal hyperintensities (T2H), and regional tissue loss in basal ganglia, lateral and mesial temporal structures, and cortical regions. Studies in bipolar subjects have also found structural changes in the same neuroanatomic circuit (LCSPT) as in major depression but these changes have been less consistent and often involved increases in structure volumes rather than decreases. In addition there is a substantial body of evidence for manic-like affective disorders following focal brain damage.

Most MRI studies have not found generalized cortical gray matter volume loss in bipolar disorder (6 ,74 ,75 ,76 ,77 and 78). A recent study in geriatric bipolar disorder (79) found increased cortical sulcal widening which was related to age of illness onset. In a small study in middle-aged bipolar subjects, Lim and co-workers (80) also found enlarged cortical sulci. In the same study, bipolar subjects had generalized decreased cortical gray volume that was intermediate between control and schizophrenia values. This finding could result from differences in segmentation algorithms, differences in covarying the data, and a more chronically ill population. In addition, lateral ventricles were enlarged; however, the difference was not significant in this small sample. Lateral ventricle findings from other groups included increases (45 ,74 ,81) and no difference from control (75 ,82).

Several groups have investigated the question of an overall

change in temporal lobe volume. Altshuler and colleagues (83) found bilateral temporal lobe reductions comparing bipolar patients to controls. Swayze also found temporal lobe abnormalities in bipolar disorder, but in symmetry rather than in volume—he found that bipolar subjects did not have the usual right greater than left volume asymmetry found in normals (16). In contrast, Harvey and associates (75) found increased left temporal lobe volume in bipolar disorder compared with controls and Johnstone and associates (84) found no differences.

In examining localized structural abnormalities, Strakowski and colleagues examined the circuit comprised of the prefrontal cortex, thalamus, hippocampus, amygdala, globus pallidus, and striatum (LCSPT) and found a significant difference in the overall structural changes. Some of these changes were increases (amygdala, striatum), whereas others were decreases (prefrontal cortex, hippocampus). Others have also examined structural changes in amygdala in bipolar disorder, finding larger (85), smaller (77), or equal (16) volumes. As mentioned, amygdala volumes are difficult to compare between studies. Prefrontal cortex volume decreases (74 ,76 ,86) generally have been small but significant in bipolar disorder and are supported by postmortem findings of decreased glia in prefrontal cortex in bipolar subjects (27). Results have been mixed for the basal ganglia. Larger caudate volumes were found in men (87) and larger globus pallidus volumes but not striatal volumes were found in another study (88) in bipolar patients compared to healthy individuals. Other MRI studies did not find any significant differences in bipolar subjects compared to controls in caudate, putamen, or lenticular nuclei (6 ,16 ,74). Results have also been mixed for hippocampus (16 ,85 ,89) and thalamus (6 ,74).

A finding that may shed some light in interpreting the contradictory findings is the recent report that chronic lithium treatment is neuroprotective and may prevent volume loss in treated patients (90). Lithium up-regulates the neurotrophic protein Bcl-2 in rat frontal cortex, hippocampus, and striatum. In analyzing reports of volume loss in bipolar patients it may be critical to know the cumulative medication history, especially regarding lithium.

The relationship between bipolar disorder and increased hyperintensities seen on T2-weighted MRI scans (T2H) is complex. The presence of hyperintensities has been associated with hypertension; however, T2H also are increased in asymptomatic elderly. In manic patients who developed bipolar disorder after age 50, Fujikawa and associates (91) found that compared with age- and sex-matched subjects who had developed affective illness prior to age 50 there was a significantly higher incidence of T2H, comparable to the incidence in subjects with late-onset depression. These results are similar to those of McDonald and colleagues (82), who found a higher incidence of subcortical hyperintensities in late-onset bipolar disorder. In younger subjects with new onset bipolar illness Strakowski and colleagues (74) found a rate of subcortical hyperintensities 1.7 times higher than control subjects, but this was not significant. Aylward and associates (87) also found a higher rate of hyperintensities in bipolar subjects, 34% versus 3% in controls; however, the bipolar subjects were 12 years older on average. In contrast, Figiel and colleagues (45) and Dupont (92) found higher rates of hyperintensities and no age differences. One study did not find differences between bipolar subjects and controls (93).

Most MRS studies have found increased levels of choline-containing molecules in basal ganglia of bipolar subjects compared with controls (94 ,95); however, Ohara and associates (96) did not find significant differences. Similarly, PET studies have also found functional abnormalities in basal ganglia in bipolar subjects (97 ,98).

Conclusion

Although advances in technology have allowed high-resolution studies of individual brain structures, understanding the organization of brain systems has been limited by the lack of noninvasive investigation of neuronal connections between functional regions. The recent development of noninvasive neuronal fiber tracking using water diffusion properties (99) will allow increasingly sophisticated reconstruction of fiber trajectories throughout the brain. The continuing development of automated tissue segmentation methods allowing determination of gray and white matter volumes using computer-generated algorithms will provide faster and more standardized volume measures. It will be important to combine structural studies with functional studies to determine the functional significance of brain structure changes. Combining MRI and functional studies such as PET, SPECT, and fMRI has the potential to more precisely localize abnormalities in blood flow/metabolism, and neurotransmitter receptors. This integrated perspective will allow further development of a structural–functional model of depression.

Studies in high-risk populations, such as first-degree relatives of affected individuals, will assist in determining whether focal structural and functional changes are genetic/neurodevelopmental or acquired and whether they predate or follow the development of depression. Additional postmortem studies in larger samples with careful clinical screening for comorbidity are also needed to examine ultrastructural correlates of volumetric and functional changes. Neuroprotective strategies aimed at preventing the damage associated with depression are likely to be an important future direction for research. Preclinical studies provide preliminary strategies for preventing stress-induced damage. These include for example prevention of stress-induced decreases in brain-derived neurotrophic factor (BDNF) with antidepressants (100 ,101 and 102), prevention of stress-induced excitotoxic injury with phenytoin (Dilantin) (103), prevention of stress-induced decreases in neurogenesis with antidepressants (104 ,105),

and increase in dendritic branching with serotonin reuptake inhibitors (106).

FUNCTIONAL STUDIES

Part of "74 - Structural and Functional Imaging of Affective Disorders "

Functional imaging extends the sensitivity and specificity of structural imaging. As it can safely be assumed that genetic, molecular, and biochemical changes precede changes in structure, the promise of functional imaging in affective disorder is to more accurately define the pathophysiology of affective diseases, better predict potential treatments and, in general, further our knowledge of mood regulation by the human brain. Despite its development over a decade ago, functional imaging has only begun to address these primary issues. The principal reason for this slow progress is the need for extensive methodologic development in both major divisions of functional imaging. In the mapping of brain function, the imaging techniques per se have not been a major factor. Rather, the limitation has been in the development and validation of relevant affective “tasks” to selectively activate the brain regions of interest. However, the widespread lack of suitable tracers and probes is the limiting factor in molecular imaging. Sufficiently selective and sensitive probes for the different receptors, enzymes, and transporters that are putatively implicated in affective disorder must be individually developed and validated. Often such probes fail after reaching the level of human application. However, advances in imaging the serotonergic system, for example, have been reported despite these difficulties. It is in this context that we can examine the contribution of functional imaging to our understanding of affective disorder and understand the promise it holds for the future.

Mapping Brain Function in Affective Disorders

Imaging of regional neuronal activity in affective illness has yielded intriguing, but heterogeneous, findings. Some of the variability in findings may rest in the different methodologies employed and a clear understanding of the limitations of the imaging techniques is important. Cerebral blood flow (CBF) and cerebral metabolic rate of glucose (CMRG) are well accepted as markers of general regional brain activity (107). As such, increased neuronal firing is reliably associated with increased CBF and CMRG allowing the spatial distribution of either CBF or CMRG to serve as a proxy measure for brain activity.

Some early studies of quantitative regional CBF in patients with major depressive disorder studied at rest reported global reductions (108 ,109). Regional changes have been reported as well with decreased CBF and metabolism in depressed subjects relative to controls in the dorsolateral prefrontal cortex (110 ,111). However, these regional changes are small (typically 5% to 10%) and are not consistently seen (112). One conclusion from these early studies is that no visually identifiable pattern of CBF or CMRG is associated with depression, even severe depression. This is in contrast to the altered patterns of CBF and CMRG identified for some other diseases, such as Alzheimer's disease and Huntington's chorea. For this and other reasons, functional imaging appears to have no current role in the clinical diagnosis or management of affective illness. However, an important role for functional brain mapping in depressive illness is to elucidate the neuroanatomic systems involved with the symptomatology of this disease. One example has been the characterization of decreased activity in the dorsal prefrontal cortices as being related to negative symptoms (111 ,113). This finding may best be interpreted as representing the relative hypofunction of these systems clinically in some patients with depression because there is much evidence that this brain area plays an active role in working memory and related executive functions.

This last point underscores the nonuniformity of the baseline state as the likely cause of the highly variable results in functional imaging studies of affective illnesses. Different cognitive and emotional states in control subjects are well known to result in regional brain activation. Thus, depressed patients imaged in a “resting” state could have highly variable internal ruminations, emotional states, and cognitive activities. This problem is usually addressed by large sample sizes in most clinical research in which these variations average out; however, as imaging studies are expensive most research studies have been limited to small sample sizes. Combining image data across sites is frustrated by differing instrumentation and approaches to data collection (image resolution, scan timing, etc.). Further complicating imaging studies is the large potential number of regions that can be independently sampled. Only in the last few years have techniques for multiple comparison correction been fully incorporated into data analysis strategies.

Imaging research into altered brain function in affective illnesses does not have to be based solely on a single image, or snapshot. Functional brain imaging of the depressed patient in a single state, a snapshot of brain function, can be complemented by examining the functional changes during a specific task or stimulus (114). The regional brain *responses* to a cognitive or emotional task could be highly informative in understanding the brain during depression. As detailed in Chapter 29 , the most sensitive manner in which to demonstrate such brain responses is by comparing two images, on a voxel-by-voxel basis, obtained in two different states in the same individual; effectively subtracting a control image from the test image (115 ,116). Increased or decreased neuronal activity in the test state will be reflected by increased or decreased CBF in the subtraction image; thus, the pattern of increased activity “maps” the processing areas used by the brain for the task. However, many tasks pertaining to uniquely human activities (e.g., language, declarative memory, emotion) clearly involve numerous brain systems

operating simultaneously and inseparably. Interpretation of a brain-mapping image resulting from such mental activity is problematic. Activated areas assumed to be involved in low-level processing of sensory information (e.g., visual or auditory cortex) may actually contain high-level processing or be critically modulated by other attentional and cognitive processes. Interpretation of brain-mapping studies in which complex tasks are employed require careful use of specific control tasks as opposed to simple “rest” state images. The best control tasks will differ from the task of interest in only the parameter of interest.

Unfortunately, designing tasks suitable for the PET or MRI scanning environment *and* relevant to the major depressive state have been difficult. Important work has been done, however. For example, induction of different emotions via different strategies has been accomplished and demonstrates neuroanatomic systems independent of the induction strategy (117 ,118). One difficulty with many emotional tasks is that the subject is highly aware of the investigators' efforts. An alternative to inducing a consciously perceived emotion is to present affect-laden stimuli at an unconscious level (119). This work developed a technique for presenting fearful faces with masking to prevent conscious processing of the visual stimuli. Fearful faces, in contrast to neutral or happy faces, have been reported by multiple investigators to invoke an increase in amygdala activity. By masking the briefly presented (<40 msec) fearful faces with neutral faces, Whalen and colleagues (119) developed a task that isolates the subconscious processing of affect laden stimuli without the confounding variable of individual cognitive processing. Despite the inability of the normal volunteers to report the emotions of the masked faces, amygdala activation was identified with fMRI during the fearful faces presentation. Such paradigm design may prove useful in investigation of affective illnesses such as depression and anxiety disorders.

The application of brain mapping to research in affective disease is still in its infancy because of only recent development of some appropriate task paradigms; however, the area has substantial promise. For example, functional MRI was used to map changes in brain activity in depressed patients and normal controls while viewing a film segment chosen to induce sadness (120). The depressed patients had significantly more activation in areas of the prefrontal cortex. Increased activation in the prefrontal cortex of the depressed cortex could be secondary to increased processing of the stimuli and associated emotions. For many tasks increased prefrontal activation accompanies an increase in task difficulty; however, several issues, including the potential for differences in cognitive response or attention level to the film being shown, must temper the interpretation of this work. Further fMRI work on the processing of emotional stimuli in depressed patients is ongoing and will certainly shed more light on this fascinating area of research.

Some consensus has begun to emerge pointing to abnormal function of specific limbic/paralimbic regions in the depressed patient. Drevets and colleagues (3) detected decreased activity in the subgenual region of the anterior cingulate in a group of depressed bipolar patients compared to controls using an image-wide search of PET measures of CBF. Importantly, the finding replicated with a second group of depressed bipolar patients. The finding was again seen when extended to a group of depressed bipolar using CMRG measures and, then again, when extended to a group of unipolar depressed patients with familial pattern. The authors point out that the subgenual anterior cingulate is heavily interconnected with autonomic structures (including the hypothalamus), ventral striatum, amygdala, and brainstem serotonergic systems; all systems implicated in mood and behavior regulation. Mayberg and associates (121) also detected decreased CMRG in the rostral anterior cingulate of unipolar depressed patients compared to control subjects. The actual center location of this region was approximately 15 mm more dorsal than the subgenual region found by Drevets and colleagues. However, in a highly creative approach, Mayberg and associates (122) examined the regional changes in neural activity with multiple PET scans after improvement of symptoms in depressed patients and compared those to the regional changes in neural activity induced by script-induced sadness in control subjects (see Fig. 74.2). Affective symptom remission was associated with an increased activity within dorsolateral prefrontal cortex, inferior parietal, dorsal anterior cingulate, and posterior cingulate and decreases in ventral limbic and paralimbic sites, including the subgenual anterior cingulate and posterior insula.

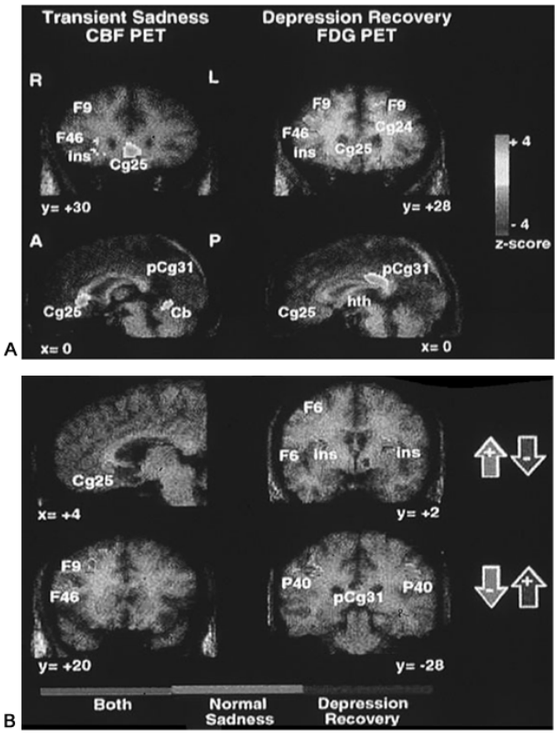


FIGURE 74.2. A: Left images show changes in regional cerebral blood flow (CBF) (with [^{15}O]water PET) associated with transient sadness in eight healthy volunteers. Right images show changes in regional glucose metabolism (with FDG PET) after six weeks of treatment in eight unipolar depressed patients. Coronal (*top row*) and sagittal (*bottom row*) views. Sadness is associated with increases in ventral paralimbic regions and decreases in dorsal frontal regions. With recovery from depression, the reverse is seen: ventral decreases and dorsal cortical increases. Slice locations are in millimeters relative to the anterior commissure. Numbers are Brodmann area designations. R, right; L, left; A, anterior; P, posterior; F, prefrontal; ins, anterior insula; Cg24, dorsal anterior cingulate; Cg25, subgenual cingulate; pCg, posterior cingulate; Cb, cerebellum; hth, hypothalamus. Color scale: red indicates increases and green indicates decreases in flow or metabolism. **B:** Logical images showing anatomical overlap of significant changes common to both experiments. *Top row:* Concordant areas with increased flow during sadness and decreased metabolism during remission. *Bottom row:* Areas with decreased flow during sadness and increased metabolism during remission. Numbers are Brodmann area designations. F, premotor; p, parietal. Red indicates changes unique to sadness; dark blue indicates changes unique to remission of depression; lighter blue indicates changes common to both. Arrows signify direction of change associated with each condition. See color version of figure.

The relevance of this particular constellation of brain regions to mood was made more certain by the findings of a similar pattern of regional brain activity, but in the opposite direction, in the normals during induced sadness. Sadness induced increases in a region of the subgenual cingulate and decreases in many of the same dorsal cortical regions. For depressed patients, the increases in dorsal cortical regions may reflect increased cognitive function in the remitted state. The decreases in the subgenual anterior cingulate as symptoms resolve appears more complicated. It is not yet possible to conclude from the data presented in Mayberg and associates (122) that the depressed state reflects an *increased* and abnormal functioning of the subgenual cingulate that returns to baseline during remission. The finding of Drevets and colleagues (3) clearly showing *decreased* metabolism in the subgenual cingulate during a depressive episode needs to be reconciled. One explanation is that the decreased activity is conceivably an actual increase in functional activity that only appears decreased on PET imaging owing to partial volume effects. This region was also reported by Drevets and colleagues to have sizable decreases in volume by MRI, a change known to decrease the measured activity. Nevertheless, the subgenual cortex and adjacent anterior

cingulate appear to play an important role in emotional regulation and expression.

Molecular Imaging of Neurotransmitter Systems in Affective Disorders

Numerous hypotheses have been proposed relating monoamine systems and depression. Serotonin, dopamine, and norepinephrine have all been implicated to various degrees in the behavioral changes and treatment of depression. Of these three, reports indicating a constellation of specific serotonergic dysfunctions in major depressive disorders have been the most convincing. For this reason, much effort in the imaging community has been focused on developing tools for imaging different aspects of the serotonergic system *in vivo*. Generally, these tools have involved the development and validation of radioligands specific for serotonin receptor subtypes (with more than 14 serotonin receptors isolated to date); however, alternative approaches to image serotonergic function have been reported as well. A report by Mann and co-workers (123) demonstrated reduced serotonin responsivity in depressed patients. In that work, fenfluramine, an indirect serotonin agonist, was used as a pharmacologic challenge and FDG PET imaging was used to image the brain responses. Depressed patients showed much less FDG uptake changes in response to fenfluramine than normal controls. Such an approach is limited, however, by the uncertainty in the relationship of the FDG changes and specific aspects of the serotonergic system.

Before reviewing the progress in this rapidly moving field, it is important to note that there is considerable difficulty in synthesizing the literature in both PET and SPECT imaging of receptor function owing to a lack of common data analysis procedures among various imaging groups. As the different data analysis techniques have specific inherent assumptions and limitations, it is useful to briefly review the pertinent issues. For the sake of brevity, we will limit this review to receptor imaging; however, the issues are often identical for imaging of other systems.

Generally, the PET or SPECT imaging device records the brain distribution of a radiotracer that binds selectively at the receptor (or transporter). In many applications a brain region is identified and a time–activity curve for that region is estimated from the images. Multiple factors directly influence the amount of tracer that accumulates in any given brain region: concentration of receptors (B_{max}), dissociation constant between the receptor and ligand (K_d), and nonspecific binding to brain tissue (corrected for in some models). In addition, the amount of uptake in the entire brain is dependent on the level of radiotracer in plasma and the degree to which the radiotracer nonspecifically binds to plasma proteins. To measure the B_{max} separately from K_d, the experiment must include some measurements with partial saturation of the receptors; this is rarely done for both logistical and ethical reasons. To address this issue, Mintun and associates (124) proposed a term binding potential (BP)

that combined both terms in the form BP = B_{max}/K_d. The binding potential reflects in a single value the ability of the regional brain receptors to bind to free radioligand during equilibrium and, importantly, is linearly related to the B_{max}. Thus, if K_d can be assumed to be constant in the experimental conditions or the different patient populations, the BP can be an adequate substitute for the regional receptor density.

The calculation of BP requires knowledge of the nonspecific binding in plasma and the brain tissue. If the nonspecific binding in plasma is measured or can be assumed, most investigators estimate the nonspecific brain binding of the tracer from a brain region that has negligible amounts of specific binding. A second requirement for calculation of BP is that the kinetics of receptor–ligand interaction be distinguishable from the transport kinetics into the brain. This is true for most tracers, with the receptor binding occurring more slowly than the transport. However, when both processes are of similar rates the BP is difficult to calculate and an alternative term, the distribution volume (DV) has been proposed (125). Conceptually, the DV of a given brain region equals the ratio of brain activity divided by plasma activity at equilibrium and includes the effects of both specific and nonspecific binding. The DV also can be calculated as a parametric image in which each pixel is assigned the appropriate calculated DV value (126). As the DV includes the effect of both specific and nonspecific binding, the DV is usually "corrected" by dividing by the DV of a region where only nonspecific binding occurs. The result, the DV_{ratio}, is a frequent form of reporting PET and SPECT receptor data. The use of the DV_{ratio} is also applied when the DV values for the areas of interest are not actually calculated from kinetic or equilibrium data, but are estimated from single images. The single image is usually standardized as being at a given time after injection of tracer but may be highly susceptible to individual differences in plasma clearance or transport into the brain.

Further complicating the use of the DV_{ratio} is that it is occasionally mistaken as being proportional to B_{max} and BP. Actually, the BP is proportional to the term DV_{ratio} - 1, which can be substantially different when the DV_{ratio} is low. Finally, the calculation of DV and DV_{ratio} is usually done without the measurement of plasma binding as it can be shown that the DV_{ratio} is mathematically independent of plasma binding. Indeed, the BP differs from the DV_{ratio} - 1 only by the incorporation of the plasma binding into the BP. Thus, if plasma binding is unknown the use of DV_{ratio} - 1 is quite appropriate and will remove all individual and group effects of plasma binding in the receptor data. What is usually not appreciated is that individual variations in *brain* nonspecific binding are *not* corrected in the DV_{ratio} - 1 calculation. This limitation of the DV_{ratio} - 1 term is rarely discussed and even more rarely quantitatively examined.

The lack of suitable radioligands for human PET imaging has significantly slowed research in affective disorders. The serotonin receptor subtypes with the highest brain densities, 5-HT_{1A} and 5-HT_{2A} receptors, do have PET radioligands with sufficient selectivity and affinity suitable for imaging (Fig. 74.3). In the 5-HT_{2A} system some initial work in depression was done using tracers with limited selectivity. Labeled spiperone (127) demonstrated alterations in binding in depressed post-stroke patients but interpretation must be limited owing to the very high affinity of this tracer for dopamine-2 receptors. 2-^{123}I-ketanserin demonstrated increased binding in the right parietal cortex of depressed patients (128), but again, the work is limited by the nonselectivity of ketanserin. Since these initial reports, methodologic development to image 5-HT_{2A} receptors has been successful using highly selective agents including [^{18}F]altanserin (129), [^{18}F]setoperone (130) and labeled MDL100,907 (131 ,132).

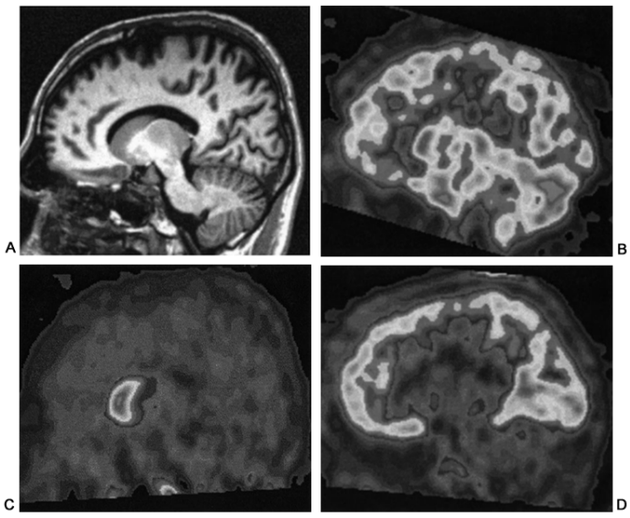


FIGURE 74.3. Parasagittal views through the head of the caudate of an anatomic magnetic resonance imaging scan and three different types of functional positron emission tomography scans. All images were collected in normal control subjects and have been converted to a standard atlas coordinate system using linear affine transformation. The upper right image is from a 60-second scan after injection with [^{15}O]water and represents the distribution of cerebral blood flow (CBF). The lower left image is from a 30-minute scan obtained 30 minutes after injection of [^{11}C]raclopride. The scan shows widespread low nonspecific binding (e.g., cerebellum) and focal increased specific binding to D_2 dopamine receptors in the caudate. The lower right image is from a 30-minute scan obtained 60 minutes after injection of [^{18}F]altanserin. Again, the cerebellum shows low nonspecific uptake, whereas the cortical gray matter shows the specific binding to the 5-HT_{2A} serotonin receptors. These images demonstrate the ability to visualize distinct aspects of brain function using different radiotracers. Relatively high CBF is seen throughout the cerebellum and cortical and subcortical gray matter. In contrast, D_2 receptors are seen to be highly concentrated in the caudate, whereas 5- HT_{2A} receptors are concentrated in the cortical gray matter with only small amount of receptor binding in basal ganglia. With the appropriate quantitative processing, these differences in radiotracer uptake can be expressed as relative or absolute receptor density. See color version of figure.

The application of these radioligand techniques in mood disorders has begun with some interesting results. Biver and associates (29) found apparent decreased [^{18}F]altanserin binding in the right insular and adjacent orbital cortices in eight medication-free depressed patients. However, as the data processing approach normalized for global [^{18}F]altanserin uptake the method cannot distinguish between a true decrease and an increase that is proportionally less than the [^{18}F]altanserin binding in the remaining brain. Nonetheless, the detection of a regional change in [^{18}F]altanserin binding suggests some type of regional difference in receptor density. Several other studies, however, have not shown altered radioligand binding to 5-HT_{2A} receptors during depression. Meyer and co-workers (133) used [^{18}F]setoperone to compare 14 medication-free depressed patients with 19 control subjects and found no difference in binding in a large prefrontal cortex region of interest. The data was reported as DV_{ratio} and quantitation using plasma activity was not done. Interestingly, the prefrontal cortex value was 1.6 in the depressed group and 1.8 in the control group. After correction of the DV_{ratio} to yield a term proportional to receptor density (the BP is proportional to DV_{ratio} - 1) this trend represents a rather large 25% decrease in the depressed group.

Yatham and colleagues (134) used [^{18}F]setoperone to measure 5-HT_{2A} binding in 20 depressed and 20 controls. All images were first normalized by cerebellar activity to account for differences in nonspecific uptake and then processed using statistical parametric mapping (SPM). Unfortunately, the actual DV image was not calculated, nor the DV_{ratio}, as only raw PET activity from a single time frame was used. SPM performed a pixel-by-pixel ANCOVA to detect pixels with significant group differences and showed generalized decreased binding in the frontal and parietal cortices of the depressed patient group. Yatham and colleagues reported that in some voxels the decrease in uptake was 20% to 27%. As the study was done without an equilibrium state or kinetic analysis, it is unknown whether transport

issues could be responsible for some of these results. In another study, Attar-Levy and associates (135) used [^{18}F]setoperone to measure 5-HT_{2A} binding in seven depressed patients and seven age-matched controls. Regions of interest over multiple cortical areas were used to measure radioligand uptake at a given time after injection and the data were normalized by first subtracting cerebellar activity and then dividing by injected dose. This result is a nonstandard term that is not corrected for variations in plasma nonspecific binding (but may be corrected for variations in brain nonspecific binding). There was little difference between depressed and control values with only the frontal region demonstrating a significant decrease in bindings (approximately 6%). However, most of the patients were concurrently being treated with benzodiazepines (six of seven), which may have further confounded the results. Finally, Meltzer and colleagues (136) used [^{18}F]altanserin to image eleven late-life depressed patients and age-matched controls. Logan graphical analysis was used to analyze regional activity data and quantitate DV and the DV_{ratio} - 1. Meltzer and colleagues reported no difference between the depressed and control groups.

The preceding data have the common thread that regional measures of 5-HT_{2A} binding appear to yield slight, nonsignificant decreases in depressed populations compared to controls. The image-wide processing methods using SPM, which had different methods of normalizing the radioligand uptake data, appeared to yield more dramatic decreases in the depressed populations. As these parametric methods are not well validated with these radiotracers, the significance of the findings will be uncertain until further work, likely using conventional analyses, is done.

The 5-HT_{1A} receptor distribution has been imaged in humans using the high-affinity antagonist [^{11}C]WAY 100,635 (137 ,138). The images are unusual because of the very high ratio of specific to nonspecific binding in the brain. Compared to images from labeled setoperone or altanserin, in which approximately one-half of all of the brain activity is not bound to the receptor of interest, the [^{11}C]WAY 100,635 PET images have less than 10% of the activity in nonspecific binding and 90% specific activity. The ratio of receptor-rich brain regions to the cerebellum, assumed to be devoid of 5-HT_{1A} receptors, is typically greater than 15, depending on the timing of the scan (compared to a ratio of 2 to 3 for altanserin or setoperone). This very high “target-to-background” allows the imaging and even quantitation of 5-HT_{1A} receptor content in very small structures, such as the midbrain raphe. Early reports show that 5-HT_{1A} receptor binding of [^{11}C]WAY 100,635 in a variety of cortical regions and in the raphe is decreased in depressed patients compared to controls (139 ,140). The observed decreases are substantial, ranging from 20% to 40% in some regions. As 5-HT_{1A} receptors are involved in widespread modulation of function in limbic and paralimbic regions, these findings are of considerable importance. Furthermore, the 5-HT_{1A} receptors are part of the autoregulation of serotonergic innervation in the raphe, increasing the significance of these findings.

Work on imaging serotonin reuptake sites, the target of the most commonly used antidepressants, is ongoing. The radioligand that has undergone the most study is [^{11}C]McNeil 5652. However, nonspecific binding with this tracer is high, and separating the receptor binding from nonspecific binding has been challenging (141 ,142). One alternative tracer is [^{123}I]B-CIT SPECT imaging. This tracer binds to other reuptake sites but has simpler kinetics and can be used in brain regions that have predominantly serotonin reuptake binding. This property was exploited in a study by Malison and co-workers of unipolar, unmedicated depressed patients in which brainstem serotonin reuptake site binding was significantly reduced by nearly 20% compared to controls (143). Another area of interest is in the *in vivo* measure of serotonin synthesis. This has been achieved using α-^{11}C-methyl-tryptophan (144). The tracer is converted to α-^{11}C-methyl-serotonin within neurons and then accumulates, unable to be degraded by monoamine oxidases. The rate of accumulation is argued to be proportional to endogenous serotonin synthesis (144). However, validation is problematic because there are few tools for verifying rates of serotonin synthesis. Also, it has been noted that the synthetic rates appear to be highly dependent on plasma tryptophan levels.

Dopaminergic innervation in depression may be altered. D'haenen and Bossuyt (145) imaged 23 depressed patients and 11 controls using SPECT and ^{123}I-iodobenzamide (IBZM), a high-affinity ligand for the D_2 receptor. The activity in the basal ganglia after normalizing with cerebellar activity was 10% greater in the depressed subjects ($p < 0.025$). The authors suggest that this increase reflects decreased dopaminergic neurotransmission. Decreased synaptic dopamine could then result in decreased occupancy of the D_2 receptors and D_2 up-regulation, both factors could lead to increased IBZM binding. In two other reports the IBZM uptake in the striatum was shown to decrease during treatment with antidepressants (146 ,147), although decreased D_2 receptor binding at baseline in depression was not seen. However, the striatal dopamine system may not be the most critical in affective disorders. With the advent of radioligands able to image extrastriatal D_2 receptors (e.g., [^{18}F]fallypride) (148), investigation of dopaminergic function in limbic cortical regions will be possible.

Conclusion

The last decade has produced numerous advances in our ability to image brain function. Although the application to affective disorders has been limited to date, the data are tantalizing. Findings have identified abnormalities in the function of limbic cortical structures and the location of these structures overlap with those areas involved with generation of emotion. With increasing sophistication of emotional paradigms, a more precise picture of the role these structures play in mood regulation may emerge. Serotonergic alterations are being further identified with existing techniques and new radioligands will be introduced over the next decade that will greatly expand our imaging capabilities. Novel methods will be explored, leading to agents able to image aspects of gene expression, perhaps even with the spatial and temporal resolution of MRI (149).

ACKNOWLEDGMENT

Part of "74 - Structural and Functional Imaging of Affective Disorders "

Yvette I. Sheline was supported in part by MH01370 and MH58444.

REFERENCES

1. Krishnan KR, McDonald W, Escalona P, et al. MRI of the caudate nuclei in depression. *Arch Gen Psychiatry* 1992;49:553–557.

2. Coffey CE, Wilkinson WE, Weiner RD, et al. Quantitative cerebral anatomy in depression: a controlled magnetic resonance imaging study. *Arch Gen Psychiatry* 1993;50:7–16.

3. Drevets WC, Price JL, Simpson JR Jr, et al. Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* 1997;386:824–827.

4. Greenwald BS, Kramer-Ginsberg E, Bogerts B, et al. Qualitative magnetic resonance imaging findings in geriatric depression. Possible link between later-onset depression and Alzheimer's disease? *Psychol Med* 1997;27:421–431.

5. Husain MM, McDonald WM, Doraiswamy PM, et al. A magnetic resonance imaging study of putamen nuclei in major depression. *Psychiatry Res* 1991;40:95–99.

6. Dupont RM, Jernigan TL, Heindel W, et al. Magnetic resonance imaging and mood disorders. Localization of white matter and other subcortical abnormalities. *Arch Gen Psychiatry* 1995;52:747–755.

7. Lenze E, Sheline Y. Absence of striatal volume differences between healthy depressed subjects and matched comparisons. *Am J Psychiatry* 1999;156(12):1989–1991.

8. Axelson DS, Doraiswamy PM, Boyko OB, et al. In vivo assessment of pituitary volume with magnetic resonance imaging and systematic stereology: relationship to dexamethasone suppression test results in patients. *Psychiatry Res* 1992;44:63–70.

9. Sheline Y, Wang P, Gado M, et al. Hippocampal atrophy in recurrent major depression. *Proc Natl Acad Sci USA* 1996;93:3908–3913.

10. Shah PJ, Ebmeier KP, Glabus MF, et al. Cortical grey matter reductions associated with treatment resistant chronic unipolar depression. Controlled magnetic resonance imaging study. *Br J Psychiatry* 1998;172:527–532.

11. Sheline Y, Sanghavi M, Mintun M, et al. Depression duration but not age predicts hippocampal volume loss in women with recurrent major depression. *J Neurosci* 1999;19:5034–504.

12. Bremner JD, Narayan M, Anderson ER, et al. Hippocampal volume reduction in major depression. *Am J Psychiatry* 2000;157:115–118.

13. Mervaala E, Fohr J, Kononen M, et al. Quantitative MRI of the hippocampus and amygdala in severe depression. *Psychol Med* 2000;30:117–125.

14. Vakili K, Pillay SS, Lafer B, et al. volume in primary unipolar major depression: a magnetic resonance imaging study. *Biol Psychiatry* 2000;47:1087–1090.

15. Sheline YI, Gado MH, Price JL. Amygdala core nuclei volumes are decreased in recurrent major depression. *Neuroreport* 1998;9:2023–2028.

16. Swayze VW, Andreasen NC, Alliger RJ, et al. Subcortical and temporal structures in affective disorder and schizophrenia: a magnetic resonance imaging study. *Biol Psychiatry* 1992;31:221–240.

17. Axelson DS, Doraiswamy PM, McDonald WM, et al. Hypercortisolemia and hippocampal changes in depression. *Psychiatry Res* 1993;47:163–173.

18. Pantel J, Schroder J, Essig M, et al. Quantitative magnetic resonance imaging in geriatric depression and primary degenerative dementia. *J Affect Disord* 1997;42:69–83.

19. Ashtari M, Greenwald BS, Kramer-Ginsberg E, et al. Hippocampal/amygdala volumes in geriatric depression. *Psychol Med* 1999;29:629–638.

20. Price J, Russchen F, Amaral D. The limbic region. II: The amygdaloid complex. In: Bjorkland A, Hokfelt T, Swanson L, eds. *Handbook of chemical neuroanatomy*, vol 5. New York: Elsevier, 1987:279–388.

21. Swerdlow NR, Koob GF. Dopamine, schizophrenia, mania and depression: toward a unified hypothesis of cortico-striato-pallido-thalamic function. *Behav Brain Sci* 1987;10:197–245.

22. Burt DB, Zembar MJ, Niederehe G. Depression and memory impairment: a meta-analysis of the association, its pattern, and specificity. *Psychol Bull* 1995;117:285–305.

23. Bowley MP, Drevets WC, Ongur D, et al. Glial changes in the amygdala and entorhinal cortex in mood disorders. *Soc Neurosci Abst* 2000;26:876.

24. Steffens DC, Krishnan KR. Structural neuroimaging and mood disorders: recent findings, implications for classification, and future directions. *Biol Psychiatry* 1998;43:705–712.

25. Coffey CE, Wilkinson WE, Parashos IA, et al. Quantitative cerebral anatomy of the aging human brain: a cross-sectional study using magnetic resonance imaging. *Neurology* 1992;42:527–536.

26. Rajkowska G, Miguel-Hidalgo JJ, Wei J, et al. Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. *Biol Psychiatry* 1999;45:1085–1098.

27. Ongur D, Drevets WC, Price JL. Glial reduction in the subgenual prefrontal cortex in mood disorders. *Proc Natl Acad Sci USA* 1998;95:13290–13295.

28. Arango V, Underwood MD, Gubbi AV, et al. Localized alterations in pre- and postsynaptic serotonin binding sites in the ventrolateral prefrontal cortex of suicide victims. *Brain Res* 1995;688:121–133.

29. Biver F, Wikler D, Lotstra F, et al. Serotonin 5-HT_2 receptor imaging in major depression: focal changes in orbito-insular cortex. *Br J Psychiatry* 1997;171:444–448.

30. Mintun MA, Sheline YI, Moerlein SM. Regional [^{18}F]altanserin binding in the treatment of major depression. *Neuroimage* 2000;11:S83.

31. Price JL. Prefrontal cortical networks related to visceral function and mood. *Ann NY Acad Sci* 1999;877:383–396.

32. Duman RS. Novel therapeutic approaches beyond the serotonin receptor. *Biol Psychiatry* 1998;44:324–335.

33. Sapolsky RM. Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Arch Gen Psychiatry* 2000;57:925–935.

34. Heim C, Ehlert U, Hellhammer DH. The potential role of hypocortisolism in the pathophysiology of stress-related bodily disorders. *Psychoneuroendocrinology* 2000;25:1–35.

35. Lupien SJ, de Leon M, deSanti S, et al. Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nat Neurosci* 1998;1:69–73.

36. White LE, Price JL. The functional anatomy of limbic status epileptics in the rat. II. The effects of focal deactivation. *J Neurosci* 1994;13:4810–4830.

37. Carmichael ST, Price JL. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol* 1995;363:615–641.

38. Szatkowski M, Attwell D. Triggering and execution of neuronal death in brain ischemia: two phases of glutamate release by different mechanisms. *Trends Neurosci* 1994;17:359–365.

39. Ransom BR, Sontheimer H. The neurophysiology of glial cells. *J Clin Neurophysiology* 1992;9:224–251.

40. Gould E, McEwen BS, Tanapat P, et al. Neurogenesis in the dentate gyrus of the adult tree shrew is regulated by psychosocial stress and NMDA receptor activation. *J Neurosci* 1997;17:2492–2498.

41. Cameron HA, Gould E. Adult neurogenesis is regulated by adrenal steroids in the dentate gyrus. *Neuroscience* 1994;61:203–209.

42. Gould E, Reeves AJ, Graziano MS, et al. Neurogenesis in the neocortex of adult primates. *Science* 1999;286:548–552.

43. Jacoby RJ, Levy R, Bird JM. Computed tomography and the outcome of affective disorder: A follow-up study of elderly patients. *Br J Psychiatry* 1981;139:288–292.

44. Coffey C, Figiel G, Djang W. Leukoencephalopathy in elderly depressed patients referred for ECT. *Biol Psychiatry* 1988;24:143–161.

45. Figiel GS, Krishnan KRR, et al. Subcortical hyperintensities on brain Magnetic Resonance Imaging: a comparison between late age onset and early-onset elderly depressed subjects. *Neurobiol Aging* 1991;26:245–247.

46. Baron M, Mendlewicz J, Klotz J. Age-of-onset and genetic transmission in affective disorders. *Acta Psychiatr Scand* 1981;64:373–380.

47. Alexopoulos GS, Meyers B, Young R, et al. Recovery in geriatric depression. *Arch Gen Psychiatry* 1996;53:305–312.

48. Simpson S, Baldwin RC, Jackson A, et al. Is subcortical disease associated with a poor response to antidepressants? Neurological, neuropsychological and neuroradiological findings in late-life depression. *Psychol Med* 1998;28:1015–1026.

49. Soares JC, Mann JJ. The anatomy of mood disorders: review of structural neuroimaging studies. *Biol Psychiatry* 1997;41:86–106.

50. Rothschild AJ, Benes F, Hebben N, et al. Relationships between brain CT scan findings and cortisol in psychotic and nonpsychotic depressed patients. *Biol Psychiatry* 1989;26:565–575.

51. Rabins PV, Pearlson GD, Aylward E, et al. Cortical magnetic resonance imaging changes in elderly inpatients with major depression. *Am J Psychiatry* 1991;148:617–620.

52. Kobayashi S, Okada K, Yamashita K. Incidence of silent lacunar lesion in normal adults and its relation to cerebral blood flow and risk factors. *Stroke* 1991;22:1379–1383.

53. Aronson S. Intracranial vascular lesions in patients with diabetes mellitus. *J Neuropathol Exp Neurol* 1973;32:183–196.

54. Starkman MN, Gebarski SS, Berent S, et al. Hippocampal formation volume, memory dysfunction, and cortisol levels in patients with Cushing's syndrome. *Biol Psychiatry* 1992;32:756–765.

55. Charness ME. Brain lesions in alcoholics. *Alcoholism: Clin Exp Res* 1993;17:2–11.

56. Folstein SE, Abbott MH, Chase GA, et al. The association of affective disorder with Huntington's disease in a case series and in families. *Psychol Med* 1983;13:537–542.

57. Starkstein SE, Robinson RG. Affective disorders and cerebral vascular disease. *Br J Psychiatry* 1989;154:170–182.

58. Burns A, Jacoby R, Levy R. Psychiatric phenomena in Alzheimer's disease. III. Disorders of mood. *Br J Psychiatry* 1990;157:81–86.

59. Cummings JL. Depression and Parkinson's disease: a review. *Am J Psychiatry* 1992;149:443–454.

60. Kumar A, Jin Z, Bilker W, et al. Late-onset minor and major depression: early evidence for common neuroanatomic substrates detected by using MRI. *Proc Natl Acad Sci USA* 1998;95:7654–7658.

61. Coffey CE, Figiel GS, Djang WT, et al. Subcortical hyperintensity on magnetic imaging: a comparison of normal and depressed elderly subjects. *Am J Psychiatry* 1990;147:187–189.

62. Zubenko GS, Sullivan P, Nelson JP, et al. Brain imaging abnormalities in mental disorders of late life. *Arch Neurol* 1990;47:1107–1111.

63. Lesser IM, Miller BL, Boone KB, et al. Brain injury and cognitive function in late-onset psychotic depression. *J Neuropsychiatry Clin Neurosci* 1991;3:33–40.

64. Howard RJ, Beats B, Forstl H, et al. White matter changes in late-onset depression: a magnetic resonance imaging study. *Int J Geriat Psychiatry* 1993;8:183–185.

65. Krishnan KR, McDonald WM, Doraiswamy PM. Neuroanatomic substrates of depression in the elderly. *Eur Arch Psychatry Clin Neurosci* 1993;243:41–46.

66. Hickie I, Scott E, Mitchell P, et al. Subcortical hyperintensities on magnetic resonance imaging: clinical correlates and prognostic significance in patients with severe depression. *Biol Psychiatry* 1995;37:151–160.

67. Guze BH, Szuba MP. Leukoencephalopathy and major depression: a preliminary report. *Psychiatry Res* 1992;45:169–175.

68. Fazekas F, Kleinert R, Offenbacher H, et al. The morphologic correlate of incidental punctate white matter hyperintensities on MR images. *AJNR Am J Neuroradiol* 1991;12:915–921.

69. Mirsen TR, Lee DH, Wong CJ, et al. Clinical correlates of white-matter changes on magnetic resonance imaging scans of the brain. *Arch Neurol* 1991;48:1015–1021.

70. Fujikawa T, Yamawaki S, Touhouda Y. Incidence of silent cerebral infarction in patients with major depression. *Stroke* 1993;24:1631–1634.

71. Fujikawa T, Yamawaki S, Touhouda Y. Background factors and clinical symptoms of major depression with silent cerebral infarction. *Stroke* 1994;25:798–801.

72. Krishnan KR, Goli V, Ellinwood EH, et al. Leukoencephalopathy in patients diagnosed as major depressive. *Biol Psychiatry* 1988;23:519–522.

73. Krishnan K, Hays J, Blazer D. MRI-Defined vascular depression. *Am J Psychiatry* 1997;154:497–501.

74. Strakowski SM, Wilson DR, Tohen M, et al. Structural brain abnormalities in first-episode mania. *Biol Psychiatry* 1993;33:602–609.

75. Harvey I, Persaud R, Ron MA, et al. Volumetric MRI measurements in bipolars compared with schizophrenics and healthy controls. *Psychol Med* 1994;24:689–699.

76. Schlaepfer TE, Harris GJ, Tien AY, et al. Decreased regional cortical gray matter volume in schizophrenia. *Am J Psychiatry* 1994;151:842–848.

77. Pearlson GD, Barta PE, Powers RE, et al. Ziskind-Somerfeld Research Award 1996. Medial and superior temporal gyral volumes and cerebral asymmetry in schizophrenia versus bipolar disorder. *Biol Psychiatry* 1997;41:1–14.

78. Zipursky RB, Seeman MV, Bury A, et al. Deficits in gray matter volume are present in schizophrenia but not bipolar disorder. *Schizophr Res* 1997;26:85–92.

79. Young RC, Patel A, Meyers BS, et al. Alpha(1)-acid glycoprotein, age, and sex in mood disorders. *Am J Geriatr Psychiatry* 1999;7:331–334.

80. Lim KO, Rosenbloom MJ, Faustman WO, et al. Cortical gray matter deficit in patients with bipolar disorder. *Schizophr Res* 1999;40:219–227.

81. Swayze VW 2d, Andreasen NC, Alliger RJ, et al. Structural brain abnormalities in bipolar affective disorder. Ventricular enlargement and focal signal hyperintensities. *Arch Gen Psychiatry* 1990;47:1054–1059.

82. McDonald WM, Krishnan KR, Doraiswamy PM, et al. Occurrence of subcortical hyperintensities in elderly subjects with mania. *Psychiatry Res* 1991;40:211–220.

83. Altshuler LL, Conrad A, Hauser P, et al. Reduction of temporal lobe volume in bipolar disorder: a preliminary report of magnetic resonance imaging. *Arch Gen Psychiatry* 1991;48:482–483.

84. Johnstone EC, Owens DG, Crow TJ, et al. Temporal lobe structure as determined by nuclear magnetic resonance in schizophrenia and bipolar affective disorder. *J Neurol Neurosurg Psychiatry* 1989;52:736–741.

85. Altshuler LL, Bartzokis G, Grieder T, et al. Amygdala enlargement in bipolar disorder and hippocampal reduction in schizophrenia: an MRI study demonstrating neuroanatomic specificity. *Arch Gen Psychiatry* 1998;55:663–664.

86. Coffman JA, Bornstein RA, Olson SC, et al. Cognitive impairment and cerebral structure by MRI in bipolar disorder. *Biol Psychiatry* 1990;27:1188–1196.

87. Aylward EH, Roberts-Twillie JV, Barta PE, et al. Basal ganglia volumes and white matter hyperintensities in patients with bipolar disorder. *Am J Psychiatr* 1994;151:687–693.

88. Strakowski SM, DelBello MP, Sax KW, et al. Brain magnetic resonance imaging of structural abnormalities in bipolar disorder. *Arch Gen Psychiatry* 1999;56(3):254–260.

89. Hauser P, Altshuler LL, Berrettini W, et al. Temporal lobe measurement in primary affective disorder by magnetic resonance imaging. *J Neuropsychiatry Clin Neurosci* 1989;1:128–134.

90. Manji HK, Moore GJ, Chen G. Lithium up-regulates the cytoprotective protein Bcl-2 in the CNS in vivo: a role for neurotrophic and neuroprotective effects in manic depressive illness. *J Clin Psychiatry* 2000;61:82–96.

91. Fujikawa T, Yamawaki S, Touhouda Y. Silent cerebral infarctions in patients with late-onset mania. *Stroke* 1995;26:946–949.

92. Dupont RM, Jernigan TL, Butters N, et al. Subcortical abnormalities detected in bipolar affective disorder using magnetic resonance imaging. Clinical and neuropsychological significance. *Arch Gen Psychiatry* 1990;47:55–59.

93. Brown FW, Lewine RJ, Hudgins PA, et al. White matter hyperintensity signals in psychiatric and nonpsychiatric subjects. *Am J Psychiatry* 1999;149:620–625.

94. Lafer B, Renshaw PF, Sachs G, et al. Proton MRS of the basal ganglia in bipolar disorder. *Biol Psychiatry* 1994;35:685.

95. Hamakawa H, Kato T, Murashita J, et al. Quantitative proton magnetic resonance spectroscopy of the basal ganglia in patients with affective disorders. *Eur Arch Psychiatry Clin Neurosci* 1998;248:53–58.

96. Ohara K, Isoda H, Suzuki Y, et al. Proton magnetic resonance spectroscopy of the lenticular nuclei in bipolar I affective disorder. *Psychiatry Res Neuroimag* 1998;84:55–60.

97. Buchsbaum MS, Wu J, DeLisi LE, et al. Frontal cortex and basal ganglia metabolic rates assessed by positron emission tomography with [^{18}F]2-deoxyglucose in affective illness. *J Affect Disord* 1986;10:137–152.

98. Kishimoto H, Takazu O, Ohno S, et al. 11C-glucose metabolism in manic and depressed patients. *Psychiatry Res* 1987;22:81–88.

99. Conturo TE, Lori NF, Cull TS, et al. Tracking neuronal fiber pathways in the living human brain. *Proc Natl Acad Sci USA* 1999;96:10422–10427.

100. Nibuya M, Morinobu S, Duman RS. Regulation of BDNF and trkB mRNA in rat brain by chronic electoconvulsive seizure and antidepressant drug treatments. *J Neurosci* 1995;15:7539–7547.

101. Nibuya M, Nestler EJ, Duman RS. Chronic antidepressant administration increases the expression of cAMP response element binding protein (CREB) in rat hippocampus. *J Neurosci* 1996;16:2365–2372.

102. Viadya VA, Duman RS. Role of 5-HT2A receptors in down-regulation of BDNF by stress. *Neurosci Lett* 1999;287:1–4.

103. Watanabe Y, Gould H, Cameron D, et al. Phenytoin prevents stress and corticosterone induced atrophy of CA3 pyramidal neurons. *Hippocampus* 1992;2:431–436.

104. Duman RS, Malberg JE. Neural plasticity in the pathophysiology and treatment of depression. *Am Coll Neuropsychopharmacol* 1998;37:261.

105. Jacobs BL, Tanapat P, Reeves AJ, et al. Serotonin stimulates the production of new hippocampal granule neurons via the 5-HT1A receptor in the adult rat. *Soc Neurosci* 1998;24: 1992.

106. Duman RS, Heninger GR, Nestler EJ. A molecular and cellular theory of depression. *Arch Gen Psychiatry* 1997;54:597–606.

107. Raichle ME. Higher function of the brain. In: Bloom F, ed. *Handbook of physiology: Section 1. The nervous system*. Bethesda, MD: American Physiological Society, 1987:643–674.

108. Sackeim HA, Prohovnik I, Moeller JR, et al. Regional cerebral blood flow in mood disorders. I. Comparison of major depressives and normal controls at rest. *Arch Gen Psychiatry* 1990;47:60–70.

109. Lesser IM, Mena I, Boone KB, et al. Reduction of cerebral blood flow in older depressed patients. *Arch Gen Psychiatry* 1994;51:677–686.

110. Baxter LR Jr, Schwartz JM, Phelps ME, et al. Reduction of prefrontal cortex glucose metabolism common to three types of depression. *Arch Gen Psychiatry* 1989;46:243–250.

111. Dolan RJ, Bench CJ, Liddle PF, et al. Dorsolateral prefrontal cortex dysfunction in the major psychoses: symptom or disease specificity? *J Neurol Neurosurg Psychiatry* 1993;56:1290–1294.

112. Drevets WC, Videen TO, Price JL, et al. A functional anatomic study of unipolar depression. *J Neurosci* 1992;12:3628–3641.

113. Bench CJ, Friston KJ, Brown RG, et al. Regional cerebral blood flow in depression measured by positron emission tomography: the relationship with clinical dimensions. *Psychol Med* 1993;23:579–590.

114. Grasby PM. Imaging strategies in depression. *J Psychopharmacol* 1999;13:346–351.

115. Fox PT, Mintun MA. Noninvasive functional brain mapping by change-distribution analysis of averaged PET images of H215O tissue activity. *J Nucl Med* 1989;30:141–149.

116. Friston KJ, Frith CD, Liddle PF, et al. Comparing functional (PET) images: the assessment of significant change. *J Cereb Blood Flow Metab* 1991;11:690–699.

117. Lane RD, Reiman EM, Ahern GL, et al. Neuroanatomic correlates of happiness, sadness, and disgust. *Am J Psychiatry* 1997;154:926–933.

118. Reiman EM, Lane RD, Ahern GL, et al. Neuroanatomic correlates of externally and internally generated human emotion. *Am J Psychiatry* 1997;154:918–925.

119. Whalen PJ, Rauch SL, Etcoff NL, et al. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *J Neurosci* 1998;18:411–418.

120. Beauregard M, Leroux JM, Bergman S, et al. The functional neuroanatomy of major depression: an fMRI study using an emotional activation paradigm. *Neuroreport* 1998;9:3253–3258.

121. Mayberg HS, Brannan SK, Mahurin RK, et al. Cingulate function in depression: a potential predictor of treatment response. *Neuroreport* 1997;8:1057–1061.

122. Mayberg HS, Liotti M, Brannan SK, et al. Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *Am J Psychiatry* 1999;156:675–682.

123. Mann JJ, Malone KM, Diehl DJ, et al. Demonstration in vivo of reduced serotonin responsivity in the brain of untreated depressed patients. *Am J Psychiatry* 1996;153:174–182.

124. Mintun MA, Raichle ME, Kilbourn MR, et al. A quantitative model for the in vivo assessment of drug binding sites with positron emission tomography. *Ann Neurol* 1984;15:217–227.

125. Koeppe RA, Holthoff VA, Frey KA, et al. Compartmental analysis of [^{11}C]flumazenil kinetics for the estimation of ligand transport rate and receptor distribution using positron emission tomography. *J Cereb Blood Flow Metab* 1991;11:735–744.

126. Frey KA, Holthoff VA, Koeppe RA, et al. Parametric in vivo imaging of benzodiazepine receptor distribution in human brain. *Ann Neurol* 1991;30:663–672.

127. Mayberg HS, Robinson RG, Wong DF, et al. PET imaging of cortical S_2 serotonin receptors after stroke: lateralized changes and relationship to depression. *Am J Psychiatry* 1988;145:937–943.

128. D'haenen H, Bossuyt A, Mertens J, et al. SPECT imaging of serotonin2 receptors in depression. *Psychiatry Res* 1992;45(4):227–237.

129. Sadzot B, Lemaire C, Maquet P, et al. Serotonin 5HT2 receptor imaging in the human brain using positron emission tomography and a new radioligand, [^{18}F]altanserin: results in young normal controls. *J Cerebr Blood Flow Metab* 1995;15:787–797.

130. Blin J, Sette G, Fiorelli M, et al. A method for the in vivo investigation of the serotonergic 5-HT_2 receptors in the human cerebral cortex using positron emission tomography and 18F-labeled setoperone. *J Neurochem* 1990;54:1744–1754.

131. Lundkvist C, Halldin C, Ginovart N, et al. [^{11}C]MDL 100907, a radioligand for selective imaging of 5-HT_{2A} receptors with positron emission tomography. *Life Sci* 1996;58:187–192.

132. Watabe H, Channing MA, Der MG, et al. Kinetic analysis of the 5-HT_{2A} ligand [11C]MDL 100,907. *J Cerb Blood Flow Metab* 2000;20:809–909.

133. Meyer JH, Kapur S, Houle S, et al. Prefrontal cortex 5-HT_2 receptors in depression: an [^{18}F]setoperone PET imaging study. *Am J Psychiatry* 1999;156:1029–1034.

134. Yatham LN, Liddle PF, Shiah IS, et al. Brain serotonin$_2$ receptors in major depression. A positron emission tomography study. *Arch Gen Psychol* 2000;57:850–858.

135. Attar-Levy D, Martinot JL, Blin J, et al. The cortical serotonin$_2$receptors studied with positron emission tomography and [^{18}F]-Septoperone during depressive illness and antidepressant treatment with clomipramine. *Biol Psychiatry* 1999;45:180–186.

136. Meltzer CC, Price JC, Mathis CA, et al. PET imaging of serotonin type 2A receptors in late-life neuropsychiatric disorders. *Am J Psychiatry* 1999;156:1871–1878.

137. Mathis CA, Simpson NR, Mahmood K, et al. [^{11}C]WAY 100635: a radioligand for imaging 5-HT1A receptors with positron emission tomography. *Life Sci* 1994;55:PL403–407.

138. Pike VW, McCarron JA, Lammertsma AA, et al. Exquisite delineation of 5-HT_{1A} receptors in human brain with PET and [*carbonyl* - 11C] WAY-100635. *Eur J Pharmacol* 1996;301:R5-R7.

139. Drevets WC, Frank E, Price JL, et al. PET imaging of serotonin 1A receptor binding in depression. *Biol Psychiatry* 1999;46:1375–1387.

140. Sargent PA, Kjaer KH, Bench CJ, et al. Brain serotonin 1A receptor binding measured by positron emission tomography with [^{11}C]WAY-100635: effects of depression and antidepressant treatment. *Arch Gen Psychiatry* 2000;57:174–180.

141. Szabo Z, Kao PF, Scheffel U, et al. Positron emission tomography imaging of serotonin transporters in the human brain. *Synapse* 1995;20:37–43.

142. Parsey RV, Kegeles LS, Hwang DR, et al. In vivo quantification of brain serotonin transporters in humans using [^{11}C]McN 5652. *J Nucl Med* 2000;41:1465–1477.

143. Malison RT, Price LH, Berman R, et al. Reduced brain serotonin transporter by [123I]-2 beta-carbomethoxy-3 beta-(4-idophenyl) tropane and single photon emission computed tomography. *Biol Psychiatry* 1998;44:1090–1098.

144. Chugani DC, Muzik O, Chakraborty P, et al. Human brain serotonin synthesis capacity measured in vivo with α-[C-11]methyl-$_L$-tryptophan. *Synapse* 1998;28:33–43.

145. D'haenen HA, Bossuyt A. Dopamine D2 receptors in depression measured with single photon emission computed tomography. *Biol Psychiatry* 1994;35:128–132.

146. Ebert D, Feistel H, Loew T, et al. Dopamine and depression. Striatal dopamine D2 receptor SPECT before and after antidepressant therapy. *Psychopharmacol* 1996;126:91–94.

147. Larisch R, Klimke A, Vosberg H, et al. In vivo evidence for the involvement of dopamine-D2 receptors in striatum and anterior cingulate gyrus in major depression. *Neuroimage* 1997;5:251–260.

148. Christian BT, Narayanan TK, Shi B, et al. Quantitation of striatal and extrastriatal D-2 dopamine receptors using PET imaging of [(^{18})F]fallypride in nonhuman primates. *Synapse* 2000;38:71–79.

149. Louie AY, Huber MM, Ahrens ET, et al. In vivo visualization of gene expression using magnetic resonance imaging. *Nat Biotechnol* 2000;18:321–325.

75

Current and Emerging Therapeutics for Depression

A. John Rush

Neal D. Ryan

A. John Rush: Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, Texas.

Neal D. Ryan: Western Psychiatric Institute & Clinic, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.

This chapter discusses critical conceptual and practical issues confronting clinicians who must distill the massive neuroscientific, psychopathologic, and clinical research information about the basis for clinical depression and its treatment and who must apply that knowledge to individual patients. This chapter does not provide an encyclopedic review of antidepressant treatments. (See refs. 1 and 2 for recent reviews.) Rather, we focus on major recent conceptual shifts in our understanding of depression and its treatment, and on practical dilemmas encountered in daily practice. The latter often calls for types of information not usually provided by standard clinical research protocols designed to obtain regulatory approval of new antidepressant agents.

After highlighting recent revisions in our knowledge about depressive disorders, we discuss the implications of that knowledge for the treatment of these conditions. We highlight the gaps in our knowledge about how to implement treatments so as to obtain optional patient outcomes. We then examine several commonly held beliefs that, rightly or wrongly, guide current treatment selection and patient care. A brief introduction to agents in development (e.g., corticotropin-releasing factor [CRF] or substance P antagonists), to alternative therapies (e.g., S-adenosyl-methionine [SAMe], St. John's wort [*hypericum perforatum*]), and to the treatment of depression in children and adolescents is provided. We conclude with suggestions for further research.

RECONCEPTUALIZING DEPRESSION AND ITS TREATMENT

Part of "75 - Current and Emerging Therapeutics for Depression "

Only two decades ago, clinical depression was seen as a transient, typically self-limited reaction to “untoward” events. Today, as detailed by Keller and Kovacs (3), clinical depressions are now recognized as far more chronic, more often recurrent (typically with a waxing and waning course), and more disabling. Historically, symptom severity has been used to distinguish different forms of depression (e.g., major depressive disorder versus dysthymia). More recent evidence, however, indicates substantial functional impairment in both “major”(4 ,5 and 6) and “minor” forms of depression (7 ,8). A more chronic course and greater symptom severity both contribute to greater levels of disability. Furthermore, prevalence rates and the degree of disability found in nonmajor forms of depression provide a basis for regarding even modest levels of nontransient depressive symptoms as a major target for treatment. Consequently, efficacy studies have been undertaken with more chronic forms of depression (9 ,10 and 11) and with “nonmajor” forms of depression (e.g., dysthymic disorder) (12 ,13 ,14 ,15 ,16 and 17).

Chronicity is often reflected in persistent residual symptoms (i.e., depressive symptoms that do not meet formal diagnostic criteria for major depressive disorder between major depressive episodes). Such residual symptoms between episodes predict a worse prognosis compared to full symptom remission between episodes (18 ,19 ,20 ,21 ,22 and 23). Second, residual symptoms are associated with increased levels of morbidity (5), as well as mortality—especially when these symptoms occur in the context of a general medical condition, such as with myocardial infarction (24 ,25), stroke (26), dementia (27), diabetes (28 ,29), or asthma (30). (See ref. 31 for review.)

Third, there is a concern (yet to fully evaluated) that more chronic forms of depression may be more likely associated with the development of treatment resistance over time. That is, a chronic course may entail the development of an underlying neurobiology that renders treatments less effective acutely or over the longer term. Such an inference is suggested by apparently longer times to develop responses and remissions in studies of chronic (10 ,11) as opposed to nonchronic forms of depression (32 ,33). This recent emphasis on the presence of modest levels of depressive symptoms,

on residual symptoms, and on a chronic course of illness has led to rethinking treatment studies. Full symptom remission and full functional restoration (not simply response) are the targets of treatment.

This conceptual shift has profound implications for practice and development and use of newer agents. It also provides a rationale for combining treatments when a monotherapy does not lead to remission. For example, subsyndromal forms of depression (i.e., those not meeting criteria for dysthymic or major depressive disorder) have recently become a focus of efficacy trials given the disabling nature of these conditions. Whether different levels of functional impairment in the context of equivalent levels of residual symptoms herald a worse prognosis has yet to be demonstrated. Furthermore, recent studies (21 ,22 and 23) suggest psychotherapy that effectively removes residual symptoms also improves prognosis, which in turn may reduce the need for long-term maintenance medication.

Treatment is divided into acute, continuation, and maintenance phases (34 ,35 ,36 and 37). Acute treatment aims at symptom remission and full functional restoration. The need to most aggressively pursue full symptom remission (also called complete response), rather than to accept a clinically significant reduction in symptoms (a response) is now accepted because of the worse prognosis and functional impairment associated with residual symptoms noted in the preceding. Earlier intervention can also be strongly recommended because major depressive episodes that have a greater duration are more likely to end with residual symptoms.

Continuation phase treatment aims at preventing a return of the most recent (index) episode—a relapse. Maintenance aims at preventing a new depressive episode—a recurrence. When continuation ends and maintenance begins for an individual is unclear, although classically recovery from the episode is estimated by when the episode would have spontaneously ended based on the duration of prior episodes, if such information is available. The need for more prolonged (i.e., continuation/maintenance phase) treatment, especially for more chronic (9 ,38 ,39 and 40) or more highly recurrent (41 ,42) forms of depression, has been recognized over the last decade (34 ,35 ,36 and 37).

Although the need for maintenance treatment for some patients with highly recurrent forms of depression is clear, exactly how long to provide antidepressant treatment remains a focus of research. Consensus strongly recommends prolonged (multiyear) maintenance treatment for those with a high likelihood of recurrence over the subsequent (1- to 3-year) period of time (34 ,35 ,36 and 37). These recommendations rest on several major, long-term, randomized, placebo-controlled, double-masked maintenance trials in adults (41 ,42) or in the elderly (43) with highly recurrent depressions, and in adults with chronic forms of major depressive disorder (9). Whether patients with two episodes of major depression *plus* a risk factor for recurrence should also be strongly encouraged toward maintenance therapy— although recommended based on clinical consensus—has not been empirically validated. In many such cases, patient preference with careful clinical monitoring, once therapy is discontinued, is recommended (34 ,35 ,36 and 37).

TRANSLATING KNOWLEDGE TO PRACTICE

Part of "75 - Current and Emerging Therapeutics for Depression "

Present practice currently relies on a trial and error approach that is only infrequently informed by well-established empirical evidence. Only a few clinical clues that recommend one treatment over another have been established scientifically. For example, the greater acute phase efficacy of monoamine oxidase inhibitors (MAOIs) as compared to tricyclic antidepressants (TCAs) in depressions with atypical symptom features is well established in double-blind, placebo-controlled trials (44 ,45). Furthermore, the greater efficacy of TCAs combined with an antipsychotic agent as compared to TCAs alone is well known for psychotic depressions (46 ,47 ,48 and 49).

On the other hand, many practical dilemmas are confronted in routine practice, yet knowledge is sparse to address these issues. This lack of practical knowledge grounded in empirical evidence, can be attributed to several factors: (a) insufficient investment in clinical research that goes beyond classic efficacy trials obtained for regulatory approval worldwide (50); (b) incomplete understanding of the neurobiological basis for clinical depressions such that specific treatment plans can be devised; and (c) substantial differences in the populations, procedures, and aims of efficacy studies conducted for regulatory purposes and representative practices.

Patients with minimally treatment-resistant, or nonchronic forms of depressions enter trials. These patients are: (a) rarely severely ill; (b) rarely inpatients; (c) never psychotic; (d) rarely encumbered by common concurrent Axis III (general medical) or other Axis I (psychiatric) disorders; (e) not affected by depressions that have been unresponsive to more than one prior medication in the current episode; and (f) without significant suicidal risk. In fact, many patients who enter efficacy randomized controlled trials (RCTs) for regulatory purposes are often symptomatic volunteers. Perhaps as a consequence, placebo response rates are often substantial (e.g., 25% to 35%) with drug effects providing a 50% to 60% response rate (34).

Notable differences exist between clinical procedures that are typically used to conduct RCTs designed regulatory approval and the procedures characteristic of current practice. Efficacy studies often: (a) use structured interviews, or highly trained staff using specific lists of diagnostic criteria for diagnostic purposes; (b) use itemized clinician-completed symptom ratings to assess treatment effects (therefore, also to adjust dosages); (c) use more frequent treatment visits; (d) limit acute phase trial durations (e.g., 6 to 8 weeks of acute phase treatment in efficacy RCTs); and (e) use

response (typically a ≥50% reduction in overall depressive symptom severity with or without residual symptoms), rather than remission (virtual absence of depressive symptoms with normalization of function) to define a clinical "success," contrary to clinical practice recommendations (34 ,35 ,36 and 37).

Routine clinical diagnoses often sharply disagree with those established by structured interviews (51) (Kashner et al., personal communication). In addition, global judgments of the severity of illness, even if codified by the Clinical Global Impression-Improvement Scale (CGI-I) (52), may relate only modestly to symptom severity ascertained by itemized clinician ratings, such as the Montgomery-Åsberg Depression Rating Scale (MADRS) (53) or the Hamilton Rating Scale for Depression (HRS-D) (54 ,55) or the Inventory of Depressive Symptomatology (IDS) (56 ,57). These differences in clinical procedures used in efficacy RCTs and those used in daily practice could well lead to radically different clinical outcomes.

Most important, however, is the fact that efficacy trials for regulatory approval are not designed to answer many key clinical questions addressed in routine practice (see the following). For example, there is a reasonably strong basis to believe that medications differ in their spectrum of actions (1). That is, some patients/depressions respond to one agent, whereas others respond to a different agent (58 ,59), especially if the medications differ regarding presumed mechanisms of action.

We do not know under what conditions one agent is preferred over another, however. Where one agent might "fit" into a multistep treatment plan is often left to tenuous inferences based on presumed mechanisms of action, to "expert" clinical opinion (34 ,35 ,36 and 37), or to marketing efforts. How to treat depressions that respond minimally or partially to one agent is now well known.

Perhaps pharmaceutical companies are reluctant to search for specific indicators of when an agent is preferred or to study agents used as second or third steps in treatment-resistant depression for fear of being "niched" by competitors. Without specific indications of when to use an agent, a "broad" spectrum of action can be claimed. Even comparative, randomized, double-masked trials to determine whether one agent is preferred over another for patients without sufficient clinical benefit to an initial agent are largely avoided by the pharmaceutical industry, perhaps in fear of finding that their agent will not fare as well as a competing agent in such treatment-resistant cases. Thus, economic forces within the industry provide a strong impetus to *not* inquire about highly salient—indeed clinically vital—information (e.g., when to select one agent over another; when to use a particular agent within a multistep therapeutic treatment program or sequence).

In spite of these knowledge gaps, the industry has developed a large number of newer antidepressants that are simpler to take, better tolerated, and safer in overdose. Antidepressant agents with new and different presumed mechanisms of action are also currently under development (see the following). Furthermore, recent regulatory and market forces have encouraged studies in depressed children, adolescents, and geriatric patients.

The National Institute of Mental Health (NIMH) has recently begun to address some of the knowledge gaps noted in the preceding by emphasizing effectiveness trials of antidepressant treatments in children/adolescents, adults, and geriatric patients. The emphasis (50) was based on the realization that efficacy trials for regulatory approval are only the first step in defining a treatment. That is, they establish the safety and efficacy of an agent in carefully conducted, highly internally valid designs. Generalizability of tolerability and efficacy findings requires different study designs. When to use one as opposed to another agent (alone or in combination) requires still different study designs.

Let us now turn to some of the vexing questions encountered in daily practice, and the knowledge available to address these issues.

CLINICAL ISSUES IN TREATING DEPRESSION

Part of "75 - Current and Emerging Therapeutics for Depression "

Clinicians routinely confront a host of practical questions that are not addressed in efficacy RCTs designed for regulatory approval. These questions include the following:

1. What is the best agent for initial treatment (i.e., predictors of response)?
2. What is the "next best" treatment following either an unsatisfactory clinical response or intolerance to the first agent?
3. What is a sufficient trial duration beyond which response is not likely (minimal duration)?
4. What is a sufficient trial duration for those benefiting from the treatment beyond which further treatment (unchanged) is unlikely to produce any more symptomatic or functional improvement (maximal duration)?
5. When is it best to augment the first agent with a second treatment? When is it best to switch from the first agent to a new, different agent?
6. What are the best ways to enhance adherence?
7. When and in what form should psychotherapy be used?
8. Do antidepressants differ in their ability to produce response or remission, and if so, for which depression is each better?
9. Do different medications differ in the time to onset of clinical benefit or time to remission?
10. What treatments are recommended if there is a return of symptoms in previously responsive patients?

Selecting the Initial Treatment

All antidepressant medications have established efficacy in major depressive disorder. Some even have placebo-controlled

evidence supporting efficacy in dysthymic disorder (sertraline) (15) or other “nonmajor” disorders, such as premenstrual dysphoric disorder, including fluoxetine (60), paroxetine (61), and sertraline (62 ,63). However, when to select one over another agent is not well defined. Clinicians use “rules of thumb” to make these judgments, but such reasonable guesses are rarely supported by prospective RCT evidence.

For example, more recently, efficacy for some antidepressants has been established for other psychiatric conditions commonly found in the presence of major depressive disorder, including: (a) venlafaxine for generalized anxiety disorder (64 ,65); (b) paroxetine (66) and sertraline (67) for posttraumatic stress disorder; (c) fluoxetine (68) and sertraline (69) for obsessive compulsive disorder; and (d) fluoxetine (70), paroxetine (71 ,72), and sertraline (73) for panic disorder. It is logical to argue that if a clinical depression is accompanied by a concurrent additional psychiatric disorder for which an antidepressant has established efficacy, then that agent is preferred (because it should be effective for both disorders) (34 ,35 ,36 and 37). However, this logical inference has not been evaluated prospectively in double-blind comparative trials.

Another clue used to select among antidepressants is cross-sectional symptom features. As noted, depressions with atypical symptom features do less well on TCAs than MAOIs (44 ,45) Depressions with psychotic features do better with a combination of an antipsychotic agent and a TCA than with a TCA alone (46 ,47 ,48 and 49).

Although atypical or psychotic symptom features are useful in selecting among the TCAs, other cross-sectional symptom features have generally not been so useful. For example, a common belief is that selective serotonin reuptake inhibitors (SSRIs) should be more effective than selective noradrenergic reuptake inhibitors (SNRIs) for major depression with marked anxiety. However, this contention has not withstood empirical study. Bupropion was as effective as sertraline in outpatients with major depressive disorder, whether pretreatment anxiety was high or low (74). Similarly, higher levels of pretreatment insomnia are *not* associated with lower efficacy for fluoxetine or with preferential response to imipramine as compared to fluoxetine (75). Reboxetine, an SNRI, is effective in both panic disorder (76) and depression (77 ,78).

Although family history of response to a MAOI or TCA should point the clinician to choose between these two classes (79), studies of family history and patient responses to newer agents are not available.

In sum, only psychotic or atypical symptom features have established value in selecting among treatments. Concurrent comorbid conditions logically recommend an initial agent, but this recommendation has not been evaluated prospectively. It would appear that other parameters such as safety in overdose, longer-term tolerability, the potential for drug–drug interactions, or likelihood of remission play a major role in selecting the first agent.

How to Select the “Next Best” Treatment following an Unsatisfactory Response (or Intolerance) to the First Agent

A major clinical problem is selecting the “next” agent if the first is ineffective, only partially effective, or not well tolerated. When TCAs fail, the MAOIs have roughly a 50% response rate based on both open and randomized trials (58 ,59).

When newer agents are used as first treatments, however, only open case series are available to define the next step (following intolerance or nonresponse). Based on open trials (80 ,81 ,82 and 83), a second SSRI is associated with a 40% to 60% response rate following failure with the first SSRI, although not all studies agree (84 ,85). Open trials (following initial SSRI failure) also support switching “out of class” to venlafaxine (86 ,87), bupropion (88 ,89), nefazodone (90 ,91), mirtazapine (92 ,93), or reboxetine (94). However, no randomized, comparative studies of a second SSRI (as compared to a non–SSRI) following nonresponse or intolerance to the first SSRI are available. Thus, both within and out of class switches following initial SSRI failure can be recommended, but the strength of the evidence is weak (95).

A recent double-masked trial, using a crossover design, in outpatients with nonpsychotic, chronic forms of major depressive disorder (96) revealed that about 50% of those who did not respond to (but did tolerate) 12 weeks of sertraline in the acute phase trial did respond to imipramine. Interestingly, similar response rates were found with sertraline for those who tolerated but who did not respond to imipramine. This large, double-blind, definitive study, provides substantial evidence for an “out-of-class” switch as a second step following unsatisfactory response to an SSRI or TCA. It also reveals that an SSRI (in this case sertraline) is effective even if a TCA (imipramine) is not—a finding that does not agree with the suggestion of greater efficacy of TCAs versus SSRIs. However, whether a within SSRI class switch (e.g., sertraline to paroxetine) would have been as effective as the out-of-class switch was not studied.

When to Augment or Switch

Inadequate benefit to an initial treatment comes in degrees that range from literally no benefit whatsoever, to a clinically significant response but without full remission (i.e., with residual symptoms). In such cases, clinicians and patients must choose between switching (i.e., discontinuing the first and starting a second treatment) and augmenting (adding a second treatment to the first). This decision, in part, rests on patient and clinician preference, desirability of simple (i.e., monotherapy) versus a more complex (i.e., two or more treatments) regimen, prior history of response/nonresponse

to other agents, and the desirability of not losing a partial, albeit modest, benefit with the first agent. With a history of no or only one prior unsuccessful treatment attempt, monotherapy (i.e., a switch) may have clinical appeal (simplicity). For more resistant depressions, even a modest benefit to the first treatment may recommend augmentation.

The best-studied augmentation methods are with lithium or thyroid used in combination with a TCA or MAOI (58 ,97). More recently, open trials or small case series suggest a benefit of adding bupropion to an SSRI (98), venlafaxine (99), mirtazapine, or nefazodone. (See ref. 97 for an extensive review of both the augmenting and switching literature.) Notable, however, is the lack of randomized comparator trials pitting one augmentation treatment against another, each used with the newer antidepressant agents. Furthermore, whether augmentation is as effective for patients who have a minimal response, as opposed to at least a partial response, or a response with residual symptoms with newer agents is not known.

What Is an Adequate Trial Duration to Reliably Declare "Failure?"

Clinicians confront two critical decision points during a treatment trial with an antidepressant. First, one wishes to stop the trial at the earliest point in time after which the patient has minimal or no chance of responding (i.e., at this point, a change in the treatment strategy—either a switch or augment is called for). Second, if some benefit has occurred, but remission has not yet been attained, then one needs to know how much more time should pass (and whether dose increases are needed) before deciding to augment or switch the treatment. That is, after what point in time are those who benefit in part unlikely to benefit any further? These two critical decision points occur at different times.

Let us consider the first critical decision point. Beyond what point in time is a clinically meaningful response *unlikely* to occur? A few post hoc analyses reveal that (a) there are both faster and slower responders in samples treated with TCAs (32), nefazodone (33), bupropion (100), MAOIs (101), fluoxetine (102 and 103), and the combination of interpersonal psychotherapy (IPT) (104) with imipramine (105) or nortriptyline (106), and likely all other antidepressants. These reports suggest that about one-fourth to one-third of depressions that do not respond by 4 weeks will do so by week 8. For example, Nierenberg and associates (103) found only 18.9% of patients treated with fluoxetine who did *not* have a less than 20% decrease in pretreatment HAM-D total score by week 4, ultimately responded (≥50% decrease in baseline HAM-D) by week 8. It seems that in some *post hoc* analyses, those with later responses are more likely to be more severely depressed at baseline, to have more Axis II disorders and possibly other psychiatric comorbidities, or to have a more chronic prior course of illness (e.g., longer episodes or residual symptoms between episodes) (32 ,106). However, some reports indicate that various agents, such as mirtazapine (107 ,108 and 109) or venlafaxine (64 ,110) may have an earlier onset of action compared to more selective agents. Whether they differ from other agents with a smaller proportion of patients evidencing later response is not yet clear.

What Is a Sufficient Trial Duration Beyond Which Further Improvement Is Unlikely?

When to decide that longer (unchanged) treatment will produce no more benefit in patients already having some symptom reduction is less clear. Remission follows response after 0 to 6 weeks (33). Thus, although response is unlikely to begin after 8 weeks of medication treatment, remission may not occur until 12 weeks (or even longer) with treatment involving a single agent. Indeed, in a recent study of outpatients with chronic major depressive disorder, 40% of acute phase responders who had residual symptoms (i.e., responders but nonremitters) at exit from a 12-week acute trial of imipramine or sertraline attained a full remission over four ensuing months of continuation phase treatment. Thus, perhaps especially for more chronically depressed, a longer trial duration—even up to more than 3 months following attainment of response—may be needed to determine if full remission will occur, or if a change in treatment is indicated.

How to Enhance Adherence?

Adherence, both in acute and later phases of treatment, is a major clinical problem (111). Clearly, better-tolerated, lower side effect, easier to use agents should increase adherence. Indeed, the newer agents (SSRIs, venlafaxine, nefazodone, bupropion, and mirtazapine) are better tolerated in acute phase trials (112). Gradual dose adjustments, as well as the sustained or extended release formulations (compared to immediate release versions) of newer agents (e.g., venlafaxine XR) (113 ,114 and 115)—enhance adherence by both creating better side-effect profiles and by reducing the number of times the medication must be taken.

A major assist in increasing adherence is patient education. Now evaluated in several randomized controlled trials, patient education clearly improves adherence, and consequently clinical outcomes as compared to minimal education (116). However, what types of education particularly benefit which patients remains to be determined.

When and How to Use Psychotherapy?

Research on psychotherapy for depressive disorders has, until recently, been focused nearly entirely on acute phase treatment studies that compare a symptom-reducing, time-limited

psychotherapy (e.g., cognitive, behavioral, interpersonal, or brief dynamic therapy) against a specific, depression-targeted medication monotherapy, the combination, or a control group. Evidence for the efficacy of acute phase psychotherapies against wait list controls is robust (2 ,35 ,117). In most trials, medication alone and psychotherapy alone have comparable efficacy (35). In a recent 10-week acute trial, Jarrett and colleagues (118) found CT to equal phenelzine and both to exceed pill placebo in outpatients with MDD and atypical symptom features.

These trials are limited, however, to outpatients with moderately severe depressive symptoms. Some argue that more severely depressed outpatients may fair better with medication as opposed to psychotherapy alone (119), whereas other data (120) suggest that depressive symptom severity is not particularly predictive of comparative treatment efficacy.

Turning to the combination of both medication and psychotherapy, six trials have not found the combination to show an advantage over either treatment alone (35). However, a very important recent 12-week acute phase trial of outpatients with chronic forms of major depressive disorder (11) was the first to find far greater acute phase efficacy for the combination of medication (nefazodone) and psychotherapy (Cognitive Behavioral Analysis System of Psychotherapy) (CBASP) (121) in chronic depression than for either nefazodone or CBASP alone. The acute response rates in the intent-to-treat sample (*n* = 662) were 48% (nefazodone), 48% (CBASP), and 73% (combination), and for patients who completed the study (*n* = 519), the response rates were 55%, 52%, and 85%, respectively. Importantly, remission rates in both the ITT and completer samples were higher with the COMB (48% and 42%, respectively) than with either nefazodone alone (29% and 22%, respectively) or CBASP alone (33% and 24%, respectively).

These findings indicate that psychotherapy increases the likelihood of responding *and* increases the magnitude of symptom reduction found with medication (nefazodone) alone and vice versa. The combination did not have a lower premature discontinuation rate than either monotherapy. In the context of prior literature on combination treatment, it appears that the combination is clearly indicated (either at the outset or in sequence) for more chronic depressions. Interestingly, approximately 50% of patients who did not respond to nefazodone acutely ultimately responded to CBASP, and vice versa in a crossover study following the acute trial (Keller, personal communication). Thus, psychotherapy may have substantial clinical utility even in those who do not respond acutely to medication.

Another role for psychotherapy may be the elimination of residual depressive symptoms for those depressions that respond, but do not remit with, medication alone. Two important recent controlled trials (21 ,22) examined the effect of adding cognitive therapy to antidepressant medication in patients with response, but with residual depressive symptoms. In both studies, cognitive therapy was compared (randomized) to treatment as usual without a formal psychotherapy. In essence, patients in the intervention group ultimately received both treatments, but in sequence with medication first. In one study (22), medications were gradually discontinued, whereas in the other (21), medications were continued while psychotherapy was provided. Both studies found a better prognosis for those who received psychotherapy. Thus, formal psychotherapy may increase the remission rates obtained with medication alone. It may also, perhaps as a consequence, improve longer-term prognosis (i.e., reduce relapse/recurrence rates) when combined with medication or during medication discontinuation.

How long to provide psychotherapy alone for those who respond to it in the acute phase has recently been evaluated (122). Continuation phase CT was associated with a lower relapse rate than no continuation phase CT for outpatients with MDD who at least responded to acute phase CT alone. This nonrandomized comparison has led to an ongoing prospective, randomized trial.

Finally, the role of IPT as a maintenance treatment alone or in combination with nortriptyline was evaluated in the elderly with major depressive disorder (43). Medication exceeded the effects of pill placebo and medication clinic visits in preventing recurrences—similar to a maintenance phase trial in adults (41). IPT had a better effect than pill placebo. The combination of IPT and nortriptyline was no better than nortriptyline alone. Psychotherapy may help to sustain medication-free periods during maintenance phase treatment (e.g., in women wishing to become pregnant).

Do Antidepressants Differ in Acute Phase Efficacy?

Recent acute phase trials have begun to examine whether medications, especially those with direct effects on multiple neurotransmitter systems, might have greater efficacy than more selective reuptake blocking agents. Note, however, that two agents may appear to have different degrees of efficacy if they are compared early in the course of acute phase treatment, whereas later (e.g., after 6 to 12 weeks of treatment), they could display equivalent efficacy (i.e., if one agent “acts more rapidly” than another). For example, the Danish University Antidepressant Group (DUAG) studies, one each lasting 5 and 6 weeks with severely depressed inpatients, revealed better outcomes for those severely depressed inpatients with nonpsychotic major depressive disorder with clomipramine than with paroxetine (123) or citalopram (124). These brief inpatient studies may have been too short in duration, however, to gauge the full benefits obtainable with longer treatments with either agent, however.

More recently, additional studies in both inpatients and outpatients have compared venlafaxine with fluoxetine (125 ,126 ,127 ,128 and 129), venlafaxine with paroxetine (130 ,131), venlafaxine

with sertraline (132), and mirtazapine with fluoxetine (109) in studies lasting 4 to 24 weeks. Several have revealed either higher response or remission rates with the dual action agents as compared to the SSRI comparator (125 ,126 ,130 ,131 and 132). Although not all studies confirm these findings (107), it would appear that for severely ill patients, dual action agents may offer somewhat better efficacy in selected patient populations. In addition, a response without remission to an SSRI might arguably lead to either the use of an augmenting medication to create a "dual action," or a switch to a dual-action agent.

Do Medications Differ in the Time to Onset of Benefit?

Some studies report that some agents have a faster onset of response (107 ,108 and 109). These findings depend on the doses used, the population under study (e.g., less versus more severe; more versus less treatment-resistant), and the length of the trials. Obviously, if one agent is dosed/titrated more gradually than another the former could *appear* to have a slower onset of action, when such might not be the case with more aggressive dosing. Several recent, double-masked, randomized trials, especially in more severely depressed patients suggest faster onset of action for mirtazapine (107 ,108 and 109) or venlafaxine (64 ,110) than comparator SSRIs. Whether more aggressive dosing of the SSRI might have produced different results remains an open question.

How to Manage Symptomatic Breakthrough?

During continuation or maintenance phase treatment, a return of clinically significant depressive symptoms, even while on medication, is not uncommon. Full relapses/recurrences range from 10% to 40% over 12 to 16 months following response to acute phase treatment. This symptom breakthrough appears to occur with all antidepressants (133). Whether it is more likely with one or another medication or medication class has not been well defined. Those who attain remission (not just response) to acute phase treatment appear more likely to remain in remission (or to at least sustain a response) over continuation phase treatment than are those who with a response but with residual symptoms at exit from acute phase treatment (40). Some studies suggest that patients with an earlier, more complete, and more sustained symptom benefit in acute treatment are less likely to encounter symptomatic breakthrough at least in the continuation phase (134 ,135 and 136).

How to manage symptom breakthrough in continuation/maintenance phases is unclear. Although clinical consensus suggests dose increases (137), others suggest dose decreases (138). Still others add a second agent (e.g., bupropion) to the first (e.g., an SSRI), whereas others recommend discontinuing the first agent and switching out of class to another agent. (See ref. 97 for review.) The question of how to manage symptomatic breakthrough is of substantial public health significance because it is commonly encountered in clinical practice. In addition, these patients likely will have increased use of the health care system, function more poorly, and will have a worse prognosis. Yet, no randomized trial data and very few case reports are available. Whether agents differ in the likelihood of symptom breakthrough has also not been studied in comparative trials.

How to Define and Manage Treatment Resistance?

The degree of treatment resistance may be based on the number of treatments (or classes of treatments) that have not led to a response or a sustained response, or to a remission or sustained remission (58). Clinically, it is often difficult to accurately define what treatments, at what doses, and used for how long produced what benefits for a particular patient, especially for patients with a history of multiple treatments, multiple providers, or longstanding depressions. The Antidepressant Treatment History Form (ATHF) provides a validated tool by which to gauge the degree of treatment resistance for research (139).

Once the level of treatment resistance is defined, however, we are still left with a large number of possible treatments, few of which have been subjected to randomized, comparative trials for treatment-resistant depression. Most studies are open trials (see the preceding). The recently launched, NIMH supported multisite, national effectiveness trial, *Sequenced Treatment Alternatives to Relieve Depression (STAR*D)*, will begin to define, using randomized comparisons, which among several treatments used as augmenting or switching strategies have greater efficacy in a large cohort of patients recruited from primary and specialty care clinical sites (for a detailed description, see http://www.edc.gsph.pitt.edu/stard).

POTENTIAL NEW ANTIDEPRESSIVE MEDICATIONS

Part of "75 - Current and Emerging Therapeutics for Depression "

Two new classes of possible antidepressant medications are under development: substance P antagonists and CRF antagonists. A single, positive, 6-week double-blind trial in outpatients with MDD study found a substance P antagonist (MK 869) as effective as paroxetine, and both exceeded the effects of pill placebo (140). The mechanism of antidepressant action of substance P antagonists is not clear; however, it seems that neither norepinephrine nor serotonin systems are directly affected. Direct effects on the substance P NK receptors, perhaps in the stratum or amygdala, to modify stress response may play a role.

Corticotropin releasing hormone (CRH) also plays a key role in modulating the neuroendocrine, autonomic, and behavioral

responses to stress (141). CRH produces signs and symptoms of depressive and anxiety disorders by activation of the CRH_1 receptors (142 ,143). These findings provide a rationale for attempts to develop medications that antagonize the CRH_1 receptor. Zobel and colleagues (144), using R121919 (an agent that binds to CRH_1 receptors) in an open trial of 20 patients, reported improvements in anxiety and depressive symptoms. CRH_1 receptor blockade did not impair corticotropin and cortisol secreting activity either at baseline or after an exogenous CRH challenge.

These early reports are tantalizing. The field awaits more definitive, placebo-controlled clinical trials for both CRH and substance P antagonists. Whether these agents will have predictable, substantial, and prolonged antidepressant or anxiolytic (e.g., posttraumatic stress disorder) is yet to be determined. In theory, if they modify stress responses, such agents may be important in the treatment of residual symptoms or symptomatic breakthroughs that occur with currently available agents. Alternatively, they may prevent the onset of a depressive episode following a stress in vulnerable individuals.

Combinations of standard medications, especially the use of atypical antipsychotic agents, alone or combined with antidepressants, have begun to be a focus of research for treatment-resistant depression (145 ,146). The rationale for the use of olanzapine combined with fluoxetine, for example, is provided by evidence of increases in norepinephrine and dopamine levels using microdialysis in the raphe prefrontal cortex. In a recently completed, 8-week, double-blind trial in 28 patients with treatment-resistant depression, large reductions in MADRS scores were obtained with the combination, as compared to either agent alone (146). Whether such findings are replicated, and/or generalize to other atypical antipsychotic agents, other SSRIs, or to other newer agents remains to be seen.

ANTIDEPRESSIVE TREATMENTS IN CHILDREN AND ADOLESCENTS

Part of "75 - Current and Emerging Therapeutics for Depression "

The psychotherapeutic approaches of cognitive behavioral therapy (CBT) and IPT appear to have specific efficacy in adolescent depression (147 ,148 ,149 ,150 ,151 and 152), whereas psychotherapeutic approaches to preschool depression are relatively less developed and have not been tested in randomized trials. (See ref. 153 for review.) The very limited available evidence does not show long-term change from these acute psychotherapeutic interventions (2).

Quite surprisingly, SSRIs appear to have adult-like efficacy in children and adolescents with major depressive disorder, whereas TCAs do not and may be no better than placebo. As reviewed elsewhere (154), RCTs of TCAs versus placebo have been uniformly negative in both children and adolescents. Many of these negative studies were relatively small. However, when considered in aggregate, the data best support the hypothesis that TCAs are either ineffective or much less effective in this age group than in adulthood. Furthermore, TCAs are particularly toxic in deliberate or accidental overdose in youth, and they are all off patent; therefore, more studies of TCAs in this population are unlikely.

To date, two large studies have found SSRI superiority to placebo (155 ,156) and they appear equally efficacious in prepubertal children as in adolescents and in both sexes (155). Encouraged by regulatory changes, randomized, placebo-controlled trials in youth are planned or underway for essentially all antidepressants now on patent.

One of us has argued, post hoc, that because the TCAs as a group are all relatively noradrenergic in youth because of their relatively more rapid metabolism, it may be that noradrenergic antidepressants as a class will prove less useful in youth, whereas serotonergic antidepressants will show efficacy throughout the lifespan. There are as yet no available RCTs on newer agents other than SSRIs in youth to further address this question (157).

Although available and ongoing work in antidepressant pharmacology in youth with MDD provides guidance for the initial treatment step for depression, the field is only now considering the arguably more important studies of chronic maintenance, combination treatment, treatment of refractory depression, and the other questions that arise in this recurrent disorder.

“ALTERNATIVE” THERAPIES

Part of "75 - Current and Emerging Therapeutics for Depression "

St. John's Wort (Hypericum perforatum)

Extracts of the plant St. John's wort, *hypericum perforatum*, appear to have antidepressant effects. These extracts contain at least 10 biologically active substances. One study has suggested that hyperforin may be the critical active constituent (158). Extracts of hypericum inhibit reuptake of serotonin, norepinephrine, and dopamine and lead to down-regulation of B-adrenoceptors and serotonin ($5\text{-}HT_2$) receptors (159). However, limiting the studies to date is the lack of a standardized concentration of the active ingredient(s) in the preparation of hypericum extract.

Several recent metaanalyses have aggregated available data (160 ,161 and 162) but did not include one additional recent study (163) to find that hypericum extract: (a) appears to be significantly better than placebo in the treatment of mild to moderate depression, and (b) appears, in general, not statistically different from relatively low doses of TCAs (e.g., amitriptyline 75 mg QD, imipramine 100 mg QD, imipramine 150 mg QD) given for a brief duration (6 weeks or less in all but the Philipp study that used 100 mg of imipramine for 8 weeks). A recent comparison of St. John's wort and the SSRI fluoxetine also found no statistically significant difference between treatments (mean decrease in

HAM-D score = 10.2 for St. John's wort versus 12.5 for fluoxetine) (164).

In a recent double-masked randomized study (165) of 201 outpatients with major depression and a baseline 17-item HAM-D score ≥20 treated for 8 weeks with either St. John's wort or placebo for 8 weeks, St. John's wort was significantly better than placebo when examining the rate of remission (defined as HAM-D score ≤7) (19.3% versus 8.7%) ($p < 0.04$). However, there was no significant difference seen between St. John's wort and placebo on any of the continuous outcome measures, including the HAM-D, CGI-I, CGI-S, or the Hamilton Rating Scale for Anxiety (HRS-A) (166). The very low rate of response both to active medication and placebo in this study are atypical when compared to most modern studies of antidepressants.

Thus, although the randomized clinical studies using placebo suggest that hypericum has some antidepressant effect in humans (and has effects in animal models for depression), the data are simply inadequate to say how well St. John's wort works compared to current standard dosages of antidepressants or data as to efficacy of this compound for maintenance. The NIMH and National Institutes of Health Office on Alternative Medicine is currently funding a study of *hypericum perforatum* versus SSRI or placebo, which may further elucidate the efficacy of this compound.

S-adenosyl-L-methionine

S-adenosyl-L-methionine (SAMe) has been tested as a potential antidepressant over the past 25 years. It is the primary methyl donor in the CNS for methyl acceptor molecules including catecholamines and phospholipids. A 1994 metaanalysis (167) examined the literature through 1992, and found six adequately controlled studies comparing SAMe (oral, intravenous, or intramuscular) to placebo with a total of 99 subjects on SAMe and 101 on placebo and seven studies comparing SAMe to TCAs with a total of 105 patients on SAMe and 96 on TCAs. Although those authors found the aggregate data to show statistically significant superiority of SAMe to placebo and equivalence of SAMe to TCAs, all but two of the comparisons with placebo and two of the comparisons with TCAs were 21 days or less in duration. Equivalence to TCA for such a short interval (before most of the TCAs effects have been realized) is unconvincing. Of the two placebo comparisons of longer duration, one trial of 30 days was positive and one of 42 days was negative.

Of studies not considered in the metaanalysis, one small study suggested more rapid onset of antidepressant response (by day 4) for SAMe + imipramine compared to placebo plus imipramine (168). A large open investigation (169) suggested, as did earlier studies, efficacy and very rapid onset of antidepressant effect of parenteral SAMe in humans. One study found imipramine-like “antidepressant” effects of SAMe compared to vehicle in a rat model of stress-induced anhedonia (170).

In aggregate, these data seem to point toward efficacy and perhaps very rapid onset of SAMe in the treatment of adult major depressive disorder. On the other hand, the total number of patients studied to date in placebo-controlled trials is quite modest, and the relative efficacy of SAMe compared to other treatments is quite unclear because almost all comparison studies were conduct for far too short an interval to reach full TCA effect. The roles of either SAMe or St. John's wort to treat residual symptoms (i.e., as augmenting agents to standard antidepressants as well) may be a useful focus of study.

Acupuncture

As reviewed in Ernst and colleagues (171), several case series and open clinical trials suggest possible efficacy of acupuncture for depression with electroacupuncture appearing to have greater effect than standard acupuncture and two randomized controlled trials have compared electro-acupuncture to amitriptyline. One study (172) compared 5 weeks of amitriptyline (average daily dose 142 mg) to electro-acupuncture in a total of 47 subjects and found no significant difference in HAM-D endpoints. A larger replication found no significant difference in outcome between amitriptyline and electro-acupuncture in a 6-week RCT in a total of 241 depressed inpatient subjects (173). As is well understood, lack of statistically significant difference of a putative treatment from a “known effective” treatment is not strong evidence for the efficacy of a treatment. In addition, the duration of treatment was relatively short, which would underestimate the maximal amitriptyline effect in these studies.

More recently, Röschke and associates (174) compared 70 adult inpatients with major depressive disorder, all treated with mianserin and then randomly assigned to the addition of acupuncture (with needling points proposed to be specific to the treatment of depression), placebo acupuncture (acupuncture at nonspecific locations) or no acupuncture. Both the specific and control acupuncture treatments showed statistical superiority to the no acupuncture group on several of the measures, although the differences were not large. There was no difference on any measure between the specific and control acupuncture treatments in this study. Thus, the results of this study are compatible with a nonspecific effect of the additional attention and expectancy and do not necessarily point toward specific efficacy of acupuncture.

In summary, data to date do not yet give a strong answer as to whether or not acupuncture has meaningful specific efficacy in the treatment of major depression. However, results of a recently published small, but randomized, controlled, and double-masked study of women with major depression ($n = 38$) suggest that acupuncture can provide

significant symptom relief from depression, at rates comparable to those of psychotherapy and pharmacotherapy (175).

FUTURE RESEARCH

Part of "75 - Current and Emerging Therapeutics for Depression "

The aim of treatment is remission, not simply response, given the better prognosis and better function associated with remission. Clinical issues raised by this recognition include: (a) Do medications truly differ in their ability to create remission, or do they differ more in the time it takes to establish an equivalent prevalence of remission? (b) Are the chances of remission increased if dual-action agents (or combinations of more selective monotherapies) are used compared to more neurotransmitter-selective agents? (c) How much time and effort should be expended to attain remission? (d) With which medication sequences is the likelihood of remission highest? (e) Does adding psychotherapy to medication increase remission rates or improve prognosis?

The importance of functional recovery, in addition to symptom remission, is now recognized. Do medications differ in their effects on day-to-day function? Some preliminary reports suggest that SNRIs (e.g., reboxetine) as opposed to SSRIs may lead to better functional recovery (176 ,177 and 178). This contention deserves more thorough and careful scrutiny in subsequent medication comparative trials.

We have yet to fully develop evidence as to when and for which patients such treatments are especially beneficial. Although several multi-step sequences, disease management pathways, or medication algorithms have been developed, only a single trial in primary care has evaluated the clinical and economic effects of using these sequences compared to treatment as usual. More research to develop evidence to establish valid pathways *and* to test their impact compared to treatment as usual is needed.

Given the range of treatments, can we better select or match a treatment to a person or to types of depression? Does a particular treatment history (or family history of treatment) recommend one versus another next step in the sequence of treatments to be provided to a particular patient?

Additionally, earlier intervention with less complex treatments, perhaps in the prodromal stages of the disorder, deserves further evaluation—especially with the potential availability of CRF antagonists. If these agents modify the stress response, could they be given quickly, close in time to the stress, before a full depressive episode appears in those with an established vulnerability to such a stress response?

Finally, we still are hampered by having to rely on symptoms and signs to gauge the adequacy of our treatments. Most of medicine can rely in part on laboratory measures to also inform clinicians about modifying the treatment plan or managing the illness. Research to find clinically obtainable measures of the disease process would remarkably improve the quality of care by better informing providers and patients.

ACKNOWLEDGMENTS

Part of "75 - Current and Emerging Therapeutics for Depression "

Supported in part by the Betty Jo Hay Distinguished Chair in Mental Health, the Rosewood Corporation Chair in Biomedical Science, and the Sara M. and Charles E. Seay Center for Basic and Applied Research in Psychiatry.

The authors thank David Savage and FastWord, Inc. (Dallas, Texas) for secretarial support in the preparation of this manuscript and Kenneth Z. Altshuler, Stanton Sharp Distinguished Chair and Professor, Department of Psychiatry for administrative support.

Dr. Rush has received research support from: Abbott Laboratories, Bristol-Myers Squibb, Cyberonics, Inc., Eli Lilly & Company, Forrest Pharmaceuticals/Parke-Davis, Glaxo Wellcome, Inc., Janssen Pharmaceutical, Robert Wood Johnson Foundation, Meadows Foundation, National Institute of Mental Health, Novartis, Organon, Inc., Pfizer, Inc., Pharmacia & Upjohn, SmithKline Beecham, Stanley Foundation, Wyeth-Ayerst, and Zeneca. In addition, Dr. Rush has served as a consultant and/or on a speaker's bureau for: Abbott Laboratories, Bristol Myers Squibb, Cyberonics, Inc., Eli Lilly & Company, Forrest Pharmaceuticals/Parke-Davis, Glaxo Wellcome, Inc., Janssen Pharmaceutica, Merck & Company, Inc., Mitsubishi Pharmaceuticals, National Institute of Mental Health, Organon, Inc., Pfizer, Inc., Pharmacia & Upjohn, Stanley Foundation, Wyeth-Ayerst, and Yamanouchi, Inc.

DISCLOSURE STATEMENT

Part of "75 - Current and Emerging Therapeutics for Depression "

A. John Rush has received $5,000 per year or more in the form of honoraria and/or research grant support for the last 3 years from the following entities: National Institute of Mental Health, Stanley Foundation, Robert Wood Johnson Foundation, Cyberonics, Inc., Abbott Laboratories, Bristol Myers Squibb, Forest Laboratories, Pfizer, Inc., Eli Lilly and Company, Glaxo Wellcome, Merck, Janssen, and Wyeth-Ayerst.

Neal Ryan has received $5,000 per year or more in the form of honoraria and/or research grant support for the last 3 years from the following entities: National Institute of Mental Health, National Library of Medicine, Forest Laboratories, Pfizer, Pharmacia & Upjohn, Solvay, SmithKline Beecham, and Wyeth-Ayerst.

REFERENCES

1. Schatzberg AF, Nemeroff CB, eds. *American Psychiatric Press textbook of psychopharmacology*, second ed. Washington, DC: American Psychiatric Press, 1998.

2. Rush AJ, Thase ME. Indications and planning of psychotherapies. In: Maj M, Sartorius N, eds. *WPA series in evidence and practice in psychiatry, vol 1. Depressive disorders*. Chichester, UK: John Wiley & Sons, 1999:161–206.

3. Keller MB, Kovacs K. Affective disorders: course over the lifetime. In: Charney DS, Coyle J, Davis K, et al, eds. *Neuropsychopharmacology: the fifth generation of progress*. Baltimore: Lippincott Williams & Wilkins, this volume.

4. Hays RD, Wells KB, Sherbourne CD, et al. Functioning and well-being outcomes of patients with depression compared with chronic general medical illness. *Arch Gen Psychiatry* 1995;52:11–19.

5. Miller IW, Keitner GI, Schatzberg AF, et al. The treatment of chronic depression. Part 3: Psychosocial functioning before and after treatment with sertraline or imipramine. *J Clin Psychiatry* 1998;59:608–619.

6. Judd LL, Akiskal HS, Zeller PJ, et al. Psychosocial disability during the long-term course of unipolar major depressive disorder. *Arch Gen Psychiatry* 2000;57:375–380.

7. Keller MB. Depression: a long-term illness. *Br J Psychiatry* 1994;165(Suppl 26):9–15.

8. Booth BM, Blow FC, Cook CA. Functional impairment and co-occurring psychiatric disorders in medically hospitalized men. *Arch Int Med* 1998;158:1551–1559.

9. Keller MB, Gelenberg AJ, Hirschfeld RM, et al. The treatment of chronic depression. Part 2: A double-blind, randomized trial of sertraline and imipramine. *J Clin Psychiatry* 1998;59:598–607.

10. Keller MB, Kocsis JH, Thase ME, et al. Maintenance phase efficacy of sertraline for chronic depression: a randomized controlled trial. *J Am Med Assoc* 1998;280:1665–1672.

11. Keller MB, McCullough JP, Klein DN, et al. A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *N Engl J Med* 2000;342:1462–1470.

12. Ravindran AV, Bialik RJ, Lapierre YD. Primary early onset dysthymia, biochemical correlates of the therapeutic response to fluoxetine. I. Platelet monoamine oxidase and the dexamethasone suppression test. *J Affect Disord* 1994;31:111–117.

13. Ravindran AV, Bialik RJ, Brown GM, et al. Primary early onset dysthymia, biochemical correlates of the therapeutic response to fluoxetine. II. Urinary metabolites of serotonin, norepinephrine, epinephrine and melatonin. *J Affect Disord* 1994;31:119–123.

14. Kocsis JH, Friedman RA, Markowitz JC, et al. Stability of remission during tricyclic antidepressant continuation therapy for dysthymia. *Psychopharmacol Bull* 1995;31:213–216.

15. Thase ME, Fava M, Halbreich U, et al. A placebo-controlled, randomized clinical trial comparing sertraline and imipramine for the treatment of dysthymia. *Arch Gen Psychiatry* 1996;53:777–784.

16. Vanelle JM, Attar-Levy D, Poirier MF, et al. Controlled efficacy study of fluoxetine in dysthymia. *Br J Psychiatry* 1997;170:345–350.

17. Versiani M. Pharmacotherapy of dysthymic and chronic depressive disorders: overview with focus on moclobemide. *J Affect Disord* 1998;51:323–332.

18. Judd LL, Akiskal HS, Maser JD, et al. A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorders. *Arch Gen Psychiatry* 1998;55:694–700.

19. Judd LL, Akiskal HS, Maser JD, et al. Major depressive disorder: a prospective study of residual subthreshold depressive symptoms as predictor of rapid relapse. *J Affect Disord* 1998;50:97–108.

20. Paykel ES. Remission and residual symptomatology in major depression. *Psychopathology* 1998;31:5–14.

21. Paykel ES, Scott J, Teasdale JD, et al. Prevention of relapse in residual depression by cognitive therapy. A controlled trial. *Arch Gen Psychiatry* 1999;56:829–835.

22. Fava GA, Rafanelli C, Grandi S, et al. Prevention of recurrent depression with cognitive therapy: preliminary findings. *Arch Gen Psychiatry* 1998;55:816–820.

23. Fava GA, Rafanelli C, Grandi S, et al. Six-year outcome for cognitive behavioral treatment of residual symptoms in major depression. *Am J Psychiatry* 1998;155:1443–1445.

24. Frasure-Smith N, Lesperance F, Talajic M. Depression following myocardial infarction. Impact on 6-month survival. *J Am Med Assoc* 1993;270:1819–1825.

25. Frasure-Smith N, Lesperance F, Talajic M. Depression and 18-month prognosis after myocardial infarction. *Circulation* 1995;91:999–1005.

26. Robinson RG, Morris PL, Fedoroff JP. Depression and cerebrovascular disease. *J Clin Psychiatry* 1990;51(Suppl):26–31.

27. Rovner BW, German PS, Brant LJ, et al. Depression and mortality in nursing homes. *J Am Med Assoc* 1991;265:993–996.

28. Lustman PJ, Griffith LS, Clouse RE, et al. Psychiatric illness in diabetes mellitus. Relationship to symptoms and glucose control. *J Nerv Ment Dis* 1986;174:736–742.

29. Popkin MK, Callies AL, Lentz RD, et al. Prevalence of major depression, simple phobia, and other psychiatric disorders in patients with long-standing type I diabetes mellitus. *Arch Gen Psychiatry* 1988;45:64–68.

30. Mancuso CA, Peterson MG, Charlson ME. Effects of depressive symptoms on health-related quality of life in asthma patients. *J Gen Int Med* 2000;15:301–310.

31. Krishnan KRR, Robinson R. Affective disorders and comorbid medical illness. In: Charney DS, Coyle J, Davis K, Nemeroff CB, eds. *Neuropsychopharmacology: the fifth generation of progress*. Baltimore: Lippincott Williams & Wilkins, this volume.

32. Rush AJ, Gullion CM, Roffwarg HP, et al. When do patients respond to tricyclic antidepressants? [Abstract] *Biol Psychiatry* 1994;35:711.

33. Trivedi MH, Rush AJ, Pan J-Y, et al. Which depressed patients respond to nefazodone and when? *J Clin Psychiatry* 2001;62:158–163.

34. Depression Guideline Panel. Clinical practice guideline, number 5. Depression in primary care, vol 1. *Detection and Diagnosis*. Rockville, MD: US Dept. of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research. AHCPR Publication No. 93-0550, 1993.

35. Depression Guideline Panel. Clinical practice guideline, number 5. Depression in primary care, volume 2. *Treatment of major depression*. Rockville, MD: US Dept. of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research. AHCPR Publication No. 93-0551, 1993.

36. American Psychiatric Association. Practice guideline for major depressive disorder in adults. *Am J Psychiatry* 1993;150(Suppl 4):1–26.

37. American Psychiatric Association. Practice guideline for the treatment of patients with major depressive disorder (revision). *Am J Psychiatry* 2000;157:4(Suppl):1–45.

38. Kocsis JH, Friedman RA, Markowitz JC, et al. Maintenance therapy for chronic depression: A controlled trial of desipramine. *Arch Gen Psychiatry* 1996;53:769–774.

39. Van Moffaert M, Bartholome F, Cosyns P, et al. A controlled comparison of sertraline and fluoxetine in acute and continuation treatment of major depression. *Hum Psychopharmacol* 1995;10:393–405.

40. Koran LM, Gelenberg AJ, Kornstein SG, et al. Sertraline versus imipramine to prevent relapse in chronic depression. *J Affect Disord* 2001;65:27–36.

41. Frank E, Kupfer DJ, Perel JM, et al. Three-year outcomes for maintenance therapies in recurrent depression. *Arch Gen Psychiatry* 1990;47:1093–1099.

42. Kupfer DJ, Frank E, Perel JM, et al. Five-year outcome for maintenance therapies in recurrent depression. *Arch Gen Psychiatry* 1992;49:769–773.

43. Reynolds CF III, Frank E, Perel JM, et al. Nortiptyline and interpersonal psychotherapy as maintenance therapies of recurrent major depression: a randomized, controlled trial in patients older than 59. *J Am Med Assoc* 1999;281:39–45.

44. Quitkin FM, Stewart JW, McGrath PJ, et al. Columbia atypical depression. A subgroup of depressives with better response to MAOI than to tricyclic antidepressants or placebo. *Br J Psychiatry* 1993;163(Suppl 21):30–34.

45. Liebowitz MR, Quitkin FM, Stewart JW, et al. Antidepressant specificity in atypical depression. *Arch Gen Psychiatry* 1988;45:129–137.

46. Glassman AH, Kantor SJ, Shostak M. Depression, delusions, and drug response. *Am J Psychiatry* 1975;132:716–719.

47. Spiker DG, Weiss JC, Dealy RS, et al. The pharmacological treatment of delusional depression. *Am J Psychiatry* 1985;142:430–436.

48. Rothschild AJ, Samson JA, Bessette MP, et al. Efficacy of the combination of fluoxetine and perphenazine in the treatment of psychotic depression. *J Clin Psychiatry* 1993;54:338–342.

49. Anton RF Jr, Burch EA Jr. Amoxapine versus amitriptyline combined with perphenazine in the treatment of psychotic depression. *Am J Psychiatry* 1990;147:1203–1208.

50. NIMH National Advisory Mental Health Council. *Bridging science and service. A report of the National Advisory Mental Health Council's Clinical Treatment and Services Research Workgroup*. Rockville, MD: NIH Publication, No. 99-4353, 1999.

51. Basco MR, Bostic JQ, Davies D, et al. Methods to improve diagnostic accuracy in a community mental health setting. *Am J Psychiatry* 2001;157:1599–1605.

52. Guy W. *ECDEU assessment manual for psychopharmacology*. Washington, DC: Superintendent of Documents, US Government Printing Office, US Department of Health, Education and Welfare Publication No. 76-338, 1976.

53. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:382–389.

54. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56–62.

55. Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* 1967;6:278–296.

56. Rush AJ, Giles DE, Schlesser MA, et al. The inventory for depressive symptomatology (IDS): preliminary findings. *Psychiatry Res* 1986;18:65–87.

57. Rush AJ, Gullion CM, Basco MR, et al. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med* 1996;26:477–486.

58. Thase ME, Rush AJ. Treatment-resistant depression. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: fourth generation of progress*. New York: Raven, 1995:1081–1097.

59. Thase ME, Rush AJ. When at first you don't succeed: sequential strategies for antidepressant nonresponders. *J Clin Psychiatry* 1997;58(Suppl 13):23–29.

60. Pearlstein TB, Stone AB, Lund SA, et al. Comparison of fluoxetine, bupropion, and placebo in the treatment of premenstrual dysphoric disorder. *J Clin Psychopharmacol* 1997;17:261–266.

61. Yonkers KA, Gullion C, Williams A, et al. Paroxetine as a treatment for premenstrual dysphoric disorder. *J Clin Psychopharmacol* 1996;16:3–8.

62. Yonkers KA, Halbreich U, Freeman E, et al. Symptomatic improvement of premenstrual dysphoric disorder with sertraline treatment. A randomized controlled trial. *J Am Med Assoc* 1997;278:983–988.

63. Freeman EW, Rickels K, Sondheimer SJ, et al. Differential response to antidepressants in women with premenstrual syndrome/premenstrual dysphoric disorder: a randomized controlled trial. *Arch Gen Psychiatry* 1999;56:932–939.

64. Gelenberg AJ, Lydiard RB, Rudolph RL, et al. Efficacy of venlafaxine extended-release capsules in nondepressed outpatients with generalized anxiety disorder: a 6-month randomized controlled trial. *J Am Med Assoc* 2000;283:3082–3088.

65. Rickels K, Pollack MH, Sheehan DV, et al. Efficacy of extended-release venlafaxine in nondepressed outpatients with generalized anxiety disorder. *Am J Psychiatry* 2000;157:968–974.

66. Marshall RD, Schneier FR, Fallon BA, et al. An open trial of paroxetine in patients with noncombat-related, chronic posttraumatic stress disorder. *J Clin Psychopharmacol* 1998;18:10–18.

67. Brady K, Pearlstein T, Asnis GM, et al. Efficacy and safety of sertraline treatment of posttraumatic stress disorder: a randomized controlled trial. *J Am Med Assoc* 2000;283:1837–1844.

68. Tollefson GD, Rampey AH Jr, Potvin JH, et al. A multicenter investigation of fixed-dose fluoxetine in the treatment of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1994;51:559–567.

69. Chouinard G. Sertraline in the treatment of obsessive compulsive disorder: two double-blind, placebo-controlled studies. *Int Clin Psychopharmacol* 1992;7(Suppl 2):37–41.

70. Schneier FR, Liebowitz MR, Davies SO, et al. Fluoxetine in panic disorder. *J Clin Psychopharmacol* 1990;10:119–121.

71. Lecrubier Y, Bakker A, Dunbar G, et al. A comparison of paroxetine, clomipramine and placebo in the treatment of panic disorder. *Acta Psychiatr Scand* 1997;95:145–152.

72. Bakker A, van Dyck R, Spinhoven P, et al. Paroxetine, clomipramine, and cognitive therapy in the treatment of panic disorder. *J Clin Psychiatry* 1999;60:831–838.

73. Pollack MH, Marzol PC. Panic: course, complications and treatment of panic disorder. *J Psychopharmacol (Oxford)* 2000;14(Suppl 1):S25–S30.

74. Rush AJ, Batey SR, Donahue RMJ, et al. Does pretreatment anxiety predict response to either buproprion SR or sertraline? *J Affect Disord* 2001;64:81–87.

75. Simon GE, Heiligenstein JH, Grothaus L, et al. Should anxiety and insomnia influence antidepressant selection: a randomized comparison of fluoxetine and imipramine. *J Clin Psychiatry* 1998;59:49–55.

76. Versiani M. Reboxetine is effective and well tolerated in the treatment of panic disorder (abstract). *Int J Neuropsychopharmacol* 2000;3(Suppl 1):S180.

77. Montgomery SA. Reboxetine: additional benefits to the depressed patient. *J Psychopharmacol (Oxford)* 1997;11(Suppl 4):S9–S15.

78. Burrows GD, Maguire KP, Norman TR. Antidepressant efficacy and tolerability of the selective norepinephrine reuptake inhibitor reboxetine: a review. *J Clin Psychiatry* 1998;59(Suppl 14):4–7.

79. Stern SL, Rush AJ, Mendels J. Toward a rational pharmacotherapy of depression. *Am J Psychiatry* 1980;137:545–552.

80. Brown WA, Harrison W. Are patients who are intolerant to one serotonin selective reuptake inhibitor intolerant to another? *J Clin Psychiatry* 1995;56:30–34.

81. Joffe RT, Levitt AJ, Sokolov ST, et al. Response to an open trial of a second SSRI in major depression. *J Clin Psychiatry* 1996;57:114–115.

82. Zarate CA, Kando JC, Tohen M, et al. Does intolerance or lack of response with fluoxetine predict the same will happen with sertraline? *J Clin Psychiatry* 1996;57:67–71.

83. Thase ME, Blomgren SL, Birkett MA, et al. Fluoxetine treatment of patients with major depressive disorder who failed initial treatment with sertraline. *J Clin Psychiatry* 1997;58:16–21.

84. Howland RH, Thase ME. Switching strategies for the treatment of unipolar major depression. In: Rush AJ, ed. *Modern problems of pharmacopsychiatry, vol 25. Systematic medication management*. Basel: Karger, 1997:56–65.

85. Howland RH, Thase ME. What to do with SSRI nonresponders? *J Pract Psychiatry Behav Health* 1999;5:216–223.

86. Nierenberg AA, Feighner JP, Rudolph R, et al. Venlafaxine for treatment-resistant unipolar depression. *J Clin Psychopharmacol* 1994;14:419–423.

87. Raynaert C, Janne P, Zdanowicz N, et al. Efficacy of venlafaxine in depressed patients after switching from prior SSRI treatment. *Int J Neuropsychopharmacol* 2000;3(Suppl 1):S234.

88. Goodnick PJ, Sandoval R, Brickman A, et al. Bupropion treatment of fluoxetine-resistant chronic fatigue syndrome. *Biol Psychiatry* 1992;32:834–838.

89. Walker PW, Cole JO, Gardner EA, et al. Improvement in fluoxetine-associated sexual dysfunction in patients switched to bupropion. *J Clin Psychiatry* 1993;54:459–465.

90. Thase ME, Zajecka J, Kornstein SG, et al. Nefazodone treatment of patients with poor response to SSRIs. Presented at the 37th Annual Meeting of the American College of Neuropsychopharmacology, San Juan, Puerto Rico, December 1998.

91. Feiger A, Kiev A, Shrivastava RK, et al. Nefazodone versus sertraline in outpatients with major depression: focus on efficacy, tolerability, and effects on sexual functional and satisfaction. *J Clin Psychiatry* 1996;57(Suppl 2):53–62.

92. Catterson ML, Preskorn SH. Double-blind crossover study of mirtazapine, amitriptyline, and placebo in patients with major depression. Presented at the 149th Annual Meeting of the American Psychiatric Association, New York, NY, May 1996.

93. Fava M, Dunner D, Griest J, et al. An open label study with mirtazapine in SSRI failure. Presented at the 152nd Annual Meeting of the American Psychiatric Association, Washington, DC, May 1999.

94. Fava M, McGrath PJ, Sheu W-P, et al. Switching fluoxetine to reboxetine: An efficacy and safety study in depressed patients resistant to fluoxetine. *Int J Neuropsychopharmacol* 2000;3(Suppl 1):S234.

95. Crismon ML, Trivedi MH, Pigott TA, et al. The Texas Medication Algorithm Project: report of the Texas consensus conference panel on medication treatment of major depressive disorder. *J Clin Psychiatry* 1999;60:142–156.

96. Thase ME, Rush AJ, Howland RH, et al. Double-blind switch study of imipramine or sertraline treatment of antidepressant-resistant chronic depression. *Arch Gen Psychiatry, in press*.

97. Rosenbaum JF, Fava M, Nierenberg AA, et al. Treatment-resistant mood disorders. In: Gabbard GO, ed. *Treatments of psychiatric disorders*, third ed. Washington, DC: American Psychiatric Press 2001:307–1383.

98. Mischoulon D, Nierenberg AA, Kisilbash L, et al. Strategies for managing depression refractory to selective serotonin reuptake inhibitor treatment: a survey of clinicians. *Can J Psychiatry* 2000;45:476–481.

99. Fatemi SH, Emamian ES, Kist DA. Venlafaxine and bupropion combination therapy in a case of treatment-resistant depression. *Ann Pharmacother* 1999;33:701–703.

100. Rush AJ, Trivedi MH, Batey SR, et al. Anxiolysis associated with antidepressant response to bupropion sustained release or sertraline [Abstract]. *Int J Neuropsychopharmacol* 2000;3(Suppl 1):S235.

101. Donovan SJ, Quitkin FM, Stewart JW, et al. Duration of antidepressant trials: clinical and research implications. *J Clin Psychopharmacol* 1994;14:64–66.

102. Schweitzer E, Rickels K, Amsterdam JD, et al. What constitutes an adequate antidepressant trial for fluoxetine? *J Clin Psychiatry* 1990;51:8–11.

103. Nierenberg AA, McLean NE, Alpert JE, et al. Early nonresponse to fluoxetine as a predictor of poor 8-week outcome. *Am J Psychiatry* 1995;152:1500–1503.

104. Klerman GL, Weissman MM, Rousaville BJ, et al. *Interpersonal psychotherapy of depression*. New York: Basic Books, 1984.

105. Frank E, Kupfer DJ, Jacob M, et al. Personality features and response to acute treatment in recurrent depression. *J Person Disord* 1987;1:14–26.

106. Karp JF, Frank E, Anderson B, et al. Time to remission in late-life depression: analysis of effects of demographic, treatment, and life events measures. *Depression* 1993;1:250–256.

107. Leinonen E, Skarstein J, Behnke K, et al. Efficacy and tolerability of mirtazapine versus citalopram: a double-blind, randomized study in patients with major depressive disorder. *Int Clin Psychopharmacol* 1999;14:329–337.

108. Benkert O, Szegedi A, Kohenen R. Rapid onset of therapeutic action in major depression: A comparative trial of mirtazapine and paroxetine. Presented at the 37th Annual Meeting of the American College of Neuropsychopharmacology, Las Croabas, Puerto Rico, December 14–18, 1998.

109. Wheatley DP, van Moffaert M, Timmerman L, et al. Mirtazapine: efficacy and tolerability in comparison with fluoxetine in patients with moderate to severe major depressive disorder. *J Clin Psychiatry* 1998;59:306–312.

110. Entsuah R, Derivan A, Kikta D. Early onset of antidepressant action of venlafaxine: pattern analysis in intent-to-treat patients. *Clin Therapeut* 1998;20:517–526.

111. Lin EH, Von Korff M, Katon W, et al. The role of the primary care physician in patients' adherence to antidepressant therapy. *Med Care* 1995;33:67–74.

112. Williams JW Jr, Trivedi MH, Chiquette E, et al. Evidence Report/Technology Assessment, Number 7. *Treatment of depression: newer pharmacotherapies*. Rockville, MD: U.S. Department of Health and Human Services, Agency for Health Care Policy and Research, AHCPR Publication 99-E014, 1999.

113. Thase ME. Efficacy and tolerability of once-daily venlafaxine extended release (XR) in outpatients with major depression. *J Clin Psychiatry* 1997;58:393–398.

114. Cunningham LA. Once-daily venlafaxine extended release (XR) and venlafaxine immediate release (IR) in outpatients with major depression. *Ann Clin Psychiatry* 1997;9:157–164.

115. Rudolph RL, Feiger AD. A double-blind, randomized, placebo-controlled trial of once-daily venlafaxine extended release (XR) and fluoxetine for the treatment of depression. *J Affect Disord* 1999;56:171–181.

116. Basco MR, Rush AJ. Compliance with pharmacotherapy in mood disorders. *Psychiatric Ann* 1995;25:269–270,276, 278–279.

117. Jarrett RB, Rush AJ. Short-term psychotherapy of depressive disorders: current status and future directions. *Psychiatry Interpersonal Biol Proc* 1994;57:115–132.

118. Jarrett RB, Schaffer M, McIntire D, et al. Treatment of atypical depression with cognitive therapy or phenelzine: a double-blind, placebo-controlled trial. *Arch Gen Psychiatry* 1999;56:431–437.

119. Elkin I, Gibbons RD, Shea MT, et al. Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *J Consult Clin Psychol* 1995;63:841–847.

120. Schulberg HC, Katon W, Simon GE, et al. Treating major depression in primary care practice: an update of the Agency for Health Care Policy and Research Practice Guidelines. *Arch Gen Psychiatry* 1998;55:1121–1127.

121. McCullough JP Jr. *Treatment for chronic depression. Cognitive behavioral analysis system of psychotherapy (CBASP)*. New York: Guilford Press, 2000.

122. Jarrett, RB, Basco MR, Risser R, et al. Is there a role for continuation phase cognitive therapy for depressed outpatients? *J Consult Clin Psychol* 1998;66:1036–1040.

123. Danish University Antidepressant Study Group. Paroxetine: a selective serotonin reuptake inhibitor showing better tolerance but weaker antidepressant effect than clomipramine in a controlled multicenter study. *J Affect Disord* 1990;18:289–299.

124. Danish University Antidepressant Study Group. Citalopram: clinical effect profile in comparison with clomipramine: a controlled multicenter study. *Psychopharmacology (Berlin)* 1986:90:131–38.

125. Clerc GE, Ruimy P, Verdeau-Pailles J. A double-blind comparison of venlafaxine and fluoxetine in patients hospitalized for major depression and melancholia. *Int Clin Psychopharmacol* 1994;9:138–143.

126. Dierick M, Ravizza L, Realini R, et al. A double-blind comparison of venlafaxine and fluoxetine for treatment of depression in outpatients. *Progr Neuropsychopharmacol Biol Psychiatry* 1996;20:57–71.

127. Tylee A, Beumont G, Bowden M, et al. A double-blind, randomised, 12 week comparison study of the safety and efficacy of venlafaxine and fluoxetine in moderate to severe major depression in general practice. *Primary Care Psychiatry* 1997;3:51–58.

128. Diaz-Martinez A, Benassinni O, Ontiveros A, et al. A randomized, open-label comparison of venlafaxine and fluoxetine in depressed outpatients. *Clin Therapeut* 1998;20:467–476.

129. Costa e Silva J. Randomized, double-blind comparison of venlafaxine and fluoxetine in outpatients with major depression. *J Clin Psychiatry* 1998;59:352–357.

130. Ballús C, Quiros G, de Flores T, et al. The efficacy and tolerability of venlafaxine and paroxetine in outpatients with depressive disorder or dysthymia. *Int Clin Psychopharmacol* 2000;15:43–48.

131. Poirier M-F, Boyer P. Venlafaxine and paroxetine in treatment-resistant depression: double-blind, randomized comparison. *Br J Psychiatry* 1999;175:12–16.

132. Mehtonen O-P, Søgaard J, Roponen P, et al. Randomized, double-blind comparison of venlafaxine and sertraline in outpatients with major depressive disorder. *J Clin Psychiatry* 2000;61:95–100.

133. Keller MB. The long-term treatment of depression. *J Clin Psychiatry* 1999;60(Suppl 17):41–45.

134. Katz MM, Koslow SH, Maas JW, et al. The timing, specificity and clinical prediction of tricyclic drug effects in depression. *Psychol Med* 1987;17:297–309.

135. Koran LM, Hamilton SH, Hertzman M, et al. Predicting response to fluoxetine in geriatric patients with major depression. *J Clin Psychopharmacol* 1995;15:421–427.

136. Nagayama H, Nagan K, Ikezaki A. Prediction of efficacy of antidepressant by a 1-week test therapy in depression. *J Affect Disord* 1991;23:213–216.

137. Fredman SJ, Fava M, Kienke AS, et al. Partial response, nonresponse, and relapse with selective serotonin reuptake inhibitors in major depression: a survey of current “next step” practices. *J Clin Psychiatry* 2000;61:403–408.

138. Cain JW. Poor response to fluoxetine: underlying depression, serotonergic overstimulation, or a “therapeutic window”? *J Clin Psychiatry* 1992;53:272–277.

139. Oquendo MA, Malone KM, Ellis SP, et al. Inadequacy of antidepressant treatment for patients with major depression who are at risk for suicidal behavior. *Am J Psychiatry* 1999;156:190–194.

140. Kramer MS, Cutler N, Feighner J, et al. Distinct mechanism for antidepressant activity by blockade of central substance P receptors. *Science* 1998;281:1640–1645.

141. Vale W, Spiess J, Rivier C, et al. Characterization of a 41-residue ovine hypothalamic peptide that stimulates secretion of corticotropin and B-endorphin. *Science* 1981;213:1394–1397.

142. Owens MJ, Nemeroff CB. The physiology and pharmacology of corticotropin-releasing factor. *Pharmacol Rev* 1991;43:425–473.

143. Holsboer F, Spengler D, Heuser I. The role of corticotropin-releasing hormone in the pathogenesis of Cushing's disease, anorexia nervosa, alcoholism, affective disorders and dementia. *Progr Brain Res* 1992;93:385–417.

144. Zobel AW, Nickel T, Künzel HE, et al. Effects of the high-affinity corticotropin-releasing hormone receptor 1 antagonist R121919 in major depression: the first 20 patients treated. *J Psychiatric Res* 2000;34:171–181.

145. Bymaster FP, Perry KW, Tollefson DT, et al. Synergistic effects of olanzapine and other antipsychotic agents in combination with fluoxetine on norepinephrine and dopamine release in rat prefrontal cortex. Presented at the XI World Congress of Psychiatry, Hamburg, Germany, August 6–11, 1999.

146. Tohen MF, Shelton RC, Tollefson GD, et al. Olanzapine plus fluoxetine: double-blind and open-label results in treatment-resistant major depressive disorder. Presented at the XI World Congress of Psychiatry, Hamburg, Germany, August 6–11, 1999.

147. Lewinsohn PM, Clarke GN, Hops H. Cognitive-behavioral treatment for depressed adolescents. *Behav Ther* 1990;21:385–401.

148. Lewinsohn PM, Clarke GN, Rohde P. A course in coping: a cognitive-behavioral approach to the treatment of adolescent depression. In: Hibbs E, Jensen PS, eds. *Psychosocial treatments for child and adolescent disorders. Empirically based strategies for clinical practice*. Washington, DC: American Psychiatric Association, 1996:109–135.

149. Wood AJ, Harrington RC, Moore A. Controlled trial of a brief cognitive-behavioural intervention in adolescent patients with depressive. *J Child Psychol Psychiatry* 1996;37:737–746.

150. Brent DA, Holder D, Kolko D, et al. A clinical psychotherapy trial for adolescent depression comparing cognitive, family, and supportive therapy. *Arch Gen Psychiatry* 1997;54:877–885.

151. Mufson L, Weissman MM, Moreau D, et al. Efficacy of interpersonal psychotherapy for depressed adolescents. *Arch Gen Psychiatry* 1999;56:573–579.

152. Birmaher B, Brent DA, Kolko D, et al. Clinical outcome after short-term psychotherapy for adolescents with major depressive disorder. *Arch Gen Psychiatry* 2000;57:29–36.

153. Harrington R, Whittaker J, Shoebridge P, et al. Systematic review of efficacy of cognitive behaviour therapies in childhood and adolescent depressive disorder. *Br Med J* 1998;316:1559–1563.

154. Ryan ND. Pharmacologic and psychotherapeutic treatments for adolescent depression. *Econ Neurosci* 2000;2:49–51.

155. Emslie GJ, Rush AJ, Weinberg WA, et al. A double-blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. *Arch Gen Psychiatry* 1997;54:1031–1037.

156. Keller MB, Ryan ND, Birmaher B, et al. Paroxetine and imipramine in the treatment of adolescent depression. In: Program and Abstracts of the 151st Annual Meeting of the American Psychiatric Association, 1998; Toronto, Ontario, Canada. Abstract NR206.

157. Ryan ND, Varma D. Child and adolescent mood disorders—experience with serotonin-based therapies. *Biol Psychiatry* 1998;44:336–340.

158. Laakmann G, Schule C, Baghai T, et al. St. John's wort in mild to moderate depression: the relevance of hyperforin for the clinical efficacy. *Pharmacopsychiatry* 1998;31(Suppl)54–59.

159. Müller WE, Rolli M, Schäfer C, et al. Effects of hypericum extract (LI 160) in biochemical models of antidepressant activity. *Pharmacopsychiatry* 1997;30(Suppl):102–107.

160. Linde K, Ramirez G, Mulrow CD, et al. St John's wort for depression—an overview and meta-analysis of randomized clinical trials. *Br Med J* 1996;313:253–258.

161. Stevison C, Ernst E. Hypericum for depression: an update of the clinical evidence. *Eur Neuropsychopharmacol* 1999;9:501–505.

162. Gaster B, Holroyd J. St. John's Wort for depression. *Arch Int Med* 2000;160:152-156.

163. Philipp M, Kohnen R, Hiller K-O. Hypericum extract versus imipramine or placebo in patients with moderate depression: randomized multicentre study of treatment for eight weeks. *Br Med J* 1999;319:1534–1538.

164. Volz HP, Laux P. Potential treatment for subthreshold and mild depression: A comparison of St. John's wort extracts and fluoxetine. *Comp Psychiatry* 2000;41(Suppl 1):133–137.

165. Shelton RC, Dunner D, Gelenberg A, et al. Placebo-controlled trial of St. John's wort in major depression. Presented at the 153rd Annual Meeting of the American Psychiatric Association, 1998; Chicago, IL.

166. Hamilton M. The assessment of anxiety states by rating. *Br J Med Psychol* 1959;32:50–55.

167. Bressa GM. S-adenosyl-l-methionine (SAMe) as antidepressant: meta-analysis of clinical studies. *Acta Neurol Scand* 1994;(Suppl):1547–1614.

168. Berlanga C, Ortega-Soto HA, Ontiveros M, et al. Efficacy of S-adenosyl-L-methionine in speeding the onset of action of imipramine. *Psychiatry Res* 1992;44:257–262.

169. Fava M, Giannelli A, Rapisarda V, et al. Rapidity of onset of the antidepressant effect of parenteral S-adenosyl-L-methionine. *Psychiatry Res* 1995;56:295–297.

170. Benelli A, Filaferro M, Bertolini A, et al. Influence of S-adenosyl-L-methionine on chronic mild stress-induced anhedonia in castrated rats. *Br J Pharmacol* 1999;127:645–654.

171. Ernst E, Rand JI, Stevinson C. Complementary therapies for depression: an overview. *Arch Gen Psychiatry* 1998;55:1026–1032.

172. Lou H, Jia Y, Zhan L. Electro-acupuncture vs amitriptyline in the treatment of depressive states. *J Trad Chin Med* 1985;5:3–8.

173. Luo H, Jia Y, Wu X, Dai W. Electro-acupuncture in the treatment of depressive psychosis. *Int J Clin Acupuncture* 1990;1:7–13.

174. Röschke J, Wolf CH, Müller MJ, et al. The benefit from whole body acupuncture in major depression. *J Affect Disord* 2000;57:73–81.

175. Allen JJB, Schnyer RN, Hitt SK. The efficacy of acupuncture in the treatment of major depression in women. *Psychol Sci* 1998;9:397–401.

176. Dubini A, Bosc M, Polin V. Do noradrenaline and serotonin differentially affect social motivation and behaviour? *Eur Neuropsychopharmacol* 1997a;7(Suppl 1):S49–S55.

177. Dubini A, Bosc M, Polin V. Noradrenaline-selective versus serotonin-selective antidepressant therapy: differential effects on social functioning. *J Psychopharmacol (Oxford)* 1997b;11(Suppl 4):S17–S23.

178. Massana J, Möller H-J, Burrows GD, et al. Reboxetine: a double-blind comparison with fluoxetine in major depressive disorder. *Int Clin Psychopharmacol* 1999;14:73–80.

76

Electroconvulsive Therapy: Sixty Years of Progress and a Comparison with Transcranial Magnetic Stimulation and Vagal Nerve Stimulation

William M. McDonald

W. Vaughn McCall

Charles M. Epstein

William M. McDonald: Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, Georgia

W. Vaughn McCall: Department of Psychiatry, Bowman Gray Medical Center, Winston Salem, North Carolina

Charles M. Epstein: Department of Neurology, Emory University, Atlanta, Georgia

The science of electroconvulsive therapy (ECT) has progressed rapidly over the last 20 years, providing new insights into the mechanisms of action of ECT, improving both the acute and long-term efficacy of the treatments, and decreasing cognitive problems associated with the treatments. The anticonvulsant hypothesis unifies many of the scientific findings in electroencephalography, neuroimaging, and clinical studies to provide a coherent testable hypothesis of the therapeutic mechanism of ECT. This hypothesis assumes that ECT enhances the transmission of inhibitory neurotransmitters and neuropeptides and that the active process of inhibiting the seizure is essential to the therapeutic action of ECT. New data are presented on improvements in the acute efficacy of ECT with suprathreshold (eight to 10 times the seizure threshold) right unilateral ECT, and three NIMH-supported studies are discussed that examine the efficacy of maintenance therapies. Decreasing cognitive side effects of ECT is another area of active research; changes in techniques and medication treatments are highlighted. Finally, two new treatments using subconvulsive stimuli, repetitive transcranial magnetic stimulation, and vagal nerve stimulation are discussed and compared to ECT.

Cerletti and Bini (1) first investigated ECT as a treatment for psychosis in 1938, theorizing that epilepsy and schizophrenia were incompatible. They hypothesized that the artificial induction of seizures in nonepileptic persons with schizophrenia would relieve psychosis. ECT was developed at approximately the same time as frontal leukotomy (2) and insulin coma therapy (3), but these treatments carried a high morbidity, and were replaced by modern psychopharmacology in the 1950s and 1960s. The indications for ECT were established during this time, and its use in conditions other than mood disorders and schizophrenia diminished.

ECT has remained as an accepted medical treatment for depression and was one of the most significant medical advances in the twentieth century. However, in 1950 the mortality and morbidity from ECT were unacceptably high and most of the early research in ECT focused on the safety and efficacy of the treatments. The death rate was approximately 0.1% (4) and the risk of fractured bones estimated to be as high as 40% (5). Over the last half century, the mortality from ECT has decreased dramatically because of a number of advances, including the widespread use of modern anesthetic agents (e.g., methohexital) (5) and succinylcholine (6), the introduction of cardiac and electroencephalogram (EEG) monitoring during treatments (7), and medications used to decrease the acute hemodynamic response during ECT (8). Abrams put the risk of mortality from ECT into perspective in 1997. He noted that ECT was ten times safer than childbirth and an order of magnitude less that the spontaneous death rate in the population (9).

Today, ECT is recognized as a safe treatment for depression and, despite the advances in pharmacotherapy in the last four decades, ECT continues to be the most effective treatment for severe melancholic depression. Four areas of research are important as we move into the twenty-first century: developing a scientific understanding of the mechanisms of action of ECT, optimization of the efficacy of acute courses of ECT, treatment of the cognitive side effects

of ECT, and continuing research into the efficacy of different ECT techniques, including novel electrode placements and continuation/maintenance ECT.

MECHANISM OF ACTION OF ECT

Part of "76 - Electroconvulsive Therapy: Sixty Years of Progress and a Comparison with Transcranial Magnetic Stimulation and Vagal Nerve Stimulation "

Salzman asserts that the psychiatric community continues to show ambivalence toward ECT and ECT research has suffered as a consequence (10). Although there are active steps within the NIMH to address these issues and provide more focused research in ECT (10), an understanding of the basic mechanisms by which ECT exerts its effect is still unclear. This fact is an irony given that ECT is one of the few treatments in psychiatry that was theoretically based (11). Although the original hypothesis that epilepsy and psychosis could not coexist in the same patient was later proven incorrect, the treatment of depression by inducing seizures in patients first with camphor monobromide and later by electrically induced convulsions is one of major medical breakthroughs of the past century.

One of the most confusing aspects of accepting ECT as a treatment for depression in the lay public and patients is the inability of clinicians to clearly explain how ECT is effective in relieving symptoms of depression. Patients have difficulty understanding how a treatment that is so seemingly toxic to the brain (i.e., causing a seizure) could actually be therapeutic. Patients readily accept that the therapeutic effect of antidepressants is to replete a serotonin deficit. However, the antidepressant medications have a number of other effects on a variety of neurotransmitters, regulatory hormones, and cellular mechanisms.

The mechanism by which a convulsive stimulus acts as one of the most powerful antidepressants is equally complex but the explanation may be as simplistic: ECT works by increasing natural brain substances that decrease the excitability of the brain. The unique therapeutic action of ECT is that the induced-seizures last only seconds and not minutes or hours. The physician’s contribution to the ECT treatment is the induction of a seizure; the patient’s contribution is turning the seizure off. Seizure termination is an active process that underlies the therapeutic mechanism of the treatment. This idea is elaborated on in the following but may give us a plausible explanation for patients on the mechanism of action of ECT.

The earliest attempts at understanding the importance of a convulsion in ECT focused on terminating the ECT-induced seizure with lidocaine to determine if the length of a seizure correlated with the efficacy of the treatment (12). The initial finding was that an ECT seizure had to continue for at least 25 seconds to be therapeutic (13) and the patient had to accumulate a minimal number of seconds of EEG seizure time during a course of ECT (14).

A popular theory of the mechanism of action of ECT, the diencephalic hypothesis (15 ,16), assumed that the degree of response to ECT was correlated with stimulation of the deep brain structures that regulate the hypothalamic pituitary–endocrine end axis activity. Stimulation of this system resulted in the release of pituitary hormones such as adrenocorticotropin hormone (ACTH), thyrotropin, prolactin, oxytocin, and vasopressin. Research using rodents administered electroconvulsive shock (ECS) and examining the CSF of patients receiving ECT has supported a relationship between increases in neuropeptides during the convulsive stimulus (17). According to the diencephalic hypothesis, ECT seizures that have a longer duration, and ECT parameters that are more effective at stimulating the diencephalic structures (i.e., bilateral [BL] greater than unilateral [UL] and high-dose greater than low-dose ECT), therefore, would be more effective in treating depression.

Both of these assumptions have been questioned recently. Sackeim and colleagues were the first to demonstrate that the therapeutic benefit of the ECT seizure was dependent on the amount above the individual patient’s seizure threshold that the stimulus was administered and not simply the duration of the seizure (18). Although the acute release of neuroendocrine markers did correlate with the type of seizure administered and seizure duration, the expected correlation between the acute surge in plasma oxytocin, vasopressin (19), or prolactin (20), and clinical response to ECT was not shown in studies of depressed patients receiving a course of ECT.

Recent research on predicting a patient’s response to ECT have focused less on the quantitative analysis of seizure duration and more on the relationship between a qualitative analysis of the ictal and postictal seizure morphology to therapeutic response (21 ,22). ECT-induced seizures have a characteristic pattern of hypersynchronous neuronal discharge with excitation of cortical neurons during the initial tonic phase, followed by alternating excitatory and inhibitory effects in the clonic phase, and finally postictal suppression owing to inhibition and neuronal hypoexcitability.

A number of treatment-related factors affect seizure morphology, including electrode placement, the percentage above the seizure threshold that the stimulus is administered, and the stimulus waveform (23). A number of features of the ictal EEG seizure that demonstrate a more intense seizure predict clinical response to ECT. (See ref. 24 for review.) Seizure intensity is measured by increased ictal EEG amplitude (high-voltage spikes and waves), symmetry or coherence of ictal EEG amplitude between the right and left hemispheres, and weaker postictal hemispheric coherence, longer duration of stereotypic ictal EEG pattern, and greater postictal suppression or flattening of the EEG. These seizure characteristics have been used to predict the efficacy of an ECT course (22 ,25 ,26 and 27), or more precisely these variables can be used to predict when a seizure is not adequate. Inadequate EEG seizures have low-amplitude waves and are not symmetric, with no clear ending of the seizure in the postictal period.

Analysis of the EEG morphology has been used to determine seizure intensity (26). Clinicians can be trained to visually inspect the EEG strips during ECT and determine the adequacy of the seizure by evaluating the amplitude of the ictal EEG relative to baseline, symmetry of right and left hemispheric EEGs, distinct spike and wave pattern, and degree of postictal suppression (28). Both the Thymatron DGx ECT device (Somatics Inc., Pine Bluff, IL) and the MECTA 5000Q ECT device (MECTA Corp., Lake Oswego, OR) provide measures of the quality of the EEG. Although further testing of the clinical use of the computer-assisted EEG analysis provided with these machines is necessary, these devices may add to stimulus dosing, particularly in patients administered UL ECT, in determining if a seizure is adequate (24 ,27 ,29).

Another promising area of investigation into the mechanism of ECT is research correlating functional brain imaging with response to ECT. Studies have shown an increase in cerebral blood flow (CBF) up to 300% of baseline values with an accompanying increased permeability of the blood–brain barrier and increased cerebral metabolic rate (CMR) up to 200% during the ictal period (30). In contrast, CBF decreased to levels below baseline (31) or returned to baseline (32) and PET scans showed a marked decrease in CMR in the postictal period (33).

Although there have not been any clinical studies correlating the increase in CBF/CMR during the ictal period with clinical response, Nobler and colleagues (34) found a correlation between decreased CBF in the immediate postictal period and clinical response. This well-designed study included 54 depressed patients imaged using the Xenon inhalation technique. Patients showed a low baseline CBF compared to matched controls, and their response to ECT was correlated with the further decrease in CBF from baseline. These changes were greatest in the anterior cortical regions and the degree of change was correlated with clinical improvement on the Hamilton Depression Rating Scale. Moreover, these reductions in CBF persisted for the 2-month follow-up in responders, was correlated with improvement in depression scores and the nonresponders continued to have an increase in CBF compared to baseline. Nobler and Sackeim (35) point out that decreased CBF in the anterior cortex supports earlier findings by Max Fink (36 ,37) and their own group (38) of a relationship between frontal delta activity on EEG and response to ECT. There are additional data suggesting that the reductions in cerebral blood flow that occur immediately after ECT may persist for days (39) to months (40) after the treatments and the most dramatic reductions occurred in the frontal cortex.

Sackeim has unified many of these EEG, CBF, and CMR findings with preclinical research in the anticonvulsant hypothesis as the mechanism of action of ECT (40). During a course of ECT, the patient’s seizure threshold is increased and seizure duration decreases during the first several treatments (41 ,42). ECT seizures usually are limited to less than 1 minute and there is an active inhibitory process in the interictal and postictal states evident by the development of slow or delta waves and decreases in the CBF and metabolic uptake of glucose. The anticonvulsant properties of ECT are hypothesized to occur because of enhanced transmission of inhibitory neurotransmitters and neuropeptides (e.g., GABA and endogenous opioid concentrations) and are an essential element of the therapeutic effect of ECT in mania and depression.

The magnitude of the seizure threshold increase is greater in more effective methods of administering ECT (i.e., high-dose UL and low-dose BL versus low-dose UL) and is correlated with clinical response (18 ,43). Clinically, Sackeim cites unpublished data that the patients who return to an acute course of ECT after relapsing have the same seizure threshold that they had at the start of treatment. However, it is unclear whether the return of the seizure threshold to baseline occurs after an acute course of ECT in all patients or only in patients who relapse. The duration of the seizure is also decreased over a series of treatments and is another indication of the anticonvulsant effect of ECT. However, seizure duration is not related to efficacy unlike seizure threshold (43).

The EEG of the patient during and immediately following therapeutic ECT treatments indicate that there is an ongoing active process in which the brain is inhibiting the seizure process. Inhibitory processes include the early onset of high amplitude slow-wave activity after the tonic phase of the seizure and bioelectric postictal suppression processes (21 ,22 ,27 ,44 ,45 and 46). The efficacy of ECT has been correlated with the early onset of these inhibitory processes, a fact that supports the anticonvulsant hypothesis. Two elements of seizure expression, seizure strength and peak amplitude of slow-wave activity, were inversely correlated with seizure threshold and the third element, postictal bioelectric suppression, was not related to seizure threshold (40). Because the seizure threshold is increasing during the treatment course, two of the essential elements of a therapeutic seizure (seizure strength and peak amplitude of slow-wave activity) are deteriorating, thereby limiting the effectiveness of subsequent seizures. This finding provides the rationale for developing EEG algorithms (see the preceding), increasing the stimulus dosing or retitrating the seizure threshold during a course of ECT in patients who are not responding.

During a course of ECT, as in epilepsy, CBF/ CMR increase dramatically during the seizure and decrease below baseline in the interictal and postictal states (35). Patients responding to ECT show a more marked global decrease in CBF as well as specific reductions in the anterior frontal cortex (34). These changes were correlated with an increase in the seizure threshold. Finally, increasing slow-wave activity in the frontal cortical regions after ECT was also associated with clinical improvement (38). Both these finding support the anticonvulsant hypothesis.

The anticonvulsant hypothesis unifies many research

findings and provides important new leads that have the potential for improving clinical outcomes and predicting patients who are at risk of relapsing after ECT. The anticonvulsant properties of ECT related to clinical outcome include an increase in the seizure threshold during a course of ECT, early onset of slow-wave activity interictally, distinct postictal suppression, and decreases in CBF/CMR and increased slow-wave activity after a series of treatments.

An anticonvulsant mechanism for ECT would be unique when compared to the antidepressant medications (which are rarely associated with seizures) and anticonvulsant medications (e.g., valproic acid and carbamazepine), which have only moderate antidepressant efficacy (47). However, there is evidence that newer anticonvulsant medications, including lamotrigine, may be more effective in the depressed phase of bipolar illness (48). The mechanism by which the anticonvulsants exert their antidepressant effects is poorly understood and is hypothesized to occur through a number of neurotransmitters, including inhibiting the presynaptic release of excitatory amino acids (e.g., glutamate) and enhancing the effect of inhibitory neurotransmitters and neuropeptides.

New research should focus on testing the anticonvulsant hypothesis and determining the neurotransmitters essential for the antidepressant properties of ECT. Examining the relationship of the anticonvulsant effects of ECT to the efficacy of the treatments by blocking or augmenting the anticonvulsant properties of ECT can test this hypothesis. For example, CSF neuropeptides associated with the anticonvulsant effects of electroconvulsive shock (ECS) have been isolated in laboratory animals (49). These neuropeptides could be given in conjunction with ECS to determine if the coadministration of these neuropeptides would block the therapeutic efficacy of ECS (50). Although ethical considerations may limit this type of research in humans, studies could be designed to investigate the relationship of the acute rise in endogenous anticonvulsant substances (including GABAergic and endogenous opioids) in the CSF of ECT responders versus nonresponders.

Clinical research should continue to concentrate on developing algorithms to determine the relationship of ECT treatment variables (e.g., seizure threshold) to ECT response or the loss of seizure efficacy during a course of ECT. These algorithms can test the anticonvulsant hypothesis and guide clinicians in administering effective treatments. Treatment variables (e.g., diminished CBF in the anterior frontal lobes) may also be investigated to predict relapse.

OPTIMIZATION OF ACUTE ECT

Part of "76 - Electroconvulsive Therapy: Sixty Years of Progress and a Comparison with Transcranial Magnetic Stimulation and Vagal Nerve Stimulation "

ECT is widely cited to have an antidepressant efficacy of greater than 80% (51). More recently, the percent of patients achieving remission after an acute course of ECT is conservatively estimated at between 50% and 80%. The reason for this apparent decline in efficacy is the increasing resistance to treatment among the patients referred for ECT in the last 15 years. Prior to the development of the serotonin reuptake inhibitors (SSRIs), the most common reason for referral for ECT was intolerance of available antidepressant medications, chiefly tricyclic antidepressants (TCAs), and monoamine oxidase inhibitors (MAOIs). Intolerance to antidepressant pharmacotherapy was particularly important in the elderly who constitute the majority of patients referred for ECT. The elderly are also the patient group most susceptible to the side effects of TCAs and MAOIs. The advent of SSRIs, serotonin, and norepinephrine reuptake inhibitors (SNRIs), 5-HT_{2A} blockers, and other novel antidepressants in the last 15 years has radically changed the medication histories of the patients referred for ECT. These newer agents are remarkably well tolerated, and a majority of patients receiving ECT have had at least one complete trial of an antidepressant without improvement. These patients who are refractory to antidepressant medications are also more refractory to ECT. The ECT response rate in antidepressant medication-refractory patients is probably no better that 50%, whereas the response rate in medication naive patients is closer to the historically quoted 80% to 90% (52). The problem of decreased acute response to ECT in this growing population of medication-refractory patients has led to new interest in the technical factors that control the response to acute ECT. Among these factors are electrode placement, stimulus dose, and possibly concurrent medications.

Despite 60 years of clinical application, the ECT field has only recently come to appreciate the relative contributions of stimulus electrode placement and stimulus intensity to the therapeutic and adverse effects of ECT. As previously discussed, clinical wisdom prior to 1987 taught that as long as the EEG seizure during ECT was ≥25 seconds, then the treatment was maximally effective. This belief was refuted by Sackeim and colleagues' report that when ECT is delivered with right unilateral (RUL) electrode placement and a stimulus just barely above the patient's seizure threshold, then the antidepressant response rate was approximately 20%, despite seizure durations ≥25 seconds (18). Further work by this group clarified that the efficacy of right unilateral (RUL) ECT was exquisitely sensitive to the magnitude of the stimulus dose expressed as a multiple of a given patient's seizure threshold (53) (Table 76.1).

	Barely Suprathreshold	2.5-Times Threshold
RUL	17%	43%
BL	65%	63%

TABLE 76.1. ANTIDEPRESSANT RESPONSE RATE BY ELECTRODE PLACEMENT AND STIMULUS DOSE

This work demonstrated that, within the stimulus dose studied, the efficacy for RUL was sensitive to dose, whereas the efficacy of bilateral (BL) ECT was insensitive to dose. Also, RUL ECT appeared inferior to BL ECT at any dose. Subsequent work confirmed that the efficacy of RUL ECT is dose dependent and that RUL ECT administered at six times the initial seizure threshold is as effective as BL ECT with fewer cognitive side effects than BL ECT (54). McCall and colleagues found a dose–response relationship in patients receiving RUL ECT that extended to 12 times the seizure threshold and, as predicted, cognitive side effects were increased as the stimulus dose increased relative to the seizure threshold (55). The response rate for RUL ECT at eight to 12 times threshold was about 70%, approximately the rates typically quoted for BL ECT.

Based on this new information, it seems reasonable to recommend that ECT be delivered with either BL electrode placement at a stimulus dose just above threshold, or with RUL placement at a stimulus dose markedly above threshold. Less intensive strategies (i.e., RUL at 2.5 times threshold) should probably be avoided for routine use. Although it is clear that the lower intensity RUL strategies have less acute amnesia side effects, this advantage is probably offset by poorer efficacy and consequently poorer quality of life (QOL), because depression severity is among the best predictors of QOL in severely depressed patients (56 ,57).

There remain three important considerations in the routine use of titrated unilateral ECT. First, many psychiatrists are hesitant to determine the seizure threshold because they feel that the procedure is medically dangerous. There is a potential risk to stimulating the vagus nerve with subconvulsive stimuli without adequate compensatory adrenergic discharge from a seizure and the possibility of prolonged asystole. However, a controlled operative setting with cardiac monitoring decreases the possibility of brief periods of bradycardia causing any significant risk for a majority of patients.

Second, some would argue that the initial treatment in a series of titrated seizures is a “wasted” seizure, adding to costs and potential cognitive side effects without any significant therapeutic benefit to the patient. The unilateral seizure at or near the seizure threshold used during the initial treatment setting to determine seizure threshold probably adds little benefit. However, the potential advantages of determining the seizure threshold and adjusting the subsequent seizures to the threshold has advantages in maximizing benefit and decreasing the potential for future ineffective treatments (if the energy delivered is too low) and cognitive side effects (if the energy is too high).

Finally, fixed high dose RUL ECT is an effective treatment for depression (55 ,58) and this data could obviate the need to determine seizure threshold. Because the majority of patients treated with ECT are older and older patients have very high seizure thresholds, dose titrations at eight to ten times threshold are generally at or above the 400-mC range used by McCall and associates. In fact, in some older patients even a suprathreshold stimulus set at 400 mC may be too low to achieve antidepressant efficacy.

Continued support for the efficacy of suprathreshold RUL ECT has raised some concern about the fact that ECT machines in the United States are restricted in the amount of energy they can deliver per treatment (504 to 576 mC maximal output). Abrams argues that the FDA mandated maximum output for ECT machines used in this country often does not allow for the administration of effective suprathreshold treatments and has called for a reexamination of these guidelines (59).

The question of whether or not to continue, discontinue, or initiate antidepressant medication during a course of acute ECT is also pertinent to the issue of improving the response rate among patients with a history of antidepressant medication resistance. Clinically, physicians are divided in their practice regarding antidepressant medications during ECT, and there are no good data to support any position. One of the developments in ECT practice in the 1990s was the occurrence of SSRI-resistance in ECT patients, and the virtual lack of TCA trials in these same patients before coming to ECT. This is especially important because there are some data to suggest that TCAs (i.e., nortriptyline) may be more effective than SSRIs in the severely depressed or hospitalized elderly (60). Possibly, the addition of an antidepressant (i.e., nortriptyline) may improve the antidepressant response to ECT in patients with a history of novel-antidepressant resistance.

PREVENTION OF RELAPSE AFTER ECT

Part of "76 - Electroconvulsive Therapy: Sixty Years of Progress and a Comparison with Transcranial Magnetic Stimulation and Vagal Nerve Stimulation "

In general, patients who receive an acute course of ECT are either subsequently resistant or intolerant of antidepressant medications. Patients diagnosed with psychotic depression are particularly susceptible to relapse after an acute course of ECT. Two studies found a relapse rare of approximately 70% in 1 year for a total of 53 patients with a diagnosis of psychotic depression (61 ,62). These studies were retrospective and could not examine compliance rates or the adequacy of either the initial (pre-ECT) or continuation medication trial.

In a prospective study, Sackeim and co-workers (63) followed 58 patients for 1 year after ECT. They examined a number of clinical variables including a retrospective review of the adequacy of the pre-ECT medication trial. The most important factor in determining relapse on maintenance medication after an acute course of ECT was the adequacy of the pre-ECT medication trial. Patients (with and without psychotic features) who were rated as receiving a therapeutic medication trial(s) prior to their acute course of ECT relapsed at a rate that was twice the rate found in patients who did not receive an adequate pre-ECT medication trial (64% versus 32%). The maintenance medication was not standardized but the results indicated that the adequacy of

the post-ECT medication trial was only marginally related to the relapse rate and then primarily in patients who did not have an adequate pre-ECT medication trial. The patients who were not determined to be medication resistant prior to ECT had a lower relapse rate when they received an adequate maintenance medication trial. Sackeim and associates argue that many of these patients may have responded to antidepressant medication prior to ECT if they had received an adequate trial. The overall relapse rate remained very high (approximately 50%) and most of the patients who relapsed did so in the first 4 months of this 1-year follow-up study. This finding indicates that the efficacy of ECT may be relatively transient. Given the increasing medication resistance found in most ECT patient populations, the clinical challenge is increasingly to determine the most effective maintenance treatments and increasingly physicians are utilizing maintenance ECT.

The American Psychiatric Association's (64) clinical guidelines for continuation ECT (C-ECT) include patients who: (a) have recurrent depressive episodes responsive to ECT; (b) demonstrate resistance/intolerance or noncompliance with antidepressant medications; (c) can comply with the overall treatment plan including behavioral restrictions (i.e., limiting driving) and providing informed consent. Continuation and maintenance ECT strategies are being used increasingly in the treatment of patients with major depression who are felt to have a high-risk for relapse (65). The guidelines for the use of C-ECT have been actively discussed but there are little prospective data on which to base recommendations on frequency of treatments and how long they should be continued or the nature of the potential cognitive side effects. A majority of the studies are case reports and many predate antidepressant medication (66). More recent reports in patients with depression (66 ,67 ,68 ,69 ,70 ,71 ,72 ,73 and 74), mania (75 ,76), and schizophrenia (77) describe a marked decrease in the number of hospitalizations, hospital days, depressive symptoms, increased functional status, and stable cognitive functioning for the period of continuation ECT.

Theoretically, there are several potential theories that may explain the efficacy of maintenance ECT over maintenance medication. First, maintenance ECT may be effective because the ECT treatments are gradually tapered rather than abruptly discontinued. During this taper, most clinical protocols decrease the interval between the maintenance treatments if the patient shows signs of relapse. A tapering schedule over this period may be critical because most patients relapse within 3 months of stopping the treatments. In 1954, prior to the use of antidepressant medications as maintenance treatment, Bourne and Long (78) coined the term "convulsion dependence" to describe psychotic patients who relapsed unless they were tapered from their ECT treatments.

The second potential therapeutic benefit may be the fact that ECT has a different mechanism of action than the antidepressant medications. Increasingly patients are presenting for ECT after multiple failed medication trials and there may be little benefit from yet another trial of an SSRI. In fact the one medication regimen that has been shown to be effective in maintenance therapy is a combination of lithium and nortriptyline (discussed in the following). The benefit of this combination therapy may be owing to the fact that few patients had been given lithium augmentation prior to their acute course of ECT.

Finally, maintenance ECT is similar to depot haloperidol in the treatment of schizophrenia and may provide prophylactic benefit from improved compliance in patients who might otherwise not comply with their maintenance medication. Most studies of maintenance ECT only report patients who complied with the treatment regimen. The experience at the Emory University Outpatient ECT program is similar to the data reported by Clarke and colleagues (67). When patients do not complete the scheduled 6-month maintenance ECT program, we found that more than half of the patients relapsed.

In 1991, Monroe (79) discussed the need for increased research into continuation and maintenance ECT. There is clear evidence that these treatments are being used increasingly in clinical practice, yet there is a lack of guidelines to establish the optimal treatment frequency, the type of patient who would benefit from maintenance ECT versus medication, or an understanding of the potential long-term side effects. The NIMH is presently funding three studies that will add significantly to our understanding of maintenance ECT. Two multicenter trials are examining the efficacy of different maintenance strategies after an acute response to ECT. Sackeim and colleagues are comparing maintenance placebo, nortriptyline versus nortriptyline and lithium after an acute response to ECT. Preliminary results from this line of investigation show that nortriptyline provides only marginally greater protection against relapse during the post-ECT period than does placebo, with relapse rates of approximately 70% during the first year. The addition of lithium to nortriptyline resulted in a further significant reduction in relapse to approximately 40% during that interval. Charles Kellner is the principal investigator on a trial comparing maintenance medication and maintenance ECT. The authors are currently examining the cost effectiveness of maintenance medication compared to maintenance ECT in elderly patients with recurrent major depression. These last two trials do not yet have preliminary data available, but together they will provide prospective data on the relative effectiveness and costs of different maintenance strategies used after an acute course of ECT.

COGNITIVE PROBLEMS ASSOCIATED WITH ECT

Part of "76 - Electroconvulsive Therapy: Sixty Years of Progress and a Comparison with Transcranial Magnetic Stimulation and Vagal Nerve Stimulation "

The one significant remaining ECT-related morbidity is the potential for cognitive side effects. Although there is little

scientific literature supporting permanent brain damage after ECT (80 ,81 and 82), memory disturbances continue to be a serious side effect after an acute course of ECT. Clearly, there have been significant advances over the last 50 years. Delirium was a clinical goal in some early protocols (so called “regressive ECT”) that equated the development of delirium with therapeutic response. This technique provided no advantage over more conservative forms of ECT and has been abandoned. Replacement of sine wave stimuli with machines that provide a brief pulse stimuli, and the elucidation of the appropriate stimulus parameters that increase the efficacy of safer treatments (suprathreshold right unilateral versus threshold bilateral ECT) have both been important contributions to the decrease in the ECT associated memory loss over the past few decades. The potential development of ultrabrief pulse ECT and research into pharmacologic treatments such as physostigmine (83), naloxone (84), and thyrotropin releasing hormone (TRH) (85) all hold promise for further decreasing ECT-associated memory loss. Double-blind placebo controlled trials using with intranasal vasopressin (86 ,87), ACTH (88 ,89), dexamethasone (90), and nimodipine (91) have not shown the active drug to be more effective than placebo.

The frequency of ECT administration is yet another way of controlling ECT-related memory loss. Twice-weekly ECT is associated with an antidepressant response as good as thrice-weekly ECT, but with less adverse memory effects (and a slower rate of antidepressant response) (92 ,93). Recently, the issue of optimal electrode placement has been reopened with isolated enthusiasm for bifrontal stimulating electrode placement or an asymmetric placement (a right fronto-temporal electrode referenced to a left frontal electrode). Preliminary reports suggest that these novel electrode placements are associated with an antidepressant efficacy comparable to standard bilateral placement, with fewer cognitive side effects, but there are insufficient data thus far to recommend them as a replacement for standard BL or RUL placements (94 ,95).

NONCONVULSIVE STIMULI IN THE TREATMENT OF DEPRESSION

Part of "76 - Electroconvulsive Therapy: Sixty Years of Progress and a Comparison with Transcranial Magnetic Stimulation and Vagal Nerve Stimulation "

Repetitive transcranial magnetic stimulation (rTMS) and vagal nerve stimulation (VNS) are two new treatments for depression that, unlike ECT, use subconvulsive stimuli to treat depressive symptoms. Compared to ECT, these treatments have the potential for targeting specific brain regions and stimulating areas that are thought to be involved in depression while sparing areas such as the hippocampus and thus reducing potential cognitive side effects.

rTMS was first used in 1985 as a noninvasive method of stimulating brain neurons (96). rTMS produces pulses of electromagnetic currents conducted by the brain that have different effects on neuronal firing, depending on the frequency in which it is administered. In general, rTMS acts by inducing action potentials in neuronal elements, but at high frequency (5 Hz) the net effect of a stimulus train is excitatory (97); similarities have been drawn between the effects of fast rTMS and long-term potentiation (LTP) produced by high frequency electrical stimulation (98). At low frequencies (1 Hz) the net effect of a stimulus train is usually inhibitory (99 ,100). In animal studies, rTMS has been demonstrated to cause down-regulation of *B*-adrenoreceptors (101) and, as has been shown in ECS, rTMS up-regulates astroglial gene expression in the CNS (102). Preliminary data support rTMS in the treatment of a number of psychiatric disorders including depression, schizophrenia, obsessive-compulsive disorder, and posttraumatic stress disorder (103).

This chapter focuses on the use of rTMS in depression and compares rTMS to ECT. From the outset it should be stressed that the efficacy of rTMS in the treatment of depression is modest and much of the research has been focused narrowly on a relatively restricted group of treatment parameters. The majority of rTMS studies use set parameters: high frequency or fast stimulations (≥10 Hz) over the left dorsolateral prefrontal cortex (DLPFC) at or around the motor threshold (e.g., 80% to 110% MT). These parameters are based on PET data (104) and ECT data (34) demonstrating hypofunctioning of the left prefrontal cortex in depression as well as two pivotal early studies demonstrating an antidepressant effect of fast left DLPFC rTMS (105 ,106). Although one recent parallel group sham controlled study found no treatment effect (107), most open (105 ,108 ,109) and double-blind studies (106 ,110 ,111) have confirmed a relatively modest antidepressant response for fast left DLPFC. When directly compared to ECT, 4 weeks of rTMS appears to be as effective as ECT in nonpsychotic, but not psychotic, depression (112). In this study there may have been a bias for the ECT group because over half of the patients in the ECT group and none of the patients in the rTMS group remained on antipsychotics and/or antidepressants. Two weeks of either RUL ECT or one ECT treatment followed by four rTMs sessions were also shown to be equivalent in the treatment of 22 depressed outpatients (113).

To date there have been no seizures reported using fast rTMS within the established safety guidelines (114) and no cognitive side effects in patient populations (109). The most common side effects are pain at the treatment site and headaches usually relieved by acetaminophen.

Realizing the potential benefits of rTMS, however, is dependent on expanding the treatment parameters including research on the variable effects of slow versus fast stimulation. Although early studies using slow TMS demonstrated poor outcomes (115 ,116 and 117), and are often cited as rationale for using fast rTMS, the stimulation site was the vertex in these three studies. More recent trials of *slow* rTMS over the *right* DLPFC have shown benefit. Klein and colleagues (118)

randomized 71 depressed patients to 2 weeks of slow, active or sham treatment over the right prefrontal cortex and 41% of active-treatment subjects had a 50% or greater drop in Hamilton depression scores compared to only 17% of sham-treatment subjects. Similarly, Tormos and associates (119) showed a significant antidepressant response to *fast left* DLPFC stimulation and *slow right* DLPFC stimulation, but not fast right DLPFC stimulation or sham. Slow left stimulation was not administered. Trials of slow TMS over the left DLPFC have also shown promise. Padberg and colleagues (120) randomized 18 nonpsychotic patients to treatment over the left DLPFC using fast, slow, or sham treatments. They showed a statistically significant (but clinically insignificant) response to fast and slow stimulation compared to placebo. In two small studies, George and Nahas demonstrated that lower frequency stimulation (5 Hz) might be more effective at 20 Hz TMS. Patients who were administered 5 Hz rTMS over the left DLPFC had a higher response rate (50% decline in the HDRS) than those administered 5 Hz rTMS over the right DLPFC (6/10 versus 3/10 responders, respectively) and placebo (0/10 responders) (121 ,122). Slow TMS (<1 Hz) has never been associated with seizures or any other adverse consequences in neurologically normal individuals; therefore, it is potentially safer than fast stimulations (114).

There are two basic clinical problems associated with rTMS. First, although half the patients respond to the treatments (defined as a 50% improvement in the HDRS), a much smaller percentage obtains remission (HDRS ≤10). Second, and perhaps of more concern, is the fact that a majority of patients (up to 100% in some studies) relapse in the month after treatment. Both the antidepressant response (M. Szuba, personal communication) and relapse rate (109) are improved by increasing the number of weeks of treatments but, as with ECT, the most severely ill patients may respond partially and relapse quickly. Combination treatment strategies (slow right and fast left) and maintenance rTMS strategies are being employed to improve response and keep patients in remission.

Beyond simply stimulation of the right or left cortex, techniques are being developed to assist in focusing the magnetic impulse on specific cortical structures. The development of new more focal coils and neuroimaging techniques to guide the placement of the rTMS stimulation (123) hold promise on targeting specific brain regions.

Using these techniques, rTMS can help in elucidating the neuronal pathways involved in depression. Initial studies using functional neuroimaging and rTMS have shown that many of the effects of rTMS occur at brain regions distant from the site of stimulation including the caudate, orbitofrontal cortex bilaterally, and cerebellum (124). These studies have also called into question the hypothesis that fast rTMS increases neuronal excitability (125 ,126) and slow rTMS inhibits neuronal firing (127 ,128).

However, it may be possible to determine if rTMS is effective in given individuals by evaluating the functional neuroimaging before and after rTMS stimulation. Speer and associates have shown that patients with hypometabolism on baseline PET scans utilizing [^{18}F]-Fluorodeoxyglucose in the left prefrontal cortex responded preferentially to 2 weeks of fast rTMS (20 Hz) as compared with patients with hypermetabolism who responded at a higher rate to slow (1Hz) rTMS (129). Individual characteristics may be a key factor in determining treatment response.

Vagal nerve stimulation (VNS) is another new and promising treatment for major depression. VNS has been successfully used to treat patients with intractable seizures since 1994 and was noticed to have positive effects on mood that were not simply secondary to the decrease in seizure frequency (130). The VN has both parasympathetic efferent fibers to the heart and GI tract and sensory afferent fibers (approximately 20% and 80% of the fibers, respectively). The afferent fibers connect the nucleus tractus solitarius to the orbital and lateral frontal cortices, and parabrachial nucleus (PN) and locus ceruleus (LC). The fibers passing through the PN/LC (which are adjacent to one another) connect to the hypothalamus and thalamus and, central to the antidepressant properties of VNS, the bed nucleus of the stria terminalis and amygdala. George and associates (131) detail the rationale for the use of VNS in treatment-resistant depression, including these VN afferent connections to the limbic system as well as PET data showing activation of limbic structures (132) and increases in CNS monoamines with VNS (133 ,134 and 135).

The first open trial of VNS by Rush and colleagues included 30 nonpsychotic patients with treatment-resistant unipolar or bipolar depression (136). As in the treatment of epilepsy, the stimulator was attached to the left vagal nerve, which can be accessed peripherally in a procedure similar to implanting a cardiac pacemaker, and compared to the right VN has fewer afferent fibers to the autonomic system controlling cardiac and gastric physiologic functions (137). Twelve of 30 patients (40%) met criteria for treatment response (≥50% decrease in HDRS) and 5 patients had a final HDRS ≤10. None of the patients discontinued treatment because of adverse events. The most common adverse events (which were similar to the AE's in the epilepsy trials) were hoarseness (60%) and throat pain (27%). Ten percent of the patients had abnormal wound healing.

VNS may be an effective treatment in resistant depression. The two variables that predicted clinical response to VNS included previous response to ECT (only one of 19 patients who had received ECT had a sustained response to VNS) and decreased stimulator output. Encouragingly, of the ten responders with available follow-up data over 4 to 9 months, all have demonstrated continued response.

VNS has been shown to be safe in the treatment of over 6,000 patients with epilepsy and may also provide relief to the 10% to 20% of depressed patients who fail or have an inadequate response to available somatic treatments.

The adverse side effect profile and costs could be dramatically reduced if a method of stimulating the VN could be achieved without surgery. The cost of the device and electrodes is $9,200, and the additional cost of the surgical procedure raises the total costs to approximately $12,000 to $25,000 (131). These costs are comparable to what might be expected for acute and 1-year maintenance treatment for ECT; however, insurance coverage may depend on the results of the pending placebo-controlled study and FDA approval for the treatment of depression. VNS is approved for the treatment of epilepsy.

Some researchers have questioned the benefit of electrical stimulation, which does not produce a seizure (138), arguing that subconvulsive seizures in ECT have not been shown to provide any clinical benefit (139). However, the anticonvulsant hypothesis assumes that the beneficial effects from ECT derive not from the convulsion, but the anticonvulsant effects of the seizure. VNS has documented anticonvulsant effects. Slow rTMS dampens neuronal excitability (100) and theoretically may be useful in treating epilepsy (140). There is also research underway at Columbia University using fast rTMS to precipitate a convulsion in animal models to determine the efficacy of TMS seizures in depression. Further investigations using animal models could potentially cause focal convulsions of specific brain regions (e.g., limbic structures) and spare brain sensitive structures that cause side effects but are not integral to the therapeutic effect (e.g., hippocampus). Together rTMS and VNS have the potential of providing valuable insight into the pathophysiology of depression and may potentially add to the treatment of resistant depression.

ACKNOWLEDGMENT

Part of "76 - Electroconvulsive Therapy: Sixty Years of Progress and a Comparison with Transcranial Magnetic Stimulation and Vagal Nerve Stimulation "

Dr. McDonald's work was supported in part by grants from grant MH56617-03 from the National Institute of Mental Health, the National Alliance for Research in Schizophrenia and Depression, and the J.B. Fuqua Foundation.

REFERENCES

1. Endler NS. The origins of electroconvulsive therapy (ECT). *Convulsiv Ther* 1988;4:5–23.

2. Freeman W, Watts JW. *Psychosurgery in the treatment of mental disorders and intractable pain* Springfield, IL: Charles C Thomas, 1952.

3. Sakel M. *Pharmacologic treatment of schizophrenia* New York: Nervous and Mental Disease Publishing, 1938.

4. Impastato D. Prevention of fatalities in electroshock therapy. *Dis Nerv Syst* 1957;18(Suppl):34–75.

5. Bennett AE. Curare: a preventive of traumatic complications in convulsive shock therapy. *Am J Psychiatry* 1941;97:1040–1060.

6. Holmberg G, Thesleff S. Succinyl-choline-iodide as a muscular relaxant in electroshock therapy. *Am J Psychiatry* 1952;108:842–846.

7. Blachy P, Gowing D. Multiple monitored electroconvulsive treatment. *Compr Psychiatry* 1966;7:100–109.

8. Figiel GS, McDonald WM, McCall WV, et al. Electroconvulsive therapy. In: Schatzberg AF, Nemeroff CB, eds. *Textbook of psychopharmacology*. Washington, DC: American Psychiatric Press, 1995:523–543.

9. Abrams R. The mortality rate with ECT. *Convuls Ther* 1997;13(3):125–127.

10. Salzman C. ECT, research, and professional ambivalence [editorial] [see comments]. *Am J Psychiatry* 1998;155(1):1–2.

11. Rudorfer MV, Goodwin FK. Introduction. In: Coffey CE, ed. *The clinical science of electroconvulsive therapy*, vol. 38. Washington, DC: American Psychiatric Press, 1993:xvii–xxi.

12. Ottosson J-O. Experimental studies of the mode of action of electroconvulsive therapy. *Acta Psychiatr Neurol Scand* 1960;35:1–141.

13. Fink M, Johnson L. Monitoring the duration of electroconvulsive therapy seizures: 'cuff' and EEG methods compared. *Arch Gen Psychiatry* 1982;39(10):1189–1191.

14. Maletzky BM. Seizure duration and clinical effect in electroconvulsive therapy. *Comp Psychiatry* 1978;19(6):541–550.

15. Abrams R, Taylor MA. Diencephalic stimulation and the effects of ECT in endogenous depression. *Br J Psychiatry* 1976;129:482–485.

16. Fink M, Ottosson JO. A theory of convulsive therapy in endogenous depression: significance of hypothalamic functions. *Psychiatry Res* 1980;2(1):49–61.

17. Mathe AA. Neuropeptides and electroconvulsive treatment. *J Ect* 1999;15(1):60–75.

18. Sackeim HA, Decina P, Kanzler M, et al. Effects of electrode placement on the efficacy of titrated, low-dose ECT. *Am J Psychiatry* 1987;144(11):1449–1455.

19. Devanand DP, Lisanby S, Lo ES, et al. Effects of electroconvulsive therapy on plasma vasopressin and oxytocin. *Biol Psychiatry* 1998;44(7):610–616.

20. Lisanby SH, Devanand DP, Prudic J, et al. Prolactin response to electroconvulsive therapy: effects of electrode placement and stimulus dosage [see comments]. *Biol Psychiatry* 1998;43(2):146–155.

21. Krystal AD, Weiner RD, McCall WV, et al. The effects of ECT stimulus dose and electrode placement on the ictal electroencephalogram: an intraindividual crossover study. *Biol Psychiatry* 1993;34(11):759–767.

22. Nobler MS, Sackeim HA, Solomou M, et al. EEG manifestations during ECT: effects of electrode placement and stimulus intensity. *Biol Psychiatry* 1993;34(5):321–330.

23. Weiner RD, Krystal AD. EEG monitoring of ECT seizures. In: Coffey CE, ed. *The clinical science of electroconvulsive therapy*, vol. 38. Washington, DC: American Psychiatric Press, 1993:93–109.

24. Krystal AD. The clinical utility of ictal EEG seizure adequacy models. *Psychiatric Ann* 1998;28(1):30–35.

25. Krystal AD, Zaidman C, Greenside HS, et al. The largest Lyapunov exponent of the EEG during ECT seizures as a measure of ECT seizure adequacy. *Electroencephalogr Clin Neurophysiol* 1997;103(6):599–606.

26. Krystal AD, Coffey CE, Weiner RD, et al. Changes in seizure threshold over the course of electroconvulsive therapy affect therapeutic response and are detected by ictal EEG ratings. *J Neuropsychiatr Clin Neurosci* 1998;10(2):178–186.

27. Krystal AD, Weiner RD, Coffey CE. The ictal EEG as a marker of adequate stimulus intensity with unilateral ECT. *J Neuropsychiatr Clin Neurosci* 1995;7(3):295–303.

28. McCall WV, Farah A, Reboussin D. Can we teach psychiatric residents to rate seizure regularity? *Convuls Ther* 1995;4:248–252.

29. Krystal AD, Weiner RD. ECT seizure therapeutic adequacy. *Convuls Ther* 1994;10(2):153–164.

30. Ingvar M. Cerebral blood flow and metabolic rate during seizures. Relationship to epileptic brain damage. *Ann NY Acad Sci* 1986;462:194–206.

31. Silfverskiold P, Rosen I, Risberg J. Effects of electroconvulsive therapy on EEG and cerebral blood flow in depression. *Eur Arch Psychiatry Neurol Sci* 1987;236(4):202–208.

32. Saito S, Yoshikawa D, Nishihara F, et al. The cerebral hemodynamic response to electrically induced seizures in man. *Brain Res* 1995;673(1):93–100.

33. Ackermann RF, Engel J Jr, Baxter L. Positron emission tomography and autoradiographic studies of glucose utilization following electroconvulsive seizures in humans and rats. *Ann NY Acad Sci* 1986;462:263–239.

34. Nobler MS, Sackeim HA, Prohovnik I, et al. Regional cerebral blood flow in mood disorders, III. Treatment and clinical response. *Arch Gen Psychiatry* 1994;51(11):884–897.

35. Nobler MS, Sackeim HA. Mechanisms of action of electroconvulsive therapy. *Psychiatr Ann* 1998;28(1):23–29.

36. Fink M, Kahn RL. Quantitative studies of slow wave activity following electroshock. *Electroenceph Clin Neurophysiol* 1956;8:158.

37. Fink M, Kahn RL. Relationship of electroencephalographic delta activity to behavioral response in electroshock. *Arch Neurol Psychiatry* 1959;78:516–525.

38. Sackeim HA, Luber B, Katzman GP, et al. The effects of electroconvulsive therapy on quantitative electroencephalograms. Relationship to clinical outcome [see comments]. *Arch Gen Psychiatry* 1996;53(9):814–824.

39. Wilson WP, Schieve JP, Scheinberg P. Effect of series of electroshock treatments on cerebral blood flow and metabolism. *Arch Neurol Psychiatry* 1952;68:651–654.

40. Sackeim HA. The anticonvulsant hypothesis of the mechanisms of action of ECT: current status. *J Ect* 1999;15(1):5–26.

41. Sackeim H, Decina P, Prohovnik I, et al. Seizure threshold in electroconvulsive therapy. Effects of sex, age, electrode placement, and number of treatments. *Arch Gen Psychiatry* 1987;44(4):355–360.

42. Shapira B, Lidsky D, Gorfine M, et al. Electroconvulsive therapy and resistant depression: clinical implications of seizure threshold [see comments]. *J Clin Psychiatry* 1996;57(1):32–38.

43. Sackeim HA, Decina P, Prohovnik I, et al. Dosage, seizure threshold, and the antidepressant efficacy of electroconvulsive therapy. *Ann NY Acad Sci* 1986;462:398–410.

44. Krystal AD, Weiner RD, Coffey CE, et al. Effect of ECT treatment number on the ictal EEG. *Psychiatry Res* 1996;62(2):179–189.

45. McCall WV, Robinette GD, Hardesty D. Relationship of seizure morphology to the convulsive threshold. *Convuls Ther* 1996;12(3):147–151.

46. Suppes T, Webb A, Carmody T, et al. Is postictal electrical silence a predictor of response to electroconvulsive therapy? *J Affect Disoord* 1996;41(1):55–58.

47. McElroy PEKaSL. Antiepileptic drugs. In: Nemeroff AFSaCB, ed. *Textbook of psychopharmacology*, second ed. Washington, DC American Psychological Press, 1998:431–454.

48. Bowden CL, Calabrese JR, McElroy SL, et al. The efficacy of lamotrigine in rapid cycling and non-rapid cycling patients with bipolar disorder [see comments]. *Biol Psychiatry* 1999;45(8):953–958.

49. Tortella FC, Long JB. Endogenous anticonvulsant substance in rat cerebrospinal fluid after a generalized seizure. *Science* 1985;228(4703):1106–1108.

50. Post RM. ECT: the anticonvulsant connection [see comments]. *Neuropsychopharmacology* 1990;3(2):89–92; discussion 97–99.

51. Abrams R. ECT technique: electrode placement, stimulus type, and treatment frequency. In: Coffey CE, ed. *The clinical science of electroconvulsive therapy*, vol. 38. Washington, DC: American Psychiatric Press, 1993:17–28.

52. Prudic J, Haskett RF, Mulsant B, et al. Resistance to antidepressant medications and short-term clinical response to ECT [see comments]. *Am J Psychiatry* 1996;153(8):985–992.

53. Sackeim HA, Prudic J, Devanand DP, et al. Effects of stimulus intensity and electrode placement on the efficacy and cognitive effects of electroconvulsive therapy [see comments]. *N Engl J Med* 1993;328(12):839–846.

54. Sackeim HA, Prudic J, Devanand DP, et al. A prospective, randomized, double-blind comparison of bilateral and right unilateral electroconvulsive therapy at different stimulus intensities [see comments]. *Arch Gen Psychiatry* 2000;57(5):425–434.

55. McCall WV, Reboussin DM, Weiner RD, et al. Titrated moderately suprathreshold vs fixed high-dose right unilateral electroconvulsive therapy: acute antidepressant and cognitive effects [see comments]. *Arch Gen Psychiatry* 2000;57(5):438–444.

56. McCall WV, Cohen W, Reboussin B, et al. Effect of mood and age on quality of life in depressed inpatients. *J Affect Disord* 1999;55:107–114.

57. McCall WV, Reboussin BA, Cohen W, et al. Electroconvulsive therapy is associated with superior symptomatic and functional change in depressed patients after psychiatric hospitalization. *J Affect Disord* 2001;63:17–25.

58. Abrams R, Swartz CM, Vedak C. Antidepressant effects of high-dose right unilateral electroconvulsive therapy. *Arch Gen Psychiatry* 1991;48(8):746–748.

59. Abrams R. Electroconvulsive therapy requires higher dosage levels: Food and Drug Administration action is required [comment]. *Arch Gen Psychiatry* 2000;57(5):445–446.

60. Roose SP, Glassman AH, Attia E, et al. Comparative efficacy of selective serotonin reuptake inhibitors and tricyclics in the treatment of melancholia. *Am J Psychiatry* 1994;151(12):1735–1739.

61. Spiker DG, Stein J, Rich CL. Delusional depression and electroconvulsive therapy: one year later. *Convuls Ther* 1985;1:167–172.

62. Aronson TA, Shukla S, Hoff A. Continuation therapy after ECT for delusional depression: a naturalistic study of prophylactic treatments and relapse. *Convuls Ther* 1987;3(4):251–259.

63. Sackeim HA, Prudic J, Devanand DP, et al. The impact of medication resistance and continuation pharmacotherapy on relapse following response to electroconvulsive therapy in major depression. *J Clin Psychopharmacol* 1990;10(2):96–104.

64. Association AP. *American Psychiatric Association Task Force on Electroconvulsive Therapy. The practice of electroconvulsive therapy: recommendations for treatment, training and privileging*. Washington, DC: American Psychiatric Press, 1978.

65. Levy SD, Albrecht E. Electroconvulsive therapy: a survey of use in the private psychiatric hospital. *J Clin Psychiatry* 1985;46:125–127.

66. Thienhaus OJ, Margletta S, Bennet JA. A study of the clinical efficacy of maintenance ECT. *J Clin Psychiatry* 1990;51:141–144.

67. Clarke TB, Coffey CE, Hoffman GW, et al. Continuation therapy for depression using outpatient electroconvulsive therapy. *Convuls Ther* 1989;5(4):330–337.

68. Decina P, Guthrie EB, Sackeim HA, et al. Continuation ECT in the management of relapses of major affective episodes. *Acta Psychiatr Scand* 1987;75(6):559–562.

69. Thornton JE, Mulsant BH, Dealy R, et al. A retrospective study of maintenance electroconvulsive therapy in a university-based psychiatric practice. *Convuls Ther* 1990;2:121–129.

70. Matzen TA, Martin RL, Watt TJ, et al. The use of maintenance electroconvulsive therapy for relapsing depression. *Jefferson J Psychiatry* 1988;6:52–58.

71. Grunhaus L, Pande A, Haskett R. Full and abbreviated courses of electroconvulsive therapy. *Convuls Ther* 1990;6:130–138.

72. Jaffe R, Dubin W, Shoyer B, et al. Outpatient electroconvulsive therapy: efficacy and safety. *Convuls Ther* 1990;6:231–238.

73. Loo H, Galinowski A, De Carvalho W, et al. Use of maintenance ECT for elderly depressed patients [letter; comment]. *Am J Psychiatry* 1991;148(6):810.

74. Dubin WR, Jaffe R, Roemer R, et al. The efficacy and safety of maintenance ECT in geriatric patients. *J Am Ger Soc* 1992;40(7):706–709.

75. Husain MM, Meyer DE, Muttakin MH, et al. Maintenance ECT for treatment of recurrent mania [letter]. *Am J Psychiatry* 1993;150(6):985.

76. Bonds C, Frye MA, Coudreaut MF, et al. Cost reduction and maintenance ECT in refractory bipolar disorder. *J Ect* 1998;14:36–41.

77. Chanpattana W. Continuation electroconvulsive therapy in schizophrenia: a pilot study. *J Med Assoc Thailand* 1997;80(5):311–318.

78. Bourne H, Long MB. Convulsion dependence. *Lancet* 1954;2:1193–1196.

79. Monroe RRJ. Mantenance electroconvulsive therapy. *Psychiatr Clin NA* 1991;14:947–960.

80. Weiner RD. Does electroconvulsive therapy cause brain damage? *Behav Brain Sci* 1984;7:1–53.

81. Devanand DP, Dwork AJ, Hutchinson ER, et al. Does ECT alter brain structure? [see comments]. *Am J Psychiatry* 1994;151(7):957–970.

82. Coffey CE, Weiner RD, Djang WT, et al. Brain anatomic effects of electroconvulsive therapy. A prospective magnetic resonance imaging study. *Arch Gen Psychiatry* 1991;48(11):1013–1121.

83. Levin Y, Elizur A, Korczyn AD. Physostigmine improves ECT-induced memory disturbances. *Neurology* 1987;37(5):871–875.

84. Prudic J, Fitzsimons L, Nobler MS, et al. Naloxone in the prevention of the adverse cognitive effects of ECT: a within-subject, placebo controlled study. *Neuropsychopharmacology* 1999;21(2):285–293.

85. Stern RA, Nevels CT, Shelhorse ME, et al. Antidepressant and memory effects of combined thyroid hormone treatment and electroconvulsive therapy: preliminary findings. *Biol Psychiatry* 1991;30(6):623–627.

86. Lerer B, Zabow T, Egnal N, et al. Effect of vasopressin on memory following electroconvulsive therapy. *Biol Psychiatry* 1983;18(7):821–824.

87. Mattes JA, Pettinati HM, Stephens S, et al. A placebo-controlled evaluation of vasopressin for ECT-induced memory impairment. *Biol Psychiatry* 1990;27(3):289–303.

88. d'Elia G, Frederiksen SO. ACTH4-10 and memory in ECT-treated and untreated patients. I. Effect on consolidation. *Acta Psychiatr Scand* 1980;62(5):418–428.

89. Small JG, Small IF, Milstein V, et al. Effects of ACTH 4-10 on ECT-induced memory dysfunctions. *Acta Psychiatr Scand* 1977;55(4):241–250.

90. Horne RL, Pettinati HM, Menken M, et al. Dexamethasone in electroconvulsive therapy: efficacy for depression and post-ECT amnesia. *Biol Psychiatry* 1984;19(1):13–27.

91. Cohen MR, Swartz CM. Absence of nimodipine premedication effect on memory after electroconvulsive therapy. *Neuropsychobiology* 1990;24(4):165–168.

92. McAllister DA, Perri MG, Jordan RC, et al. Effects of ECT given two vs. three times weekly. *Psychiatry Res* 1987;21(1):63–69.

93. Shapira B, Tubi N, Lerer B. Balancing speed of response to ECT in major depression and adverse cognitive effects: role of treatment schedule. *J Ect* 2000;16(2):97–109.

94. Bailine SH, Rifkin A, Kayne E, et al. Comparison of bifrontal and bitemporal ECT for major depression. *Am J Psychiatry* 2000;157(1):121–123.

95. Swartz CM. Asymmetric bilateral right frontotemporal left frontal stimulus electrode placement for electroconvulsive therapy. *Neuropsychobiology* 1994;29(4):174–178.

96. Barker AT, Jalinous R, Freeston IL. Non-invasive magnetic stimulation of human motor cortex [letter]. *Lancet* 1985;1(8437):1106–1107.

97. Pascual-Leone A, Valls-Sole J, Wassermann EM, et al. Responses to rapid-rate transcranial magnetic stimulation of the human motor cortex. *Brain* 1994;117(Pt 4):847–858.

98. George MS, Wassermann EM, Post RM. Transcranial magnetic stimulation: a neuropsychiatric tool for the 21st century. *J Neuropsychiatr Clin Neurosci* 1996;8(4):373–382.

99. Wassermann EM, Grafman J, Berry C, et al. Use and safety of a new repetitive transcranial magnetic stimulator. *Electroencephalogr Clin Neurophysiol* 1996;101(5):412–417.

100. Chen R, Classen J, Gerloff C, et al. Depression of motor cortex excitability by low-frequency transcranial magnetic stimulation. *Neurology* 1997;48(5):1398–403.

101. Fleischmann A, Sternheim A, Etgen AM, et al. Transcranial magnetic stimulation downregulates beta-adrenoreceptors in rat cortex. *J Neural Transm (Budapest)* 1996;103(11):1361–1366.

102. Fujiki M, Steward O. High frequency transcranial magnetic stimulation mimics the effects of ECS in upregulating astroglial gene expression in the murine CNS. *Brain Res Mol Brain Res* 1997;44(2):301–308.

103. George MS, Lisanby SH, Sackeim HA. Transcranial magnetic stimulation: applications in neuropsychiatry [comment]. *Arch Gen Psychiatry* 1999;56(4):300–311.

104. George MS, Ketter TA, Post RM. Prefrontal cortex dysfunction in clinical depression. *Depression* 1994;2:59–72.

105. George MS, Wassermann EM, Williams WA, et al. Daily repetitive transcranial magnetic stimulation (rTMS) improves mood in depression. *Neuroreport* 1995;6(14):1853–1856.

106. Pascual-Leone A, Catala MD, Pascual-Leone Pascual A. Lateralized effect of rapid-rate transcranial magnetic stimulation of the prefrontal cortex on mood. *Neurology* 1996;46(2):499–502.

107. Loo C, Mitchell P, Sachdev P, et al. Double-blind controlled investigation of transcranial magnetic stimulation for the treatment of resistant major depression. *Am J Psychiatry* 1999;156(6):946–948.

108. Figiel GS, Epstein C, McDonald WM, et al. The use of rapid-rate transcranial magnetic stimulation (rTMS) in refractory depressed patients. *J Neuropsychiatr Clin Neurosci* 1998;10(1):20–25.

109. Greene Y, McDonald W, Epstein C, et al. An open trial of repetitive transcranial magnetic stimulation in treatment-resistant depression. 1998 Annual Meeting New Research Program and Abstracts, Washington, DC: American Psychiatric Association, 1998, p. 90.

110. George MS, Wassermann EM, Kimbrell TA, et al. Mood improvement following daily left prefrontal repetitive transcranial magnetic stimulation in patients with depression: a placebo-controlled crossover trial [see comments]. *Am J Psychiatry* 1997;154(12):1752–1756.

111. Berman RM, Narasimhan M, Sanacora G, et al. A randomized clinical trial of repetitive transcranial magnetic stimulation in the treatment of major depression. *Biol Psychiatry* 2000;47(4):332–337.

112. Grunhaus L, Dannon P, Schrieber S. Effects of transcranial magnetic stimulation on severe depression: similarities with ECT (abstract). *Biol Psychiatry* 1998;43(76s):254.

113. Pridmore S. Substitution of rapid transcranial magnetic stimulation treatments for electroconvulsive therapy treatments in a course of electroconvulsive therapy. *Depress Anxiety* 2000;12:118–123.

114. Wassermann EM. Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the International Workshop on the Safety of Repetitive Transcranial Magnetic Stimulation, June 5–7, 1996. *Electroencephalogr Clin Neurophysiol* 1998;108(1):1–16.

115. Hoflich G, Kasper S, Hufnagel A, et al. Application of transcranial magnetic stimulation in the treatment of drug-resistant major depression. *Hum Psychopharmacol* 1993;8:361–365.

116. Grisaru N, Yarovslavsky U, Arbarbanel J, et al. Transcranial magnetic stimulation in depression and schizophrenia. *Eur Neuropsychopharmacol* 1994;4:287–288.

117. Kolbinger HM, Hoflich G, Hufnagel A, et al. Transcranial magnetic stimulation (TMS) in the treatment of major depression: a pilot study. *Hum Psychopharmacol* 1995;10:305–310.

118. Klein E, Kreinin I, Chistyakov A, et al. Therapeutic efficacy of right prefrontal slow repetitive transcranial magnetic stimulation in major depression: a double-blind controlled study [see comments]. *Arch Gen Psychiatry* 1999;56(4):315–20.

119. Tormos J, Pascual-Leone AP, Catala MD, et al. Antidepressant effects of repetitive transcranial magnetic stimulation to dorsolateral prefrontal cortex: (1) Hemispheric asymmetry. *Arch Gen Psychiatry* in press.

120. Padberg F, Zwanzger P, Thoma H, et al. Repetitive transcranial magnetic stimulation (rTMS) in pharmacotherapy-refractory major depression: comparative study of fast, slow and sham rTMS. *Psychiatry Res* 1999;88(3):163–171.

121. George MS, Speer AM, Molloy M, et al. Low-frequency daily left prefrontal rTMS improves mood in bipolar depression: a placebo-controlled case report. *Hum Psychopharmacol* 1998;13:271–275.

122. Nahas Z, Speer AM, Molloy M, et al. Preliminary results concerning the roles of frequency and intensity in the antidepressant effect of daily left prefrontal rTMS [abstract]. *Biol Psychiatry* 1998;43:94.

123. Bohning DE, Percheny AP, Epstein CM, et al. Mapping transcranial magnetic stimulation (TMS) fields in vivo with MRI. *Neuroreport* 1997;8:2535–2538.

124. Kimbrell TA, et al. Changes in cerebral metabolism during transcranial magnetic stimulation (abstract). *Biol Psychiatry* 1997;41:1085.

125. Paus T, Jech R, Thompson CJ, et al. Dose-dependent reduction of cerebral blood flow during rapid-rate transcranial magnetic stimulation of the human sensorimotor cortex. *J Neurophysiol* 1997;79:1102–1107.

126. Stallings LE, Speer AM, Spicer KM, et al. Combining SPECT and repetitive transcranial magnetic stimulation (rTMS): left prefrontal stimulation decreases relative perfusion locally in a dose-dependent manner (abstract). *Neuroimage* 1997;5:S521.

127. Fox P, Ingham R, George MS, et al. Imaging human intra-cerebral connectivity by PET during TMS. *Neuroreport* 1997;8(12):2787–2791.

128. Bohning DE, Shastri A, McConnell KA, et al. A combined TMS/fMRI study of intensity-dependent TMS over motor cortex. *Biol Psychiatry* 1999;45(4):385–394.

129. Kimbrell TA, Little JT, Dunn RT, et al. Frequency dependence of antidepressant response to left prefrontal repetitive transcranial magnetic stimulation (rTMS) as a function of baseline cerebral glucose metabolism. *Biol Psychiatry* 1999;46(12):1603–1613.

130. Harden CL, Pulver MC, Nikolov B, et al. Effect of vagus nerve stimulation on mood in adult epilepsy patients. *Neurology* 1999;52(Suppl 2):A238–P03122.

131. George MS, Sackeim HA, Rush AJ, et al. Vagus nerve stimulation: a new tool for brain research and therapy [see comments]. *Biol Psychiatry* 2000;47(4):287–295.

132. Henry TR, Bakay RA, Votaw JR, et al. Brain blood flow alterations induced by therapeutic vagus nerve stimulation in partial epilepsy. I. Acute effects at high and low levels of stimulation. *Epilepsia* 1998;39(9):983–990.

133. Ben-Menachem E, Hamberger A, Hedner T, et al. Effects of vagus nerve stimulation on amino acids and other metabolites in the CSF of patients with partial seizures. *Epilepsy Res* 1995;20(3):221–227.

134. Krahl SE, Clark KB, Smith DC, et al. Locus coeruleus lesions suppress the seizure-attenuating effects of vagus nerve stimulation. *Epilepsia* 1998;39(7):709–714.

135. Walker BR, Easton A, Gale K. Regulation of limbic motor seizures by GABA and glutamate transmission in nucleus tractus solitarius. *Epilepsia* 1999;40(8):1051–1057.

136. Rush AJ, George MS, Sackeim HA, et al. Vagus nerve stimulation (VNS) for treatment-resistant depression: a multicenter study [see comments]. *Biol Psychiatry* 2000;47(4):276–286.

137. Schachter SC, Saper CB. Vagus nerve stimulation. *Epilepsia* 1998;39(7):677–686.

138. Swartz CM. Subconvulsive magnetic brain stimulation no replacement for ECT [letter]. *Am J Psychiatry* 1997;154(5):716–717.

139. Fink M, Kahn RI, Green MA. Experimental studies in the electroshock process. *Disord Nerv Syst* 1958;19:113–118.

140. Epstein CM. TMS in epilepsy. In: George MS, Belmaker RH, eds. *Transcranial magnetic stimulation in psychiatry*. Washington, DC: American Psychiatric Press, 2000:173–184.

77

Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder

Paul E. Keck Jr.

Husseini K. Manji

Paul E. Keck, Jr.: Department of Psychiatry, University of Cincinnati College of Medicine, Cincinnati, Ohio.

Husseini K. Manji: Laboratory of Molecular Pathophysiology, National Institute of Mental Health, Bethesda, Maryland.

The range of available medications for the acute treatment of bipolar mania and maintenance treatment of bipolar disorder (BD) has expanded rapidly in recent years. Data regarding medications with established antimanic efficacy are growing and a number of new agents with putative mood-stabilizing properties are under study. These developments are fortunate because recent studies also indicate that the long-term outcome of many patients with BD remains poor (43 ,45 ,100).

Data from randomized, controlled clinical trials supporting the efficacy of lithium, valproate (VPA), carbamazepine (CBZ), and typical antipsychotics as antimanic and mood-stabilizing agents are reviewed in the following. Studies of two other important drug classes under active study for patients with BD, atypical antipsychotics and novel antiepileptics, are also reviewed (62). Finally, the development of signal transduction modifiers and regulators of neuroplasticity and cellular resiliency as truly novel agents for the treatment of BD is discussed.

LITHIUM

Part of "77 - Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder "

Acute Mania

Lithium was superior to placebo in the treatment of acute mania in five controlled studies (8 ,31 ,55 ,85 ,93). Only one of these studies utilized a parallel design (8). The remaining four studies were crossover trials of varying duration (31 ,55 ,85 ,93). These crossover studies may have been vulnerable to carry-over, period, and abrupt treatment discontinuation effects, which may have deflated placebo response rates, and contamination of the study blind (16 ,49). In addition, two studies utilized nonrandom assignment (31 ,93). Finally, many of the early landmark lithium studies used diagnostic criteria to define BD that may not be comparable to those of DSM-III-R1 or DSM-IV (2 ,10 ,62). Although Bowden and associates used lithium as an active control (8), data from this parallel-design study are the most methodologically rigorous. In this study, 17 (49%) of 35 lithium-treated patients displayed more than 50% reduction in manic symptoms as measured by the Mania Rating Scale (MRS) total score from the Schedule for Affective Disorders and Schizophrenia (SADS-C) compared with 24% of placebo-treated patients at 3 weeks. The lithium-placebo effect size was a moderate 0.4 (48). Further analysis of data from this study revealed that mania characterized by predominantly elevated or elated mood was associated with lithium response, whereas depressive symptoms during mania and multiple prior affective episodes were associated with poor response to lithium (95 ,96). These findings are similar to those of earlier reports, which found that patients with mixed mania had a lower likelihood of lithium response compared with classic mania (64). In studies in which response of psychotic symptoms was also assessed, lithium also produced significant improvement in these symptoms (8 ,31 ,55 ,93).

Lithium has also been compared with standard antipsychotic agents in nine controlled trials in patients with acute mania (30 ,40 ,75 ,78 ,86 ,87 ,90 ,97). Of these studies, three found lithium to be comparable to chlorpromazine (40 ,90) or haloperidol (86) over treatment intervals ranging from 1 to 3 weeks; four found lithium superior to chlorpromazine (40 ,75 ,87 ,97) or haloperidol (86) over 1 to 5 weeks; and one study found haloperidol plus placebo and haloperidol plus lithium superior to lithium plus placebo after 1 and 2 weeks (30). In the largest and most rigorous study comparing

lithium and a typical antipsychotic in acute bipolar mania, Prien and colleagues (78) assessed the efficacy of lithium versus chlorpromazine in 225 patients divided into “highly active” or “mildly active” groups. In the mildly active group, both medications produced significant and comparable improvement. However, chlorpromazine-treated patients experienced more frequent and severe side effects. In contrast, chlorpromazine produced more rapid reduction in measures of agitation, grandiosity, hostility, and psychotic disorganization than lithium in the highly active group over the first week of treatment. In this latter group, dropouts were higher for lithium (38%) than chlorpromazine (8%). At the end of 3 weeks of treatment, both drugs were significantly and comparably effective.

In summary, the studies reviewed indicate that lithium is superior in efficacy to placebo and comparable or possibly superior to standard antipsychotics for the normalization of affective symptoms in mania. These studies also indicate that lithium exerts antipsychotic effects in patients with psychotic mania. On the other hand, standard antipsychotics appear to have a more rapid onset of action and, therefore, may be more effective initially, especially in severely manic or agitated patients.

Maintenance

The prevention of affective episodes is the essential goal of the maintenance treatment of patients with BD. Double-blind, placebo-controlled studies conducted during the late 1960s and 1970s demonstrated that lithium was superior to placebo in preventing recurrent affective episodes (3 ,19 ,21 ,25 ,27 ,38 ,41 ,67 ,79 ,83 ,92). A recent analysis of six parallel group, placebo-controlled lithium maintenance treatment trials in patients with bipolar I disorder found an odds ratio favoring relapse in patients receiving placebo of 4.1 (95% CI: 2.1, 7.7) at both 6 and 12 months of treatment (49). However, more recent data from five naturalistic long-term studies of patients maintained on lithium for greater than 1 year suggest that a substantial number of patients do not respond adequately to lithium prophylaxis (36 ,56 ,69 ,73 ,81). Baldessarini and Tondo recently reviewed 24 reports on long-term lithium treatment published between 1970 and 1996 and also analyzed its clinical efficacy in 360 patients at a single center who began lithium maintenance monotherapy after 1970 (4). There were no significant differences in lithium response, defined either as percentage of patients displaying more than 50% improvement or percentage episode-free, among patients treated in the 1970s, 1980s, or 1990s. The specific response rates are of clinical interest: Approximately 66% of patients displayed more than 50% improvement and only 33% were episode-free. Not surprisingly, patients with mixed episode, psychotic features or rapid cycling were less treatment responsive. Other studies also found that mixed mania or rapid cycling were associated with a poor lithium response (24 ,64).

VALPROATE

Part of "77 - Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder "

Acute Mania

The efficacy of VPA in the treatment of acute bipolar mania was established in six double-blind, randomized controlled trials (8 ,11 ,26 ,28 ,37 ,76). These studies include comparisons of VPA versus placebo in crossover trials (11 ,26) versus placebo in a parallel group trial in lithium-refractory or intolerant patients (76), versus placebo and lithium in a parallel group trial8 and against lithium in two parallel group trials (28 ,37), one of which also compared rapid loaded of divalproex (30 mg/kg per day) with gradual titration (37). Two of these trials led to approval by the United States Food and Drug Administration (USFDA) of divalproex for the treatment of manic episodes associated with BD (8 ,76). In these two studies, response to divalproex as defined by more than 50% reduction in total manic symptoms was comparable, with 53% (76) and 48% (8) of patients responding, respectively. Similarly, among responders in these two studies significant improvement was evident by day 7 and 10 of treatment, respectively, using a gradual divalproex titration schedule beginning with an initial dose of 750 mg per day. Divalproex exerted therapeutic effects on psychotic and manic symptoms in both studies. In the study by Bowden and colleagues (8), all patients with rapid cycling (n = 8) were randomly assigned to divalproex; four (50%) displayed at least 50% improvement on the MRS, which was comparable to the overall response of the divalproex-treated group. This finding is consistent with two other reports from open trials suggesting efficacy of divalproex in patients with rapid cycling (13 ,63). In a subsequent analysis of response according to several definitions of depressive (mixed) mania, Swann and associates found that the presence of even mild depressive symptoms was associated with a poor antimanic response to lithium, but had no significant effect on VPA response (95). Divalproex was well tolerated in this study and significantly more lithium-treated patients dropped out of this trial owing to side effects than patients receiving divalproex or placebo.

VPA and lithium were comparable in efficacy in two other head-to-head comparison trials, although the relatively small sample sizes in these trials makes them vulnerable to a type II error (28 ,37). As in the Bowden and colleagues study (8), Freeman and colleagues also found that the presence of depressive symptoms during mania was associated with a greater likelihood of VPA than lithium response (28). VPA was compared with haloperidol in an open randomized trial in patients with psychotic mania (65). In that study, 36 inpatients with bipolar I disorder in a manic or mixed episode with psychotic features received either divalproex rapid loading (20 mg/kg per day) or haloperidol (0.2 mg/kg per day) for 6 days. Divalproex and haloperidol were equally effective in reducing both manic and psychotic symptoms (8 ,76). The improvement in psychosis

with divalproex treatment was consistent with findings in two other studies and with improvement in manic psychosis produced by lithium in earlier studies (8 ,31 ,55 ,93). Several other studies investigated the efficacy and tolerability of divalproex rapid loading (20 to 30 mg/kg per day) (37 ,46 ,65) and one recent pilot 5-day trial explored intravenous VPA administration (1,200 to 1,800 mg per day) (35). Although all four studies reported rapid improvement (within 3 days) in manic symptoms among responders, only one of these studies was a double-blind, randomized trial (37). In this controlled trial, the study was designed to assess the tolerability of divalproex loading (30 mg/kg per day × 2 days, then 20 mg/kg per day) and was not powered sufficiently to detect differences in efficacy. VPA has also been compared against placebo as adjunctive therapy to standard antipsychotics in acute mania (68). In a multicenter, double-blind, parallel design, 3-week trial conducted in Europe, 136 patients receiving standard antipsychotics were randomized to VPA or placebo. By study termination, significantly more VPA-treated patients displayed a decrease in concomitant antipsychotic treatment. In summary, these studies suggest that VPA has a broad spectrum of efficacy in acute mania, mixed mania, and rapid cycling, and appears to be comparable to lithium and haloperidol in overall antimanic efficacy.

Maintenance

Open maintenance trials of VPA in patients with BD reported that approximately 45% to 50% of patients experienced a recurrent affective episode in follow-up periods ranging from 6 to 24 months (15 ,52 ,82). A randomized, open comparison of lithium and VPA found generally good efficacy for both drugs over an 18-month period (52). Bowden and associates recently reported the results of the largest, prospective, double-blind, randomized maintenance trial of pharmacological treatment in patients with BD using survival analysis to assess time to and rates of relapse (9). In this study, 38% of patients receiving placebo relapsed, compared with 31% on lithium and 24% on divalproex (differences not significant). It is instructive to compare the results of this study with those of the other large maintenance study that compared lithium to placebo (79). In this latter study, 68% of patients receiving placebo relapsed compared with 36% of patients on lithium by 1 year (79). Thus, the drug relapse rates were very similar between studies but the placebo relapse rate was much lower in the Bowden and co-worker study (9). This disparity in placebo relapse rates is probably owing to several factors. First, although both studies standardized enrollment by an index manic episode, it is likely that patients in the Bowden and associates study were less severely ill because only 18% had been hospitalized during the index episode, whereas all patients in the Prien and colleagues study had been hospitalized. Second, the definition of relapse differed between the studies. In the Prien and colleagues study (79), a relapse was an event that required hospitalization or supplemental drug treatment; in the Bowden and associates study (9), it was the occurrence of any affective episode. Finally, the Prien and colleagues study consisted of a more homogeneous sample of lithium responders diagnosed by more conservative diagnostic criteria (80).

As described, many patients with BD have symptomatic or syndromic recurrences on monotherapy with lithium or VPA. The addition of a second mood-stabilizer is a common strategy to enhance maintenance treatment efficacy. Unfortunately, only one controlled trial has examined this approach using the combination of lithium and divalproex in maintenance treatment (89). In this pilot study, 12 patients with bipolar I disorder receiving lithium plasma concentrations (0.8 to 1.0 mEq/L) were randomized to adjunctive maintenance treatment with divalproex (plasma concentrations 50 to 125 mg/L) or placebo and followed for 1 year. Patients who received the combination of lithium and divalproex were significantly less likely to experience a relapse but significantly more likely to suffer at least one moderate or severe adverse event. It is possible that adverse events might be reduced by using doses of lithium and/or divalproex at the lower end of the therapeutic range for each agent.

CARBAMAZEPINE

Part of "77 - Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder "

Acute Mania

Although 14 double-blind controlled studies have found CBZ to be effective in the treatment of acute mania, only five of these studies are not confounded by the use of concomitant agents with antimanic effects (reviewed in Keck and associates 1992) (44). In the only placebo-controlled trial, 19 patients were crossed over between CBZ or placebo (5). During CBZ treatment, 63% of patients displayed significant improvement on global nursing measures of mania, depression, anxiety, anger, and psychosis. Relapse typically occurred on placebo.

Two studies compared CBZ with lithium (53 ,88) In the first study, 34 inpatients were randomized to lithium or CBZ for up to 4 weeks (53). Twenty-eight patients (14 per treatment group) completed the study and were included in the data analysis. There were no significant differences in improvement between the two drugs on the BPRS and the Beigel-Murphy Manic State Rating Scale. However, lithium-treated patients showed significantly greater improvement in CGI change scores. In addition, only four (29%) of 14 CBZ-treated patients were considered responders, whereas 11 (79%) of 14 lithium-treated patients responded. In the second lithium comparison study (88) 70% of 52 hospitalized patients randomized to lithium or CBZ dropped out of the trial by 8 weeks owing to lack of efficacy.

Of the study completers, 36% who received CBZ and 37% who received lithium were rated as improved (defined as at least partial remission of manic symptoms). Two studies compared CBZ with chlorpromazine in the treatment of acute mania (34 ,71). In the first comparison trial, 60 acutely manic patients were randomized to either agent in a 6-week trial (71). There were no significant differences in efficacy between the two drugs, with moderate to marked improvement in 70% of the 32 patients treated with CBZ and 60% of 28 patients receiving chlorpromazine. In the second study (34 ,37) patients were randomized to CBZ (*n* = 15) or chlorpromazine (*n* = 19) in a 3-week trial. Response was assessed in 26 patients who completed the trial. Patients treated with CBZ (*n* = 15) or chlorpromazine (*n* = 11) had comparable improvement. In the CBZ group, 67% were rated as displaying at least moderate improvement and 59% for the chlorpromazine group. In summary, data from these five trials suggest that CBZ is superior to placebo and comparable to lithium and chlorpromazine in the treatment of acute bipolar mania.

Maintenance

The efficacy of CBZ in the prevention of recurrent affective episodes has been a matter of controversy (22 ,77). This controversy rests in part on the heterogeneity among the early controlled maintenance studies (22 ,54 ,74 ,101), and the availability of only one placebo-controlled maintenance trial (70). Interpretation of this latter study is also limited by the use of adjunctive rescue medications other than lithium and CBZ to treat breakthrough symptoms. The liberal use of these adjunctive treatments thus limits the degree to which relapse rate can be directly attributed to CBZ or placebo in this study.

Two recent large prospective, double-blind, long-term maintenance studies provide new data comparing the efficacy of CBZ with lithium (23 ,33). In the first study (33), 144 patients were randomized to lithium (*n* = 74; mean ±SD serum level, 0.6 ±1 mEq/L) or CBZ (*n* = 70; mean ±SD) dose, 621 ±186 mg per day) and followed for 2.5 years. Affective relapse, hospitalization, need for supplemental medication, and adverse events requiring treatment discontinuation were used to define treatment failure. Using survival analysis, there were no statistically significant differences between the two treatment groups in time to episode recurrence or hospitalization. However, significantly more patients receiving CBZ required supplemental medications for symptomatic recurrences and experienced adverse events requiring treatment discontinuation. In a secondary analysis of predictors or response, patients with classical features (bipolar I patients without mood-incongruent delusions and comorbidity) had a lower rehospitalization rate with lithium than with CBZ (32). For patients with nonclassical features (mixed states, bipolar II, and NOS) a trend in favor of CBZ was found.

In the second lithium comparison study, 52 outpatients with bipolar I and II disorder were randomly assigned for an initial year of treatment with lithium or CBZ, a crossover to the alternate drug in the second year, followed by a third year on the combination (23). Among evaluable patients, 13 (31%) of 42 lithium-treated patients relapsed within 1 year compared with 13 (37%) of 35 CBZ-treated patients. Seven (24%) of 29 remaining patients relapsed on combination therapy. As in the previous study, a higher percentage of patients receiving CBZ withdrew because of adverse events. The percentage of patients who had moderate or marked improvement on the CGI was not significantly different: 33% on lithium, 31% on CBZ, 55% on the combination; however, on a variety of measures of mania, lithium was more effective than CBZ. Finally, patients with a history of rapid cycling responded significantly better to the combination (56% response) compared with lithium (28%) or CBZ (19%).

STANDARD ANTIPSYCHOTICS

Part of "77 - Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder "

Acute Mania

There is one double-blind, placebo-controlled study of a standard antipsychotic in the treatment of acute bipolar mania (50). In that study, 13 patients were randomized to chlorpromazine (1,200 mg per day), imipramine (300 mg per day), or placebo for 7 weeks. Response was assessed using a global scale ranging from -9 to +9. Chlorpromazine was significantly superior to placebo and imipramine on global outcome (6.1 versus 2.0 and -2.8, respectively). Studies comparing standard antipsychotics with lithium, VPA, and CBZ are reviewed in the preceding.

Maintenance

There are no parallel group, double-blind randomized maintenance trials of standard antipsychotics in the maintenance treatment of patients with BD.

ATYPICAL ANTIPSYCHOTICS

Part of "77 - Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder "

Clozapine

Acute Mania

No double-blind, randomized controlled trials of clozapine in the treatment of acute bipolar mania have been published. Clozapine has been reported to be an effective antimanic agent in three open-label, prospective studies (6 ,14 ,18). The first two studies evaluated treatment-refractory bipolar patients. In the first (6), 13 (87%) of 15 acutely manic patients who had failed to respond to a minimum 6-week antipsychotic trial of 500 mg per day chlorpromazine-equivalents

in combination with lithium (mean serum level 0.8 mEq/L) were rated as moderate or marked responders to clozapine. In the second study (14), 25 acutely manic patients with bipolar or schizoaffective disorder, all of whom had failed to respond to or tolerate lithium, VPA, or CBZ, and at least two standard antipsychotics, received a 13-week trial of clozapine (mean ±SD, 194 ± 145 mg per day) following a 1-week washout. Eighteen (72%) patients displayed marked improvement on the YMRS and eight (32%) on the BPRS, defined as more than 50% reduction in total score on either scale. Bipolar patients compared with schizoaffective patients, and non-rapid cyclers compared with rapid cyclers, exhibited significantly greater improvement on the BPRS. In the third study (18), 30 hospitalized acutely manic patients were randomized to clozapine (mean 166 mg per day) or chlorpromazine (mean 310 mg per day) in a 3-week trial. Although clozapine-treated patients displayed significantly lower YMRS scores after 2 weeks, there were no significant differences between the two groups at the end of the trial.

Maintenance

There are no randomized double-blind controlled studies of clozapine in the maintenance treatment of BD. Suppes and co-workers randomized 85 patients with DSM-IV criteria for schizoaffective or BD with treatment-refractory illness to add-on clozapine or treatment as usual (94). After 1 year, patients receiving clozapine displayed significant reductions in measures of mania, psychosis, and global improvement compared with patients receiving treatment as usual. In addition, total medication use decreased significantly in the clozapine group. This open-label, randomized trial confirmed earlier reports suggesting that clozapine exerted long-term mood-stabilizing effects in patients with treatment-refractory BD (102).

Olanzapine

Acute Mania

The efficacy of olanzapine in the treatment of acute bipolar mania has been established in three double-blind, controlled trials (7 ,98 ,99). In the first of two placebo-controlled studies, 139 inpatients with bipolar I disorder were randomized to olanzapine (*n* = 70) or placebo (*n* = 69) for up to 3 weeks (99). Olanzapine was begun at 10 mg per day and adjusted by 5-mg per day increments within a range of 5 to 20 mg per day; the median modal dose was 15 mg per day. Concomitant lorazepam up to 4 mg per day was permitted for the first 7 days as needed for agitation; 2 mg per day was permitted for the subsequent 3 days. The olanzapine group displayed significant improvement on the YMRS, CGI-BP severity of mania, and the PANSS total and positive symptom scores compared with the placebo group. Olanzapine also produced a significantly higher response rate (defined as ≥50% improvement in YMRS score) of 49% compared with placebo at 24%. However, olanzapine did not separate from placebo on the YMRS until the third week of treatment.

In the second placebo-controlled trial, 115 bipolar I inpatients were randomized to olanzapine (*n* = 55) or placebo (*n* = 60) for up to 4 weeks (98). This study used a higher initial starting dose of olanzapine, 15 mg per day, and concomitant lorazepam use was further restricted to 2 mg per day for the first 10 days. The olanzapine group again displayed significant improvement on the YMRS, CGI-BP severity of mania, PANSS total and positive symptom scores compared with the placebo group. These differences were evident by week one (the time of the first rating) and sustained throughout the trial. Olanzapine-treated patients also exhibited significantly higher response (65% versus 43%, respectively) and remission (61% versus 36%, respectively) rates than placebo-treated patients. In addition, patients presenting with prominent depressive symptoms during mania (HamD-21 scores ≥ baseline) showed a significant reduction in HamD-21 total scores with olanzapine compared with placebo.

In the third controlled trial, 30 inpatients with acute mania were randomized to olanzapine (*n* = 15) or lithium (*n* = 15) for 4 weeks (7). Olanzapine was administered as a fixed dose of 10 mg per day and lithium at 400 mg BID (mean serum level 0.74 mEq/L). Concomitant lorazepam 4 to 12 mg per day was permitted throughout the 4-week trial. There were no significant differences between the two treatment groups on any of the primary outcome measures—the Mania Scale and BPRS total scores and the CGI improvement scale. However, olanzapine-treated patients displayed significantly greater improvement than lithium-treated patients on the CGI severity scale at the end of 4 weeks of treatment.

Maintenance

There are no controlled trials of olanzapine in the maintenance treatment of BD published to date. In an open-label 52-week extension trial following the initial placebo-controlled acute mania study, none of the 98 patients who participated developed tardive dyskinesia during long-term treatment (99).

Risperidone

Acute Mania

There are two double-blind randomized active comparator studies of risperidone in the treatment of acute bipolar mania (84 ,86). In the first study, 45 inpatients were randomized to risperidone 6 mg per day (*n* = 15), haloperidol 10 mg per day (*n* = 15), or lithium 800 to 1,200 mg per

day (with levels ranging from 0.6 to 1.2 mEq/L) (*n* = 15) for up to 4 weeks (86). There were no significant differences among the three treatment groups in reductions on the YMRS, BPRS, CGI, and GAF from baseline to endpoint (LOCF). In the second trial, 158 inpatients receiving lithium or VPA were randomized to adjunctive therapy with risperidone 1 to 6 mg per day (*n* = 52), haloperidol (*n* = 53), or placebo (*n* = 51) for up to 3 weeks (84). Patients receiving risperidone or haloperidol displayed significantly greater improvement at 1 week, 2 weeks, and at endpoint (LOCF), but not at 3 weeks of treatment on the YMRS compared with patients receiving placebo. There were no significant differences between the risperidone and placebo groups from baseline to endpoint in BPRS and HamD total scores. There are no controlled trials of risperidone in the maintenance treatment of patients with BD.

Ziprasidone

Acute Mania

Ziprasidone has been studied in placebo-controlled trials in the treatment of acute mania in patients with BD and schizoaffective disorder, bipolar type (42 ,47). In a 3-week acute treatment trial of inpatients with bipolar I disorder, manic or mixed, 199 patients were randomized to ziprasidone 80 to 160 mg per day or placebo (42). Ziprasidone produced significant reductions on the Manic Rating Scale (MRS) total score (from the SADS-C) at day two and throughout the remainder of the trial compared with placebo. Based on a ≥50% reduction in MRS total scores, significantly more ziprasidone-treated patients responded (50%) compared with placebo-treated patients (36%). This study confirmed the findings of an earlier study that found that ziprasidone produced significant reductions in manic symptoms compared with placebo in patients with schizoaffective disorder (47). There are no controlled maintenance trials yet published of ziprasidone in BD. Finally, there are no controlled trials of quetiapine in the acute or maintenance treatment of BD.

NEW ANTIEPILEPTICS

Part of "77 - Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder "

Four new antiepileptic agents, gabapentin, lamotrigine, topiramate, and tiagabine are being investigated as potential antimanic agents (62). To date, no controlled trials have been reported for topiramate or tiagabine. However, two controlled studies evaluated gabapentin in the treatment of bipolar mania (29 ,72). In the first study, 117 outpatients with bipolar I disorder who displayed breakthrough manic symptoms (defined as a YMRS total score ≥12), whereas on therapeutic doses of lithium, VPA, or the combination, were randomized to adjunctive treatment with gabapentin 600 to 3,600 mg per day (*n* = 55) or placebo (*n* = 59) (72). Patients receiving placebo exhibited significantly greater improvement in YMRS total scores compared with patients receiving gabapentin.

In the second controlled trial, 28 patients with bipolar I (*n* = 13) or bipolar II (*n* = 15) disorder received 6-week crossover trials of gabapentin, lamotrigine, or placebo (29) The reduction in manic symptoms as measured by the CGI-BP was not significantly different among the three treatments. However, manic symptoms were quite low at baseline, raising the possibility that meaningful differences among the three groups might not have been detected. Lamotrigine was also compared with lithium in a 4-week double-blind, randomized trial in 30 inpatients with bipolar I disorder (39). Patients received lithium 800 mg per day or lamotrigine 25 mg per day for the first week, 50 mg per day during the second, then 100 mg per day for the final 2 weeks of the trial. At the conclusion of the study, both agents produced significant reductions in mean Mania Rating Scale, BPRS, and CGI total scores from baseline to endpoint. There were no significant differences between the two treatment groups. The small sample size, low lithium dose, use of as needed lorazepam throughout the trial, and absence of a placebo control group limit the results of this trial.

Lamotrigine is the only new antiepileptic agent studied to date in a randomized controlled trial in the maintenance treatment of bipolar disorder (12). In this study, bipolar I (*n* = 130) and bipolar II (*n* = 52) patients who were stabilized on initial open-label lamotrigine monotherapy were randomized to lamotrigine or placebo in a 26-week prevention trial. There were no significant differences between the lamotrigine and placebo groups in time to drop out for any reason and time to need for additional medication among bipolar I patients. However, in bipolar II patients, treatment with lamotrigine was associated with significantly lower relapse rates on these measures compared with placebo.

THE RATIONAL DEVELOPMENT OF TRULY NOVEL TREATMENTS FOR BD

Part of "77 - Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder "

Signal Transduction Modifiers

Although the large number of anticonvulsants and atypical antipsychotic agents in the pharmacopeia has greatly enhanced our ability to treat patients with BD, there is still clearly a real need to develop truly novel agents to adequately treat this devastating illness, and modify the long-term outcome for millions of sufferers. A major impediment in the development of truly novel agents has been the dearth of knowledge pertaining to the underlying pathophysiology of the illness. A true understanding of the pathophysiology of an illness as complex as BD must clearly address its neurobiology at different physiologic levels (i.e., molecular, cellular, systems, and behavioral). Abnormalities in gene expression

undoubtedly underlie the neurobiology of the disorder at the molecular level; this will become evident as we identify the susceptibility and protective genes for BD in the coming years. Once this has been accomplished, however, the even more difficult work must begin to examine the impact of the faulty expression of these gene products (proteins) on integrated cell function. It is at these levels that critical signaling molecules recently have been identified as candidate targets for the development of truly novel agents for the treatment of BD.

Multicomponent, cellular signaling pathways interact at various levels, thereby forming complex signaling networks that allow the cell to receive, process, and respond to information (58). The high degree of complexity generated by these signaling provides neurons with the flexibility to generate the wide range of responses observed in the nervous system. These pathways are undoubtedly involved in regulating such diverse vegetative functions as mood, appetite, and wakefulness, as well as higher cognitive functions—systems that are all affected in BD—and thus represent attractive putative mediators of mood stabilization. Over the last decade, there have been major advances in our understanding of the critical role of the protein kinase C (PKC) signaling pathway as a therapeutically relevant target for the long-term actions of mood stabilizers (57 ,59). The preponderance of the data indicates that chronic lithium attenuates PKC responses and down-regulates specific PKC isozymes (57). Studies in rodents and cultured cells have demonstrated that chronic (but not acute) lithium produces an isozyme-selective reduction in PKC α and ε. Moreover, the structurally highly dissimilar antimanic agent VPA produces strikingly similar effects on the PKC signaling pathway, as does lithium (17 ,57). In view of the pivotal role of the PKC signaling pathway in the regulation of neuronal excitability, and neurotransmitter release (57), it was postulated that the attenuation of PKC activity might play a major role in the antimanic effects of lithium and VPA. There is currently only one relatively selective PKC inhibitor available for human use—tamoxifen (20). Tamoxifen, a synthetic nonsteroidal antiestrogen, is also a potent PKC inhibitor at therapeutically relevant concentrations (20). Therefore, a pilot study was initiated to investigate the efficacy of tamoxifen in the treatment of acute mania, and it was found that tamoxifen did indeed possess antimanic activity (57 ,58). Clearly, the results have to be considered preliminary owing to the small sample size thus far. Nevertheless, the significant (and in some cases rapid and striking) results that have observed are intriguing. In view of the preliminary data suggesting the involvement of the PKC signaling system in the pathophysiology of BD (17 ,57), these results suggest that PKC inhibitors may be very useful agents in the treatment of BD. Larger, double-blind placebo-controlled studies of tamoxifen and novel selective PKC inhibitors in the treatment of mania are clearly warranted.

In recent years, a hitherto completely unexpected target for the action of lithium has been identified. Klein and Melton (51) were the first to demonstrate that lithium, at therapeutically relevant concentrations, inhibits glycogen synthase kinase 3B (GSK3B). GSK3B is now known to play a critical role in the CNS, by regulating various cytoskeletal processes, synaptic plasticity, and long-term gene expression (51 ,59). Interestingly, VPA (but not CBZ) also concentration-dependently inhibits GSK-3B *in vitro*, with significant effects observed at concentrations of VPA similar to those attained clinically (17 ,60). Most recently, it has been demonstrated that the chronic (3- to 4-week) administration of lithium and VPA also increase B-catenin levels in rodent brain (Chen and Manji, unpublished observations), compatible with inhibition of GSK3B during chronic *in vivo* administration of the agents under therapeutic paradigms. Overall, GSK-3B appears to play a critical role in regulating neuroplastic events in the CNS, and there is considerable excitement about the possibility of developing novel GSK-3B modulators as potential new therapeutics for both BD and neurodegenerative diseases (60).

NEUROTROPHIC AND NEUROPROTECTIVE AGENTS FOR THE OPTIMAL LONG-TERM TREATMENT OF BD

Part of "77 - Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder "

Recent studies investigating potential structural brain changes in mood disorders have demonstrated reductions in regional CNS volume and cell numbers (both neurons and glia) in many patients (61). It is thus noteworthy that lithium and VPA have recently been demonstrated to robustly increase the expression of the cytoprotective protein bcl-2 in the CNS (59 ,61). Chronic lithium not only exerts neuroprotective effects in several preclinical paradigms, but also enhances hippocampal neurogenesis (61). VPA robustly promotes neurite outgrowth and activates the ERK MAP kinase pathway, a signaling pathway utilized by many endogenous neurotrophic factors (61). Consistent with its preclinical neurotrophic/neuroprotective effects, chronic lithium treatment of patients with BD increases brain N-acetylaspartate (NAA, a putative marker of neuronal viability and function) levels, effects that are localized almost exclusively to gray matter (61). To determine if lithium was producing neuropil increases, quantitative three-dimensional MRI studies were undertaken that revealed that chronic lithium significantly increases total gray matter volume in the human brain of patients with BD (61). The evidence demonstrating the neurotrophic effects of lithium, VPA, and antidepressants, the enhancement of hippocampal neurogenesis in the adult mammalian brain, as well as the growing appreciation that mood disorders are associated with cell loss and atrophy, suggest that these effects may be very relevant for the long-term treatment of mood disorders. It is perhaps useful to conceptualize the cell death and atrophy that occurs in mood disorders as arising from an impairment

of “cellular resiliency.” In this context, it is noteworthy that a variety of strategies to enhance cellular resiliency via effects on molecules involved in cell survival and cell death pathways are currently under investigation.

CONCLUDING REMARKS

Part of "77 - Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder "

The growing body of data implicating signaling pathways in the pathophysiology and treatment of BD suggests that compounds with potent effects distal to the receptor may represent truly novel treatments for BD. The rapid technologic advances in molecular and cellular biology have greatly enhanced our ability to understand the complexities of the regulation of neuronal function; these advances are leading to an increasing number of avenues to modulate transmembrane signaling pathways, and hold much promise for the development of truly novel and improved therapeutics. Emerging results from a variety of clinical and preclinical experimental and naturalistic paradigms also suggest that optimal long-term treatment of BD may only be achieved by the early use of agents with neurotrophic/neuroprotective effects, irrespective of the primary, symptomatic treatment. The development of new treatments that regulate molecules involved in cell survival and cell death pathways, such as CREB, BDNF, Bcl-2, and MAP kinases remains an exciting prospect for the future.

ACKNOWLEDGMENTS

Part of "77 - Current and Emerging Treatments for Acute Mania and Long-Term Prophylaxis for Bipolar Disorder "

The authors’ research is supported by NIMH, the Theodore and Vada Stanley Foundation, NARSAD, and Joseph Young Senior Research Awards.

Dr. Keck has received research support from the following companies: Merck, Inc., Pfizer, Inc., Abbott Laboratories, Eli Lilly & Company, Janssen, and Glaxo-Wellcome. In addition, he has served as a consultant for: Abbott Laboratories, Eli Lilly & Company, Astra-Zeneca, Pfizer, Inc., Wyeth-Ayerst, Parke-Davis, Shire Pharmaceuticals, Pharmacia-Upjohn, and Janssen Pharmaceutica.

Dr. Manji has served as a consultant and/or has received research support from Abbott Laboratories, Eli Lilly & Company, and Janssen Pharmaceutica.

REFERENCES

1. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington, DC: American Psychiatric Press, 1994.

2. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, third ed revised. Washington, DC: American Psychiatric Press, 1987.

3. Baastrup PC, Paulsen JC, Schor M, et al. Prophylactic lithium: double-blind discontinuation in manic-depressive and recurrent-depressive disorders. *Lancet* 1970;2:376–330.

4. Baldessarini RJ, Tondo L. Does lithium treatment still work? Evidence of stable responses over three decades. *Arch Gen Psychiatry* 2000;57:157–190.

5. Ballenger JL, Post RM. Therapeutic effects of CBZ in affective illness: a preliminary report. *Commun Psychopharmacol* 1978;1:159–175.

6. Barbini B, Scherillo P, Benedetti F, et al. Response to clozapine in acute mania is more rapid than that of chlorpromazine. *Int Clin Psychopharmacol* 1997;12:109–112.

7. Berk M, Ichim L, Brook S. Olanzapine compared to lithium in mania: a double-blind randomized controlled trial. *Int Clin Psychopharmacol* 1999;14: 339–343.

8. Bowden CL, Brugger AM, Swann AC, et al. Efficacy of divalproex vs. lithium and placebo in the treatment of mania. *JAMA* 1994;271:915–924.

9. Bowden CL, Calabrese JR, McElroy SL, et al. A randomized, placebo-controlled 12-month trial of divalproex and lithium in treatment of outpatients with bipolar I disorder. *Arch Gen Psychiatry* 2000;57:481–489.

10. Bowden CL, Calabrese JR, Wallin BA, et al. Who enters therapeutic trials? Illness characteristics of patients in clinical drug studies of mania. *Psychopharmacol Bull* 1995;31:103–109.

11. Brennan MIW, Sandyk R, Borsook D. Use of sodium VPA in the management of affective disorders: basic and clinical aspects. In: Emrich HM, Okuma T, Muller AA, eds. *Anticonvulsants in affective disorders*. Amsterdam: Excerpta Medica, 1984:56–65.

12. Calabrese JR, Suppes T, Bowden CL, et al. A double-blind, placebo-controlled prophylaxis study of lamotrigine in rapid-cycling bipolar disorder. *J Clin Psychiatry* 2000;61:841–850.

13. Calabrese JR, Delucchi GA. Spectrum of efficacy of VPA in 55 patients with rapid-cycling BD. *Am J Psychiatry* 1990;147:431–434.

14. Calabrese JR, Kimmel SE, Woyshville MJ, et al. Clozapine for treatment-refractory mania. *Am J Psychiatry* 1996;153:759–762.

15. Calabrese JR, Markovitz PJ, Kimmel SE, et al. Spectrum of efficacy of VPA in 78 rapid-cycling bipolar patients. *J Clin Psychopharmacol* 1992;12[Suppl]:53–56.

16. Calabrese JR, Rapport DJ. Mood stabilizers and the evolution of maintenance study designs in bipolar I disorder. *J Clin Psychiatry* 1999;60[Suppl 5]:5–13.

17. Chen G, Hasanat KA, Bebchuk JM, et al. Regulation of signal transduction pathways and gene expression by mood stabilizers and antidepressants. *Psychosom Med* 1996a;61(5):599–617.

18. Cole JO, Banov MD, Green A, et al. Clozapine in the treatment of refractory acute mania [abstract]. New Research Program and Abstracts of the 146th Annual Meeting of the American Psychiatric Association. San Francisco, CA; Abstract NR 455;174.

19. Coppen A, Nuguera R, Bailey J. Prophylactic lithium in affective disorders. *Lancet* 1971;2:275–279.

20. Couldwell WT, Weiss MH, DeGiorgio CM, et al. Clinical and radiographic response in patients with recurrent malignant gliomas treated with high-dose tamoxifen. *Neurosurgery* 1993;32:485–489.

21. Cundall RL, Brooks PW, Murray LG. Controlled evaluation of lithium prophylaxis in affective disorders. *Psychol Med* 1972;2:308–311.

22. Dardennes R, Even C, Bange F, et al. Comparison of CBZ and lithium in the prophylaxis of BDs. A meta-analysis. *Br J Psychiatry* 1995;166:375–381.

23. Denicoff KD, Smith-Jackson EE, Disney ER, et al. Comparative prophylactic efficacy of lithium, CBZ and the combination in BD. *J Clin Psychiatry* 1997;58:470–478.

24. Dunner DL, Fieve RR. Clinical factors in lithium carbonate prophylaxis failure. *Arch Gen Psychiatry* 1974;30:229–233.

25. Dunner DL, Stallone F, Fieve RR. A double-blind study of prophylaxis of depression in bipolar illness. *Arch Gen Psychiatry* 1976;33:117–121.

26. Emrich HM, von Zerssen D, Kissling W. On a possible role of GABA in mania: therapeutic efficacy of sodium VPA. In: Costa E, Dicharia G, Gessa GL, eds. *GABA and benzodiazepine receptors*. New York: Raven, 1981:287–296.

27. Fieve RL, Kambaraci T, Dunner DL. Lithium prophylaxis of depression in bipolar I, bipolar II and unipolar patients. *Am J Psychiatry* 1976;133:925–929.

28. Freeman TW, Clothier JL, Pazzaglia P, et al. A double-blind comparison of valproic acid and lithium in the treatment of acute mania. *Am J Psychiatry* 1992;149:247–250.

29. Frye MA, Ketter TA, Kimbrell TA, et al. A placebo controlled study of lamotrigine and gabapentin monotherapy in refractory mood disorders. *J Clin Psychopharmacol* 2000;20:607–614.

30. Garfinkel PE, Stancer HC, Persad E. A comparison of haloperidol, lithium carbonate and their combination in the treatment of mania. *J Affect Disord* 1980;2:279–288.

31. Goodwin FK, Murphy DL, Bunney WE Jr. Lithium carbonate treatment in depression and mania. *Arch Gen Psychiatry* 1969;21:486–496.

32. Griel W, Kleindienst N, Erazo N, Muller-Oerlinghausen B. Differential response to lithium and CBZ in the prophylaxis of BD. *J Clin Psychopharmacol* 1998;18:455–460.

33. Greil W, Ludwig-Mayerhofer W, Erazo N, et al. Lithium versus CBZ in the maintenance treatment of BDs: a randomized study. *J Affect Disord* 1997;43:151–161.

34. Grossi E, Sacchetti E, Vita A, et al. CBZ versus chlorpromazine in mania: a double-blind trial. In: Emrich HM, Okuma T, Müller AA, eds. *Anticonvulsants in affective disorders*. Amsterdam: Excerpta Medica, 1984:177–187.

35. Grunze H, Erfurth A, Amann B, et al. Intravenous VPA loading in acutely manic and depressed bipolar I patients. *J Clin Psychopharmacol* 1999;19:303–309.

36. Harrow M, Goldberg, JF, Grossman LS, et al. Outcome in manic disorders: a naturalistic follow-up study. *Arch Gen Psychiatry* 1990;47:665–671.

37. Hirschfeld RMA, Allen ML, McEvoy J, et al. Safety and tolerability of oral loading divalproex sodium in acutely manic bipolar patients. *J Clin Psychiatry* 1999;60:815–818.

38. Hullin RP, McDonald R, Allsopp MNE. Prophylactic lithium in recurrent affective disorders. *Lancet* 1972;1:1044–1046.

39. Ichim L, Berk M, Brook S. Lamotrigine compared with lithium in mania: a double-blind randomized controlled trial. *Ann Clin Psychiatry* 2000;12:5–10.

40. Johnson G, Gershon S, Burdock EI, et al. Comparative effects of lithium and chlorpromazine in the treatment of acute manic states. *Br J Psychiatry* 1971;119:267–276.

41. Kane JM, Quitkin FM, Rifkin A, et al. Lithium carbonate and imipramine in the prophylaxis of unipolar and bipolar II illness. *Arch Gen Psychiatry* 1982;39:1065–1069.

42. Keck, PE, Jr, Ice K, and the Ziprasidone Study Group. Controlled treatment of acute mania with ziprasidone [abstract]. American Psychiatric Association Annual Meeting. Chicago, IL, May 13–15, 2000. New Research Abstract.

43. Keck PE, Jr, McElroy SL. Outcome in the pharmacologic treatment of BD. *J Clin Psychopharmacol* 1996;16[Suppl]:15–23.

44. Keck PE, Jr, McElroy SL, Nemeroff CB. Anticonvulsants in the treatment of BD. *J Neuropsychiatry Clin Neurosci* 1992;4:595–405.

45. Keck PE Jr, McElroy SL, Strakowski SM, et al. Twelve-month outcome of bipolar patients following hospitalization for a manic or mixed episode. *Am J Psychiatry* 1998;155:646–652.

46. Keck PE Jr, McElroy SL, Tugrul KC, et al. VPA oral loading in the treatment of acute mania. *J Clin Psychiatry* 1993;54:305–308.

47. Keck PE Jr, Reeves KL, Harrigan EP. Ziprasidone in the short-term treatment of patients with schizoaffective disorder: results from two double- blind, placebo-controlled, multicenter studies. *J Clin Psychopharmacol* 2001;21:27–35.

48. Keck PE Jr, Welge JA, McElroy SL, et al. Placebo effect in randomized, controlled studies of acute bipolar mania and depression. *Biol Psychiatry* 2000;47:748–755.

49. Keck PE Jr, Welge JA, Strakowski SM, et al. Placebo effect in randomized, controlled maintenance studies of patients with bipolar disorder. *Biol Psychiatry* 2000;47:756-761.

50. Klein DF. Importance of psychiatric diagnosis in prediction of clinical drug effects. *Arch Gen Psychiatry* 1967;16:118–126.

51. Klein PS, Melton DA. A molecular mechanism for the effect of lithium on development. *Proc Natl Acad Sci USA* 1996;93:8455–8459.

52. Lambert PA, Venaud G. Comparative study of valpromide versus lithium in the treatment of affective disorders. *Nervure* 1992;5:57–65.

53. Lerer B, Moere N, Meyendorff E, et al. CBZ versus lithium in mania: a double blind study. *J Clin Psychiatry* 1987;48:89–93.

54. Lusznat RM, Murphy DP, Nunn CM. CBZ vs. lithium in the treatment and prophylaxis of mania. *Br J Psychiatry* 1988;153:198–204.

55. Maggs R. Treatment of manic illness with lithium carbonate. *Br J Psychiatry* 1963;109:56–65.

56. Maj M, Pirozzi R, Kemali D. Long-term outcome of lithium prophylaxis in bipolar patients. *Arch Gen Psychiatry* 1991;48:772.

57. Manji HK, Lenox RH. Protein kinase C signaling in the brain: molecular transduction of mood stabilization in the treatment of manic-depressive illness. Ziskind-Somerfeld Award Paper. *Biol Psychiatry* 1999;46:1328–1351.

58. Manji HK, Lenox RH. Signaling: cellular insights into the pathophysiology of bipolar disorder. *Biol Psychiatry* 2000;48:518–530.

59. Manji HK, Moore GJ, Chen G. Lithium at 50: have the neuroprotective effects of this unique cation been overlooked? *Biol Psychiatry* 1999;46:929–940.

60. Manji HK, Chen G, Hsiao JK, et al. Regulation of signal transduction pathways by mood stabilizing agents: implications for the pathophysiology and treatment of bipolar affective disorder. In: Manji HK, Bowden CL, Belmaker RH, eds. *Bipolar medications: mechanisms of action*. Washington, DC:American Psychiatric Press, 2000:129–177.

61. Manji HK, Moore GJ, Chen G. Clinical and preclinical evidence for the neurotrophic and effects of mood-stabilizing agents: implications for the pathophysiology and treatment of manic-depressive illness. *Biol Psychiatry* 2000;48:740–54.

62. McElroy SL, Keck PE Jr. Pharmacologic agents for the treatment of acute bipolar mania. *Biol Psychiatry* 2000;48:539–557.

63. McElroy SL, Keck PE Jr, Pope HG Jr. VPA in the treatment of rapid cycling BD. *J Clin Psychopharmacol* 1988;8:275–279.

64. McElroy SL, Keck PE Jr, Pope HG Jr, et al. Clinical and research implications of the diagnosis of dysphoric or mixed mania or hypomania. *Am J Psychiatry* 1992;149:1633–1644.

65. McElroy SL, Keck PE Jr, Stanton SP, et al. A randomized comparison of divalproex oral loading versus haloperidol in the initial treatment of acute psychotic mania. *J Clin Psychiatry* 1996;57:142–146.

66. McElroy SL, Keck PE Jr, Tugrul KC, et al. VPA as a loading treatment in acute mania. *Neuropsychobiology* 1993;27:146–149.

67. Melia PI. Prophylactic lithium: a double-blind trial in recurrent affective disorders. *Br J Psychiatry* 1970;116:621–624.

68. Muller-Oerlinghausen B, Retzow A. VPA as adjunct to neuroleptic medication in the treatment of acute episodes of mania. *Pharmacopsychiatry* 1997;30:202–206.

69. O'Connell RA, Mayo JA, Flatlow L, et al. Outcome of BD on long-term treatment with lithium. *Br J Psychiatry* 1991;159:123–129.

70. Okuma T, Inanaga K, Otsuki S, et al. A preliminary double-blind study on the efficacy of CBZ in prophylaxis of manic-depressive illness. *Psychopharmacology* 1981;73:95–96.

71. Okuma T, Inanaga K, Otsuki S, et al. Comparison of the antimanic efficacy of CBZ and chlorpromazine: a double-blind controlled study. *Psychopharmacology* 1979;66:211–217.

72. Pande AC, Crockett JG, Janney CA, et al. Gabapentin in bipolar disorder: a placebo-controlled trial of adjunctive therapy. Gabapentin Bipolar Disorder Study Group. *Bipolar Disord* 2000;2:249–255.

73. Peselow ED, Fieve RR, Defaglia C, et al. Lithium prophylaxis of bipolar illness. The value of combination treatment. *Br J Psychiatry* 1994;164:208–214.

74. Placidi GF, Lenzi A, Lazzerini F, et al. The comparative efficacy and safety of CBZ versus lithium: a randomized, double-blind 3-year trial in 83 patients. *J Clin Psychiatry* 1986;47:490–494.

75. Platman SR. A comparison of lithium carbonate and chlorpromazine in mania. *Am J Psychiatry* 1970;127:351–353.

76. Pope HG Jr, McElroy SL, Keck PE Jr, et al. VPA in the treatment of acute mania: a placebo-controlled study. *Arch Gen Psychiatry* 1991;48:62–68.

77. Post RM, Denicoff KD, Frye MA, et al. Re-evaluating CBZ prophylaxis in BD. *Br J Psychiatry* 1997;170:202–204.

78. Prien RF, Caffey EM Jr, Klett CJ. Comparison of lithium carbonate and chlorpromazine in the treatment of mania: report of the Veterans Administration and National Institute of Mental Health Collaborative Study Group. *Arch Gen Psychiatry* 1972;26:146–153.

79. Prien RF, Caffey EM Jr, Klett CJ. Prophylactic efficacy of lithium carbonate in manic-depressive illness. Report of the Veterans Administration and National Institute of Mental Health Collaborative Study Group. *Arch Gen Psychiatry* 1973;28:337–341.

80. Prien RF, Klett CJ, Caffey EM Jr. Lithium carbonate and imipramine in prevention of affective episodes. *Arch Gen Psychiatry* 1973;29:420–425.

81. Prien RF, Potter WZ. NIMH workshop on treatment of BD. *Psychopharmacology Bull* 1990;26:409–427.

82. Puzynski S, Klosrewicz L. Valproic acid amide as a prophylactic agent in affective and schizoaffective disorders. *Psychopharmocol Bull* 1984;20:151–159.

83. Quitkin F, Rifkin A, Kane J. Prophylactic efficacy of lithium and imipramine in unipolar and bipolar II patients: a preliminary report. *Am J Psychiatry* 1978;135:570–572.

84. Sachs GS and the Risperidone Bipolar Study Group. Safety and efficacy of risperidone vs. placebo as add-on therapy to mood stabilizers in the treatment of manic phase of BD [abstract]. American College of Neuropsychopharmacology 38th Annual Meeting. Acapulco, Mexico, December 12–16, 1999.

85. Schou M, Juel-Nielson N, et al. The treatment of manic psychoses by the administration of lithium salts. *J Neurol Neurosurg Psychiatr* 1954;17:250–260.

86. Segal J, Berk M, Brook S. Risperidone compared with both lithium and haloperidol in mania: a double-blind randomized controlled trial. *Clin Neuropharmacol* 1998;21:176–180.

87. Shopsin B, Gerson S, Thompson H, et al. Psychoactive drugs in mania: a controlled comparison of lithium carbonate, chlorpromazine and haloperidol. *Arch Gen Psychiatry* 1975;32:34–42.

88. Small JG, Klapper MH, Milstein V, et al. CBZ compared with lithium in the treatment of mania. *Arch Gen Psychiatry* 1991;48:915–921.

89. Solomon DA, Ryan CE, Keitner GI, et al. A pilot study of lithium carbonate plus divalproex sodium for the continuation and maintenance treatment of patients with bipolar I disorder. *J Clin Psychiatry* 1997;58:95–99.

90. Spring G, Schweid D, Gray C, et al. A double-blind comparison of lithium and chlorpromazine in the treatment of manic states. *Am J Psychiatry* 1970;126:1306–1310.

91. Stallone F, Mendlewicz J, Fieve RR. How blind is double-blind? An assessment in a lithium-prophylaxis study. *Lancet* 1974;1:619–620.

92. Stallone F, Shelley E, Mendlewicz J, et al. The use of lithium in affective disorders, III: a double-blind study of prophylaxis in bipolar illness. *Am J Psychiatry* 1973;130:1006–1010.

93. Stokes PE, Stoll PM, Shamoian CA, et al. Efficacy of lithium as acute treatment of manic-depressive illness. *Lancet* 1971;1:1319–1325.

94. Suppes T, Webb A, Paul B, et al. Clinical outcome in a randomized one-year trial of clozapine versus treatment as usual for patients with treatment-resistant illness and a history of mania. *Am J Psychiatry* 1999;156:1164–1169.

95. Swann AC, Bowden CL, Moris D, et al. Depression during mania. Treatment response to lithium or divalproex. *Arch Gen Psychiatry* 1997;54:37–42.

96. Swann AC, Bowden CL, Calabrese JR, et al. Differential effect of previous episodes of affective disorder on response to lithium or divalproex in acute mania. *Am J Psychiatry* 1999;156:1264–1266.

97. Takahashi R, Sakuma A, Itoh K, et al. Comparison of efficacy of lithium carbonate and chlorpromazine in mania: report of collaborative study group on treatment of mania in Japan. *Arch Gen Psychiatry* 1975;32:1310–1318.

98. Tohen M, Jacobs TG, Grundy SL, et al. Efficacy of olanzapine in acute bipolar mania: a double-blind, placebo-controlled study. The Olanzipine HGGW Study Group. *Arch Gen Psychiatry* 2000;57:841–849.

99. Tohen M, Sanger TM, McElroy SL, et al. Olanzapine versus placebo in the treatment of acute mania. *Am J Psychiatry* 1999;156:702–709.

100. Tohen M, Waternoux CS, Tsuang MT. Outcome in mania. A 4-year prospective follow-up on 75 patients utilizing survival analysis. *Arch Gen Psychiatry* 1990;47:1106–1111.

101. Watkins SE, Cullender K, Thomas DR, et al. The effect of CBZ and lithium on remission from affective illness. *Br J Psychiatry* 1987;150:180–182.

102. Zarate CA Jr, Tohen M, Baldessasini RJ. Clozapine in severe mood disorders. *J Clin Psychiatry* 1995;56:411–417.

78

Cost-Effectiveness of the Newer Generation of Antidepressants

Scott W. Woods

C. Bruce Baker

Scott W. Woods and C. Bruce Baker: Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.

In the last decade or so, several new antidepressants have been introduced into practice, including the selective serotonin reuptake inhibitors fluoxetine (1987), sertraline (1991), paroxetine (1992), and citalopram (1998), in addition to bupropion (1985), venlafaxine (1993), nefazodone (1994), and mirtazapine (1996). Two of these medications are available in delayed-release formulations: bupropion SR (1996) and venlafaxine XR (1997). Only one of these medications is currently available in a generic formulation (bupropion, 1999).

Each of these medications is more expensive in terms of acquisition costs than the older generation of tricyclics and heterocyclics and monoamine oxidase inhibitors. Together, the newer generations of antidepressants accounted for approximately $7.7 billion in retail sales in the United States in 1999 (1), and antidepressants are second only to antibiotics in prescription sales by drug category.

Prices are of course subject to variation, but Table 78.1 shows that the "average wholesale price" for most of these medications is more than $2 per day, according to the 2000 Drug Topics "Red Book" (2). Actual costs to health care systems or patients can be substantially lower or higher than the figures in Table 78.1 would indicate, depending on a number of factors. Certain health care systems are able to qualify for discounted prices for certain medications, and patients can sometimes receive medications free of charge via physician samples or indigent care programs. Acquisition costs are also lowered by patient noncompliance. In addition, some of the newer antidepressants are available as scored tablets (Table 78.1), so that patients taking lower doses have the opportunity to reduce costs substantially with only the modest inconvenience entailed by breaking the pills. However, higher doses of some medications lead to higher acquisition costs when more than one pill is required for a particular dose. Regardless of these caveats, it is clear that the newer generation of antidepressants is more expensive to purchase than are the older generations.

Antidepressant	Strength (mg)	Dosing	Average Wholesale Price ($)[a]	
			Brand	Generic
Fluoxetine	10[b]	QD	76.60	NA
	20		73.14	
	40		158.45	
Sertraline	25[b]	QD	65.99	NA
	50[b]		66.50	
	100[b]		67.73	
Paroxetine	10	QD	66.95	NA
	20[b]		56.97	
	30		71.95	
	40		76.00	
Citalopram	20[b]	QD	60.51	NA
	40[b]		63.14	
Bupropion	75	TID	65.97	64.87
	100		74.30	86.54
Bupropion SR	100	BID	85.51	NA
	150		91.64	
Venlafaxine	25[b]	BID/TID	70.00/104.99	NA
	37.5[b]		69.72/104.58	
	50[b]		74.24/111.36	
	75[b]		78.71/118.07	
	100[b]		83.43/125.14	
Venlafaxine XR	37.5	QD	62.18	NA
	75		69.65	
	150		75.86	
Nefazodone	50	BID	74.11	NA
	100[b]		74.00	
	150[b]		74.00	
	200		74.11	
	250		74.00	
Mirtazapine	15[b]	QD	69.72	NA
	30[b]		71.83	
	45		76.50	

BID, twice daily; NA, not applicable; QD, daily; TID, three times daily.
[a]Least expensive price across suppliers, including repackaging houses. For QD dosing, the lowest cost for 30 pills was used, and 60 pills for BID dosing and 90 pills for TID dosing. When few or no suppliers offered lots of 30, 60, or 90 pills, the lowest-price 100-pill lot was multiplied by 0.3, 0.6, or 0.9, respectively, or the lowest-price 30-pill lot was multiplied by 2 or 3, respectively, whichever was lower. Unit dosing price excluded.
[b]Scored.

TABLE 78.1. AVERAGE WHOLESALE PRICE FOR A 30-DAY SUPPLY OF NEWER ANTIDEPRESSANTS

Given their higher acquisition costs, it is important to determine whether these new, more expensive medications are cost-effective as first-line treatment in comparison with the older, less expensive antidepressants. In other words, in lay terms, are the newer medications worth the prices charged? Despite their higher acquisition costs, the newer antidepressants could be more cost-effective if they resulted in greater increases in quality of life and functioning, or in a reduction of the nonmedication costs of illness in comparison with the older antidepressants sufficient to offset their higher acquisition costs.

The impairments in quality of life and functioning and the nonmedication treatment costs associated with depression are large, and any improvements in these areas achieved with the newer antidepressants would clearly offset their higher purchase price. The World Health Organization has calculated that depression was the fourth leading cause of disease burden throughout the world in 1990, and projected that it would be the second leading cause by 2020 (3). The costs of depression to society as a result of productivity lost because of morbidity and mortality have been estimated at $14.2 billion (in 1980) (4) and $31.3 billion (in 1990) (5) in the United States. Moreover, direct treatment costs for depression, exclusive of medication acquisition costs, have been estimated at $2.0 billion (in 1980) (4) and $11.2 billion (in 1990) (5) in the United States. Most of these costs were related to hospitalization. The better tolerability of the newer versus the older antidepressants might well lead to reductions in these expensive treatment services.

A formal determination of whether the higher acquisition price of the newer antidepressants relative to the older antidepressants is offset by savings in other areas or increased benefits is traditionally conducted with a cost-effectiveness

analysis. Cost-effectiveness is represented as a ratio between direct costs, the numerator, and changes in health status, the denominator. The relative cost-effectiveness of newer versus older antidepressants is represented as the incremental or marginal difference between the cost-effectiveness ratios determined for the newer and older antidepressants.

A cost-effectiveness model depends on many parameters, such as the effectiveness of alternative initial treatments, effectiveness of switching to secondary treatments, postulated lengths of treatment, and costs and health effects included.

The following are the major potential categories of costs and health effects (6). Direct costs are the resources consumed in providing the intervention, in this case the treatment of depression, which includes dealing with side effects and other consequences. Direct costs are further subdivided into four major categories. The first category encompasses changes in the use of health care resources (e.g., the costs of medication acquisition, physicians and other personnel, laboratory and other services, and the appropriately apportioned capital costs of buildings and equipment). The second category of direct costs encompasses changes in the use of other resources (e.g., transportation costs). The final two categories encompass changes in the use of informal caregiver time and in the use of patient time for treatment.

Health effects are divided into two major categories. In

the first category, the intrinsic value of changes in health status, a value is placed on achieving or avoiding a specific health state. The health state may be characterized by using a single domain or multiple domains (e.g., changes in clinical status, functioning, and quality of life). The outcomes measured in any one of these domains can be intermediate (e.g., changes in the Hamilton Depression Scale) or distant (e.g., years of life gained). In practice, when intermediate outcomes are used, the health state and cost-effectiveness ratio is sometimes denoted simply in the native units of a single domain (e.g., cost per patient remitted from depression), and value weights are not assigned. In fuller analyses, weights are assigned to the benefits, and the weights are denominated in more comprehensive generic units that can be compared and combined across domains. The most common generic unit is the quality-adjusted life year (QALY).

The second category major category of health effects, indirect costs or productivity effects, refers to resource consumption attributable to changes in productivity caused by changes in morbidity or mortality.

In the most comprehensive cost-effectiveness analysis, these cost and health effects categories are applied to all sectors of health care, even if the specific intervention falls within a limited sector (e.g., treatment of depression within the mental health specialty sector). In more limited analyses, the categories are applied only in the specialty sector.

We should note that *costs* are not the same thing as *prices*. From an economic perspective, the term *costs* refers to the value of the resources consumed in providing/producing a service such as treatment of depression, most ideally calculated in terms of the consumed resources next-best use. The many types of *prices* that can be assigned to resources, and which are used in most studies, may or may not reflect the economic value of the resources consumed.

The conclusions suggested by any given cost-effectiveness analysis depend heavily on each of the factors we have listed: overall structure of the model, cost categories and specific cost values used, health effects categories, method of measuring health effects, and weights assigned to outcomes. The conclusions of the analysis also depend on its perspective—that is, for whom is the treatment cost-effective? The perspective determines which costs, benefits, and outcomes are potentially relevant and what weights are appropriate. Clarity about perspective is critical because in most contexts, various combinations of cost and benefits are borne by or accrue to different entities. For example, in a highly simplified and hypothetical case, if an HMO pays for prescriptions completely, and if the choice of a particular antidepressant results in higher total expenditures for drug purchase but allows patients to be less dependent on family members, the cost is borne by the HMO but the benefit is gained by the patient's family. In this case, the antidepressant might be cost-effective from the perspective of the patient's family or even from the broader perspective of society, but not from the perspective of the HMO.

Some of the perspectives commonly discussed or used include the following: patient or patient/family, employer/payer, individual health care institution (e.g., an HMO), national health care specialty sector (e.g., specialty mental health), national health care comprehensive system (i.e., including all health care sectors), and global societal (i.e., including *all* costs and *all* health effects) perspectives.

In considering whether the available studies suggest that newer antidepressants are cost-effective, we will limit ourselves to addressing the question from the two perspectives most commonly used in studies. First, we ask, "Are newer antidepressants cost-effective as first-line treatment from a health care system perspective?" In addressing this question with evidence from the available studies, one must appreciate that the studies to be reviewed have utilized multiple conceptualizations of cost-effectiveness. Some studies implicitly or explicitly assume equal effectiveness of newer and older antidepressants and ask whether the first-line use of newer antidepressants produces savings to the health care system in the direct treatment, nonmedication costs of treating depression that are sufficient to offset total direct treatment costs. Others model or measure clinical benefit and calculate average or incremental cost-effectiveness ratios. We report the authors' conclusions and discuss the limitations in design and methods. Most of the evidence regarding this first question is based on the perspective of a national health care comprehensive system, not merely of a mental health sector; consequently, the health care system perspective we address refers to all of health care, not just mental health care.

Second, we ask, "Are newer antidepressants cost-effective as first-line treatment from a global societal perspective?" Again, we review studies that utilize multiple conceptualizations of cost-effectiveness from this general perspective.

We also examine studies reporting relative rates of cost-effectiveness of the newer antidepressants.

To address our two major questions, we reviewed the recently published (July 1, 1995 to June 30, 2000) literature in English on antidepressants and cost analysis, focusing particularly on the newer antidepressants and updating our previous review (7). Relevant publications were identified by a search of Medline, Current Contents, and HealthSTAR computer databases and by manual bibliographic review. Studies available only as abstracts were not included. Other reviews of this topic have also been published recently (8 ,9 ,10 ,11 ,12 ,13 and 14).

The evidence most centrally relevant to the cost-effectiveness of antidepressants comes from studies that can be grouped into four methodologies: efficacy study metaanalyses, cost-effectiveness simulations, retrospective analysis of administrative databases, and prospective cost-effectiveness experiments. We review the data from each of these in turn. We also review briefly the data on whether a decrease in deaths from suicide is a benefit that favors the newer antidepressants.

EFFICACY AND TOLERABILITY METAANALYSES

Efficacy and tolerability studies provide information on expected percentages of responders and dropouts, which are central parameters in cost-effectiveness calculations, in addition to information on side effect burden. Cost-effectiveness simulations (see next section) often use data from efficacy metaanalyses.

Numerous metaanalyses of randomized short-term trials of selective serotonin reuptake inhibitors (SSRIs) versus tricyclic antidepressants (TCAs) are now available (Table 78.2), as are several metaanalyses of trials comparing non-SSRI newer antidepressants with older, control antidepressants (Table 78.3). These metaanalyses include several monumental efforts with careful attention to the unbiased inclusion of studies and minimization of publication bias. To simplify the presentation, Table 78.2 shows the original authors' conclusions about the identified principal efficacy and tolerability measures. These authors often considered treatment continuation as an efficacy measure, and treatment discontinuation for side effects as a tolerability measure. A metaanalysis of placebo-controlled comparisons in 49 studies from 1966 through 1995 that included an investigational antidepressant and a reference antidepressant (15) and two other metaanalyses (16 ,17) are not included in Table 78.2 because it is not clear whether the reported effect sizes for TCAs and SSRIs are restricted to the trials internally comparing TCAs with SSRIs. The results of the metaanalyses in Table 78.2 are remarkably consistent. Perhaps the consistency is not surprising given that the articles report on highly overlapping sets of clinical trials. The metaanalyses almost uniformly conclude that these two classes of antidepressant are quite similar in regard to efficacy. Only one analysis found evidence of greater efficacy for the SSRIs in comparison with a subgroup of older TCAs; another found

evidence for greater efficacy of TCAs when attention was restricted to inpatients, and the other eight metaanalyses reported similar efficacy for TCAs and SSRIs.

Reference	Newer AD	Control AD	Inclusion Criteria	No. Studies[a]	Reported Advantage: Efficacy	Reported Advantage: Tolerability
Song et al., 1993 (18)	SSRIs	TCAs and related ADs	Double-blinded published	58	NC	N>C
Montgomery et al., 1994 (19)	SSRIs	TCAs	Double-blinded published	42	—	N>C
Anderson and Torrenson, 1995 (20)	SSRIs	TCAs	Double-blinded published	62	NC	N>C
Hotopf et al., 1997 (65)	SSRIs	Older TCAs	Randomized published	51	N>C	—
		Newer TCAs		24	NC	
		Heterocyclic ADs		17	NC	
Steffens et al., 1997 (66)	SSRIs	TCAs	Double-blinded published	36	NC	N>C
Anderson, 1998 (67)	SSRIs	TCAs and related ADs	Double-blinded inpatient	25, 23	C>N	N>C
Trindade et al., 1998 (68)	SSRIs	TCAs	Double-blinded	84	—	NC[b]
Mulrow et al., 1998 (69)	SSRIs	TCAs	Randomized	43, 76	NC	N>C
Bech et al., 2000 (70)	Fluoxetine	TCAs	Randomized	25	NC	N>C
Geddes et al., 2000 (71)	SSRIs	TCAs	Double-blinded	71	NC	—
Anderson, 2000 (72)	SSRIs	TCAs	Randomized published	102, 95	NC	N>C
Williams et al., 2000 (73)	SSRIs	Older TCAs	Randomized	38	NC	N>C
		Newer TCAs		5	NC	N>C
		Tetracyclics		7	NC	NC

AD, antidepressant; C, control; N, nerver; SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant.
[a]When two numbers are shown, the first is the number of studies in the efficacy analysis and the second the number in the tolerability analysis.
[b]Significant for adult outpatient studies but not for the entire group.

TABLE 78.2. METAANALYSES OF STUDIES COMPARING SSRIs WITH OLDER CONTROL ANTIDEPRESSANTS FOR MAJOR DEPRESSION

Reference	Newer AD	Control AD	Inclusion Criteria	No. Studies[a]	Reported Advantage: Efficacy	Reported Advantage: Tolerability
Stahl et al., 1997 (74)	Mirtazapine	Amitriptyline	Sufficiently similar	4	NC	N>C
Srisurapanont, 1998 (75)	Nefazodone	TCAs	Randomized	17, 13	N>C	NC[b]
	Mirtazapine	nTCAs				
	Venlafaxine	SSRIs				
Mulrow et al., 1998 (69)	SNRIs[c]	TCAs	Randomized	8, 9	NC	N>C
Williams et al., 2000 (73)	SNRIs[d]	Older TCAs	Randomized	6	NC	NC
	5-HT2 antagonists	Older TCAs		5	NC	—

5-HT, serotonin; AD, antidepressant; C, control; N, nerver; nTCA, nontricyclic antidepressant; SNRI, selective norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant.
[a]When two numbers are shown, the first is the number of studies in the efficacy analysis and the second the number in the tolerability analysis.
[b]N>C for comparisons vs. TCAs, NC for comparisons vs. nTCAs and SSRIs.
[c]Includes venlafaxine and mirtazapine.
[d]Includes venlafaxine, mirtazapine, and milnacipran.

TABLE 78.3. METAANALYSES OF STUDIES COMPARING NON-SSRI NEWER ANTIDEPRESSANTS WITH OLDER CONTROL ANTIDEPRESSANTS FOR MAJOR DEPRESSION

The metaanalyses also conclude almost uniformly that the SSRIs have a small but consistent tolerability advantage over the TCAs in these short-term randomized trials. Nine of ten studies investigating tolerability found evidence of greater tolerability for the newer agents, but in general, the magnitude of the effect was relatively small. For example, dropout rates attributed to side effects were 15.4% versus 18.8% (18), 14.9% versus 19.0% (19), and 14.4% versus 18.8% (20) for SSRIs versus TCAs in three of the early metaanalyses.

The small size of the SSRI acceptability advantage in the efficacy studies is surprising, given the widespread entry of the SSRIs into clinical practice. Studies of antidepressant use in naturalistic practice often find a more pronounced tolerability advantage (21 ,22 ,23 ,24 ,25 and 26). Alternative explanations for this discrepancy are possible. Studies suggest that patients who enroll in randomized trials may differ from the general clinical population (27 ,28), and it is possible that patients who enroll in clinical trials are more tolerant of side effects than average. On the other hand, because these naturalistic studies are not double-blinded, patients in clinical practice could be influenced by expectancy effects.

COST-EFFECTIVENESS SIMULATIONS

Part of "78 - Cost-Effectiveness of the Newer Generation of Antidepressants "

Cost-effectiveness simulations most commonly construct mathematical models of clinical practice based on decision analysis. Usually, pathways of branching alternative treatment options and outcomes are created. Costs, probabilities, and in some cases value weights or utilities for outcomes are assigned to the alternatives. Typically, the values for these parameters are derived from metaanalyses, literature reviews, administrative databases, or expert panels. In addition to medication acquisition costs, these models can include any subset of the costs we defined above.

A substantial number of simulations have been published. Table 78.4 shows data from 19 published simulations comparing SSRIs with older control antidepressants. Table 78.5 shows data from nine simulations comparing newer non-SSRI antidepressants with controls. For each study, Table 78.4 and Table 78.5 show the medications compared, duration of the simulation, authors' apparent principal cost-effectiveness outcomes, authors' conclusions about relative cost-effectiveness, and a brief summary of methodologic limitations. Two additional mirtazapine simulations from U.K. and Swedish health care system perspectives are described in a review (11). Results were similar to those of the French and Austrian simulations (Table 78.4 and Table 78.5).

Study (Reference)	Newer AD(s), Population	Control AD(s)	Treatment Length[a]	Principal Outcome	Analysis Favors[b] detail	Methodologic Limitations
Boyer and Feighner, 1993 (76)	fluoxetine sertraline, first episode	amitriptyline imipramine nortriptyline desipramine	6 mo	direct cost per pt	Newer	1. model not fully transparent 2. short treatment durations 3. initial success rates overly different
Jonsson 1993 (34) Jonsson and Bobbington, 1994 (35)	paroxetine, U.K. primary care	imipramine	a–12 wk m–12 wk	direct cost per tx success	Newer	1. no success for switch tx 2. short treatment durations 3. initial success rates overly different 4. equal tx delivery costs
Stewart, 1994 (38)	sertraline paroxetine, primary care	amitriptyline imipramine	a–12 wk m–26 wk	direct cost per tx success	i>a>p>s Control	1. tx failure costs possibly somewhat low 2. equal tx delivery costs
McFarland, 1994 (37)	paroxetine, primary care	imipramine	1:6 mo 2:12 mo	primary care cost per pt	1:ip 2:i>p Control	1. limited perspective 2. initial success rates
Hatziandreu et al., 1994 (77)	sertraline, young female recurrent	dothiepin	a–3 mo m–2 Y+	lifetime direct costs per discounted QALY saved	Newer	1. maintenance vs. episodic tx confounded across drug
Lapierre et al., 1995 (31)	paroxetine, Canada primary care	imipramine	a–6 mo m–4 mo	direct cost per pt	Newer	1. no dc of switch tx when failure 2–4. similar to (34,35)
Anton and Revicki, 1995 (78)	fluoxetine, young female recurrent, Canada	imipramine	9 mo[c]	discounted direct cost per QALY	Newer	1. multiple unreported probabilities 2. health state utility generation methods not described
Revicki et al., 1995 (79)	fluoxetine, young female recurrent, U.S.	imipramine	9 mo[d]	discounted direct cost per QALY	Newer	1. variability for utilities not reported
Nuijten et al., 1995 (80)	citalopram, unspecified	three TCAs	12 mo	direct cost per pt	Newer	1. similar to (77) 2. treatment duration unspecified
Einarsen et al., 1995 (81)	SSRIs inpt start	TCAs HCAs[e]	300 d	direct cost per symptom-free day	H>T>S Control	1. high delivery costs for TCAs 2. initial success rates overly different
Einarsen, et al., 1995 (81)	SSRIs, outpt start	TCAs HCAs[e]	300 d	direct cost per symptom-free day	H>S>T Control	1. high delivery costs for TCAs 2. initial success rates overly different
Hylan et al., 1996 (32)	fluoxetine, primary care	imipramine	6 mo	direct plus indirect cost per cured pt	Newer	1–3. similar to (34,35)
Bentkover and Feighner, 1996 (30)	paroxetine, primary care and psychiatry	imipramine	a–6 mo m–none	direct cost per pt	Newer	1–4. similar to (31)
Woods and Rizzo, 1997 (82)	paroxetine, primary care	imipramine	a–34 wk m–78 wk	direct cost per tx success	Control	1. improved version of (34,35)
Einarsen et al., 1997 (83)	SSRIs, Ontario inpt start	TCAs	6 mo	direct cost per tx success	Tie	1. metaanalysis comparisons indirect 2. short tx duration
Einarsen et al., 1997 (83)	SSRIs, Ontario outpt start	TCAs	6 mo	direct cost per tx success	Newer	1. metaanalysis comparisons indirect 2. short tx duration
Canadian Office, 1998 (84)	SSRIs, Canada, site unspecified	TCAs alone TCAs/switch	a–3 mo m–6 mo	various direct cost ratios	Control	1. report brief
Brown, et al., 1999 (85)	fluoxetine, Austria, moderate severe	amitriptyline	~6 mo	direct cost per pt and tx success, lost productivity	Tie	1. mirtazapine vs. fluoxetine based on one study 2. reliance on Delphic panel 3. hospitalization estimates high

[a]Duration of specific treatment phases is shown if available. If not specified, the overall simulation length is shown. a, duration of successful acute plus continuation treatment; wk, weeks; m, duration of successful maintenance treatment; tx, treatment; dc, discontinuation; mo, months; pt, patient.
[b]Authors' conclusions. AD, antidepressant(s); T, TCA; S, SSRI; C, Control AD(s).
[c]Nine months for first recurrence, then lifetime if second recurrence.
[d]Nine months for first recurrence, then 27 months if second recurrence.
[e]Heterocyclic antidepressants maprotiline and trazodone.
QALY, quality-adjusted life year; >, more cost-effective

TABLE 78.4. COST-EFFECTIVENESS SIMULATION STUDIES COMPARING SSRIs WITH OLDER CONTROL ANTIDEPRESSANTS

Study (Reference)	Newer AD(s), Population	Control AD(s)	Treatment Lengths[a]	Principal Outcome	Analysis Favors[b] detail	Methodologic Limitations
Anton and Revicki, 1995 (78)	nefazodone, young female recurrent, Canada	imipramine fluoxetine	9 mo[c]	discounted direct cost per QALY	n>f>i Newer	1. multiple unreported probabilities 2. health state utility generation methods not described
Revicki et al., 1995 (79)	nefazodone, young female recurrent, U.S.	imipramine fluoxetine	9 mo[d]	discounted direct cost per QALY	n>f>i Newer	1. variability for utilities not reported
Einarsen et al., 1995 (81)	venlafaxine, input start	TCAs HCAs[e] SSRIs	300 d	direct cost per symptom-free day	v>H>T>S Newer	1. high delivery costs for TCAs 2. initial success rates overly different
Einarsen et al., 1995 (81)	venlafaxine, output start	TCAs HCAs[e] SSRIs	300 d	direct cost per symptom-free day	H>v>S>T Control	1. high delivery costs for TCAs 2. initial success rates overly different
Montgomery et al., 1996 (33)	nefazodone, U.K. primary care	imipramine	a = 12 wk m = 12 wk	direct cost per tx success	Newer	1–4. similar to (34,35)
Einarsen et al., 1997 (83)	venlafaxine, Ontario input start	TCAs SSRIs	6 mo	direct cost per tx success	v>f,T Newer	1. metaanalysis somewhat indirect 2. short treatment duration
Einarsen et al., 1997 (83)	venlafaxine, Ontario output start	TCAs SSRIs	6 mo	direct cost per tx success	v,f>T Newer	1. metaanalysis somewhat indirect 2. short treatment duration
Brown et al., 1999 (86)	mirtazapine, France, moderate/severe	amitriptyline	28 wk	direct cost per tx success, lost productivity	Newer	1. reliance on Delphic panel 2. brief duration
Brown et al., 1999 (85)	mirtazapine, Austria, moderate/severe	amitriptyline fluoxetine	~6 mo	direct cost per pt and tx success, lost productivity	m>a m>f Newer	1. mirtazapine vs. fluoxetine based on one study 2. reliance on Delphic panel 3. hospitalization estimates

Legend: see Table 78.4.

TABLE 78.5. COST-EFFECTIVENESS SIMULATION STUDIES COMPARING NEWER NON-SSRI WITH CONTROL ANTIDEPRESSANTS

A simple “vote counting” across studies in Table 78.4 shows that most of the simulations concluded that SSRIs are more cost-effective than TCAs (10 favor SSRIs, six favor TCAs, and two are ties). Table 78.5 shows that eight of nine simulations favor the cost-effectiveness of newer non-SSRI antidepressants over older agents. Simple vote counting is problematic because it ignores numerous methodologic limitations of the individual studies. At least equally as problematic is the fact that the methodologic problems may far exceed those that are apparent; often, the published models are not “transparent,” meaning that they fail to specify

clearly the inputs to the model and exactly how the computations were made. This is critical because, as we noted above, the results of any simulation are depend entirely on the many details of the model. When these details are not stated, it is not possible for the reader independently to look for errors in critical assumptions or independently to conduct “sensitivity analyses” of the values of input parameters to evaluate the resulting impact on the models’ conclusions.

Other concerns arise about the simulations as a consequence of their sensitivity analyses. Generally, these studies report that results are not sensitive to any of the variations that they show in inputs. This raises concerns because if a simulation is properly designed, and if it contains no calculation errors, it ought to be sensitive to at least extreme variations in some inputs. One approach is to show what input values would be required for the medications to “break even” (29). The reader can then come to an opinion about whether the break-even inputs are reasonable possibilities or not.

The input values required to reverse the cost-effectiveness conclusion may be unreasonably high or low, but demonstration that the model is sensitive to input variation increases confidence in the integrity of the model and in the reported lack of sensitivity to less extreme variations. For example, if it is not possible to demonstrate the cost-effectiveness of TCAs when the acquisition cost of SSRIs is increased 1,000-fold, something is wrong with the model. In many of the decision analytic simulations concluding that the newer antidepressants are more cost-effective (30 ,31 ,32 and 33), the design and assumptions were very similar to those in an early model of SSRIs versus TCAs (34 ,35). This simulation was reported very explicitly and so is transparent and could be replicated by others. When the model was replicated, a design flaw was discovered and unrealistic assumptions were identified that drove the results (29). Correction of the design flaw and substitution of longer treatment lengths recommended by practice guidelines reversed the findings and yielded a cost-effectiveness advantage in favor of the TCAs. These same corrections could be applied to the other simulations that depended on the early example. The Australian Pharmaceutical Benefits Scheme reported similar significant problems with 67% of the pharmacoeconomic simulations it received in support of efforts to meet regulatory requirements for registration of new drugs in that country (36).

This limitation of short time horizons for the simulations is relatively common in the studies shown in Table 78.4 and Table 78.5 . In general, the longer treatment with antidepressants is continued, the less cost-effective the newer antidepressants as first-line treatment are likely to be. A longer treatment period progressively increases the medication acquisition costs associated with newer antidepressants. By contrast, much of the greater cost of treatment delivery of the older drugs is expended early in treatment, in visits for dose titration and management of side effects. Longer treatment periods progressively dilute this early cost over time. Of the only two simulations in which sensitivity analysis of treatment length was performed (29 ,37), both showed cost-effectiveness advantages for the inexpensive drug as length of treatment length increased, and one simulation utilizing intermediate treatment lengths favored the TCAs (38).

Given the subtle but powerful effect of the many details of cost-effectiveness simulations, many have expressed concern that these simulations may harbor critical biases that are difficult to expose. For example, one leading journal has taken the stance that cost-effectiveness simulations are more vulnerable to conflict of interest than other types of research, and it declines to publish any cost-effectiveness simulations (39).

In this regard, it may be noted that all the 15 simulations funded by industry shown in Tables 78.4 and 78.5 reported their own products to be more cost-effective than older control antidepressants. Of the studies sponsored by companies manufacturing newer non-SSRI antidepressants, four of six found the SSRIs to be less cost-effective than or tied with older antidepressants, and four of six found their own products to be more cost-effective than SSRIs. In both of the two studies funded by government, the SSRIs were less cost-effective than the TCAs when provisions were made for patients intolerant to TCAs to switch. The source of funding was not identified in three studies.

RETROSPECTIVE ANALYSES OF ADMINISTRATIVE DATABASES

Retrospective administrative database studies are a source of data on antidepressant costs and efficacy in actual clinical practices. In these studies, computerized pharmacy and service utilization records are used to analyze cost outcomes as a function of clinical assignment to antidepressant. Retrospective studies are less expensive than prospective trials and can be conducted more quickly. However, they are much more vulnerable to questions about the interpretation of results for several reasons. These studies are vulnerable to “selection bias.” Patients are not randomly assigned to treatment; therefore, it is likely that the constructed groups may not be comparable in some important way at baseline. They are also vulnerable to “cohort effects.” The retrospective groups may be drawn from different time frames. Apparent differences between treatments may in fact be a consequence of changing trends in practice over time (40).

Database studies generally lack any direct measure of clinical outcome. As a result, they generally assume a worst case of equivalent outcomes for the newer antidepressants and the older antidepressants and then defined the more cost-effective care as the treatment associated with lower overall costs of health care. A newer antidepressant can be

associated with lower overall costs of health care if the higher acquisition costs are more than offset by lower costs for other services. This type of cost-effectiveness analysis is known as *cost minimization*.

A few of the administrative database studies have constructed proxy outcome measures based on pharmacy refill data, such as "number of prescriptions refilled" (41). For example, one study used pharmacy claims to determine the duration of antidepressant treatment and then held that longer care is likely to be more beneficial. This study found fluoxetine to be associated with longer continuation on medication and costs similar to those of the comparison groups. The authors concluded that fluoxetine is cost-effective because adherence to treatment guidelines is better with no increment in cost. Other retrospective analyses have reported similar natural course of therapy findings but base a judgment of cost-effectiveness on finding a reduction in overall "depression-related" health care costs (42 ,43).

We briefly review the designs and results of available retrospective administrative database studies in Table 78.6 , Table 78.7 , and Table 78.8 . Table 78.6 lists studies comparing SSRIs with older antidepressants. Table 78.7 lists studies making comparisons among SSRIs, and Table 78.8 presents one study comparing a newer non-SSRI with control antidepressants. These tables indicate for each study the sample size in the administrative database, the time interval over which data were sampled, the type of patient population, the newer and control antidepressants analyzed, the stated principal economic outcome measure, the overall results on that outcome measure as interpreted by the authors, and a brief discussion of methodologic limitations. One small pilot study is not included in Table 78.6 (44), and a published retrospective database study not included in Table 78.6 reportedly found fluoxetine to be cost-effective in comparison with TCAs (Skaer et al., 1996; cited in ref. 12).

Reference, *N*, time frame	Population	Newer AD(s)	Control AD(s)	Principal Outcome	Analysis Favors[b] detail	Methodologic Limitations
Sclar et al., 1998 (42) *N* = 550 7/1/88–12/31/91	U.S. HMO single-episode depressed patients initiating AD	fluoxetine	amitriptyline nortriptyline	depression-related 12-mo health care expenditures	f>T Newer	1. determination of depression relatedness of concern 2. very high doses of n 3. no breakdown of high TCA costs 4. possible cohort effect
Sclar et al., 1994 (87) *N* = 701 1/1/89–10/31/93	U.S. HMO depressed pts who stayed on initial AD for 12-mo	fluoxetine	amitriptyline nortriptyline desipramine	as for (42)	f>T Newer	1–4. as for (42) 5. possible uncorrected selection bias
Skaer et al., 1995 (88) *N* = 823 1/1/89–6/30/94	U.S. HMO single-episode depressed pts who stayed on initial AD for 12-mo	sertraline	amitriptyline nortriptyline desipramine	as for (42)	s>T Newer	1–5. as for (87)
Forder et al., 1996 (89) *N* = 398 time frame not stated	U.K. general practice pts selected from a clinical trial and matched controls	sertraline	TCAs	direct and partial indirect 12-mo cost per successfully treated pt	s>T Newer	1. subsample selection not random 2. shorter treatment length for s (168 d) than T (278 d) reduces cost of s 3. selection bias[a] or cohort effect
Croghan et al., 1997 (41,90) *N* = 1,242 1990–1992	U.S. privately insured depressed patients treated in primary care	fluoxetine	4 TCAs trazodone	total medical charges for 12-mo	fC Tie	1. high cost of other drugs for trazodone 2. possible uncorrected selection bias 3. possible cohort effect 4. ?effects of previous treatment 5. ?no adjustment for baseline costs
Thompson et al., 1998 (91) *N* = 1,661	New England private insurer depressed patients	SSRIs	TCAs	nonpsychiatric health care costs for 12-mo	Tie	1. possible uncorrected selection bias 2. possible cohort effect 3. ?effects of previous treatment 4. no adjustment for baseline costs
Hylan et al., 1998 (92) *N* = 2,693 1990–1994	U.S. fee-for-service depressed patients, insured by 20 large employers	fluoxetine sertraline paroxetine	TCAs	total health care costs for 12-mo	f>T 1/3 Newer	1. possible uncorrected selection bias 2. possible cohort effect 3. ?effects of previous treatment 4. no adjustment for baseline costs
Crown et al., 1998 (93) *N* = 3,439 1990–1994	U.S. fee-for-service depressed patients, insured by large employers	fluoxetine sertraline paroxetine	TCAs	total health care costs for 12-mo	S>T Newer	1. possible uncorrected selection bias 2. possible cohort effect on costs 3. possible effect of previous treatment on costs 4. no adjustment for baseline costs
Croghan et al., 2000 (46) *N* = 2,557 1990–1994	U.S. fee-for-service depressed patients, insured by ~20 large employers	fluoxetine sertraline	TCAs	psychiatric hospitalization for 12-mo	f>T sT 1/2 Newer	1. possible uncorrected selection bias 2. possible cohort effect 3. possible effect of previous treatment
Simon 1998 (47) *N* = 5,169 1/1/92–6/30/94	U.S. HMO depressed patients	fluoxetine	imipramine desipramine	total health care costs for 6-mo	fT Tie	1. possible uncorrected selection bias 2. possible cohort effect
Smith 1996 (94) *N* = 152 1994	U.S. HMO, ICD-9 major depression, received AD 3 mo, excluded if switched drug	SSRIs	TCAs	as for (87)	Tie	1–2. as for (87) 3. possible selection bias
Sullivan 2000 (95,96) *N* = 981 1993–1997	9 U.S. health plans second-line treatment	SSRIs	TCAs	total health care costs for 12-mo	Tie	1. possible uncorrected selection bias 2. possible cohort effect
Sclar 1999 (43) *N* = 1,339 1/1/96–4/30/99	as for (42)	fluoxetine sertraline paroxetine citalopram	amitriptyline	depression-related 6-mo health care expenditures	f,p,c>s,a 3/4 Newer	1. *N* = 71 for citalopram 2. possible uncorrected selection bias 3. possible cohort effect

Legend: see Table 78.4.

TABLE 78.6. RETROSPECTIVE COST-EFFECTIVENESS ANALYSES OF ADMINISTRATIVE DATABASES: SSRIs VERSUS OLDER ANTIDEPRESSANTS

Reference, *N*, time frame	Population	Newer AD(s)	Principal Outcome	Authors' Conclusions	Methodologic Limitations
Sclar et al., 1995 (97) *N* = 744 1/1/89–3/31/94	U.S. HMO single-episode depressed pts who stayed on initial AD and received at least 3 Rxs	fluoxetine sertraline paroxetine	as for (87)	f>s,p	1. possible uncorrected selection bias 2. possible cohort effect
Hylan et al., 1998 (92) *N* = 2,693 1990–1994	U.S. fee-for-service depressed patients, insured by 20 large employers	fluoxetine sertraline paroxetine	total health care costs for 12 mo	f>s	1. possible uncorrected selection bias 2. possible cohort effect 3. ?effects of previous treatment 4. no adjustment for baseline costs
Croghan et al., 2000 (46) *N* = 2,557 1990–1994	U.S. fee-for-service depressed patients, insured by ~20 large employers	fluoxetine sertraline	psychiatric hospitalization for 12 mo	f>s	1. possible uncorrected selection bias 2. possible cohort effect 3. possible effect of previous treatment
Russell et al., 1999 (98) *N* = 2,342 1995–1996	U.S. depressed patients contributing data to a publicly available claims database	fluoxetine sertraline paroxetine	total health care costs for 12 mo	f s p	1. possible uncorrected selection bias 2. possible effect of previous treatment
Sclar et al., 1999 (43) *N* = 1,339 1/1/96–4/30/99	as for (42)	fluoxetine sertraline paroxetine citalopram	depression-related 6 mo health care expenditures	f,p,c>s	1. *N* = 71 for citalopram 2. possible uncorrected selection bias 3. possible cohort effect

Legend: see Table 78.4.

TABLE 78.7. RETROSPECTIVE COST-EFFECTIVENESS ANALYSES OF ADMINISTRATIVE DATABASES: COMPARISONS AMONG SSRIs

Reference, *N*, time frame	Population	Newer AD(s)	Control AD(s)	Principal Outcome	Authors' Conclusions	Methodologic Limitations
Sullivan et al., 2000 (95,96) *N* = 981 1993–1997	9 U.S. health plans second-line treatment	venlafaxine other – nefazodone bupropion trazodone	TCAs SSRIs	total health care costs for 12 mo	v,o,S,T	1. possible uncorrected selection bias 2. possible cohort effect

Legend: see Table 78.4.

TABLE 78.8. RETROSPECTIVE COST-EFFECTIVENESS ANALYSES OF ADMINISTRATIVE DATABASES: NON-SSRIs VERSUS CONTROL ANTIDEPRESSANTS

Simple vote counting across the studies in Table 78.6 shows that the majority have concluded that SSRIs are more cost-effective than TCAs (seven favor at least one of the studied SSRIs, none favor TCAs, and five are ties). Again, simple vote counting is unsatisfactory because these studies are subject to numerous methodologic limitations.

The most important limitation of the studies in Table 78.6 is a possible cohort effect; the distribution of starts of different antidepressants may have changed during the study interval (40 ,45). During this time period, important influences on clinical practice totally unrelated to which antidepressant was used may have changed, so that the influence of starts on one type of antidepressant versus another may have been confounded with the effect of changes in clinical influences. For example, during the period encompassed by the first study in Table 78.6 , fluoxetine progressively gained market share, while at the same time health care organizations independently reduced expenditures through tighter management. Thus, a higher proportion of TCA starts may have occurred early in the study period, when care was not so firmly managed, and a higher proportion of fluoxetine starts may have occurred later in the study period, when visits and hospitalizations were more carefully scrutinized. Thus, cost savings in later years could erroneously be attributed to fluoxetine that are really a consequence of tighter management. The distribution of fluoxetine and TCA starts within the study period was not reported.

In relation to the problem of cohort effects, recent studies have included time of the antidepressant start within the study interval as an explanatory variable in the analysis, but they appear to have restricted attention primarily to its effect on initial selection of antidepressant. No study presents data indicating whether health care costs associated with antidepressant starts were increasing or decreasing during the study interval, or how a secular cost trend, if present, may have interacted with the distribution of starts of individual antidepressants during the study interval.

Other important limitations in the studies in Table 78.6 include a tendency not to adjust the analysis for baseline costs in some of them.

Table 78.7 shows the results of administrative database studies comparing cost-effectiveness among SSRIs. The first three studies, which sampled data from 1989 to 1994, found fluoxetine to be more cost-effective than sertraline or more cost-effective than sertraline and paroxetine. A type of selection bias that has been termed *launch bias* may have affected these findings (46). The time frames of these studies overlapped with the first year or two after launch of sertraline and paroxetine. It is possible that a new antidepressant is prescribed for a different type of patient in the early years after its launch than after it has been on the market for several years. Patients selected by their physicians to receive a brand-new antidepressant may generally be more severely ill than other patients. Recent analyses have attempted to control for initial severity. However, another possibility is that patients selected by their prescribers to receive a brand-new antidepressant may on average have been more resistant to previous treatment than other patients. Because treatment resistance correlates only partially with severity, adjustment for severity of illness may only partially correct a launch bias effect driven by treatment resistance. Most studies attempting to control for previous treatment eliminate patients who received antidepressants in the 4 to 6 months before the start of the index antidepressant. Although this exclusion is somewhat reassuring, it still does not prevent patients from being included in the analysis who were taking an antidepressant at the start of the study interval, then stopped antidepressant therapy for 4 to 6 months, perhaps because of ineffectiveness, then started a different, recently launched antidepressant identified as the index. Such patients would be predicted to be relatively unlikely to respond to treatment and relatively likely to incur treatment costs subsequent to the antidepressant start. To the extent that

such patients are included in an analysis, a launch bias may exist that is unfavorable to the second antidepressant.

In the first study in Table 78.7 , the sample was restricted to DSM-IV 296.2 “single-episode” depressed patients. This restriction was intended to reduce the possibility of a selection bias if the patients chosen to receive the newest medications were more refractory as a group than patients chosen to receive more established medications. This possibility may still have influenced the analysis because in subsequent studies, it was found that some patients with “single-episode” depression had had previous episodes that were treated with antidepressants (42 ,43).

The fourth study in Table 78.7 is consistent with a launch bias interpretation of the three earlier studies. In this study, the time horizon (1995 through 1996) was 3 to 4 years after the launch of sertraline and paroxetine. This study found the three SSRIs to be equally cost-effective. Limited retrospective data suggest that venlafaxine and other newer non-SSRI antidepressants are similar to SSRIs and TCAs in cost-effectiveness (Table 78.8).

The retrospective database method may be especially vulnerable to publication bias. Because the retrospective studies are inexpensive in comparison with prospective trials, and because the number of potential study sites is large, the possibility is greater with retrospective studies that multiple analyses are conducted but only a limited number published.

PROSPECTIVE COST-EFFECTIVENESS TRIALS

Part of "78 - Cost-Effectiveness of the Newer Generation of Antidepressants "

Prospective randomized cost-effectiveness experiments offer a potential “gold standard” methodology for investigating cost-effectiveness because of the internal validity arising from the randomization. In addition, they directly collect data on both outcomes and costs. The randomization permits the investigator to ascribe any observed differences in cost-effectiveness among treatment groups to the treatment itself and not to unmeasured baseline differences among the groups. The major difficulty with prospective randomized cost-effectiveness experiments, in addition to their expense and the time required to complete them, is the question of external validity. Are the patients who consent to random assignment representative of the entire group of patients in routine practice, or are they different in some important way? Relatively little attention has been paid to this issue in depression research, although in one study, the patients participating in a randomized trial of depression had significantly fewer comorbid diagnoses than did excluded patients and were more likely to have a single episode of depression (27).

At present, only two prospective pharmacoeconomic studies examining the cost-effectiveness of newer antidepressant treatment have been published. The initial report from the first study included data up to 6 months after randomization (4). Patients were followed for 2 years after randomization, and the long-term data were reported recently (26). Patients were enrolled from participating primary care clinics in a large HMO in the United States. Patient out-of-pocket copayment prescription expenses were waived. Patient identification depended on primary care physician referral. Physicians were asked to refer patients whom they were starting on an antidepressant for depression when both patient and physician were willing to consider random assignment. Of 621 patients referred, 579 (93%) were eligible, and 536 (93%) consented and were randomized. At baseline, 67% of randomized patients met DSM-III-R criteria for major depression; the remainder met criteria for either minor depression or dysthymia. The average score on the Hamilton Depression Scale at baseline was below 14. Patients were randomly assigned to receive fluoxetine (*N* = 173) or the commonly prescribed TCAs imipramine (*N* = 182) or desipramine (*N* = 181). After randomization, the patients were free to switch antidepressants. Evaluators but not patients or prescribing physicians were blinded to the initial treatment assignment. Randomized patients were evaluated at baseline and at 1, 3, 6, 9, 12, 18, and 24 months with measures of symptoms, quality of life, and service utilization.

At the 6-month follow-up, the proportion of patients continuing on the original antidepressant was nearly 60%

for fluoxetine, less than 40% for imipramine, and approximately 30% for desipramine. At the 24-month follow-up, the proportion of patients continuing on the original antidepressant was roughly 35% for fluoxetine and 10% to 15% for imipramine and desipramine. These data suggest a substantial acceptability advantage for the SSRI over the TCAs, at least when patients and prescribers are aware of the identity of the medication. However, the proportion of patients continuing to take any antidepressant medication was approximately equal at 6 months and at all the subsequent evaluations for the three groups. These data suggest that patients who find TCAs unacceptable generally agree to treatment with a second medication. Rates of symptoms and quality of life showed similar improvement at all time points, although some evidence was found at or near the trend level for the fluoxetine group to be slightly more improved at the 1-month time point only. These data indicate that the clinical outcomes in actual practice are essentially equivalent whether patients are initially assigned to an SSRI or a TCA. If average improvement is slightly faster when an SSRI is the initial choice, perhaps because fewer patients switch and start over, any difference is no longer apparent at 3 months or thereafter.

Among patients remaining on the initial antidepressant at 1 month, adverse effects were significantly lower in the group assigned to fluoxetine. The method of measurement of adverse effects is not described in detail. Differences in adverse events between the groups were not reflected in the measures of quality of life.

Cost-of-treatment data showed, as expected, that antidepressant medication costs were roughly double for the group initially assigned to fluoxetine ($217 vs. $97 for imipramine and $123 for desipramine during the first 6 months and $609 vs. $324 for imipramine and $376 for desipramine for the entire 24 months). Outpatient costs and inpatient medical and inpatient psychiatric costs were lower, although not significantly so, in the fluoxetine group ($1,750 vs. $2,008 for imipramine and $2,238 for desipramine during the first 6 months and $6,092 vs. $6,459 for imipramine and $6,381 for desipramine for the entire 24 months). These effects resulted in total direct costs across the groups that were not significantly different.

In this study, clinical outcomes were almost identical, whether patients were prescribed fluoxetine or a tricyclic first, when patients were permitted to switch medications freely. Patients who found tricyclics unacceptable generally agreed to treatment with a second medication and “caught up” in terms of clinical outcome. The newer medication “broke even” on costs, in that the higher acquisition costs of fluoxetine were balanced by the lower costs of other services, but did not result in an overall savings to the health care system. On the other hand, a formulary policy of requiring failure of a tricyclic before access to the newer medication was granted would not have saved money in this particular primary care practice. The authors concluded that the data provide no clear guidance in the initial selection of antidepressant medications and that patient and physician preference therefore provide an appropriate basis for treatment selection.

Interestingly, this group conducted a retrospective database analysis of patients during a similar period of time who did not participate in the randomized trial (47). Reassuringly, the cost-effectiveness results (Table 78.6) were very similar to those from the randomized study.

In the discussion, the authors point out that these conclusions may not apply to other practices. In particular, similar studies in psychiatric specialty practices are needed. The depression of patients in psychiatric specialty practices is generally more severe, and the consequences of delay in treatment response related to a need to switch medications and start over may be more worrisome.

The second randomized prospective antidepressant cost-effectiveness study was conducted in a primary care setting in France (2). Outpatients meeting DSM-IV criteria for major depression were randomized to sertraline (50 to 150 mg/d; *N* = 122) or fluoxetine (20 to 60 mg/d; *N* = 120) in double-blinded fashion for 6 months. Both groups improved significantly from baseline on measures of symptoms and quality of life, and analyses comparing the groups showed no significant differences; however, patients treated with fluoxetine utilized more medical resources. Analyses comparing groups in regard to work and productivity losses were not significant. Cost comparisons (converted to dollars) from the societal perspective favored sertraline over fluoxetine ($1,551 vs. $1,735), but neither the variability within groups nor statistical significance of the comparisons was reported.

COSTS OF AVERTING SUICIDE

Part of "78 - Cost-Effectiveness of the Newer Generation of Antidepressants "

Suicide is fortunately a fairly rare event, even in depressed patients (48 ,49). As a result, many cost-effectiveness studies have not included a consideration of suicide. We touch on the issue only briefly.

The relative safety of the newer antidepressants in overdose is well-known. The use of SSRIs is associated not only with fewer deaths from antidepressant overdose but also with reduced costs of treating overdose (50 ,51 and 52)

Despite clear reductions in mortality from overdoses with the newer medications in comparison with TCAs (53), controversial data from a general practice research database involving 4 million residents of the United Kingdom indicated the rate of death by suicide in patients receiving fluoxetine to be no lower than the rate of death by suicide in patients receiving TCAs (54). Fluoxetine-treated patients appeared to substitute violent methods or carbon monoxide poisoning for overdoses. Such complete method substitution would suggest that SSRIs do not save lives and that saving lives cannot therefore be used to justify their added

expense. These data, however, are based on only 11 suicides among a limited number of fluoxetine cases.

Freemantle et al. (55) modeled the cost per life year saved and cost per life saved that would result from the routine first-line use of SSRIs rather than TCAs in general practice in the U.K. in comparison with current U.K. practice patterns. This analysis produced a very wide range of estimates, primarily because of uncertainty regarding whether SSRI-treated patients would substitute other methods of suicide for overdoses. As part of a simulation, another study included suicide as part of the cost of treatment dropout (56). However, the expert panel estimates of suicide rates in this latter study were very high, and few costs were considered in the analysis.

Whether newer antidepressants do in fact save lives is crucial in a consideration of their cost-effectiveness. More data are needed about the extent to which patients, knowing that the pills are not lethal, might substitute methods of suicide that are even more deadly than TCA overdoses. However, recently emerging epidemiologic data appear to suggest that newer antidepressants may have a favorable impact on death by suicide when all methods, not just overdoses, are taken into account (57 ,58 ,59 and 60).

CONCLUSION

Part of "78 - Cost-Effectiveness of the Newer Generation of Antidepressants "

The available data that may be confidently brought to bear on the two cost-effectiveness questions posed in the introduction are surprisingly sparse. Only two prospective randomized studies have been carried out, both in primary care. More prospective studies are clearly needed. Most of the retrospective studies and the simulations contain methodologic limitations sufficient to generate significant concern about their conclusions. Additionally, the studies include diverse variations in almost all the elements of cost-effectiveness analysis, so that cross-comparisons and aggregate conclusions are very difficult to make. However, if we must draw conclusions from the current data, we would suggest the following tentative conclusions.

Based on the limited evidence available, the provisional summary from this review regarding our first question, “Are newer antidepressants cost-effective as first-line treatment from the health care system perspective?” is that first-line use of the newer antidepressants within primary care practice in the United States may be roughly equally effective and also cost-neutral in terms of direct medical resource costs to the health care system. The recently published long-term data from the only randomized study support this view (26), and the simulations and retrospective studies, with all their limitations, do not contradict it. However, because the data are sparse and contain multiple methodologic problems, health care organizations or systems feeling the pinch of the high acquisition costs of the newer medications would be well advised to conduct their own randomized studies. This is especially true for psychiatric practices and for practices in countries other than the United States.

The data for the primary care treatment of depression are sparse, but those for the cost-effectiveness of the newer antidepressants in psychiatric practice are even more scarce. No randomized studies have been published, and few of the simulations and none of the administrative database studies focus exclusively on psychiatric practice. Many features distinguish the treatment of depression in primary care from the treatment of depression in psychiatric practice, and these could potentially lead to different conclusions about cost-effectiveness. For example, the direct costs of treatment failure may be higher in psychiatric practice than in primary care (61), and this consideration would favor the cost-effectiveness advantages of the better tolerated newer agents. On the other hand, the tendency of psychiatrists to use higher and possibly more effective doses of TCAs than primary care prescribers do would likely favor the cost-effectiveness of TCAs over SSRIs in psychiatric practice. Similarly, it may be less likely in psychiatric practice than in primary care that the greater tolerability of SSRIs and the reduced requirement for dose titration would offset costs by decreasing the need for outpatient visits; for depressed patients in psychiatric practice, with relatively severe depression and higher levels of comorbidity, frequent visits may be necessary independently of these considerations to monitor for increased suicide risk. Prospective randomized cost-effectiveness experiments in psychiatric practice could address the substantially different environment of specialty mental health care.

Similarly, although the data on the treatment of depression in the United States are limited, those on the cost-effectiveness of the newer antidepressants in other countries, especially developing countries, are still more limited. No randomized studies outside the United States have compared newer and older antidepressants. Some of the simulations and none of the administrative database studies focus on other developed countries, such as Canada and the European nations. The many ways in which the treatment of depression differs across countries and economies could potentially lead to different conclusions about cost-effectiveness (62). Acquisition costs for the newer medications are generally lower in countries other than the United States (63). Nevertheless, price may still put the newer antidepressants out of reach for most of the population in some developing countries (64). The organization of health care systems varies greatly, and the potential of the newer antidepressants to offset costs could also vary greatly across countries. Prospective randomized cost-effectiveness experiments in countries other than the United States would make it possible to evaluate whether cost-effectiveness conclusions are widely applicable.

The second question we posed in the introduction was, “Are newer antidepressants cost-effective as first-line treatment from the global societal perspective?” This perspective

includes *all* costs and *all* health effects. Fewer studies have attempted to address societal indirect/productivity costs, and none of the studies is prospective. Indirect/productivity costs were not comprehensive in some studies, being limited to family burden or absence from work. In the studies reporting QALYs, the outcome rates were taken from expert opinion panels, or utility determinations were uncertain. Most of these studies have numerous other methodologic limitations. Better studies are needed, particularly on the substitution of suicide methods, enhancement of work productivity, and reduction of absenteeism and family burden. However, if the newer antidepressants are in fact health care resource cost-neutral in the health care system, the chance is significant that the newer antidepressants are cost-effective in society. Health care system health resource cost neutrality clearly suggests similar cost neutrality in total health resource costs to society because the total health care resource costs to society also are borne by the health care system. It is likely that society reaps benefits not seen from the health care system perspective, including decreased use of informal caregiver time, decreased use of patient time, and perhaps decreased use of resources other than health care resources, in addition to positive changes such as increased productivity.

Lastly, as newer antidepressants begin to come off patent, their cost eventually should go down as a result of generic competition. When this occurs, the cost-effectiveness of the newer medications will increase (45).

DISCLOSURE

Part of "78 - Cost-Effectiveness of the Newer Generation of Antidepressants "

Dr. Woods has received honoraria from Janssen Pharmaceutica and Wyeth-Ayerst Pharmaceuticals for speaking engagements and invited publications.

REFERENCES

1. Special report. Top 200 brand-name drugs by retails sales in 1999. *Drug Topics* 2000 (March 6):64.

2. Boyer P, Danion JM, Bisserbe JC, et al. Clinical and economic comparison of sertraline and fluoxetine in the treatment of depression. A 6-month double-blind study in a primary-care setting in France. *Pharmacoeconomics* 1998;13(1 Pt 2):157–169.

3. Murray CJL, Lopez AD. *The global burden of disease*. World Health Organization. Cambridge, MA: Harvard University Press, 1996.

4. Stoudemire A, Frank R, Hedemark N, et al. The economic burden of depression. *Gen Hosp Psychiatry* 1986;8:387–394.

5. Rice DP, Miller LS. The economic burden of affective disorders. *Br J Psychiatry Suppl* 1995; 27:34–42.

6. Gold MR, Siegel JE, Russell LB, et al., eds. *Cost-effectiveness in health and medicine*. New York: Oxford University Press, 1996.

7. Woods SW, Baker CB. Cost-effectiveness of newer antidepressants. *Curr Opin Psychiatry* 1997;10:95–101.

8. Jones MT, Cockrum PC. A critical review of published economic modeling studies in depression. *Pharmacoeconomics* 2000;17:555–583.

9. Rosenbaum JF, Hylan TR. Costs of depressive disorders: a review. *Depressive Disord* 1999;1:401–449.

10. Crott R, Gilis P. Economic comparisons of the pharmacotherapy of depression: an overview. *Acta Psychiatr Scand* 1998;97:241–252.

11. Holm KJ, Jarvis B, Foster RH. Mirtazapine—a pharmacoeconomic review of its use in depression. *Pharmacoeconomics* 2000;17:515–534.

12. Wilde MI, Benfield P. Fluoxetine—a pharmacoeconomic review of its use in depression. *Pharmacoeconomics* 1998;13:543–561.

13. Davis R, Wilde MI. Sertraline—a pharmacoeconomic evaluation of its use in depression—reply. *Pharmacoeconomics* 1997;11:378–380.

14. Wilde MI, Whittington R. Paroxetine—a pharmacoeconomic evaluation of its use in depression. *Pharmacoeconomics* 1995;8:62–81.

15. Joffe R, Sokolov S, Streiner D. Antidepressant treatment of depression: a metaanalysis [see Comments]. *Can J Psychiatry* 1996;41:613–616.

16. Einarson TR, Arikian SR, Casciano J, et al. Comparison of extended-release venlafaxine, selective serotonin reuptake inhibitors, and tricyclic antidepressants in the treatment of depression: a meta-analysis of randomized controlled trials. *Clin Ther* 1999;21:296–308.

17. Greenberg RP, Bornstein RF, Zborowski MJ, et al. A meta-analysis of fluoxetine outcome in the treatment of depression. *J Nerv Ment Dis* 1994;182:547–551.

18. Song F, Freemantle N, Sheldon TA, et al. Selective serotonin reuptake inhibitors: meta-analysis of efficacy and acceptability [see Comments]. *BMJ* 1993;306:683–687.

19. Montgomery SA, Henry J, McDonald G, et al. Selective serotonin reuptake inhibitors: meta-analysis of discontinuation rates [published erratum appears in *Int Clin Psychopharmacol* 1994;9:296]. *Int Clin Psychopharmacol* 1994;9:47–53.

20. Anderson IM, Tomenson BM. Treatment discontinuation with selective serotonin reuptake inhibitors compared with tricyclic antidepressants: a meta-analysis. *BMJ* 1995;310:1433–1438.

21. Lawrenson RA, Tyrer F, Newson RB, et al. The treatment of depression in UK general practice: selective serotonin reuptake inhibitors and tricyclic antidepressants compared. *J Affect Disord* 2000;59:149–157.

22. Dunn RL, Donoghue JM, Ozminkowski RJ, et al. Longitudinal patterns of antidepressant prescribing in primary care in the UK: comparison with treatment guidelines. *J Psychopharmacol* 1999;13:136–143.

23. Donoghue JM, Tyler A. The treatment of depression: prescribing patterns of antidepressants in primary care in the UK. *Br J Psychiatry* 1996;168:164–168.

24. MacDonald TM, McMahon AD, Reid IC, et al. Antidepressant drug use in primary care: a record linkage study in Tayside, Scotland [see Comments]. *BMJ* 1996;313:860–861.

25. Simon GE, VonKorff M, Heiligenstein JH, et al. Initial antidepressant choice in primary care. Effectiveness and cost of fluoxetine vs. tricyclic antidepressants. *JAMA* 1996;275:1897–1902.

26. Simon GE, Heiligenstein J, Revicki D, et al. Long-term outcomes of initial antidepressant drug choice in a “real world” randomized trial. *Arch Fam Med* 1999;8:319–325.

27. Partonen T, Sihvo S, Lonnqvist JK. Patients excluded from an antidepressant efficacy trial. *J Clin Psychiatry* 1996;57:572–575.

28. Woods SW, Ziedonis DM, Sernyak MJ, et al. Characteristics of participants and nonparticipants in medication trials for treatment of schizophrenia. *Psychiatr Serv* 2000;51:79–84.

29. Woods SW, Rizzo JA. Cost-effectiveness of antidepressant treatment reassessed [see Comments]. *Br J Psychiatry* 1997;170:257–263.

30. Bentkover JD, Feighner JP. Cost analysis of paroxetine vs. imipramine in major depression. *Pharmacoeconomics* 1996;8:223–232.

31. Lapierre Y, Bentkover J, Schainbaum S, et al. Direct costs of depression: analysis of treatment costs of paroxetine versus imipramine in Canada. *Can J Psychiatry* 1995;40:370–377.

32. Hylan TR, Kotsanos JG, Andersen JS, et al. Comparison of a decision analytic model with results from a naturalistic economic clinical trial: an application of evaluating alternative antidepressants. *Am J Managed Care* 1996;2:1211–1223.

33. Montgomery SA, Brown RE, Clark M. Economic analysis of treating depression with nefazodone vs. imipramine. *Br J Psychiatry* 1996;168:768–771.

34. Jonsson B, Bebbington P. Economic studies of the treatment of depressive illness. In: Jonsson B, Rosenbaum J, eds. *Health economics of depression*. New York: John Wiley and Sons, 1993:35–48.

35. Jonsson B, Bebbington PE. What price depression? The cost of depression and the cost-effectiveness of pharmacological treatment [see Comments]. *Br J Psychiatry* 1994;164:665–673.

36. Hill SR, Mitchell AS, Henry DA. Problems with the interpretation of pharmacoeconomic analyses—a review of submissions to the Australian pharmaceutical benefits scheme. *JAMA* 2000;283:2116–2121.

37. McFarland BH. Cost-effectiveness considerations for managed care systems: treating depression in primary care. *Am J Med* 1994;97[Suppl 6A]:47S–58S.

38. Stewart A. Anti-depressant pharmacotherapy: cost comparison of SSRIs and TCAs. *Br J Med Econ* 1994;7:67–79.

39. Kassirer JP, Angell M. The journal's policy on cost-effectiveness analyses [Editorial] [see Comments]. *N Engl J Med* 1994;331:669–670.

40. Woods SW, Baker CB. Cost effectiveness of newer antidepressants. *Curr Opin Psychiatry* 1997;10:95–101.

41. Croghan TW, Lair TJ, Engelhart L, et al. Effect of antidepressant therapy on health care utilization and costs in primary care. *Psychiatr Serv* 1997;48:1420–1426.

42. Sclar DA, Skaer TL, Robison LM, et al. Economic outcomes with antidepressant pharmacotherapy: a retrospective intent-to-treat analysis. *J Clin Psychiatry* 1998;59:13–17.

43. Sclar DA, Skaer TL, Robison LM, et al. Economic appraisal of citalopram in the management of single-episode depression. *J Clin Psychopharmacol* 1999;19[5 Suppl 1]:47S–54S.

44. Melton ST, Kirkwood CK, Farrar TW, et al. Economic evaluation of paroxetine and imipramine in depressed outpatients. *Psychopharmacol Bull* 1997;33:93–100.

45. Woods SW, Baker CB. Cost-effectiveness of treatment for depression: methods and policies. *Depressive Disord* 1999;1:456–458.

46. Croghan TW, Kniesner TJ, Melfi CA, et al. Effect of antidepressant choice on the incidence and economic intensity of hospitalization among depressed individuals. *Administration and Policy in Mental Health* 2000;27:183–195.

47. Simon GE, Fishman P. Cost implications of initial antidepressant selection in primary care. *Pharmacoeconomics* 1998;13:61–70.

48. Simon GE, VonKorff M. Suicide mortality among patients treated for depression in an insured population. *Am J Epidemiol* 1998;147:155–160.

49. Blair-West GW, Mellsop GW, Eyeson-Annan ML. Down-rating lifetime suicide risk in major depression [see Comments]. *Acta Psychiatr Scand* 1997;95:259–263.

50. D'Mello DA, Finkbeiner DS, Kocher KN. The cost of antidepressant overdose. *Gen Hosp Psychiatry* 1995;17:454–455.

51. Revicki DA, Palmer CS, Phillips SD, et al. Acute medical costs of fluoxetine versus tricyclic antidepressants. A prospective multicentre study of antidepressant drug overdoses. *Pharmacoeconomics* 1997;11:48–55.

52. Kapur N, House A, Creed F, et al. Costs of antidepressant overdose: a preliminary study. *Br J Gen Pract* 1999;49:733–734.

53. Henry JA, Alexander CA, Sener EK. Relative mortality from overdose of antidepressants [see Comments] [published erratum appears in *BMJ* 1995;310:911]. *BMJ* 1995;310:221–224.

54. Jick SS, Dean AD, Jick H. Antidepressants and suicide. *BMJ* 1995;310:215–218.

55. Freemantle N, House A, Song F, et al. Prescribing selective serotonin reuptake inhibitors as strategy for prevention of suicide. *BMJ* 1994;309:249–253.

56. Le Pen C, Levy E, Ravily V, et al. The cost of treatment dropout in depression. A cost-benefit analysis of fluoxetine vs. tricyclics. *J Affect Disord* 1994;31:1–18.

57. Ohberg A, Vuori E, Klaukka T, et al. Antidepressants and suicide mortality. *J Affect Disord* 1998;50:225–233.

58. Isacsson G, Holmgren P, Druid H, et al. The utilization of antidepressants—a key issue in the prevention of suicide: an analysis of 5281 suicides in Sweden during the period 1992–1994. *Acta Psychiatr Scand* 1997;96:94–100.

59. Barbui C, Campomori A, D'Avanzo B, et al. Antidepressant drug use in Italy since the introduction of SSRIs: national trends, regional differences and impact on suicide rates. *Soc Psychiatry Psychiatr Epidemiol* 1999;34:152–156.

60. Rihmer Z, Rutz W, Pihlgren H, et al. Decreasing tendency of seasonality in suicide may indicate lowering rate of depressive suicides in the population. *Psychiatry Res* 1998;81:233–240.

61. Baker CB, Woods SW. The cost of treatment failure for major depression: direct costs of continued treatment. *Administration and Policy in Mental Health* 2001;28:263–277.

62. Shah A, Jenkins R. Mental health economic studies from developing countries reviewed in the context of those from developed countries. *Acta Psychiatr Scand* 2000;101:87–103.

63. Sasich L, Torrey EF, Wolfe SM, et al. Average cost of newer medication psychotropic, antidepressant drugs twice as high in US as in Europe, study finds. *Psychiatr Serv* 1998;49:1248.

64. Razali SM, Hasanah CI. Cost-effectiveness of cyclic antidepressants in a developing country. *Aust N Z J Psychiatry* 1999;33:283–284.

65. Hotopf M, Hardy R, Lewis G. Discontinuation rates of SSRIs and tricyclic antidepressants: a meta-analysis and investigation of heterogeneity [see Comments]. *Br J Psychiatry* 1997;170:120–127.

66. Steffens DC, Krishnan KR, Helms MJ. Are SSRIs better than TCAs? Comparison of SSRIs and TCAs: a meta-analysis. *Depression Anxiety* 1997;6:10–18.

67. Anderson IM. SSRIs versus tricyclic antidepressants in depressed inpatients: a meta-analysis of efficacy and tolerability. *Depression Anxiety* 1998;7[Suppl 1]:11–17.

68. Trindade E, Menon D, Topfer LA, et al. Adverse effects associated with selective serotonin reuptake inhibitors and tricyclic antidepressants: a meta-analysis [see Comments]. *Can Med Assoc J* 1998;159:1245–1252.

69. Mulrow CD, Williams JW, Trivedi M, et al. Evidence report on treatment of depression—newer pharmacotherapies. *Psychopharmacol Bull* 1998;34:409–795.

70. Bech P, Cialdella P, Haugh MC, et al. Meta-analysis of randomised controlled trials of fluoxetine vs. placebo and tricyclic antidepressants in the short-term treatment of major depression. *Br J Psychiatry* 2000;176:421–428.

71. Geddes JR, Freemantle MN, Mason J, et al. SSRIs versus other antidepressants for depressive disorder. *Cochrane Database of Systematic Reviews* 2000:1.

72. Anderson IM. Selective serotonin reuptake inhibitors versus tricyclic antidepressants: a meta-analysis of efficacy and tolerability. *J Affect Disord* 2000;58:19–36.

73. Williams JW, Mulrow CD, Chiquette E, et al. A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. *Ann Intern Med* 2000;132:743–756.

74. Stahl S, Zivkov M, Reimitz PE, et al. Meta-analysis of randomized, double-blind, placebo-controlled, efficacy and safety studies of mirtazapine versus amitriptyline in major depression. *Acta Psychiatr Scand Suppl* 1997;391:22–30.

75. Srisurapanont M. Response and discontinuation rates of newer antidepressants: a meta-analysis of randomized controlled trials in treating depression. *J Med Assoc Thailand* 1998;81:387–392.

76. Boyer WF, Feighner JP. The financial implications of starting treatment with a selective serotonin reuptake inhibitor or tricyclic antidepressant in drug-naïve depressed patients. In: Jonsson B, Rosenbaum J, eds.*Health economics of depression*. New York: John Wiley and Sons, 1993:65–75.

77. Hatziandreu EJ, Brown RE, Revicki DA, et al. Cost utility of maintenance treatment of recurrent depression with sertraline versus episodic treatment with dothiepin. *Pharmacoeconomics* 1994;5:249–264.

78. Anton SF, Revicki DA. The use of decision analysis in the pharmacoeconomic evaluation of an antidepressant: a cost-effectiveness study of nefazodone. *Psychopharmacol Bull* 1995;31:249–258.

79. Revicki DA, Brown RE, Palmer W, et al. Modelling the cost effectiveness of antidepressant treatment in primary care. *Pharmacoeconomics* 1995;8:524–540.

80. Nuijten MJC, Hardens M, Soutre E. A Markov process analysis comparing the cost effectiveness of maintenance therapy with citalopram versus standard therapy in major depression. *Pharmacoeconomics* 1995;8:159–168.

81. Einarson TR, Arikian S, Sweeney S, et al. A model to evaluate the cost-effectiveness of oral therapies in the management of patients with major depressive disorders. *Clin Ther* 1995;17:136–153.

82. Woods SW, Rizzo JA. Cost-effectiveness of antidepressant treatment reassessed [see Comments]. *Br J Psychiatry* 1997;170:257–263.

83. Einarson TR, Addis A, Iskedjian M. Pharmacoeconomic analysis of venlafaxine in the treatment of major depressive disorder. *Pharmacoeconomics* 1997;12:286–296.

84. Canadian Coordinating Office for Health Technology Assessment. Selective serotonin uptake inhibitors (SSRIs) for major depression. Part II: The cost effectiveness of SSRIs in treatment of depression. August 1997. Report 4E. *Evidence-Based Med* 1998;3:87.

85. Brown MC, Nimmerrichter AA, Guest JF. Cost-effectiveness of mirtazapine compared to amitriptyline and fluoxetine in the treatment of moderate and severe depression in Austria. *Eur Psychiatry* 1999;14:230–244.

86. Brown MCJ, van Loon JMT, Guest JF. Cost-effectiveness of mirtazapine relative to amitriptyline in the treatment of moderate and severe depression in France. *Eur J Psychiatry* 1999;13:197–208.

87. Sclar DA, Robison LM, Skaer TL, et al. Antidepressant pharmacotherapy: economic outcomes in a health maintenance organization. *Clin Ther* 1994;16:715–730.

88. Skaer TL, Sclar DA, Robison LM, et al. Economic evaluation of amitriptyline, desipramine, nortriptyline, and sertraline in the management of patients with depression. *Curr Ther Res* 1995;56:556–567.

89. Forder J, Kavanagh S, Fenyo A. A comparison of the cost-effectiveness of sertraline versus tricyclic antidepressants in primary care. *J Affect Disord* 1996;38:97–111.

90. Obenchain RL, Melfi CA, Croghan TW, et al. Bootstrap analyses of cost effectiveness in antidepressant pharmacotherapy. *Pharmacoeconomics* 1997;11:464–472.

91. Thompson D, Hylan TR, McMullen W, et al. Predictors of a medical-offset effect among patients receiving antidepressant therapy. *Am J Psychiatry* 1998;155:824–827.

92. Hylan TR, Crown WH, Meneades L, et al. Tricyclic antidepressant and selective serotonin reuptake inhibitors antidepressant selection and health care costs in the naturalistic setting: a multivariate analysis. *J Affect Disord* 1998;47:71–79.

93. Crown WH, Hylan TR, Meneades L. Antidepressant selection and use and healthcare expenditures—an empirical approach. *Pharmacoeconomics* 1998;13:435–448.

94. Smith W, Sherrill A. A pharmacoeconomic study of the management of major depression: patients in a TennCare HMO. *Med Interface* 1996;9:88–92.

95. Sullivan EM, Griffiths RI, Frank RG, et al. One-year costs of second-line therapies for depression. *J Clin Psychiatry* 2000;61:290–298.

96. Griffiths RI, Sullivan EM, Frank RG, et al. Medical resource use and cost of venlafaxine or tricyclic antidepressant therapy following selective serotonin reuptake inhibitor therapy for depression. *Pharmacoeconomics* 1999;15:495–505.

97. Sclar DA, Robison L, Skaer TL, et al. Antidepressant pharmacotherapy: economic evaluation of fluoxetine, paroxetine, and sertraline in a health maintenance organization. *J Int Med Res* 1995;23:395–412.

98. Russell JM, Berndt ER, Miceli R, et al. Course and cost of treatment for depression with fluoxetine, paroxetine, and sertraline. *Am J Managed Care* 1999;5:597–606.

79

Mechanism of Action of Antidepressants and Mood Stabilizers

Robert H. Lenox

Alan Frazer

Robert H. Lenox: Head CNS Group, Aventis Pharmaceuticals, Bridgewater, New Jersey.

Alan Frazer: Department of Pharmacology, University of Texas Health Science Center, San Antonio, Texas.

Bipolar disorder (BPD), the province of mood stabilizers, has long been considered a recurrent disorder. For more than 50 years, lithium, the prototypal mood stabilizer, has been known to be effective not only in acute mania but also in the prophylaxis of recurrent episodes of mania and depression. By contrast, the preponderance of past research in depression has focused on the major depressive episode and its acute treatment. It is only relatively recently that investigators have begun to address the recurrent nature of unipolar disorder (UPD) and the prophylactic use of long-term antidepressant treatment. Thus, it is timely that we address in a single chapter the most promising research relevant to the pharmacodynamics of both mood stabilizers and antidepressants.

As we have outlined in Fig. 79.1 , it is possible to characterize both the course and treatment of bipolar and unipolar disorder in a similar manner. Effective treatments exist for the acute phases of both disorders; maintaining both types of patients on such drugs on a long-term basis decreases the likelihood and intensity of recurrences. Further, because the drugs are given long-term, they produce a cascade of pharmacologic effects over time that are “triggered” by their acute effects. Both classes of psychotropic drugs incur a lag period for therapeutic onset of action, even in the acute phase; therefore, studies during the past two decades have focused on the delayed (subchronic) temporal effects of these drugs over days and weeks. Consequently, it is widely thought that the delayed pharmacologic effects of these drugs are relevant for either the *initiation* of behavioral improvement or the *progression* of improvement beyond that initiated by acute pharmacologic actions. The early realization that lithium is effective prophylactically in BPD and the more recent understanding that antidepressants share this property in UPD have focused research on long-term events, such as alterations in gene expression and neuroplasticity, that may play a significant role in stabilizing the clinical course of an illness. In our view, behavioral improvement and stabilization stem from the acute pharmacologic effects of antidepressants and mood stabilizers; thus, both the acute and longer-term pharmacologic effects of both classes of drugs are emphasized in this chapter.

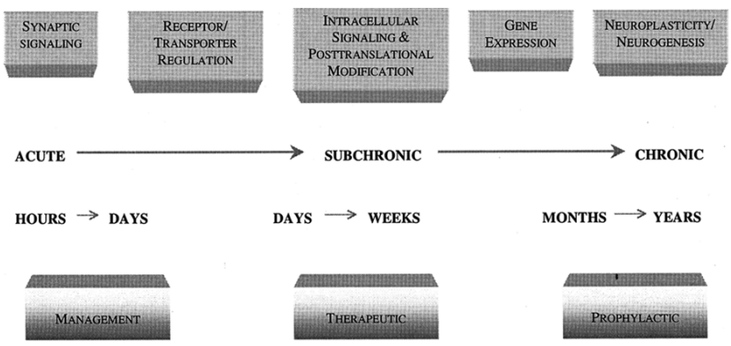


FIGURE 79.1. Mood stabilizers and antidepressant actions: short-term and long-term events. Lithium and antidepressants have acute effects on synaptic signaling that serve to trigger progressively longer-term events in signal transduction; these in turn lead to changes in gene expression and plastic changes in brain. The acute affects in critical regions of the brain result in changes in certain behavioral and physiologic symptomatology (e.g., activation, sleep, appetite) that facilitate the acute clinical management of mania or depression. Subchronic effects lead to amelioration of symptoms related directly to mood, whereas it is thought that the longer-term (chronic) effects underlie the prophylactic properties of these drugs to prevent recurrent affective episodes in both unipolar and bipolar disorders.

MOOD STABILIZERS

Part of "79 - Mechanism of Action of Antidepressants and Mood Stabilizers "

The term *mood stabilizer* within the clinical setting is commonly used to refer to a class of drugs that treat BPD. However, for the purpose of our discussion, it is important to differentiate the three clinical phases of BPD—acute mania, acute depression, and long-term prophylactic treatment for recurrent affective episodes. Although a variety of drugs are used to treat BPD (i.e., lithium, anticonvulsants, antidepressants, benzodiazepines, neuroleptics), we suggest that only a drug with properties of prophylaxis should be referred to as a mood stabilizer and included in this chapter. Significant evidence supports a therapeutic action for lithium, both in acute mania and prophylactically in a major subset of patients with BPD 1. However, the data for long-term prophylaxis with anticonvulsants (i.e., valproate, carbamazepine), although supported in part in clinical practice, remains less well established scientifically (see Chapter 77). In the absence of a suitable animal model, an experimental approach, used to ascribe *therapeutic relevance* to any observed biochemical finding, is the identification of shared biochemical targets that are modified by drugs belonging to the same therapeutic class (e.g., antimanic agents) but possessing distinct chemical structures (e.g., lithium and valproate). Although unlikely to act via identical mechanisms, such common targets may provide important clues

about molecular mechanisms underlying mood stabilization in the brain. Thus, in our discussion, we use studies of lithium as a prototypal mood stabilizer and cross-reference evidence for the anticonvulsants when the data are available. Furthermore, it is important to note that drugs that are useful in the treatment of acute mania or depression may not necessarily have prophylactic properties (1) and, as in the case of antidepressants that are effective in treating BPD, may actually serve to destabilize the illness. Although it is likely that the targets for lithium action early in treatment trigger its long-term properties of mood stabilization, to what extent the biological mechanisms underlying long-term lithium prophylaxis contribute to the efficacy of lithium in acute mania remain to be demonstrated.

Studies through the years have proposed multiple sites for the action of lithium in the brain, and such research has paralleled advances in the field of neuroscience and the experimental strategies developed during the past half-century. For the most part, proper interpretation of these data has at times been limited by experimental design, which has often ignored not only the clinically relevant therapeutic range of concentrations and onset of action of lithium, but also critical control studies defining its specificity of action in comparison with other monovalent cations and classes of psychopharmacologic agents. While the targets for the action of lithium have shifted from ion transport and presynaptic neurotransmitter-regulated release to postsynaptic receptor regulation, to signal transduction cascades, to gene expression and neuroplastic changes in the neuropil, the research strategy has evolved from a focus on a class of neurotransmitter to the ability of the monovalent cation to alter the pattern of signaling in critical regions of the brain in a unique manner. It is in this context that we highlight the most current thinking regarding putative sites for the therapeutic action of lithium in the brain, which is heuristic and sets the stage for future research directions.

Ion Transport

Ion-gated channels, which are driven by either adenosine triphosphate (ATP) or the net free energy of transmembrane concentration gradients, regulate the distribution of lithium across the cell membrane. These transport systems are critical for the regulation of resting lithium in the bulk cytoplasm in that they regulate steady-state intracellular ion concentrations that set the threshold for depolarization in excitable cells. Lithium exchanges readily with sodium; however, by virtue of its high energy of hydration, it can also substitute for the divalent cations calcium and magnesium, which may account for some of its major biochemical sites of action. Much of the anticonvulsant properties of valproate, carbamazepine, and lamotrigine have been attributed to their ability to inhibit sustained repetitive firing by prolonging the recovery of voltage-gated sodium channels from inactivation (2). However, it is important to note that anticonvulsant activity appears to be neither necessary nor sufficient for mood stabilization because lithium has proconvulsant properties outside its narrow therapeutic range.

Although some membrane transport systems specifically recognize lithium and regulate its transmembrane concentration (e.g., a gradient-dependent Na-Li exchange process) (3 ,4), it is likely that the primary regulation of lithium is affected by transport systems that accept the lithium ion as a substitute for their normal ionic substrates. The Na, K-ATPase pump has been extensively studied in relation to the membrane transport of lithium and the therapeutic effect of lithium (see refs. 5 and 6 for review). Based on measurements of lithium in peripheral neurons and synaptosomal membrane fractions from brain, long-term lithium treatment was found to decrease Na, K-ATPase activity, particularly in hippocampus (7). Various groups have studied Na, K-ATPase activity in patients with mood disorders and have reported alterations in the erythrocyte-to-plasma ratio of lithium in patients with BPD as a function of clinical state and genetic loading. Despite the fact that clinical studies through the years have been constrained by relatively small and often variable findings, evidence has been found that Na, K-ATPase activity may be reduced, especially in the depressed phase of both UPD and BPD, and is associated with an increase in sodium retention (see refs. 1 ,6 for review). Furthermore, long-term lithium treatment has been observed to result in an increased accumulation of lithium and activity of Na, K-ATPase in erythrocyte membranes, with concomitant reduction of sodium and calcium within erythrocytes in patients with BPD. Because the concentration of free calcium ion tends to parallel the concentration of free sodium ion, this finding may account for observations that intracellular calcium is increased in patients with BPD (8). Interestingly, when patients with BPD were treated with lithium, Na, K-ATPase activity was found to be increased, consistent with observations of reduced Ca^{2+} after treatment. However, such evidence from blood cells must be interpreted with caution; more recent data support the evolution of specific gene products for Na, K-ATPase expressed and uniquely regulated after translation, not only in neurons but in brain regions (9 ,10).

Although a balance of resting lithium conductance and net transport/efflux mechanisms regulates lithium homeostasis, the ligand gating of ion channels on the time scale of channel activity may play a more significant role in the regulation of intracellular lithium concentration within regulatory sites of an excitable cell such as the neuron. In the local environment of a dendritic spine, the surface area-to-volume ratio becomes relatively large, such that the lithium component of a synaptic current may result in significant (as much as fivefold to 10-fold) increases in intracellular lithium concentration following a train of synaptic stimuli (11). Such an activity-dependent mechanism for creating focal, albeit transient, increases of intracellular lithium at sites of high synaptic activity may play a role in the therapeutic specificity of lithium and its ability to regulate synaptic function in the brain.

Neurotransmitter Signaling/Circadian Rhythm

In search of a link between the mechanism of action of lithium and neurotransmission, the effect of lithium has been extensively studied in virtually every neurotransmitter system. Earlier studies focused on the modulation of presynaptic components, including the synthesis, release, turnover, and reuptake of neurotransmitters. In recent years, the focus has shifted to postsynaptic events, such as the regulation of signal transduction mechanisms (see refs. 12 ,13 and 14 for review). Despite the fact that some of the results of the presynaptic and postsynaptic investigations are not in full agreement, at present the evidence supports the action of lithium at multiple sites that modulate neurotransmission. Lithium appears to reduce presynaptic dopaminergic activity and acts postsynaptically to prevent the development of receptor up-regulation and supersensitivity. In the cholinergic system, lithium enhances receptor-mediated responses at neurochemical, electrophysiologic, and behavioral levels. Long-term lithium treatment increases GABAergic inhibition and has been shown to reduce excitatory glutamatergic neurotransmission. It is of interest that valproate has been shown to enhance γ-aminobutyric acid (GABA) signaling, and the anticonvulsant lamotrigine has been shown to reduce glutamatergic neurotransmission. (2). It is currently thought that the effect of lithium on the spectrum of neurotransmitter systems may be mediated through its action at intracellular sites, with the net effect of long-term lithium attributed to its ability to alter the balance among neurotransmitter/neuropeptide signaling pathways.

One of the unique and most robust properties of lithium is its ability to lengthen the circadian period across species—unicellular organisms, plants, invertebrates, and vertebrates (including primates)—so that a phase delay in the circadian cycle often results (see refs. 15 ,16 for review). These effects are noted following long-term but not acute exposure and occur within the range of concentrations used in humans to treat BPD (0.6 to 1.2 m*M*). It has long been recognized that a dysregulation of circadian rhythms is associated with the clinical manifestation of recurrent mood disorders in patient populations (see refs. 17 ,18 for review). In fact, it appears to be the interaction between the circadian pacemaker and the sleep–wake cycle that determines variations in sleepiness, alertness, cognitive performance, and mood (19 ,20 ,21 and 22). The early morning awakening, shortened latency in rapid-eye-movement (REM) sleep, and advances in hormonal and temperature regulation of many depressed patients, including those with BPD, are thought by some investigators to indicate a phase advance of the central pacemaker within the suprachiasmatic nucleus of the hypothalamus relative to other internal oscillators or external zeitgebers (23 ,24 ,25 ,26 and 27). Lithium may achieve its therapeutic and prophylactic effects by altering the balance of neurotransmitter signaling in critical regions of the brain, such as the

hypothalamus, and resynchronizing the physiologic systems underlying recurrent affective illness (1 ,28 ,29 and 30).

Signal Transduction

Phosphoinositide Cycle

Since it was discovered that lithium is a potent inhibitor of the intracellular enzyme inositol monophosphatase (IMPase) (K_i = 0.8 m*M*), which converts inositol monophosphate to inositol (31 ,32), receptor G protein-coupled phosphoinositide (PI) hydrolysis has been extensively investigated as a site for the action of lithium as a mood stabilizer (see ref. 33 for review) (Fig. 79.2). The "inositol-depletion hypothesis" posited that lithium produces its therapeutic effects via a depletion of neuronal *myo*-inositol levels. Furthermore, because the mode of enzyme inhibition of IMPase is uncompetitive, likely through interaction with Mg^{2+} binding sites (34), the preferential site of action for lithium was proposed to be on the most overactive receptor-mediated neuronal pathways undergoing the highest rate of phosphatidylinositol 4,5 bisphosphate (PIP_2) hydrolysis (35 ,36). It is also of interest that a number of structurally similar phosphomonoesterases that require magnesium have also been found to be inhibited by lithium at K_i values below 1 m*M* (37 ,38).

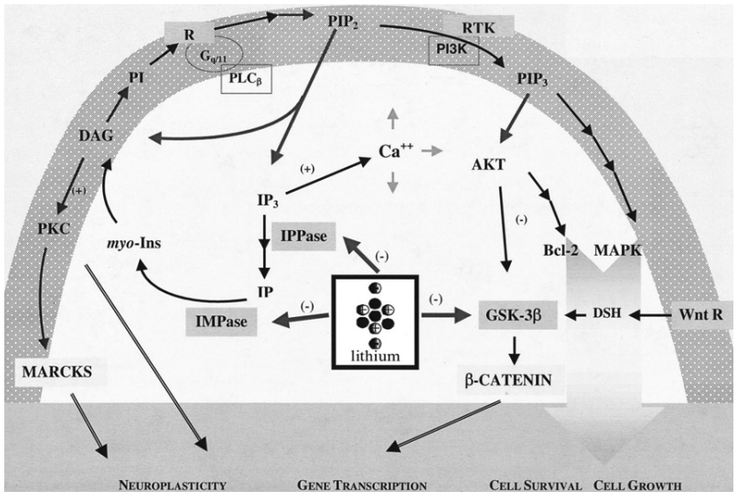


FIGURE 79.2. Molecular targets for lithium in phosphoinositide (PI) signaling. Pathways depicted within the figure are three major sites for an inhibitory action of lithium: inositol 1-monophosphatase (IMPase); inositol polyphosphate 1-phosphatase (IPPase); and glycogen synthase kinase 3ß (GSK-3ß). Inhibition of IMPase and IPPase can result in a reduction of *myo*-inositol (*myo*-Ins) and subsequent changes in the kinetics of receptor-activated phospholipase C (PLC) breakdown of phosphoinositide-4,5-bisphosphate to diacylglycerol (DAG) and inositol-1,4,5-trisphosphate. Alteration in the distribution of inositol phosphates can affect mechanisms mediating presynaptic release. DAG directly activates protein kinase C (PKC), and this activation results in downstream post-translational changes in proteins that affect receptor complexes and ion channel activity and in transcription factors that alter gene expression of proteins such as MARCKS (myristoylated alanine-rich C-kinase substrate), which are integral to long-term neuroplastic changes in cell function. Inhibition of GSK-3ß within the wnt-receptor (wnt-R) pathway alters gene transcription and neuroplastic events through an increased expression of downstream proteins such as ß-catenin. In addition, this inhibition can indirectly affect phosphoinositide 3 kinase pathways and intermediate factors (e.g., Bcl-2 and MAP kinases), which are thought to mediate cell growth and survival.

In cell systems and in cerebral cortical slices of chronically

treated rats, the effects of lithium on receptor-coupled PI signaling (39 ,40 ,41 and 42) and the down-regulation of myristoylated alanine-rich C-kinase substrate (MARCKS) protein (discussed below) can be prevented or reversed by a high concentration of *myo*-inositol. Recent genetic data from *Drosophila* indicate a role for the upstream inositol polyphosphatase (IPPase) as an additional target for lithium (43) (Fig. 79.2). *Drosophila* harboring a null mutation for the IPPase gene demonstrate aberrant firing of the neuromuscular junction, an effect that is mimicked by the treatment of wild-type flies with lithium. Although studies during the past several years have provided evidence that *myo*-inositol clearly plays a role in the action of lithium, it is evident that lithium-induced *myo*-inositol reduction may depend on cell type (39) and that sites other than PI signaling may be lithium targets, depending on the physiologic system under investigation. In studies examining the *in vivo* physiologic effects of lithium, such as polyuria or enhancement of cholinergically induced seizures, the addition of *myo*-inositol reduced but did not fully reverse the lithium-induced effects (44 ,45). Furthermore, the effect of long-term lithium on developmental polarity in the *Xenopus* embryo is rescued in the presence of *myo*-inositol (46), but this effect may not be totally attributable to a direct effect of lithium on IMPase (47) (see below).

Although the lithium-induced reduction in agonist-stimulated PIP_2 hydrolysis in rat brain slices has often been small and inconsistent, probably secondary to the size of the signaling-dependent PIP_2 pool (33 ,48), a recent study of patients with BPD in which proton magnetic resonance spectroscopy was used has demonstrated a significant lithium-induced reduction in *myo*-inositol levels in the right frontal lobe (49). However, the reduction in *myo*-inositol preceded the improvement in mood symptoms, indicating a temporal dissociation between changes in *myo*-inositol and clinical improvement. Consequently, these and other studies suggest that although inhibition of IMPase may represent an initial effect of lithium, reducing *myo*-inositol levels *per se* may be more important in the specificity of the cellular site of action for lithium than in the actual therapeutic response, which may be mediated by a cascade of downstream changes in signal transduction and gene expression (see below).

Adenylyl Cyclase

The other major receptor-coupled second-messenger system in which lithium has been shown to have significant effects is the adenylyl cyclase system. The cyclic AMP (cAMP) generating system plays a major role in the regulation of neuronal excitability and has been implicated in the pathophysiology of seizure disorders (50 ,51 and 52) and BPD (53). Studies in a variety of cell systems and in human brain have demonstrated that lithium attenuates receptor-coupled activation of the cAMP pathway at concentration that inhibits 50% (IC_{50}) values ranging from 1 to 5m*M* (see ref. 1 for review). Lithium *in vitro* inhibits adenylyl cyclase activity stimulated by guanosine triphosphate (GTP) or calcium/calmodulin, both of which interact directly with adenylyl cyclase (54 ,55 and 56). These inhibitory effects of lithium are antagonized by Mg^{2+}, which suggest that the action of lithium on the adenylyl cyclase system is mediated by direct competition with Mg^{2+} (55). However, attenuation of adenylyl cyclase activity following long-term lithium treatment in rat cortical membranes was not antagonized by Mg^{2+} alone but was reversed by increasing concentrations of GTP, which implies that the effect of long-term lithium treatment may be mediated at the level of G proteins (54 ,56).

Recent studies have examined the effects of valproate on components of the ß- adrenergic receptor (ßAR)-coupled cAMP generating system (57). Long-term valproate at a clinically relevant concentration has been shown to produce a significant alteration of the ßAR-coupled cAMP generating system in cultured cells *in vitro*. In contrast to long-term lithium (discussed above), long-term valproate was found to produce a significant reduction in the density of ßARs. Data generated during the past two decades reveal that carbamazepine inhibits basal and forskolin-stimulated activity of purified adenylyl cyclase and also basal and stimulated adenylyl cyclase in rodent brain and neural cells in culture (57 ,58 ,59 and 60). In addition, carbamazepine has been reported to reduce elevated cAMP in the cerebrospinal fluid (CSF) of manic patients (61). It appears that carbamazepine inhibits cAMP production by acting directly on adenylyl cyclase or through factor(s) that co-purify with adenylyl cyclase.

Lithium may have dual effects on the intracellular generation of cAMP. Whereas lithium decreases receptor-coupled stimulation of adenylyl cyclase, lithium increases basal levels of cAMP formation in rat brain (62 ,63). In addition, long-term lithium has been found to increase not only cAMP levels (64) but also levels of adenylyl cyclase type I and type II messenger RNA (mRNA) and protein levels in frontal cortex (65 ,66), which suggests that the net effect of lithium may derive from a direct inhibition of adenylyl cyclase, up-regulation of adenylyl cyclase subtypes, and effects on G proteins. Thus, it has been suggested that the action of lithium on the adenylyl cyclase system depends on state of activation; under basal conditions, in which tonic inhibition of cAMP formation through $G_{\alpha i}$ is predominant, levels of cAMP are increased, whereas during receptor activation of adenylyl cyclase mediated by $G_{\alpha s}$, cAMP formation is attenuated. It has been suggested that such a “bimodal model” for the mechanism of action of lithium may account for its therapeutic efficacy in both depression and mania (12). Although this would appear to be overly simplistic, it may bear clinical relevance to side effects of lithium, such as nephrogenic diabetes insipidus and subclinical hypothyroidism, which have generally been attributed to inhibition of vasopressin or thyrotropin-sensitive adenylyl cyclase.

G Proteins

As noted above, considerable evidence indicates that lithium attenuates receptor-mediated second-messenger generation in the absence of consistent changes in receptor density (see refs. 1 ,67 for review). Although lithium has been reported to reduce PI signaling via alteration in G-protein function in cell preparations (68 ,69 and 70), these data have not been replicated in rat or monkey brain (71 ,72). Although it appears that long-term lithium administration affects G-protein *function* (12 ,73), the preponderance of data suggest that lithium, at therapeutically relevant concentrations, does not have any direct effects on G proteins (1 ,74). A number of studies have reported modest changes in the *levels* of G-protein subunits; however, the effects of long-term lithium on signal transducing properties occur in the absence of changes in the levels of G-protein subunits *per se* (1 ,63 ,65 ,75). At the mRNA level, some evidence suggests that $G_{\alpha s}$, $G_{\alpha i1}$, and $G_{\alpha i2}$ may be down-regulated in rat cerebral cortex following long-term lithium (65 ,75 ,76). Again, however, these effects are small, and their physiologic significance is still unclear. Interestingly, the valproate-induced reduction in the density of ßARs (noted above) was accompanied by an even greater decrease in receptor- and post-receptor-mediated cAMP accumulation, which suggests that long-term valproate may exert effects at the ßAR/G_s interaction, or at post-receptor sites (e.g., G_s, adenylyl cyclase). A subsequent study has reported a reduction in the levels of $G_{\alpha s}$-45 but not in the levels of any of the other G-protein α subunits examined ($G_{\alpha s}$-52, $G_{\alpha i1/2}$, $G_{\alpha o}$, $G_{\alpha q/11}$) following long-term exposure to valproate (77).

Long-term lithium treatment has been shown to produce a significant increase in pertussis toxin-catalyzed [^{32}P]adenosine diphosphate (ADP)-ribosylation in rat frontal cortex and human platelets (1). Because pertussis toxin selectively ADP-ribosylates the undissociated, inactive Gαßγ heterotrimeric form of G_i, these data are consistent with a stabilization of G_i in the inactive conformation and an elevation in basal adenylyl cyclase activity. In this context, it is noteworthy that lithium appeared to increase the levels of endogenous ADP-ribosylation in C6 glioma cells and rat brain, whereas anticonvulsants either reduced ADP-ribosylation or had no effect (78 ,79). Currently, it is thought that the effects of long-term lithium may in part be mediated through post-translational modifications of G proteins that in turn may alter its coupling to receptor and second-messenger systems (1). However, given the relative abundance of G proteins, the physiologic impact of the level of post-translational changes induced by therapeutic levels of lithium on the balance of receptor-mediated signaling in brain is yet to be determined.

Protein Kinases and Protein Kinase Substrates

Based on the action of lithium in the PI signaling pathway, as discussed earlier, it was hypothesized that long-term prophylactic effects of lithium might be mediated via the diacylglycerol (DAG) arm of the PIP_2 hydrolytic pathway consequent to relative depletion of *myo*-inositol and subsequent DAG-mediated action on the regulation of protein kinase C (PKC) and specific phosphoprotein substrates (80 ,81) (Fig. 79.2). Studies during the past several years have provided evidence that PKC plays a crucial role in mediating the action of long-term lithium in a variety of cell systems, including primary and immortalized neurons in culture, and in rat brain (see refs. 12 ,33 ,81 ,82 for review). PKC represents a large family of at least 12 isozymes that are closely related in structure but differ in several ways—intracellular and regional distribution in the brain, second-messenger activators, specificity of association with the RACK (receptor for activated C-kinase) proteins, and substrate affinities—all of which suggest distinct cellular functions for these isozymes. PKC plays a major role in the regulation of neuronal excitability, neurotransmitter release, and long-term alterations in gene expression and plasticity. In fact, PKC activity has been implicated in processes underlying amygdala kindling and behavioral sensitization, putative animal models for BPD (83 ,84). PKC isozymes are highly expressed in the brain, with the γ isoform expressed exclusively, and are localized both presynaptically and postsynaptically. PKC is located in the cytoplasmic and membrane compartments of cells, and its activation requires translocation from the cytosol to RACK proteins within the membrane. Translocation from the cytosol to the membrane is most often associated with phosphorylation and activation of the enzyme, which is followed by autocatalysis and down-regulation of the enzyme on prolonged activation.

Studies of long-term lithium administration in the rat have demonstrated a reduction in membrane-associated PKC-α and PKC-ε in the subiculum and in CA1 regions of the hippocampus (85 ,86). In brain slices from lithium-treated rats exposed to phorbol ester, a known activator of PKC, a marked reduction was noted in the translocation of PKC activity from the cytoplasm to the membrane, and this was accompanied by a reduction in phorbol esterinduced serotonin release (87). Studies of long-term lithium in both C6 glioma cells and immortalized hippocampal cells in culture also demonstrate a reduction in the expression of these same PKC isozymes (see ref. 88 for review). This is interesting in light of data demonstrating an enhancement of PKC activity in platelets of patients during a manic episode (88). Moreover, administration of *myo*-inositol to rats was reported to reverse the down-regulation of PKC-ε in brain following long-term lithium, consistent with a role of *myo*-inositol in the downstream action of lithium on regulation of PKC by DAG. It is of note that valproate produces effects on the PKC signaling pathway similar to that reported for lithium (33 ,89). Long-term lithium and valproate appear to regulate PKC isozymes by distinct mechanisms, however, with the effects of valproate appearing to be largely independent of *myo*-inositol. These studies have

led to a pilot clinical study of the use of tamoxifen, a drug known to inhibit PKC *in vitro*, in the treatment of acute mania (90 ,91). Although the preliminary results appear consistent with the hypothesis, the sample size was small, and it is not known whether this drug *in vivo* inhibits PKC isozymes or whether its other properties (i.e., anti-estrogenic) play a role.

The activation of PKC results in the phosphorylation of a number of membrane-associated phosphoprotein substrates, the most prominent of which in brain is the MARCKS protein. Direct activation of PKC by phorbol esters in immortalized hippocampal cells effectively down-regulates the MARCKS protein (92). Long-term lithium administered to rats during a period of 4 weeks in clinically relevant concentrations dramatically reduces the expression of MARCKS protein in the hippocampus, and these findings have been replicated and extended in immortalized hippocampal cells in culture (39 ,93 ,94). Studies in hippocampal cells have demonstrated that the extent of down-regulation of MARCKS expression after long-term lithium exposure (1 m*M*) depends on both the inositol concentration and activation of receptor-coupled PI signaling, consistent with the hypothesis as stated above. Recent studies provide evidence for the regulation of transcription as a major site for the action of long-term lithium on MARCKS expression in brain (95). Moreover, this action of lithium in the brain and hippocampal cells is apparent only after *long-term* administration and persists beyond abrupt discontinuation of the drug for an extended period of time, paralleling the clinical time course of the therapeutic effects of lithium during initial treatment and discontinuation. Subsequent studies have discovered that this property of reducing the expression of MARCKS in hippocampal cells is shared by the anticonvulsant valproic acid, but not by other classes of psychotropic agents (96). Additionally, therapeutic concentrations of combined lithium and valproate have induced an additive reduction in MARCKS, also consistent with experimental findings that the two drugs work through different mechanisms on the PKC system and the clinical observation of the additivity of the two drugs in treatment responses (96). The altered expression of MARCKS further supports the role of PI signaling and PKC in the action of long-term lithium in the brain and may serve to provide insight regarding a role for neuroplasticity in the long-term treatment of BPD, as discussed below.

A crucial component of cAMP signaling is protein kinase A (PKA), which is a principal mediator of cAMP action in the central nervous system. Long-term lithium treatment has been shown to increase the regulatory and catalytic subunits of PKA in rat brains, an effect that appears to result in increased cAMP binding (97). Consistent with a lithium-induced increase in basal cAMP and adenylyl cyclase levels, a more recent study has reported that platelets from lithium-treated euthymic patients with BPD demonstrated an enhanced basal and the cAMP-stimulated phosphorylation of Rap1 (a PKA substrate) and a 38-kilodalton phosphoprotein not observed in healthy controls (98). The effects of lithium on the phosphorylation and activity of cAMP response element binding (CREB) protein, however, have been examined in rodent brain and in cultured human neuroblastoma cells, with somewhat conflicting results (99 ,100). Postmortem studies of brains of patients with BPD have shown changes in cAMP binding and in PKA activity in temporal cortex (101 ,102). These findings suggest that alterations in PKA activity may be associated with the action of lithium. It is of interest in this regard that carbamazepine attenuates forskolin-induced phosphorylation of CREB in C6 glioma cells (57).

It is well-known that lithium ion can have a significant effect on the development of a variety of organisms (103). In *Xenopus*, lithium significantly alters the ventral–dorsal axis of the developing embryo (104). One hypothesis regarding this action of lithium was based on its inhibition of IMPase and alteration in the dorsal–ventral balance of PI signaling in the embryo (105 ,106). Support for this hypothesis was derived from the observation that exposure to high concentrations of *myo*-inositol can reverse the effect of lithium (107). However, lithium ion has been shown to inhibit the activity of glycogen synthase kinase 3β (GSK-3β) (K_i = 2.1 m*M*) directly, thereby antagonizing the wnt signaling pathway, known to be instrumental in normal dorsal–ventral axis development in the *Xenopus* embryo (108 ,109 ,110 and 111). Furthermore, studies in which an embryo expressing a dominant negative form of GSK-3 was used have demonstrated that *myo*-inositol can reverse the resulting aberrant axis development in *Xenopus*, which suggests that *myo*-inositol reversal of dorsalization of the embryonic axis by lithium may be mediated, at least in part, by events independent of IMPase inhibition (47). Substrates for GSK-3β in cells include not only glycogen synthase but also β-catenin and microtubule-associated proteins (MAPs), both of which have been implicated in cytoskeletal restructuring; further, β-catenin is known to play a role in the expression of transcription factors [e.g., lymphoid enhancer factor (LEF) and T cell factor (TCF)]. Recent studies in human neuroblastoma cells have demonstrated that valproic acid also inhibits GSK-3β, after which levels of β-catenin increase (112). Thus, GSK-3β may contribute to our understanding of an action for long-term lithium observed in events associated with apoptosis and neuroplasticity, as discussed below.

Gene Expression

The clinical data indicating that onset of the therapeutic effect of lithium requires days to weeks of lag time and that reversal of the therapeutic effect on discontinuation occurs during a period of weeks to months suggest that the therapeutically relevant action of lithium in the brain involves long-term neuroplastic changes mediated by gene regulation. Evidence has accumulated that lithium can regulate

gene expression via nuclear transcriptional factors. One of the immediate early genes, c-*fos*, works as a master switch of gene regulation through interactions with *cis*-acting elements and other transcriptional factors. Lithium has been shown to alter the expression of c-*fos* in various cell systems (113) and in the brain (114 ,115 and 116); however, its effects have varied depending on brain region, cell type, and time course examined. (112 ,117 ,118 and 119).

It is known that c-*fos* interacts with *jun* family members to form activator protein 1 (AP-1), which binds to a common DNA site. Studies in both cell culture and rat brain following long-term lithium exposure *in vivo* demonstrate an enhancement of AP-1 DNA binding activity (99 ,120). Subsequent studies in cells with an AP-1-coupled reporter gene have confirmed a time- and concentration-dependent increase in transcriptional activity in the presence of lithium (121 ,122). These studies have also noted increases in the protein levels of c-*fos*, c-*jun*, and phosphorylated CREB. It is of interest that phosphorylation of c-*jun* inhibits DNA binding, whereas phosphorylation of CREB activates gene expression; both are substrates for GSK-3ß activity, which is inhibited by lithium. However, when AP-1 binding activity was measured following receptor activation, lithium treatment attenuated the induced AP-1 DNA binding activity (123 ,124). These seemingly contradictory findings suggest that the effect of lithium on gene transcription may depend on the activity level of the neurons. It has been suggested that by increasing AP-1 binding activity at the basal level, but decreasing it during stimulation, lithium can constrain the overall magnitude of fluctuations of gene expression as a function of neuronal activity (125). Valproic acid has been shown to have similar effects on the activity of AP-1 (120 ,122 ,126), which lends support to the possibility that gene regulation through AP-1 may represent a target for mood stabilizers. In addition, carbamazepine has been shown to inhibit forskolin-induced c-*fos* gene expression in cultured pheochromocytoma (PC12) cells (127). It must be kept in mind, however, that AP-1 binding activity is responsive to a multitude of signals and is unlikely to define the specific action underlying the therapeutic effect of lithium in BPD. Future studies may fruitfully examine a potential role for lithium in the regulation of newly discovered candidate genes linked to BPD (128), in addition to those implicated in its pathophysiology (129).

Lithium-induced alterations in gene expression may also account for recent findings of a neuroprotective effect in some cell systems. A number of groups have demonstrated a neuroprotective effect of lithium in systems both *in vivo* and *in vitro* against a variety of insults, including glutamate-induced excitatory apoptosis (130 ,131 and 132). It is well established that neuronal survival during apoptosis or programmed cell death depends on the relative expression of “executioner” proteins and “protector” proteins and the presence of neurotrophic factors. The B-cell lymphoma/leukemia 2 gene (*bcl2*), abundantly present in mammalian neurons, encodes one of the protector proteins that inhibits apoptosis and cell death under variety of circumstances. Recent studies in rat brain have demonstrated that long-term exposure to lithium or valproate increases the expression of the polyomavirus enhancer-binding protein 2ß gene (*PEBP2B*), a regulator of *bcl2* expression (133). Subsequent studies in rat brain have demonstrated an increase in cells immunoreactive for Bcl-2 in layers II and III of frontal cortex, dentate gyrus, and striatum after long-term lithium (134). In cultured cerebellar granule cells, long-term treatment with lithium induces a concentration-dependent decrease in *p53* and *bax* (apoptotic genes) mRNA and protein, with a concomitant increase in *bcl2* at both the mRNA and protein levels (135). It is of interest that these actions of lithium have been attributed to an enhancement of the PI_3 kinase pathway, in which GSK-3ß plays a prominent role (136) (Fig. 79.2). To what extent this neuroprotective effect may be related to the long-term prophylactic effect of lithium in stabilizing the course of BPD and the putative role of cellular loss in the pathophysiology of affective disorders remains to be demonstrated (137).

Neuroplasticity and Cytoskeletal Remodeling

Recent studies in a number of laboratories have provided evidence that long-term lithium treatment may alter molecular substrates underlying neuroplastic changes in brain that mediate alterations in interneuronal connectivity. As noted above, developmental studies in the *Xenopus* embryo have recently provided evidence that lithium can act as an inhibitor of GSK-3ß, a component of the wnt signaling pathway, at concentrations that may be relevant to clinical treatment (110). Several groups have reported that inhibition of GSK-3ß by lithium reduces phosphorylation of tau protein in different cell systems, the effect of which is to enhance the binding of tau to microtubules and promote microtubule assembly (110 ,138 ,139 and 140). Lithium treatment also decreases phosphorylation of MAP-1ß, a microtubule-associated protein involved in microtubule dynamics within the growth cone and axonal outgrowth (141). Lithium-induced dephosphorylation of MAP-1ß reduces its ability to bind to microtubules; in cerebellar granule neurons, this effect was accompanied by axonal spreading and increases in growth cone area and perimeter (142 ,143). Thus, it is possible under the appropriate conditions that inhibition of GSK-3ß by lithium can induce significant changes in microtubule assembly that result in changes in the association dynamics among cytoskeletal proteins mediating neuroplastic changes in regions of the brain.

The significance of actin-membrane remodeling in the long-term action of lithium is also supported by a series of studies demonstrating that long-term lithium down-regulates the expression of the PKC substrate MARCKS in brain, as noted previously. MARCKS is a complex protein

that binds calmodulin in a calcium-dependent manner; it also binds and cross-links filamentous actin, thereby conferring focal rigidity to the plasma membrane. Following phosphorylation of its phosphorylation site domain in the presence of activated PKC, MARCKS translocates from the plasma membrane and neither binds calmodulin nor cross-links actin. Thus, this protein is in a key position to transduce extracellular signals to alterations in the conformation of the actin cytoskeleton, which are critical to cellular processes underlying development and signaling, including morphogenesis and secretion. MARCKS is enriched in neuronal growth cones, developmentally regulated, and necessary for normal brain development (144 ,145 and 146). MARCKS expression remains elevated in specific regions of the hippocampus and limbic-related structures, which retain the potential for plasticity in the adult rat (147 ,148) and human brain (149), and its expression is induced in the mature central nervous system during axonal regeneration (150). Recent studies support a role for MARCKS in plastic events associated with learning and memory. Induction of long-term potentiation, thought to be a physiologic component of learning and memory, elevates MARCKS phosphorylation in hippocampus (151). Moreover, adult mutant mice expressing MARCKS at 50% exhibit significant spatial learning deficits that are reversed in the presence of a MARCKS transgene (144). These data reveal that MARCKS plays an important role in the mediation of neuroplastic processes in the developing and mature central nervous system. Thus, by virtue of its action in signaling pathways utilizing PI/PKC and GSK-3ß cascades (Fig. 79.2), long-term lithium administration may alter presynaptic and postsynaptic membrane structure to stabilize aberrant neuronal activity in critical regions of the brain involved in the regulation of mood (92).

ANTIDEPRESSANTS

Part of "79 - Mechanism of Action of Antidepressants and Mood Stabilizers "

Neurotransmitter Signaling

Antidepressants are usually classified according to structure [e.g., tricyclic antidepressants (TCAs)] or function [e.g., monoamine oxidase inhibitors (MAOIs), selective serotonin reuptake inhibitors (SSRIs)]. However, it may be more useful to classify them according to the acute pharmacologic effects that are presumed to trigger behavioral improvement. If this is done, the antidepressants can be grouped in four categories (Table 79.1). First are the drugs that selectively block the reuptake of norepinephrine (NE). These include certain TCAs and TCA-like compounds (maprotiline). Another drug that falls into this category is reboxetine, although it is distinct structurally from the TCAs and TCA-like compounds (152). It is currently available as an antidepressant in European and South American countries but is not yet marketed in the United States. Second are the SSRIs, which, as their class name implies, selectively block the reuptake of serotonin [5-hydroxytryptimine (5-HT)] *in vivo*.

Category	Mechanism	Examples	Current Classification (If Any)
I	Selective blockade of NE reuptake (SNRIs)	DMI, NT amoxapine, maprotiline reboxetine	TCAs TCA-like —
II	Selective blockade of 5-HT reuptake (SSRIs)	Citalopram, fluoxetine, paroxetine, sertraline	SSRIs
III	Nonselective enhancement of NE and 5-HT transmission	IMI, AMI phenelzine, tranylcypromine venlafaxine mirtazapine	TCAs MAOIs (sometimes with SSRIs) —
IV	Unknown potent stimulatory effects on NE or 5-HT	trimipramine bupropion nefazodone, trazodone	TCA — —

5-HT, 5-hydroxytryptamine (serotonin); AMI, amitriptyline; DMI, desipramine; IMI, imipramine; MAOI, monoamine oxidase inhibitor; NE, norepinephrine; NT, nortriptyline; SNRI, selective norepinephrine reuptake inhibitor; SSRI, selective serotonin reuptake inhibitor; TCA, tricyclic antidepressant.

TABLE 79.1. MECHANISM-BASED CLASSIFICATION FOR ANTIDEPRESSANTS

Third are the drugs that act nonselectively on noradrenergic and serotoninergic neurons with a resultant enhancement of synaptic transmission. Some TCAs are in this category, as are the MAOIs. Some novel drugs are also in this category. One of these is venlafaxine, discussed in more detail later. Another is mirtazapine. Mirtazapine is not a potent inhibitor of the reuptake of either NE or 5-HT (153). It is a relatively potent antagonist, though, of inhibitory α_2 autoreceptors on noradrenergic nerves. By blocking such autoreceptors, mirtazapine removes their inhibitory influence on noradrenergic transmission. Thus, even though it is not a reuptake inhibitor, mirtazapine can directly enhance NE-mediated transmission (154 ,155 and 156). In this respect, then, it might be appropriate to place mirtazapine in the first

category. However, mirtazapine may also enhance serotoninergic transmission, albeit *indirectly* (157 ,158 and 159). This enhancement is caused in part by NE activation of α_1 noradrenergic receptors located on serotoninergic soma and dendrites to increase cell firing and the release of 5-HT (160 ,161). Mirtazapine may also block inhibitory α_2 adrenoceptors located on serotoninergic terminals (i.e., heteroceptors) (154 ,162). However, some recent data call into question the likelihood that mirtazapine enhances serotoninergic transmission (163). Whether mirtazapine increases serotoninergic transmission may depend on the state of activation of the central noradrenergic system when the drug is administered. Further research is needed to clarify this point. At this time, we have placed mirtazapine in the third category.

In the fourth and final heterogeneous group are drugs without known potent, acute pharmacologic effects that result in enhancement of noradrenergic or serotoninergic transmission. In other words, their mechanisms of action are unknown. Drugs in this category include the TCA trimipramine and also bupropion, nefazodone, and trazodone. It has been speculated that bupropion acts through dopaminergic mechanisms because it is the only antidepressant that more potently blocks the reuptake of dopamine than that of either NE or 5-HT (164). However, in reality, bupropion and its metabolites are very weak inhibitors of the reuptake of all three biogenic amines, with potencies in the micromolar range (164). Perhaps this is why the data regarding whether bupropion inhibits dopamine reuptake in patients at clinically relevant doses are at best conflicting (165). Some data indicate an as yet ill-defined effect of bupropion or its hydroxylated metabolite on noradrenergic function (164), but the efficacy of bupropion cannot at this time be attributed to effects on noradrenergic transmission.

The most potent acute effect of nefazodone and trazodone on serotoninergic or noradrenergic systems is their antagonism of 5-HT_{2A} receptors (166). They are very weak inhibitors of NE reuptake and relatively weak inhibitors of 5-HT reuptake (167). If enhancement of serotoninergic transmission is a mechanism that ultimately leads to clinical efficacy, it is not clear how antagonism of the 5-HT_{2A} receptor produces such enhancement. Some data indicate that 5-HT_2-receptor antagonism enhances 5-HT_{1A}-receptor responsivity (168 ,169), or that 5-HT_2-receptor antagonists share discriminative stimulus properties with 5-HT_{1A}receptor antagonists (170). However, not everyone finds such effects (171), and whether such an effect would enhance endogenous serotoninergic transmission is uncertain. Thus, acute pharmacologic properties that contribute to the efficacy of the drugs in the fourth category remain unknown.

Originally, brain tissue from rats was used to measure the potencies of drugs *in vitro* to block the reuptake of ^{3}H-NE or ^{3}H-5-HT. Subsequently, radioligand binding techniques were developed such that the potencies of antidepressants to displace the specific binding of ligands to the norepinephrine transporter (NET) or serotonin transporter (SERT) could be measured. These studies were also carried out in brain tissue, usually from rats. The potencies of drugs to produce such effects were thought to be reflective of their potencies at blocking NE or 5-HT uptake clinically. The cloning of the SERT and NET in the early 1990s enabled many types of studies not possible heretofore (172). Among these are studies in which the human NET (*h*NET) or human SERT (*h*SERT) is transfected, often stably, into cells that normally do not have any NET or SERT. These cells can be maintained in cell culture systems and used to measure the uptake of ^{3}H-NE and ^{3}H-5-HT by the *h*NET and *h*SERT, respectively, and the binding of radioligands to the *h*NET and *h*SERT. Further, such cells can be used to measure the potencies of antidepressants to block such effects. The advantage of such systems, obviously, is that potencies are measured directly on human transporters. The disadvantages of such systems are equally obvious—namely, they are artificial, and a variety of factors can influence results (173). As Kenakin (173) has written, "Transfecting the cDNA of a receptor protein into a foreign cell and expecting a physiologic system can be likened to placing the Danish King Hamlet on the moon and expecting Shakespeare to emerge."

It might be illustrative to compare potencies of antidepressants obtained with the different preparations and approaches. This is done in Tables 79.2 and 79.3. Irrespective of the noradrenergic parameter chosen (Table 79.2), the orders of potency are almost identical, especially for the most potent compounds (i.e., desipramine > nortriptyline > amitriptyline = imipramine > paroxetine). Also, citalopram is the least potent drug on all measures. Perhaps the value that most stands out quantitatively from the others is that for ^{3}H-NE uptake by *h*NET. In general, these values tend to be sixfold to 10-fold higher (i.e., potencies are less) than those found to inhibit such uptake into rat brain synaptosomes. An interesting specific difference is seen with venlafaxine; its potency to inhibit ^{3}H-NE uptake by rat brain is five to eight times greater than its potency on the other noradrenergic parameters. For serotoninergic parameters also, the rank order of potencies appears reasonably similar irrespective of the specific parameter—namely, paroxetine > sertraline ≥ citalopram ≥ fluoxetine ≥ imipramine ≥ venlafaxine ≥ amitriptyline > nortriptyline ≥ desipramine ≥ nefazodone. However, the potencies found for most of the drugs to inhibit *h*SERT binding are greater than those to inhibit ^{3}H-5-HT uptake by the *h*SERT, oftentimes eightfold to 20-fold greater (Table 79.3).

TABLE 79.2. VALUES (nM) OF THE INHIBITION CONSTANT (K_i)

Drug	^{3}H-NE Uptake (Rat)	*r*NET Binding	^{3}H-NE Uptake (Human)	*h*NET Binding
Amitriptyline	14	9	102	27
Citalopram	>3,000[a]	>3,000	>30,000	>5,500
Desipramine	0.6	0.3	3.5	0.7
Fluoxetine	143	473	2186	508
Imipramine	14	11	142	28
Nefazodone	570	555	713	489
Nortriptyline	2	1	21	3
Paroxetine	33	59	328	62
Sertraline	220	1597	1716	618
Venlafaxine	210	1067	1644	1664

Potencies of these drugs for blocking the uptake of ^{3}H-NE or ^{3}H-5-HT into rat brain synaptosomes were taken primarily from Bolden-Watson and Richelson, 1993 (167). These values tend to be in good agreement with those reported by others. Potencies for drugs to inhibit the binding of radioligands to the NET or SERT in rat brain synaptosomes were taken from Owens et al., 1997 (175) for the same reason. Potencies of drugs to inhibit the binding of selective radioligands to the *h*NET and *h*SERT were averaged from results in Owens et al., 1997 (175) and Tatsumi et al., 1997 (176). In general, the results obtained in these two studies are in remarkably close agreement. Finally, potencies of drugs to inhibit uptake of ^{3}H-NE and ^{3}H-5-HT by the *h*NET and *h*SERT, respectively, were taken from Owens et al., 1997 (175). Such values tend to be in good agreement with those obtained by others using transfected cell systems, such as Eshleman et al., 1999 (177).
5-HT, 5-hydroxytryptamine (serotonin); NET, norepinephrine transporter; SERT, serotonin transporter.
[a]From Hyttel and Larsen, 1985 (174).

TABLE 79.3. VALUES (nM) OF THE INHIBITION CONSTANT (K_i)

Drug	^{3}H-5-HT Uptake (Rat)	*r*SERT Binding	^{3}H-5-HT Uptake (Human)	*h*SERT Binding
Amitriptyline	84	16	36	4
Citalopram	1.4[a]	0.8	9	1
Desipramine	180	129	163	20
Fluoxetine	14	2	20	0.9
Imipramine	41	9	20	1
Nefazodone	137	220	549	330
Nortriptyline	154	60	279	16
Paroxetine	0.7	0.05	0.8	0.1
Sertraline	3	0.3	3	0.2
Venlafaxine	39	19	102	8

Potencies of these drugs for blocking the uptake of ^{3}H-NE or ^{3}H-5-HT into rat brain synaptosomes were taken primarily from Bolden-Watson and Richelson, 1993 (167). These values tend to be in good agreement with those reported by others. Potencies for drugs to inhibit the binding of radioligands to the NET or SERT in rat brain synaptosomes were taken from Owens et al., 1997 (175) for the same reason. Potencies of drugs to inhibit the binding of selective radioligands to the *h*NET and *h*SERT were averaged from results in Owens et al., 1997 (175) and Tatsumi et al., 1997 (176). In general, the results obtained in these two studies are in remarkably close agreement. Finally, potencies of drugs to inhibit uptake of ^{3}H-NE and ^{3}H-5-HT by the *h*NET and *h*SERT, respectively were taken from Owens et al., 1997 (175). Such values tend to be in good agreement with those obtained by others using transfected cell systems, such as Eshleman et al., 1999 (177).
5-HT, 5-hydroxytryptamine (serotonin); NET, norepinephrine transporter; SERT, serotonin transporter.
[a]Calculated from Hyttel, 1978 (178).

It is important to recognize that potencies obtained *in vitro* for any pharmacologic effect give only some indication of whether the pharmacologic effect in question could occur clinically. High potency *in vitro* (e.g., ≤ 10 n*M*) certainly increases the likelihood that an effect will occur clinically, and low potency (e.g., > 500 n*M*) decreases the probability. However, as emphasized by others (179), whether or not a

specific effect occurs clinically depends on how much drug reaches its presumed site(s) of action (i.e., a function of pharmacokinetics). Because these drugs must act on brain to exert their beneficial effects, a factor that substantially influences how much reaches the brain is the extent to which they are protein-bound. Because of the blood–brain barrier, the amount of drug in the extracellular fluid of brain (i.e., CSF) tends to be equivalent at steady state to the concentration of non–protein-bound drug in plasma (i.e., "free" drug). Normal CSF contains so little protein that it may be regarded as an ultrafiltrate of serum. Because most, but not all, antidepressants are extensively bound to plasma proteins (180 ,181), their concentration in CSF is only a small fraction of the total concentration in serum.

Table 79.4 shows the percentages of protein binding of certain antidepressants. Also shown are steady-state total plasma concentrations of drug and concentrations in CSF. It is apparent that drug actually measured in CSF approximates what would be calculated to be the non–protein-bound concentration in plasma. For this reason, also shown in Table 79.4 are some antidepressants with concentrations in CSF that have not been reported. It is possible, then, to compare these CSF concentrations of drugs with their K_i values for the inhibition of uptake or ligand binding, shown in Table 79.2 and Table 79.3 . For a drug such as citalopram, it is obvious that its concentration in CSF is much greater than that required to inhibit serotoninergic uptake or binding to the SERT, irrespective of whether one is obtaining measurements with rat synaptosomes or *h*SERT. It is also obvious that citalopram does not reach sufficient concentration in CSF to block NE reuptake, again irrespective of the noradrenergic parameter or type of tissue. Considerable data indicate that citalopram maintains selectivity as a 5-HT uptake inhibitor *in vivo* (162). It is also apparent that desipramine reaches concentrations in CSF sufficient to block NE uptake, irrespective of the parameter or tissue used for measurement. However, the potency of desipramine to block ligand binding to the *h*SERT (20 n*M*) is sufficient to indicate that it may have a substantial effect on 5-HT uptake in the brain of patients. Although treatment with desipramine does lower concentrations of the serotonin metabolite 5-hydroxyindolacetic acid (5-H1AA) in the CSF of patients (192),

it is unlikely that this observation directly reflects the ability of the drug to block 5-HT uptake. Rather, it is more likely to be some indirect effect. The clinical efficacy of desipramine does not appear to depend on the availability of 5-HT (194). Furthermore, considerable preclinical data demonstrate little to no effect of desipramine on serotoninergic function (195 ,196 and 197). Given this, any of the other three serotoninergic values for desipramine, which are quite similar, would seem to be a better indicator of what happens clinically. The situation with nortriptyline appears to be similar to that with desipramine. Even at low concentrations, nortriptyline is likely to block NE uptake. Nortriptyline maintains reasonable selectivity *in vivo* as an inhibitor of NE reuptake (198 ,199 ,200 and 201). Because its concentration in CSF exceeds its potency to block ligand binding to the *h*SERT (and approaches its potency to inhibit binding at the *r*SERT), it seems doubtful that these potencies have relevance for functional inhibition of 5-HT uptake by nortriptyline clinically. Irrespective of the parameters used to assess the effects of fluoxetine, the interpretation would be the same—namely, levels of fluoxetine in CSF are likely sufficient to block 5-HT uptake, but not NE uptake. Again, considerable clinical and preclinical data indicate that fluoxetine maintains selectivity *in vivo* as a 5-HT reuptake inhibitor (202 ,203 and 204).

TABLE 79.4. TOTAL PLASMA AND CEREBROSPINAL FLUID CONCENTRATIONS OF SOME ANTIDEPRESSANTS

Drug	Protein Binding (%)[a]	Concentration (nM) in Plasma	CSF (measured)	CSF (estimated)	Reference
Amitriptyline	95	512	33		Hanin et al., 1985 (182)
(Nortriptyline)[b]	92	524	48		
Citalopram	50	40–750[c]	—	20–375	
Fluoxetine	95	854	26		Martensson et al., 1989 (183)
(Norfluoxetine)[b]	—	1,006	17		
Imipramine	90	433	40		Muscettola et al., 1978 (184)
(Desipramine)[b]	82–92	431	56		
Imipramine	90	475	36		Hanin et al., 1985 (182)
(Desipramine)[b]	82–92	642	79		
Nortriptyline	92	443	39		Nordin et al., 1985 (185)
Paroxetine	95	275	7		Lundmark et al., 1994 (186)
Venlafaxine[d]	27–30	370–3,000	—	100–850	

These numbers do not take into account concentrations of hydroxylated metabolites in CSF, which can have pharmacologic activity [Nordin and Bertilsson, 1995 (190); Nordin et al., 1987 (191); Potter et al., 1979 (192)]. Even though such hydrophylic metabolites may have diminished lipid solubility, the penetration of some hydroxylated metabolites into CSF may be somewhat greater than that of the parent compounds, presumably because of decreased protein binding [Nordin et al., 1985 (185); Sallee and Pollock, 1990 (181)]. Nevertheless, such metabolites more often than not are more weakly potent than their parent compounds, so it is not likely as a rule that such metabolites contribute substantially to pharmacologic activity in brain.
[a]Values taken from van Harten, 1993 (180); Sallee and Pollock, 1990 (181); and Benet et al., 1996 (187).
[b]Parentheses indicate measurements were taken of the drug as a metabolite of the parent antidepressant.
[c]Values taken from Bjerkenstedt et al., 1985 (188) and Fredricson-Overo, 1982 (189).
[d]Range of values for venlafaxine refers to venlafaxine plus *O*-desmethylvenlafaxine.

Given the great potency of paroxetine on any of the serotoninergic parameters in Table 79.3 , its concentration in CSF is sufficient to cause functional blockade of 5-HT uptake. Also, concentrations of paroxetine in CSF appear insufficient to cause much, if any, functional blockade of NE reuptake. Its greatest potency on a noradrenergic parameter (33 n*M* for inhibition of ^{3}H-NE uptake by rat brain synaptosomes) is considerably higher than its highest concentration in CSF (Table 79.2). Considerable data *in vivo* indicate that paroxetine maintains selectivity as an inhibitor of 5-HT reuptake (205 ,206 and 207). Other than its ability to decrease concentrations of methoxyhydroxyphenylglycol (MHPG) in the CSF of depressed patients (186), an effect produced by all SSRIs, including citalopram (186 ,188 ,208), no other clinical data are available from which it may be inferred that paroxetine blocks NE reuptake.

Venlafaxine (and its metabolite *O*-desmethylvenlafaxine) appears to be likely to inhibit 5-HT uptake, even at the low end of its concentration in CSF. This conclusion is reached by comparing its concentration in CSF with any of its "serotoninergic" potencies, with the exception of its potency in inhibiting ^{3}H-5-HT uptake by *h*SERT. Its low concentration in CSF is only equivalent to this potency. Interestingly, one might conclude that venlafaxine blocks NE uptake at higher concentrations only if one considers its potency in blocking ^{3}H-NE uptake by rat brain synaptosomes (210 n*M*). Even its highest concentration in CSF (850 n*M*) is only about half its potency on either of the *h*NET parameters and comparable with its potency to inhibit ligand binding to the *r*NET. Data show that venlafaxine is likely to block NE uptake *in vivo*, especially at higher doses (206 ,209). Thus, it seems reasonable to speculate that the most clinically relevant potency for venlafaxine at a noradrenergic parameter is its potency to block ^{3}H-NE uptake by rat brain synaptosomes.

For many of the drugs, then, the same conclusion is reached about selectivity (or nonselectivity) *in vivo* based on concentrations achieved in CSF and any of the noradrenergic or serotoninergic parameters. For some of the drugs, though, the parameter chosen influences the prediction of what will happen clinically. If one examines potencies to inhibit ligand binding to the *h*SERT, one might predict that both desipramine and nortriptyline are nonselective inhibitors of both NE and 5-HT reuptake *in vivo*. This does not seem to be so (194). For these drugs, then, the values for the *h*SERT should be viewed cautiously. As discussed, the clinical situation with venlafaxine causes some concern about its potency to inhibit ^{3}H-5-HT uptake by *h*SERT or to inhibit ligand binding to either *r*NET or *h*NET. This analysis demonstrates that K_i values measured *in vitro* allow only a prediction of what will occur *in vivo*—they offer no proof. Experiments must be carried out *in vivo* to prove (or disprove) the predictions.

Regulatory Effects

The pharmacologic effects of uptake inhibitors, just described in detail, are acute and direct effects of the drugs. As mentioned previously, the optimal behavioral effects of antidepressants on mood may not be evident immediately after initiation of treatment; rather, they are delayed from 2 to 3 weeks (210 ,211), although some symptoms may show early improvement (212 ,213 and 214). Furthermore, until this past decade, even though UPD was recognized as a recurrent illness in some patients, antidepressants were used primarily on a short-term basis (e.g., 2 to 4 months). However, evidence accumulated during the past several years has caused a fundamental shift in the treatment of UPD, so that prophylaxis is emphasized in addition to acute treatment. Such evidence includes the following: (a) UPD is widely recurrent, with more than 50% of patients having a recurrence sometime during their lifetime (215); (b) long-term (years) treatment of patients with recurrent UPD with different classes of antidepressants is effective in preventing depressive recurrences (215 ,216); (c) antidepressants (SSRIs, venlafaxine, mirtazapine) have been developed that have a much better side effect (and toxicity) profile than the TCAs, so that they are much better tolerated by patients (217). Such realizations have led to an extensive study of the longer-term and more slowly developing effects of antidepressants, particularly on central monoamine systems. It is beyond the scope of this chapter to review this area in detail. The interested reader can find more exhaustive reviews elsewhere (218 ,219 ,220 ,221 and 222). Rather, we emphasize the long-term effects of antidepressants that would be expected to continue or markedly

enhance the increase in serotoninergic and noradrenergic transmission initiated by the inhibition of uptake. Further, we emphasize some issues we believe to be important in long-term studies of antidepressant effects in laboratory animals such as rats.

Receptors/Transporters

Clearly, a key assumption is that an understanding of delayed pharmacologic effects and the mechanisms that produce them can lead to the development of drugs (or drug combinations) that produce such effects earlier, with consequent earlier clinical improvement. For example, early research showed that although uptake inhibitors do acutely block uptake, they also rapidly decrease the firing rate of serotoninergic or noradrenergic soma (200 ,223). For this reason, it was speculated that appreciable enhancement of neurotransmission does not occur with these drugs early in treatment. With serotonin, this was thought to be a consequence of a rise in 5-HT in the raphe nucleus during 5-HT uptake inhibition, which activates inhibitory somatodendritic autoreceptors so as to restrain the rise in serotonin in terminal fields (224 ,225 and 226). A similar mechanism was thought to underlie the decrease in firing in the locus ceruleus (227 ,228 and 229).

With time, though, regulatory changes occur that can enhance transmission, especially in the presence of continued inhibition of uptake. Perhaps chief among these is a time-dependent desensitization of inhibitory serotoninergic autoreceptors. In general, the consensus is that long-term administration of inhibitors of 5-HT uptake cause a desensitization of somatodendritic 5-HT_{1A} receptors (230 ,231 and 232), although whether terminal autoreceptors become desensitized is more controversial (233 ,234 ; see references in 235). The time-dependent desensitization of inhibitory somatodendritic autoreceptors enhances serotoninergic neurotransmission in terminal fields during the long-term administration of 5-HT uptake inhibitors (231 ,236 ,237). Such observations led to the idea that concomitant administration of pindolol, a 5-HT_{1A}-receptor antagonist, with an SSRI would enhance the rate of response. Unfortunately, data about whether this happens are controversial (232 ,238 ,239). The data are somewhat more contradictory regarding whether desensitization of inhibitory somatodendritic noradrenergic α_2 autoreceptors occurs after long-term administration of NE reuptake inhibitors (227 ,229 ,240 ,241). Less attention has been focused on whether concomitant administration of an α_2-adrenoceptor antagonist would improve the speed of response of a noradrenergic reuptake inhibitor.

Many of the studies of long-term effects of antidepressants have focused on presynaptic or postsynaptic changes in 5-HT and NE receptors (such as those just described) and their physiologic or behavioral consequences. Even though the SERT and NET are the initial cellular targets for reuptake inhibitors, few early studies examined whether such treatment has regulatory effects on these proteins. However, the cloning of the SERT and NET in the early 1990s (172) made it possible to determine whether these integral plasma membrane proteins exhibit plasticity. This work culminated in studies of mice with knockouts of these transporters. In heterozygotes, in which transporter density is reduced by 50%, the impact on transporter capacity is marginal, which suggests that powerful post-transcriptional events regulate transporter function (242). Studies of the mechanisms of such events have in general been carried out *in vitro*, either with cells that naturally express these transporters or with cells into which the transporters have been stably transfected. The realization that transporters can be regulated stimulated considerable research during the last decade to determine whether long-term treatment of rats with reuptake inhibitors produces regulatory effects on the SERT or NET. Unfortunately, no consistent picture has emerged (243). We think some of the inconsistency may be a consequence of such factors as route of drug administration and tissue preparation. Because this area has not been reviewed in detail previously, we do so here.

In 1990, Marcusson and Ross (244) reviewed the literature about the effect of antidepressants on the SERT. At that time, the approach to measuring SERT function was to measure ^{3}H-5-HT uptake *in vitro*. Factors that may contribute to regulatory effects can be lost during tissue preparation (slices, synaptosomes) and the use of artificial incubation media. Other techniques, such as binding a radioligand to the SERT or quantifying mRNA for the SERT, do not measure SERT function. Another important factor that can affect results is how the drugs are administered. Assessment of the literature and the experience of one of the authors (A. F.) with sertraline caused us to believe that sustained antagonism of the SERT throughout the day is needed to demonstrate down-regulation of the SERT. When sertraline was administered intraperitoneally to rats at a dose of 5 mg/kg twice daily for 21 days (245), quantitative autoradiographic analysis of ^{3}H-cyanoimipramine (^{3}H-CN-IMI) binding to the SERT in 23 areas of brain revealed small (15% to 21%) decreases in binding in only four areas. By contrast, when the drug was administered subcutaneously by minipump for 21 days (at a daily dose of 7.5 mg/d, which is even less than that used in the previous study), a large (70% to 75%) decrease in the binding of ^{3}H-CN-IMI was seen throughout brain (246). Given that the analytic methodology was essentially identical in the two studies, as was the time of drug administration, the factor that most likely accounts for the difference in results is the route of drug administration.

In general, the metabolism of drugs is faster in rats than in humans. Most uptake inhibitors (or their active metabolites) have half-lives in humans that average about 1 day or even longer (247). Given this, even the administration of antidepressants once a day with doses producing recommended “therapeutic” plasma concentrations (usually measured

at the trough of the daily variation in concentration) is sufficient to maintain consistent blockade of the transporter. In the rat, though, sertraline has a half-life of about 4.5 hours (248). Thus, twice-daily and especially once-daily injections of this drug may not be sufficient to maintain adequate occupancy of the SERT in brain throughout the day, and continuous adequate occupancy may be necessary to obtain regulatory effects. Such considerations are even more important with a drug such as citalopram, which has an elimination half-life in the rat of 3 to 5 hours (189 ,249), or venlafaxine (or its metabolite *O*-desmethylvenlafaxine), which has a half-life of about 1 hour (250). All these considerations are relevant in a review of effects of long-term administration of uptake inhibitors on the SERT (251).

In general, studies of long-term citalopram given intraperitoneally either once or twice daily have found little to no regulatory effects on the SERT (234 ,245 ,252 ,253 and 254). In addition to the lack of effect of long-term intraperitoneal administration of citalopram on SERT parameters, comparable administration of this drug has produced inconsistent effects on somatodendritic 5-HT_{1A}-receptor sensitivity (233 ,254 ,255 and 256). A recent report is illustrative (235). The investigators speculated that positive results with long-term citalopram administration might be obtained if the administration and dosage were adequate to maintain plasma levels in a “therapeutic” range. This was achieved by giving the drug subcutaneously by minipump at a dose of 20 mg/kg per day for 15 days. This regimen produced stable plasma citalopram levels of about 300 n*M*. After a 48-hour washout, when analytic experiments were carried out, plasma levels of citalopram had dropped to levels that were not pharmacologically active. At this time, evidence for marked subsensitivity of somatodendritic 5-HT_{1A} receptors was obtained. Their conclusion was that if a proper, pharmacokinetically validated, long-term regimen with citalopram is used, 5-HT_{1A}-autoreceptor desensitization can be observed. We think that these observations account for why those who gave long-term paroxetine by minipump consistently obtained evidence of decreases in SERT function (246 ,257 ,258 and 259).

Fluoxetine has a half-life in the rat of just 5 to 8 hours, but that of its metabolite, norfluoxetine, is about 15 hours (260). Inconsistent results have been obtained in studies of the effects of long-term administration of this drug on SERT parameters. Gobbi et al. (234) reported no effect in rats of 3 weeks of treatment with fluoxetine (15 mg/kg orally twice daily) on either ^{3}H-5-HT uptake into synaptosomes or ^{3}H-citalopram binding. In this study, measurements were made after 7 days of drug washout. Similar results were obtained by Dean et al. (261), who used homogenates; they gave the drug at a dose of 10 mg/kg per day intraperitoneally for either 10 or 28 days. By contrast, in the study of Durand et al. (262), the same dose of fluoxetine, when given for 21 days, markedly lowered ^{3}H-citalopram binding in brain homogenates. Berton et al. (263) also reported a significant but more modest (25%) reduction in ^{3}H-citalopram binding in the midbrain of rats treated once daily with 7.5 mg of fluoxetine per kilogram for 21 days. It is interesting that such dose schedules for fluoxetine produce inconsistent results on SERT measures because comparable schedules cause consistent effects on other measures of serotoninergic function, such as 5-HT_{1A}-receptor sensitivity (231 ,264 ,265). It does appear, then, that stable, nonfluctuating plasma levels of 5-HT uptake inhibitors over time are needed to show regulatory effects on the SERT, whereas this may not be so for other serotoninergic parameters.

Several investigators have examined mRNA for the SERT at different times after long-term administration of 5-HT uptake inhibitors. Here, also, the results have been inconsistent (246 ,266 ,267 ,268 ,269 ,270 and 271). This may be a consequence not only of the route of drug administration but also of the duration. Although no change in mRNA for the SERT after 21 days of treatment with paroxetine or sertraline by minipump has been reported (246 ,272), changes in mRNA are found earlier in treatment, with peak effects after about 10 days (272). It seems that only if the drugs are given in a regimen that causes decreases in SERT binding can changes in its mRNA be observed, and even then only if one looks at the proper times (i.e., early in treatment).

As with the SERT, a number of investigators have studied the effect of long-term treatment of rats with NE uptake inhibitors on NET binding sites. In more recent work, ^{3}H-nisoxetine, a radioligand that binds specifically to the NET, has been used (273 ,274). In the study of Bauer and Tejani-Butt (275), 21 days of treatment with desipramine (10 mg/kg intraperitoneally once daily) caused modest (20% to 40%) but significant decreases in ^{3}H-nisoxetine binding in some areas of brain, but not in others. More robust and widespread changes were obtained when desipramine was given by osmotic minipump for 21 days (272). We think it likely that the greater effect seen is a result of giving the drug by minipump to obtain consistent daily plasma levels in the “therapeutic” range. It does seem likely that long-term desipramine treatment can down-regulate the NET; its addition *in vitro* to PC12 cells in culture at concentrations above 100 n*M* caused a decrease in the B_{max} of (i.e., the maximum density of binding sites) ^{3}H-nisoxetine binding, with a maximum effect occurring after 3 days of exposure (276). The uptake of ^{3}H-NE was also decreased. These effects occurred after exposure of the cells to desipramine for as little as 4 hours, but always after desipramine had been removed from the incubation medium. In a follow-up study in which cells transfected with *h*NET were used, similar results were obtained with desipramine; in addition, less NET protein was measured in the desipramine-treated cells. Interestingly, desipramine did not cause any change in mRNA for the NET in these cells. In contrast, in the study of Zavosh et al. (277), addition of desipramine to the human neuroblastoma cell line SK-N-SHSY 5Y not only decreased the B_{max} of ^{3}H-nisoxetine binding by 24 hours

but also decreased message, although only after 72 hours of treatment, not after 24 hours. For this reason, Zavosh et al. (277) concluded that the decreases in ligand binding were not a consequence of changes in message. Interestingly, Szot et al. (278) found that both 2-day and 4-week treatment of rats with desipramine (10 mg/kg intraperitoneally once daily) increased mRNA for the NET in the locus ceruleus.

The antidepressant-induced loss of SERT binding sites (and presumably also NET binding sites) may have important functional consequences relevant to the behavioral improvement produced by reuptake inhibitors. The changes that acute local administration of an SSRI has on the clearance of 5-HT *in vivo*, measured by chronoamperometry, are significant but modest. Variable and quite small effects are produced on the peak amplitude of the electrochemical signal caused by 5-HT, and clearance of the indolalkylamine is inhibited by 30% to 40% (246). However, when antidepressant treatment causes a marked reduction in SERT binding sites, then the peak amplitude of the 5-HT signal is substantially increased and the clearance time is more than doubled (246). Similar effects on the 5-HT chronoamperometric signal were also observed in rats treated with a serotoninergic neurotoxin to cause more than a 70% loss of SERT binding sites (279). Thus, acute blockade of the SERT by SSRIs may not produce the same enhancement of serotoninergic transmission as that caused by the loss of SERT after longer-term administration of drug. Because sertraline treatment has been found to cause a marked loss of SERT binding sites only after 10 to 15 days of treatment (272), which corresponds to the time when drug-induced behavioral improvement becomes obvious, it may be that such loss of SERT binding sites is among the effects necessary to obtain marked enhancement of serotoninergic transmission and consequent behavioral improvement.

Signal Transduction

In recent years, there has been considerable speculation that the beneficial behavioral effects of antidepressants are a consequence of changes in the intracellular signaling pathways linked to noradrenergic or serotoninergic receptors. In other words, behavioral improvement is not a direct consequence of antidepressant-induced receptor activation (which may occur quickly); rather, it results when such receptor activation alters signaling pathways to cause more slowly developing changes in gene expression. Two major areas have been studied. One deals with effects of antidepressants on second messenger-regulated protein kinases in brain. The other deals with changes in activities of protein kinases that result in changes in gene expression and perhaps even neurogenesis. Such effects are reviewed in this section.

Protein Kinases

Phosphorylation of proteins may well be the primary regulatory mechanism for intracellular events. Such phosphorylation is controlled by protein kinases, which catalyze the binding of phosphate groups to substrate proteins, or by protein phosphatases, which catalyze the removal or release of such groups. Most often, these enzymes are the primary sites of action of the intracellular second messengers in many signaling cascades. Importantly, many kinases can regulate different and independent functions within a cell, presumably by selective co-localization with necessary substrates (280). This implies that drug effects on translocation of kinases, in addition to direct effects on their activity, may have important functional consequences.

The long-term administration of either fluoxetine or desipramine decreases the basal activity of both soluble and particulate PKC in cerebral cortex and hippocampus (281). Because PKC may be involved in the desensitization of 5-HT_{2A} receptors (282) and cell surface expression of the SERT (283), antidepressant-induced effects on PKC activity may cause changes in 5-HT_{2A}-receptor sensitivity or SERT expression (246 ,257).

Long-term, but not acute, treatment of rats with various antidepressants activates two other protein kinases in brain, namely PKA in the microtubule fraction and calcium/calmodulin-dependent protein kinase II (CaMKII) in the synaptic vesicle fraction. Such activation results in phosphate incorporation into selected substrates (284 ,285). With respect to PKA, Nestler et al. (286) had earlier reported a result consistent with the idea that antidepressants cause a translocation of PKA. They found that long-term antidepressant administration decreases the activity of PKA in the cytosol but increases enzyme activity in the nuclear fraction. Other data have also been reported suggestive of an antidepressant-induced translocation of PKA within intracellular compartments (287). Interestingly, long-term desipramine treatment increases the phosphorylation of MAP-2, a substrate for PKA (288); the increased phosphorylation is coupled to inhibition of the microtubule assembly. The effects on PKA may be caused when long-term treatment with antidepressants increases the binding of cAMP to the regulatory II subunit of PKA in brain homogenates (287 ,289).

Thus, one site affected by long-term antidepressant treatment may be cAMP-dependent phosphorylation, mediated by PKA, in microtubules. It may be speculated that such phosphorylation causes cytoskeletal changes that result in a modification of neurotransmission and antidepressantinduced changes in gene expression (see below), as PKA translocation to the nucleus is microtubule-dependent (290). Antidepressant-induced activation of PKA is interesting in light of findings of decreased PKA activity in cultured fibroblasts of melancholic patients with major depression (291), perhaps a consequence of reduced binding of cAMP to the regulatory subunit of PKA (292). Given the established facilitative function of CaMKII on neurotransmitter release (293), the effect of long-term antidepressant treatment on CMKII, with increased phosphorylation of substrates such as synapsin 1 and synaptotagmin, may underlie

the facilitation of monoamine transmitter release produced by these drugs.

Gene Expression/Neuroplasticity

Although the precise mechanism is not understood, long-term but not acute treatment with antidepressants has effects on the expression of specific genes that may be a consequence of the activation of protein kinases, particularly PKA. It is known, for example, that PKA can phosphorylate the transcription factor CREB. CREB binds to specific promoter sites (cAMP response elements) to produce changes in the expression of specific genes, such as those for brain-derived neurotrophic factor (BDNF) and its receptor, trkB. Relevant, then, to this signaling cascade is the observation that long-term but not acute administration of various types of antidepressants increases the mRNA for CREB in addition to CREB protein in brain (294 ; see 221 ,295). More recently, it was shown that such treatments increase CREB expression and CREB phosphorylation, indicative of functional activation of CREB (296). Furthermore, long-term antidepressant treatment increases BDNF and trkB expression in hippocampus (294 ,297). The increase in BDNF expression is likely to be a consequence of the increase in CREB expression (298 ,299). Finally, exogenous BDNF has been shown to have antidepressant-like activity in behavioral tests sensitive to antidepressant treatment (301).

Such results are viewed as evidence that long-term antidepressant treatment causes sustained activation of the cAMP system and the intracellular events described. The noradrenergic and serotoninergic receptors producing increases in cAMP are ß-adrenoceptors and 5-HT_4, 5-HT_6, and 5-HT_7 receptors. If such intracellular effects are responsible for clinical improvement, then these receptors may be the important ones triggering such improvement.

Finally, these slowly developing intracellular effects of antidepressants have been put into an interesting hypothesis to explain antidepressant actions (221 ,295). The hypothesis is fundamentally different from earlier views of antidepressant action, in which depression was a problem of synaptic transmission and drugs acted within the synapse to improve behavior directly (301 ,302). The new view is more morphologic in nature. It posits that depression may be caused by chronic, stress-induced atrophy of neurons in certain areas of brain, particularly the hippocampus. It is well established that such atrophy occurs (303 ,304). Further, more recent data indicate a decrease in hippocampal volume in some depressive patients (305 ,306 and 307), although this need not indicate a loss of neurons. However, postmortem studies have revealed a loss of glia and neurons in the cortex of depressed patients (137 ,308). The theory then proceeds to make use of the fact that BDNF is known to be involved not only in the differentiation and growth of neurons in the developing brain but also in neuron maintenance and survival in the adult brain (309 ,310 and 311). Thus, the antidepressant-induced increase in BDNF can oppose and perhaps overcome the stress-induced cell death pathway. Indeed, long-term antidepressant treatment has been shown recently to increase neurogenesis of dentate gyrus granule cells (312).

CONCLUSION

Part of "79 - Mechanism of Action of Antidepressants and Mood Stabilizers "

Our understanding of the mechanism of action of drugs that treat mood disorders such as depression and manic-depressive illness derives for the most part from their interaction with known signaling systems within the brain. It is evident that intracellular effects initiated by antidepressant or mood stabilizers in synaptic physiology may trigger subsequent neuroplastic changes that result in the long-term regulation of signaling in critical regions of the brain. Although much more research is needed to test this hypothesis and establish whether and how such long-term changes are of physiologic significance, current evidence suggests that such changes in brain may be quite important for the now well-established prophylactic effects of mood stabilizers and antidepressants in the treatment of recurrent mood disorders.

With the advent of new molecular biological strategies that use gene expression arrays, we have the opportunity to examine multiple targets in the brain, both known and unknown, for the action of these drugs. Within this chapter, we have tried to identify the most promising of the candidate targets of mood stabilizers and antidepressants. However, research to determine which current and future targets constitute a profile that is most relevant to the therapeutic action of these agents will continue to be hampered by a lack of animal models for these complex behavioral disorders that have strong construct and predictive validity. Although the field of antidepressant research has used animal models with some of these properties for the development of “like” agents, the development of animal models with which new mood stabilizers can be discovered has proved more challenging. We suggest that the creation of models with both construct and predictive validity to permit the discovery of novel targets directly related to therapeutic efficacy will be significantly enhanced by the identification of susceptibility and protective genes for these illnesses.

REFERENCES

1. Lenox RH, Manji HK. Lithium. In: Nemeroff CB, Schatzberg AF, eds. *American Psychiatric Association textbook of psychopharmacology*, second ed. Washington DC: American Psychiatric Association, 1998:303–349.

2. McNamara JO. Drugs effective in the therapy of the epilepsies. In: Hardman JG, Limbird LE, eds. *Goodman and Gilman’s the pharmacological basis of therapeutics*, ninth ed. New York: McGraw-Hill, 1996:461–486.

3. Hitzemann R, Mark C, Hirschowitz J, et al. RBC lithium transport in the psychoses. *Biol Psychiatry* 1989;25:296–304.

4. Sarkadi B, Alifimoff JK, Gunn RB, et al. Kinetics and stoichiometry of Na-dependent Li transport in human red blood cells. *J Gen Physiol* 1978;72:249–265.

5. Lenox RH, McNamara RK, Papke RL, et al. Neurobiology of lithium: an update. *J Clin Psychiatry* 1998;59:37–47.

6. El-Mallakh RS. *Lithium actions and mechanisms*. Washington, DC: American Psychiatric Association, 1996.

7. Swann AC, Marini JL, Sheard MH, et al. Effects of chronic dietary lithium on activity and regulation of [Na^+, K^+]-adenosine triphosphatase in rat brain. *Biochem Pharmacol* 1980;29:2819–2823.

8. Dubovsky SL. Calcium channel antagonists as novel agents for manic-depressive disorder. In: Nemeroff C, Shatzberg A, eds. *American Psychiatric Association textbook of psychopharmacology*, second ed. Washington, DC: American Psychiatric Association, 1998:455–472.

9. Arystarkhoua E, Sweadner KJ. Isoform-specific monoclonal antibodies to Na, K-ATPase alpha-subunit: evidences for a tissue-specific post-translational modification of the alpha-subunit. *J Biol Chem* 1996;271:23407–23417.

10. Munzer JS, Daly SE, Jewell-Motz EA, et al. Tissue- and isoform-specific kinetic behavior of the Na, K-ATPase. *J Biol Chem* 1994;269:16668–16676.

11. Kabakov AY, Karkanias NB, Lenox RH, et al. Synapse-specific accumulation of lithium in intra-cellular microdomains: a model for uncoupling coincidence detection in the brain. *Synapse* 1998;28:271–279.

12. Jope RS. Anti-bipolar therapy: mechanism of action of lithium. *Mol Psychiatry* 1999;4:117–128.

13. Williams RSB, Harwood, AJ. Lithium therapy and signal transduction. *Trends Pharmacol Sci* 2000;21:61–64.

14. Lenox RH, Hahn HK. Overview of the mechanism of action of lithium in the brain: fifty-year update. *J Clin Psychiatry* 2000;61[Suppl 9]:5–15.

15. Klemfuss H. Rhythms and the pharmacology of lithium. *Pharmacol Ther* 1992;56:53–78.

16. Healy D, Waterhouse JM. The circadian system and the therapeutics of the affective disorders. *Pharmacol Ther* 1995;65:241–263.

17. Goodwin FK, Jamison KR. *Manic-depressive illness*. New York: Oxford University Press, 1990.

18. Goodwin FK, Ghaemi SN. Bipolar disorder: state of the art. *Dialogues Clin Neurosci* 1999;1:41–51.

19. Dijk DJ, Duffy, JF, Czeisler CA. Circadian and sleep/wake aspects of subjective alertness and cognitive performance. *J Sleep Res* 1992;1:112–117.

20. Boivin DB, Duffy JF, Kronauer RE, et al. Dose–response relationships for resetting of human circadian clock by light. *Nature* 1996;379:540–542.

21. Hiddinga AE, Beersma DGM, Van den Hoofdakker RH. Endogenous and exogenous components in the circadian variation of core body temperature in humans. *J Sleep Res* 1997;6:156–163.

22. van den Haffaker RH. Total sleep deprivation: clinical and theoretical aspects. In: Honig AV, Praag HM, eds. *Depression: neurobiology, psychopathological and theoretical advances*. New York: John Wiley and Sons, 1997:564–589.

23. Kupfer DJ, Foster FG, Coble PA, et al. The application of EEG sleep for the differential diagnosis of affective disorders. *Am J Psychiatry* 1978;135:64–74.

24. Gillin JC, Duncan W, Pettigrew KD, et al. Successful separation of depressed, normal, and insomniac subjects by EEG sleep data. *Arch Gen Psychiatry* 1979;36:85–90.

25. Wehr TA, Goodwin FK. Can antidepressants cause mania and worsen the course of affective illness? *Am J Psychiatry* 1987;144:1403–1411.

26. Wirz-Justice A. Biological rhythms in mood disorders. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:999–1017.

27. Bicakova-Rocher A, Gorceix A, Reinberg A, et al. Temperature rhythm of patients with major affective disorders: reduced circadian period length. *Chronbiol Int* 1996;13:47–57.

28. Lenox RH, Mangi HK. Lithium. In: Nemeroff C, Schatzberg A, eds. *American Psychiatric Association textbook of psychopharmacology*. Washington, DC: American Psychiatric Association, 1995:303–349.

29. Williams MB, Jope RS. Circadian variation in rat brain AP-1 DNA binding activity after cholinergic stimulation: modulation by lithium. *Psychopharmacology* 1995;122:363–368.

30. Ikonomov O, Manji HK. Molecular mechanisms underlying mood stabilization in manic-depressive illness: the phenotype challenge. *Am J Psychiatry* 1999;156:1506–1514.

31. Hallcher LM, Sherman WR. The effects of lithium ion and other agents on the activity of *myo*inositol-1-phosphatase from bovine brain. *J Biol Chem* 1980;255:10896–10901.

32. Sherman WR, Gish BG, Honchar MP, et al. Effects of lithium on phosphoinositide metabolism *in vivo. Fed Proc* 1986;45:2639–2646.

33. Manj HK, Lenox RH. Protein kinase C signaling in the brain: molecular transduction of mood stabilization in the treatment of manic-depressive illness. *Biol Psychiatry* 1999;46:1328–1351.

34. Nahorski SR, Ragan CI, Challiss RAJ. Lithium and the phosphoinositide cycle: an example of uncompetitive inhibition and its pharmacological consequences. *Trends Pharmacol Sci* 1991;12:297–303.

35. Nahorski SR, Jenkinson S, Challiss RA. Disruption of phosphoinositide signaling by lithium. *Biochem Soc Trans* 1992;20:430–434.

36. Berridge MJ, Downes CP, Hanley MR. Lithium amplifies agonist-dependent phosphatidylinositol responses in brain and salivary glands. *Biochem J* 1982;206:587–595.

37. York JD, Ponder JW, Majerus PW. Definition of a metaldependent/Li^+-inhibited phosphomonoesterase protein family based upon a conserved three-dimensional core structure. *Proc Natl Acad Sci U S A* 1995;92:5149–5153.

38. Phiel CJ, Klein PS. Molecular targets of lithium action. *Annu Rev Pharmacol Toxicol* 2000;41.

39. Watson DG, Lenox RH. Chronic lithium-induced down-regulation of MARCKS in immortalized hippocampal cells: potentiation by muscarinic receptor activation. *J Neurochem* 1996;67:767–777.

40. Kendall DA, Nahorski SR. Acute and chronic lithium treatments influence agonist- and depolarization-stimulated inositol phospholipid hydrolysis in rat cerebral cortex. *J Pharmacol Exp Ther* 1987;241:1023–1027.

41. Kennedy ED, Challiss RJ, Nahorski SR. Lithium reduces the accumulation of inositol polyphosphate second messengers following cholinergic stimulation of cerebral cortex slices. *J Neurochem* 1989;53:1652–1655.

42. Kennedy ED, Challiss RJ, Ragan CI, et al. Reduced inositol polyphosphate accumulation and inositol supply induced by lithium in stimulated cerebral cortex slices. *Biochem J* 1990;267:781–786.

43. Acharya JK, Labarca P, Delgado R, et al. Synaptic defects and compensatory regulation of inositol polyphosphate-1-phosphatase mutants. *Neuron* 1998;20:1219–1229.

44. Kofman O, Belmaker RH, Grisaru N, et al. *Myo*-inositol attenuates two specific behavioral effects of acute lithium in rats. *Psycopharmacol Bull* 1991;27:185–190.

45. Tricklebank MD, Singh L, Jackson A, et al. Evidence that a proconvulsant action of lithium is mediated by inhibition of *myo*-inositol phosphatase in mouse brain. *Brain Res* 1991;558:145–148.

46. Becchetti A, Whitaker M. Lithium blocks cell cycle transitions in the first cell cycles of sea urchin embryos, an effect rescued by *myo*-inositol. *Development* 1987;124:1099–1107.

47. Hedgepeth CM, Conrad LJ, Zhang J, et al. Activation of the Wnt signaling pathway: a molecular mechanism for lithium action. *Dev Biol* 1997;185:82–91.

48. Jope RS, Williams MB. Lithium and brain signal transduction systems. *Biochem Pharmacol* 1994;77:429–441.

49. Moore GJ, Bebchuk JM, Parrish JK, et al. Temporal dissociation between lithium-induced CNS *myo*-inositol changes and clinical response in manic-depressive illness. *Am J Psychiatry* 1999;156:1902–1908.

50. Kuriyama K, Kakita K. Cholera toxin-induced epileptogenic focus: an animal model for studying roles of cyclic AMP in the establishment of epilepsy. *Prog Clin Biol Res* 1980;39:141–55.

51. Ludvig N, Moshe SL. Different behavioral and electrographic effects of acoustic stimulation and dibutyryl cyclic AMP injection into the inferior colliculus in normal and in genetically epilepsy-prone rats. *Epilepsy Res* 1989;3:185–190.

52. Ludvig N, Mishra PK, Jobe PC. Dibutyryl cyclic AMP has epileptogenic potential in the hippocampus of freely behaving rats: a combined EEG–intracerebral microdialysis study. *Neurosci Lett* 1992;141:187–191.

53. Warsh JJ, Young LT, Li PP. Guanine nucleotide binding (G) protein disturbances in bipolar disorder. In: Manji HK, Bowden CL, Belmaker RH, eds. *Bipolar medications: mechanisms of action*. Washington, DC: American Psychiatric Association, 2000:299–329.

54. Newman M, Klein E, Birmaher B, et al. Lithium at therapeutic concentrations inhibits human brain noradrenaline-sensitive cyclic AMP accumulation. *Brain Res* 1983;278:380–381.

55. Mork A, Geisler A. Effects of lithium *ex vivo* on the GTP-mediated inhibition of calcium-stimulated adenylate cyclase activity in rat brain. *Eur J Pharmacol* 1989;168:347–354.

56. Newman ME, Belmaker RH. Effects of lithium *in vitro* and *ex vivo* on components of the adenylate cyclase system in membranes from the cerebral cortex of the rat. *Neuropharmacology* 1987;26:211–217.

57. Chen G, Hawver D, Wright C, et al. Attenuation of adenylyl cyclases by carbamazepine *in vitro. J Neurochem* 1996;67:2079–2086.

58. Ferrendelli JA, Kinscherf DA. Inhibitory effects of anticonvulsant drugs on cyclic nucleotide accumulation in brain. *Ann Neurol* 1979;5:533–538.

59. Elphick M, Anderson SM, Hallis KF, et al. Effects of carbamazepine on 5-hydroxytryptamine function in rodents. *Psychopharmacology* 1990;100:49–53.

60. Van Calker D, Steber R, Klotz KN, et al. Carbamazepine distinguishes between adenosine receptors that mediate different second messenger responses. *Eur J Pharmacol* 1991;206:285–290.

61. Post RM, Ballenger JC, Uhde TW, et al. Effect of carbamazepine on cyclic nucleotides in CSF of patients with affective illness. *Biol Psychiatry* 1982;17:1037–1045.

62. Ebstein RP, Hermoni M, Belmaker RH. The effect of lithium on noradrenaline-induced cyclic AMP accumulation in rat brain: inhibition after chronic treatment and absence of supersensitivity. *J Pharmacol Exp Ther* 1980;213:161–167.

63. Masana MI, Bitran JA, Hsiao JK, et al. *In vivo* evidence that lithium inactivates G[I] modulation of adenylate cyclase in brain. *J Neurochem* 1992;59:200–205.

64. Wiborg O, Kruger T, Jakobsen SN. Region-selective effects of long-term lithium and carbamazepine administration on cyclic AMP levels in rat brain. *Pharmacol Toxicol* 1999;84:88–93.

65. Colin SF, Chang HC, Mollner S, et al. Chronic lithium regulates the expression of adenylate cyclase and G_i-protein α subunit in rat cerebral cortex. *Proc Natl Acad Sci USA* 1991;88:10634–10637.

66. Jensen JB, Mikkelsen JD, Mork A. Increased adenylyl cyclase type 1 mRNA, but not adenylyl cyclase type 2 in the rat hippocampus following antidepressant treatment. *Eur Neuropsychopharmacol* 2000;10:105–111.

67. Mork A, Geisler A, Hollund P. Effects of lithium on second messenger systems in the brain. *Pharmacol Toxicol* 1992;71[Suppl 1]:4–17.

68. Avissar S, Schreiber G. The involvement of guanine nucleotide binding proteins in the pathogenesis and treatment of affective disorders. *Biol Psychiatry* 1992;31:435–459.

69. Avissar S, Schreiber G, Danon A, et al. Lithium inhibits adrenergic and cholinergic increases in GTP binding in rat cortex. *Nature* 1988;331:440–442.

70. Song L, Jope RS. Chronic lithium treatment impairs phosphatidylinositol hydrolysis in membranes from rat brain regions. *J Neurochem* 1992;58:2200–2206.

71. Dixon JF, Los GV, Hokin LE. Lithium stimulates glutamate "release" and inositol-1,4,5-triphosphate accumulation via activation of the *N*-methy-D-aspartate receptor in monkey and mouse cerebral cortex slices. *Proc Natl Acad Sci USA* 1994;91:8358–8362.

72. Jope RS, Song L, Kolasa K. Inositol trisphosphate, cyclic AMP, and cyclic GMP in rat brain regions after lithium and seizures. *Biol Psychiatry* 1992;31:505–514.

73. Wang HY, Friedman E. Effects of lithium on receptor-mediated activation of G proteins in rat brain cortical membranes. *Neuropharmacology* 1991;38:403–414.

74. Ellis J, Lenox RH. Receptor coupling to G proteins: interactions not affected by lithium. *Lithium* 1991;2:141–147.

75. Li PP, Tam YK, Young LT, et al. Lithium decreases Gs, Gi-1 and Gi-2 alpha-subunit mRNA levels in rat cortex. *Eur J Pharmacol* 1991;206:165–166.

76. Li PP, Young LT, Tam YK, et al. Effects of chronic lithium and carbamazepine treatment on G-protein subunit expression in rat cerebral cortex. *Biol Psychiatry* 1993;34:162–170.

77. Manji HK, Lenox RH. Signaling: cellular insights into the pathophysiology of bipolar disorder. *Biol Psychiatry* 2000;48:518–530.

78. Young LT, Woods CM. Mood stabilizers have differential effects on endogenous ADP ribosylation in C6 glioma cells. *Eur J Pharmacol* 1996;309:215–218.

79. Nestler EJ, Terwilliger RZ, Duman RS. Regulation of endogenous ADP-ribosylation by acute and chronic lithium in rat brain. *J Neurochem* 1995;64:2319–2324.

80. Lenox RH. Role of receptor coupling to phosphoinositide metabolism in the therapeutic action of lithium. In: Ehrlich, et al., eds. *Molecular mechanisms of neuronal responsiveness*. New York: Plenum Publishing, 1987:515–530 (*Advances in Experimental Biology and Medicine*, vol 221).

81. Manji HK, Lenox RH. Long-term action of lithium: a role for transcriptional and posttranscriptional factors regulated by protein kinase C. *Synapse* 1994;16:11–28.

82. Hahn C-G, Friedman E. Abnormalities in protein kinase C signaling and the pathophysiology of bipolar disorder. *Bipolar Disord* 1999;2:81–86.

83. Daigen A, Akiyama K, Otsuki S. Long-lasting change in the membrane-associated protein kinase C activity in the hippocampal kindled rat. *Brain Res* 1991;545:131–136.

84. Kantor L, Gnegy ME. Protein kinase C inhibitors block amphetamine-mediated dopamine release in rat striatal slices. *J Pharmacol Exp Ther* 1998;284:592–598.

85. Manji HK, Etcheberrigaray R, Chen G, et al. Lithium dramatically decreases membrane-associated PKC in the hippocampus: selectivity for the alpha isozyme. *J Neurochem* 1993;61:2303–2310.

86. Chen G, Rajkowska G, Du F, et al. Enhancement of hippocampal neurogenesis by lithium. *J Neurochem* 2000;75:1729–1734.

87. Friedman E, Wang HY. Effect of chronic lithium treatment on 5-hydroxytryptamine autoreceptors and release of 5-[^{3}H]hydroxytryptamine from rat brain cortical, hippocampal, and hypothalamic slices. *J Neurochem* 1998;50:195–201.

88. Friedman E, Wang H-Y, Levinson D, et al. Altered protein kinase C activity in bipolar affective disorder, manic episode. *Biol Psychiatry* 1993;33:520–525.

89. Chen G, Manji HK, Hawver DB, et al. Chronic sodium valproate selectivity decreases protein kinase C α and ε *in vitro. J Neurochem* 1994;63:2361–2364.

90. Manji HK. A preliminary investigation of a protein kinase C inhibitor in the treatment of acute mania. *Arch Gen Psychiatry* 2000;57:95–97.

91. Bebchuk JM, Arfken CL, Dolan-Manji S, et al. A preliminary investigation of a protein kinase C inhibitor (tamoxifen) in the treatment of acute mania. *Arch Gen Psychiatry* 2000;57:95–97.

92. Lenox RH, Watson DG. Lithium and the brain: a psychopharmacological strategy to a molecular basis for manic-depressive illness. *Clin Chem* 1994;40:309–314.

93. Lenox RH, Watson DG, Patel J, et al. Chronic lithium administration alters a prominent PKC substrate in rat hippocampus. *Brain Res* 1992;570:333–340.

94. Manji HK, Bersudsky Y, Chen G, et al. Modulation of protein kinase C isozymes and substrates by lithium: the role of *myo*-inositol. *Neuropsychopharmacology* 1996;15:370–381.

95. Wang LX, Liu, Lenox RH. Transcriptional down-regulation of MARCKS gene expression in immortalized hippocampal cells by lithium. *J Neurochem* 2001; in press.

96. Watson DG, Watterson JM, Lenox RH. Sodium valproate down-regulates the myristoylated alanine-rich c kinase substrate (MARCKS) in immortalized hippocampal cells: a property unique to PKC-mediated mood stabilization. *J Pharmacol Exp Ther* 1998;285:307–316.

97. Mori S, Tardito D, Dorigo A, et al. Effects of lithium on cAMP-dependent protein kinase in rat brain. *Neuropsychopharmacology* 1998;19:233–240.

98. Perez J, Tardito D, Mori S, et al. Abnormalities of cAMP signaling in affective disorders: implications for pathophysiology and treatment. *Bipolar Disord* 2000;2:27–36.

99. Ozaki N, Chuang D-M. Lithium increases transcription factor binding to AP-1 and cyclic AMP-responsiveness element in cultured neurons and rat brain. *J Neurochem* 1997;69:2336–2344.

100. Wang HY, Markowitz P, Levinson D, et al. Increased membrane-associated protein kinase C activity and translocation in blood platelets from bipolar affective disorder patients. *J Psychiatr Res* 1999;33:171–179.

101. Rahman S, Li PP, Young LT, et al. Reduced [^{3}H] cyclic AMP binding in postmortem brain from subjects with bipolar affective disorder. *J Neurochem* 1997;68:297–304.

102. Fields A, Li PP, Kish SJ, et al. Increased cyclic AMP-dependent protein kinase activity in postmortem brain from patients with bipolar affective disorder. *J Neurochem* 1999;73:1704–1710.

103. Stachel SE, Grunwald DJ, Myers PZ, et al. Lithium perturbation and goosecoid-expression identify a dorsal specification pathway in the pregastrula zebrafish. *Development* 1993;117:1261–1274.

104. Kao KR, Masui U, Elinson RP, et al. Lithium-induced respecification of pattern in *Xenopus* laevic embryos. *Nature* 1986;322:371–373.

105. Berridge MJ, Downes CP, Hanley MR. Neural and developmental actions of lithium: a unifying hypothesis. *Cell* 1989;59:411–419.

106. Ault KT, Durmwicz G, Galione A, et al. Modulation of *Xenopus* embryo mesoderm-specific gene expression and dorsoanterior patterning by receptors that activate the phosphatidylinositol cycle signal transduction pathway. *Development* 1996;122:2033–2041.

107. Busa WB, Gimlich RL. Lithium-induced teratogenesis in frog embryos prevented by a polyphosphoinositide cycle intermediate or a diacylglycerol analog. *Dev Biol* 1989;132:315–324.

108. He X, Saint-Jeannet JP, Woodgett JR, et al. Glycogen synthase kinase-3 and dorsoventral patterning in *Xenopus* embryos. *Nature* 1995;374:617–622.

109. Harwood AJ, et al. Glycogen synthase kinase 3 regulates cell fate in *Dictyostelium Cell* 1995;80:139–148.

110. Klein PS, Melton DA. A molecular mechanism for the effect of lithium on development. *Proc Natl Acad Sci USA* 1996;93:8455–8459.

111. Stambolic V, Laurent R, Woodgett JR. Lithium inhibits glycogen synthase kinase-3 activity and mimics wingless signaling in intact cells. *Curr Biol* 1996;6:1664–1668.

112. Chen B, Wang JF, Hill BC, et al. Lithium and valproate differentially regulate brain regional expression of phosphorylated CREB and c-*fos. Mol Brain Res* 1999;70:45–53.

113. Kalaspudi VD, Sheftel G, Divish MM, et al. Lithium augments c-*fos* protooncogene expression in PC 12 pheochromocytoma cells: implications for therapeutic action of lithium. *Brain Res* 1990;521:47–54.

114. Leslie RA, Moorman JM, Grahame-Smith DG. Lithium enhances 5-HT_{2A} receptor-mediated c-*fos* expression in rat cerebral cortex. *Neuroreport* 1993;5:241–244.

115. Miller JC, Mathe AA. Basal and stimulated c-*fos* mRNA expression in the rat brain: effect of chronic dietary lithium. *Neuropsychopharmacology* 1997;16:408–418.

116. Mathe AA, Miller JC, Stenfors C. Chronic dietary lithium inhibits basal c-*fos* mRNA expression in rat brain. *Prog Neuropsychopharmacol Biol Psychiatry* 1995;19:1177–1187.

117. Arenander AT, de Veliis J, Herschman HR. Induction of c-*fos* and *TIS* genes in cultured rat astrocytes by neurotransmitters. *J Neurosci Res* 1989;24:107–114

118. Thiele TE, Seeley RJ, D'Alessio D, et al. Central infusion of glucagon-like peptide-1-(7-36) amide (GLP-1) receptor antagonist attenuates lithium chloride-induced c-Fos induction in rat brainstem. *Brain Res* 1998;801:164–170.

119. Lee Y, Hamamura T, Ohashi K, et al. The effect of lithium on methamphetamine-induced regional Fos protein expression in the rat brain. *Neuroreport* 1999;10:895–900.

120. Chen G, Yuan PX, Jiang YM, et al. Valproate robustly enhances AP-1-mediated gene expression. *Mol Brain Res* 1999;64:52–58

121. Manji HK, Moore GJ, Chen G. Lithium at 50: have the neuroprotective effects of this unique cation been overlooked? *Biol Psychiatry* 1999;46:929–940.

122. Yuan PX, Chen G, Huang LD, et al. Lithium stimulates gene expression through the AP-1 transcription factor pathway. *Mol Brain Res* 1998;58:225–230.

123. Unlap MT, Jope RS. Lithium attenuates nerve growth factor-induced activation of AP-1 DNA binding activity in PC12 cells. *Neuropsychopharmacology* 1997;17:12–17.

124. Jope RS, Song L. AP-1 and NF-kappaB stimulated by carbachol in human neuroblastoma SH-SY5Y cells are differentially sensitive to inhibition by lithium. *Brain Res* 1997;50:171–180.

125. Jope RS. A bimodal model of the mechanism of action of lithium. *Mol Psychiatry* 19994:21–25.

126. Asghari V, Wang JF, Reiach JS, et al. Differential effects of mood stabilizers on Fos/Jun proteins and AP-1 DNA binding activity in human neuroblastoma SH-SY5Y cells. *Mol Brain Res* 1998;58:95–102

127. Divish MM, Sheftel G, Boyle A, et al. Differential effect of lithium on *fos* protooncogene expression mediated by receptor and postreceptor activators of protein kinase C and cyclic adenosine monophosphate: model for its antimanic action. *J Neurosci* 1991;28:40–48.

128. Berrettini W, Ferraro T, Choi H, et al. Linkage studies of bipolar illness. *Arch Gen Psychiatry* 1997;54:32–39.

129. King DP, Takahashi JS. Molecular genetics of circadian rhythms in mammals. *Annu Rev Neurosci* 2000;23:713–742.

130. Nonaka S, Hough CJ, Chuang DM. Chronic lithium treatment robustly protects neurons in the central nervous system against excitotoxicity by inhibiting *N*-methyl-D-aspartate receptor-mediated calcium influx. *Proc Natl Acad Sci USA* 1998;95:2642–2647.

131. Alvarez G, Munoz-Montano JR, Satrustegui J, et al. Lithium protects cultured neurons against beta-amyloid-induced neurodegeneration. *FEBS Lett* 1999;453:260–264.

132. Khodorov B, Pinelis V, Vinskaya, et al. Li^+ protects nerve cells against destabilization of Ca^{2+} homeostasis and delayed death caused by removal of external Na^+. *FEBS Lett* 1999;448:173–176.

133. Chen G, Zeng WZ, Yuan PX, et al. The mood-stabilizing agents lithium and valproate robustly increase the levels of the neuroprotective protein bcl-2 in the CNS. *J Neurochem* 1999;72:879–882.

134. Manji HK, Bebchuk JM, Moore GJ, et al. Modulation of CNS signal transduction pathways and gene expression by mood-stabilizing agents: therapeutic implications. *J Clin Psychiatry* 1999;60[Suppl 2]:27–39

135. Chen RW, Chuang DM. Long-term lithium treatment suppresses *p53* and *Bax* expression but increases *Bcl-2* expression: a prominent role in neuroprotection against excitotoxicity. *J Biol Chem* 1999;274:6039–6042.

136. Vanhaesebroeck B, Leevers S, et al. Phosphoinositide 3-kinases: a conserved family of signal transducers. *Trends Biol Sci* 1997;22:267–272.

137. Rajkowska G, Miguel-Hidalgo JJ, Wei J, et al. Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. *Biol Psychiatry* 1999;45:1085–1098.

138. Munoz-Montano JR, Moreno FJ, Avila J, et al. Lithium inhibits Alzheimer's disease-like tau protein phosphorylation in neurons. *FEBS Lett* 1997;411:183–188.

139. Hong M, Chen DCR, Klein PS, et al. Lithium reduces tau phosphorylation by inhibition of glycogen synthase kinase-3. *J Biol Chem* 1997;272:25326–25332.

140. Lovestone S, Davis DR, Webster MT, et al. Lithium reduces tau phosphorylation: effects in living cells and in neurons at therapeutic concentrations. *Biol Psychiatry* 1999;45:995–1003.

141. Ahlawalia P, Grewaal DS, Singhal RL. Brain GABAergic and dopaminergic systems following lithium treatment and withdrawal. *Neuropharmacology* 1981;20:483–487.

142. Lucas FR, Goold RG, Gordon-Weeks PR, et al. Inhibition of GSK-3β leading to the loss of phosphorylated MAP-1B is an early event of axonal remodeling induced by WNT-7α or lithium. *J Cell Sci* 1998;111:1351–1361.

143. Garcia-Perez J, Avila J, Diaz-Nido J. Implication of cyclindependent kinases and glycogen synthase kinase 3 in the phosphorylation of microtubule-associated protein 1B in developing neuronal cells. *J Neurosci Res* 1998;52:445–452.

144. McNamara RK, Stumpo DJ, Lewis MH, et al. Effects of alterations in MARCKS expression on spatial learning in mutant mice: transgenic rescue and interaction with gene background. *Proc Natl Acad Sci USA* 1998;95:14517–14522.

145. Stumpo DJ, Bock CB, Tuttle JS, et al. MARCKS deficiency in mice leads to abnormal brain development and perinatal death. *Proc Natl Acad Sci USA* 1995;92:944–948.

146. Swierczynski SL, Blackshear PJ. Myristoylation-dependent and electrostatic interactions exert independent effects on the membrane association of the myristoylated alanine-rich protein kinase C substrate protein in intact cells. *J Biol Chem* 1996;271: 3424–23430.

147. McNamara RK, Lenox RH. Comparative distribution of myristoylated alanine-rich C kinase substrate (MARCKS) and *F1/GAP-43* gene expression in the adult rat brain. *J Comp Neurol* 1997;379:2–27.

148. McNamara RK, Lenox RH. Distribution of protein kinase C substrates MARCKS and MRP in the postnatal developing rat brain. *J Comp Neurol* 1998;397:337–356.

149. McNamara RK, Hyde TM, Kleinman JE, et al. Expression of the myristoylated alanine-rich C kinase substrate (MARCKS) and MARCKS-related protein (MRP) in the prefrontal cortex and hippocampus of suicide victims. *J Clin Psychiatry* 1999;60 [Suppl 2]:21–26.

150. McNamara RK, Jiang Y, Streit WJ, et al. Facial motor neuron regeneration induces a unique spatial and temporal pattern of myristoylated alanine-rich C kinase substrate (MARCKS) expression. *Neuroscience* 2000;97:581–589.

151. Ramakers GMJ, McNamara RK, Lenox RH, et al. Differential changes in the phosphorylation of the protein kinase C substrates MARCKS and GAP 43/B-50 following Schaffer collateral long-term potentiation and long-term depression. *J Neurochem* 1999;73:125–130.

152. Wong EHF, Sonders MS, Amara SG, et al. Reboxetine: a pharmacologically potent, selective, and specific norepinephrine reuptake inhibitor. *Biol Psychiatry* 2000;47:818–829.

153. DeBoer T, Maura G, Raiteri M, et al. Neurochemical and autonomic pharmacological profiles of the 6-AZA-analogue of mianserin, ORG 3770 and its enantiomers. *Neuropharmacology* 1988;27:399–408.

154. Gobbi M, Frittoli E, Mennini T. Further studies on α_2-adrenoceptor subtypes involved in the modulation of [^{3}H]noradrenaline and [^{3}H]5-hydroxytryptamine release from rat brain cortex synaptosomes. *J Pharm Pharmacol* 1993;45:811–814.

155. Gobert A, Rivet JM, Cistarelli L, et al. Potentiation of the fluoxetine-induced increase in dialysate levels of serotonin (5-HT) in the frontal cortex of freely moving rats by combined blockade of 5-HT_{1A} and 5-HT_{1B} receptors with WAY 100,635 and GR 127,935. *J Neurochem* 1997;68:1159–1163.

156. Kawahara Y, Kawahara H, Westerink BHC. Tonic regulation of the activity of noradrenergic neurons in the locus coeruleus of the conscious rat studied by dual-probe microdialysis. *Brain Res* 1999;823:42–48.

157. DeBoer T, Nefkens E, Helvoirt A, et al. Differences in modulation of noradrenergic and serotoninergic transmission by the alpha$_2$-adrenoceptor antagonists mirtazapine, mianserin, and idazoxan. *J Pharmacol Exp Ther* 1996;277:852–860.

158. Haddjeri N, Blier P, de Montigny C. Effect of the α2-adrenoceptor antagonist mirtazapine on the 5-hydroxytryptamine system in the rat brain. *J Pharmacol Exp Ther* 1996;277:861–871.

159. Haddjeri N, Blier P, de Montigny C. Effects of long-term treatment with the α_2-adrenoceptor antagonist mirtazapine on 5-HT neurotransmission. *Naunyn Schmiedebergs Arch Pharmacol* 1997;355:20–29.

160. Baraban JM, Aghajanian GK. Suppression of firing activity of 5-HT neurons in the dorsal raphe by alpha-adrenoceptor antagonists. *Neuropharmacology* 1980;19:355–363.

161. Hjorth S, Bengtsson HJ, Milano S, et al. Studies on the role of 5-HT_{1A} autoreceptors and α_1-adrenoceptors in the inhibition of 5-HT release I. BMY7378 and prazosin. *Neuropharmacology* 1995;34:383–392.

162. Tao R, Hjorth S. α_2-Adrenoceptor modulation of rat ventral hippocampal 5-hydroxytryptamine release *in vivo*. *Naunyn Schmiedebergs Arch Pharmacol* 1992;345:137–143.

163. Millan MJ, Gobert A, Rivet JM, et al. Mirtazapine enhances frontocortical dopaminergic and corticolimbic adrenergic, but not serotoninergic, transmission by blockade of α_2-adrenergic and serotonin$_{2C}$ receptors: a comparison with citalopram. *Eur J Neurosci* 2000;12:1079–1095.

164. Ascher JA, Cole JO, Colin J-N, et al. Bupropion: a review of its mechanism of antidepressant activity. *J Clin Psychiatry* 1995;56:395–401.

165. Golden RN, Rudorfer MV, Sherer MA, et al. Bupropion in depression, I: biochemical effects and clinical response. *Arch Gen Psychiatry* 1988;45:139–143.

166. Cusack B, Nelson A, Richelson E. Binding of antidepressants to human brain receptors: focus on newer-generation compounds. *Psychopharmacology* 1994;114:559–565.

167. Bolden-Watson C, Richelson E. Blockade by newly developed antidepressants of biogenic amine uptake into rat brain synaptosomes. *Life Sci* 1993;52:1023–1029.

168. Gudelsky GA, Koenig JI, Meltzer HY. Thermoregulatory responses to serotonin (5-HT) receptor stimulation in the rat. Evidence for opposing roles of 5-HT$_2$ and 5-HT$_{1A}$ receptors. *Neuropharmacology* 1986;25:1307–1313.

169. Backus LI, Sharp T, Grahame-Smith DG. Behavioral evidence for a functional interaction between central 5-HT$_2$ and 5-HT$_{1A}$ receptors. *Br J Pharmacol* 1990;100:793–799.

170. Herremans AH, van der Hayden JAM, van Drimmelen M, et al. The 5-HT$_{1A}$ receptor agonist flesinoxan shares discriminative stimulus properties with some 5-HT$_2$ receptor antagonists. *Pharmacol Biochem Behav* 1999;64:389–395.

171. Hensler JG, Truett KA. Effect of chronic serotonin-2 receptor agonist or antagonist administration on serotonin-1A receptor sensitivity. *Neuropsychopharmacology* 1998;19:354–364.

172. Barker EL, Blakely RD. Norepinephrine and serotonin transporters. Molecular targets of antidepressant drugs. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:321–333.

173. Kenakin T. Human recombinant receptor systems. In: Kenakin T. *Pharmacologic analysis of drug–receptor interaction*, third ed. Philadelphia: Lippincott-Raven Publishers, 1997:57–77.

174. Hyttel J, Larsen JJ. Serotonin-selective antidepressants. *Acta Pharmacol Toxicol* 1985;56[Suppl 1]:146–153.

175. Owens MJ, Morgan WN, Plott SJ, et al. Neurotransmitter receptor and transporter binding profile of antidepressants and their metabolites. *J Pharmacol Exp Ther* 1997;283:1305–1322.

176. Tatsumi M, Groshan K, Blakely RD, et al. Pharmacological profile of antidepressants and related compounds at human monoamine transporters. *Eur J Pharmacol* 1997;340:249–258.

177. Eshleman A, Carmolli M, Cumbay M, et al. Characteristics of drug interactions with recombinant biogenic amine transporters expressed in the same cell type. *J Pharmacol Exp Ther* 1999;289:877–885.

178. Hyttel J. Effect of a specific 5-HT uptake inhibitor, citalopram (Lu 10-171), on ^{3}H-5-HT uptake in rat brain synaptosomes *in vitro. Psychopharmacology* 1978;60:13–18.

179. Preskorn SH. De-spinning *in vitro* data. *J Pract Psychiatry Behav Health* 1999;283–287.

180. van Harten J. Clinical pharmacokinetics of selective serotonin reuptake inhibitors. *Clin Pharmacokinet* 1993;24:203–220.

181. Sallee FR, Pollock BG. Clinical pharmacokinetics of imipramine and desipramine. *Clin Pharmacokinet* 1990;18:346–364.

182. Hanin I, Koslow SH, Kocsis JH, et al. Cerebrospinal fluid levels of amitriptyline, nortriptyline, imipramine and desmethylimipramine: relationship to plasma levels and treatment outcome. *J Affect Disord* 1985;9:69–78.

183. Martensson B, Nyberg S, Toresson G, et al. Fluoxetine treatment of depression. *Acta Psychiatr Scand* 1989;79:586–596.

184. Muscettola G, Goodwin FK, Potter WZ, et al. Imipramine and desipramine in plasma and spinal fluid. *Arch Gen Psychiatry* 1978;35:621–625.

185. Nordin C, Bertilsson L, Siwers B. CSF and plasma levels of nortriptyline and its 10-hydroxy metabolite. *Br J Clin Pharmacol* 1985;20:411–413.

186. Lundmark J, Walinder J, Alling C, et al. The effect of paroxetine on cerebrospinal fluid concentrations of neurotransmitter metabolites in depressed patients. *Eur Neuropsychopharmacol* 1994;4:1–6.

187. Benet LZ, Oie S, Schwartz JB. Design and optimization of dosage regimens: pharmacokinetic data. In: Hardman JG, Linbird L, eds. *Goodman and Gilman's the pharmacological basis of therapeutics*, ninth ed. New York: McGraw-Hill, 1996:1707–1792.

188. Bjerkenstedt L, Flyckt L, Fredrickson Overo K, et al. Relationship between clinical effects, serum drug concentration and serotonin uptake inhibition in depressed patients treated with citalopram. *Eur J Clin Pharmacol* 1985;28:553–557.

189. Fredricson Overo OK. Kinetics of citalopram in man: plasma levels in patients. *Prog Neuropsychopharmacol Biol Psychiatry* 1982;6:311–318.

190. Nordin C, Bertilsson L. Active hydroxymetabolites of antidepressants. Emphasis on E-10-hydroxy-nortriptyline. *Clin Pharmacokinet* 1995;28:26–40.

191. Nordin C, Bertilsson L, Siwers B. Clinical and biochemical effects during treatment of depression with nortriptyline: the role of 10-hydroxynortriptyline. *Clin Pharmacol Ther* 1987;42:10–19.

192. Potter WZ, Calil HM, Manian AA, et al. Hydroxylated metabolites of tricyclic antidepressants: preclinical assessment of activity. *Biol Psychiatry* 1979;14:601–613.

193. Deleted in press.

194. Delgado PL, Miller HL, Salomon RM, et al. Tryptophan-depletion challenge in depressed patients treated with desipramine or fluoxetine: implications for the role of serotonin in the mechanism of antidepressant action. *Soc Biol Psychiatry* 1999;46:212–220.

195. Manias B, Taylor DA. Inhibition of *in vitro* amine uptake into rat brain synaptosomes after *in vivo* administration of antidepressants. *Eur J Pharmacol* 1983;95:305–309.

196. Hensler JG, Kovachich GB, Frazer A. A quantitative autoradiographic study of serotonin$_{1A}$ receptor regulation. *Neuropsychopharmacology* 1991;4:131–144.

197. Page ME, Detke MJ, Dalvi A, et al. Serotoninergic mediation of the effects of fluoxetine, but not desipramine, in the rat forced swimming test. *Psychopharmacology (Berl)* 1999;147:162–167.

198. Sacerdote P, Brini A, Mantegazza P, et al. A role for serotonin and beta-endorphin in the analgesia induced by some tricyclic antidepressant drugs. *Pharmacol Biochem Behav* 1987;26:153–158.

199. Winslow JT, Insel TR. Serotoninergic and catecholaminergic reuptake inhibitors have opposite effects on the ultrasonic isolation calls of rat pups. *Neuropsychopharmacology* 1990;3:51–59.

200. Nyback HV, Walters JR, Aghahanian GK, et al. Tricyclic antidepressants: effects on the firing rate of brain noradrenergic neurons. *Eur J Pharmacol* 1975;32:302–312.

201. Ross SB, Renyi AL. Tricyclic antidepressant agents. II. Effect of oral administration on the uptake of ^{3}-H-noradrenaline and ^{14}-C-5-hydroxytryptamine in slices of the midbrain-hypothalamus region of the rat. *Acta Pharmacol Toxicol* 1975;36:395–408.

202. Fuller RW, Perry KW, Molloy BB. Effect of 3-(*p*-trifluoromethylphenoxy)-*N*-methyl-3-phenylpropylamine on the depletion of brain serotonin by 4-chloroamphetamine. *J Pharmacol Exp Ther* 1975;193:796–803.

203. Lemberger L, Rowe H, Carmichael R, et al. Fluoxetine, a selective serotonin uptake inhibitor. *Clin Pharmacol Ther* 1978;23:421–429.

204. Stark P, Fuller RW, Wong DT. The pharmacologic profile of fluoxetine. *J Clin Psychiatry* 1985;46:7–13.

205. Hassan SM, Wainscott G, Turner P. A comparison of the effect of paroxetine and amitriptyline on the tyramine pressor response test. *Br J Clin Pharmacol* 1985;19:705–706.

206. Abdelmawla AH, Langley RW, Szabadi E, et al. Comparison of the effects of venlafaxine, desipramine, and paroxetine on noradrenaline- and methoxamine-evoked constriction of the dorsal hand vein. *Br J Clin Pharmacol* 1999;48:345–354.

207. Bitsios P, Szabadi E, Bradshaw CM. Comparison of the effects of venlafaxine, paroxetine and desipramine on the pupillary light reflex in man. *Psychopharmacology* 1999;143:286–292.

208. Little JT, Ketter TA, Mathe AA, et al. Venlafaxine but not bupropion decreases cerebrospinal fluid 5-hydroxyindoleacetic acid in unipolar depression. *Biol Psychiatry* 1999;45:285–289.

209. Harvey AT, Rudolph RL, Preskorn SH. Evidence of the dual mechanisms of action of venlafaxine. *Arch Gen Psychiatry* 2000;57:503–509.

210. Quitkin FM, Rabkin JG, Ross D, et al. Identification of true drug response to antidepressants. *Arch Gen Psychiatry* 1984;41:782–786.

211. Quitkin FM, McGrath PJ, Stewart JW, et al. Can the effects of antidepressants be observed the first two weeks? *Neuropsychopharmacology* 1996;15:390–394.

212. Stassen HH, Delini-Stula A, Angst J. Time course of improvement under antidepressant treatment: a survival-analytical approach. *Eur Neuropsychopharmacol* 1993;3:127–135.

213. Montgomery SA. Rapid onset of action of venlafaxine. *Int Clin Psychopharmacol* 1995;10[Suppl 2]:21–27.

214. Katz MM, Koslow SH, Frazer A. Onset of antidepressant activity: reexamining the structure of depression and multiple actions of drugs. *Depression Anxiety* 1996/1997;4:257–267.

215. Keller MB. The long-term treatment of depression. *J Clin Psychiatry* 1999;60:41–45.

216. Frank E, Kupfer DJ, Perel JM, et al. Three-year outcomes for maintenance therapies in recurrent depression. *Arch Gen Psychiatry* 1990;47:1093–1099.

217. Frazer A. Pharmacology of antidepressants. *J Clin Psychopharmacol* 1997;17:2S–18S.

218. Mongeau R, Blier P, deMontigny C. The serotoninergic and noradrenergic systems of the hippocampus: their interactions and the effects of antidepressant treatment. *Brain Res* 1997;23:145–195.

219. Pineyro G, Blier P. Autoregulation of neurons: role in antidepressant drug action. *Pharmacol Rev* 1999;51:533–591.

220. Duman RS. Novel therapeutic approaches beyond the 5-HT receptor. *Biol Psychiatry* 1998;44:324–335.

221. Duman RS, Malberg J, Thome J. Neural plasticity to stress and antidepressant treatment. *Biol Psychiatry* 1999;46:1181–1191.

222. Popoli M, Venegoni A, Vocaturo C, et al. Long-term blockade of serotonin reuptake affects synaptotagmin phosphorylation in the hippocampus. *Mol Pharmacol* 2000;51:19–26.

223. Sugrue MF. Current concepts on the mechanisms of action of antidepressant drugs. *Pharmacol Ther* 1981;13:219–247.

224. Bel N, Artigas F. Fluvoxamine preferentially increases extracellular 5-hydroxytryptamine in the raphe nuclei: an *in vivo* microdialysis study. *Eur J Pharmacol* 1992;229:101–103.

225. Invernizzi R, Belli S, Samanin R. Citalopram's ability to increase the extracellular concentrations of serotonin in the dorsal raphe prevents the drug's effect in the frontal cortex. *Brain Res* 1992;584:322–324.

226. Hervás I, Queiroz CMT, Adell A, et al. Role of uptake inhibition and autoreceptor activation in the control of 5-HT release in the frontal cortex and dorsal hippocampus of the rat. *Br J Pharmacol* 2000;130:160–166.

227. Svensson TH, Usdin T. Feedback inhibition of brain noradrenaline neurons by tricyclic antidepressants: α-receptor mediation. *Science* 1978;202:1089–1091.

228. Mateo Y, Pineda J, Meana JJ. Somatodendritic α_2-adrenoceptors in the locus coeruleus are involved in the *in vivo* modulation of cortical noradrenaline release by the antidepressant desipramine. *J Neurochem* 1998;71:790–798.

229. Linnér L, Arborelius L, Nomikos GG, et al. Locus coeruleus neuronal activity and noradrenaline availability in the frontal cortex of rats chronically treated with imipramine: effect of α_2-adrenoceptor blockade. *Biol Psychiatry* 1999;46:766–774.

230. Blier P, deMontigny C, Chaput Y. A role for the serotonin system in the mechanism of action of antidepressant treatments: preclinical evidence. *J Clin Psychiatry* 1990;51:14–21.

231. Kreiss DS, Lucki I. Effects of acute and repeated administration of antidepressant drugs on extracellular levels of 5-hydroxytryptamine measured *in vivo*. *J Pharmacol Exp Ther* 1995;274:866–876.

232. Artigas F, Bel N, Casanovas JM, et al. Adaptive changes of the serotoninergic system after antidepressant treatments. *Adv Exp Med Biol* 1996;398:51–59.

233. Chaput Y, deMontigny C, Blier P. Effects of a selective 5-HT reuptake blocker, citalopram, on the sensitivity of 5-HT autoreceptors: electrophysiological studies in the rat brain. *Naunyn Schmiedebergs Arch Pharmacol* 1986;333:342–348.

234. Gobbi M, Crespi D, Foddi MC, et al. Effects of chronic treatment with fluoxetine and citalopram on 5-HT uptake, 5-HT_{1B} autoreceptors, 5-HT_3 and 5-HT_4 receptors in rats. *Naunyn Schmiedebergs Arch Pharmacol* 1997;356:22–28.

235. Cremers TIFH, Spoelstra EN, DeBoer P, et al. Desensitisation of 5-HT autoreceptors upon pharmacokinetically monitored chronic treatment with citalopram. *Eur J Pharmacol* 2000;397:351–357.

236. Bel N, Artigas F. Chronic treatment with fluvoxamine increases extracellular serotonin in frontal cortex but not in raphe nuclei. *Synapse* 1993;15:243–245.

237. Wikell C, Apelqvist G, Carlsson B, et al. Pharmacokinetic and pharmacodynamic responses to chronic administration of the selective serotonin reuptake inhibitor citalopram in rats. *Clin Neuropharmacol* 1999;22:327–336.

238. Maes M, Libbrecht I, Van Hunsel F, et al. Pindolol and mianserin augment the antidepressant activity of fluoxetine in hospitalized major depressed patients, including those with treatment resistance. *J Clin Psychopharmacol* 1999;19:177–182.

239. Berman RM, Anand A, Cappiello A, et al. The use of pindolol with fluoxetine in the treatment of major depression: final results from a double-blind, placebo-controlled trial. *Biol Psychiatry* 1999;45:1170–1177.

240. Lacroix D, Blier P, Curet O, et al. Effects of long-term desipramine administration on noradrenergic neurotransmission: electrophysiological studies in the rat brain. *J Pharmacol Exp Ther* 1991;257:1081–1090.

241. Brady LS. Stress, antidepressant drugs and the locus coeruleus. *Brain Res Bull* 1994;35:545–556.

242. Blakely RD, Ramamoorthy S, Schroeter S, et al. Regulated phosphorylation and trafficking of antidepressant-sensitive serotonin transporter proteins. *Biol Psychiatry* 1998;44:169–178.

243. Owens MJ, Nemeroff CB. The serotonin transporter and depression. *Depression Anxiety* 1998;8[Suppl 1]:5–12.

244. Marcusson JO, Ross SB. Binding of some antidepressants to the 5-hydroxytryptamine transporter in brain and platelets. *Psychopharmacology* 1990;102:145–155.

245. Kovachich GB, Aronson CE, Brunswick DJ. Effect of repeated administration of antidepressants on serotonin uptake sites in limbic and neocortical structures of rat brain determined by quantitative autoradiography. *Neuropsychopharmacology* 1992;7:317–324.

246. Benmansour S, Cecchi M, Morilak DA, et al. Effects of chronic antidepressant treatments on serotonin transporter function, density, and mRNA level. *J Neurosci* 1999;19:10494–10501.

247. Frazer A, Morilak D. Drugs for the treatment of affective (mood) disorders. In: Brody TM, Larner J, Minneman KP, eds. *Human pharmacology: molecular to clinical*, third ed. St. Louis: Mosby, 1998:349–363.

248. Tremaine LM, Welch WM, Ronfeld RA. Metabolism and disposition of the 5-hydroxytryptamine uptake blocker sertraline in the rat and dog. *Drug Metab Dispos* 1989;17:542–550.

249. Melzacka M, Rurak A, Adamus A, et al. Distribution of citalopram in the blood serum and in the central nervous system of rats after single and multiple dosage. *Pol J Pharmacol Pharm* 1984;36:675–682.

250. Howell SR, Hicks DR, Scatina JA, et al. Pharmacokinetics of venlafaxine and *O*-desmethylvenlafaxine in laboratory animals. *Xenobiotica* 1994;24:315–327.

251. Graham D, Tahraoui L, Langer SZ. Effect of chronic treatment with selective monoamine oxidase inhibitors and specific 5-hydroxytryptamine uptake inhibitors on [^{3}H]paroxetine binding to cerebral cortical membranes of the rat. *Neuropharmacology* 1987;26:1087–1092.

252. Johanning J, Plenge P, Mellerup E. Serotonin receptors in the brain of rats treated chronically with imipramine or RU24969: support for the 5-HT$_{1B}$ receptor being a 5-HT autoreceptor. *Pharmacol Toxicol* 1992;70:131–134.

253. Gundlah C, Hjorth S, Auerbach SB. Autoreceptor antagonists enhance the effect of the reuptake inhibitor citalopram on extracellular 5-HT: this effect persists after repeated citalopram treatment. *Neuropharmacology* 1997;36:475–482.

254. Arborelius L, Nomikos GG, Grillner P, et al. 5-HT$_{1A}$ receptor antagonists increase the activity of serotoninergic cells in the dorsal raphe nucleus in rats treated acutely or chronically with citalopram. *Naunyn Schiedebergs Arch Pharmacol* 1995;352:157–165.

255. Hjorth A, Auerbach SB. Lack of 5-HT$_{1A}$ autoreceptor desensitization following chronic citalopram treatment, as determined by *in vivo* microdialysis. *Neuropharmacology* 1994;33:331–334.

256. Hjorth S, Auerbach SB. Autoreceptors remain functional after prolonged treatment with a serotonin reuptake inhibitor. *Brain Res* 1999;835:224–228.

257. Pineyro G, Blier P, Dennis T, et al. Desensitization of the neuronal 5-HT carrier following its long-term blockade. *J Neurosci* 1994;14:3036–3047.

258. Blier P, Bouchard C. Modulation of 5-HT release in the guinea-pig brain following long-term administration of antidepressant drugs. *Br J Pharmacol* 1994;113:485–495.

259. Mansari M, Bouchard C, Blier P. Alteration of serotonin release in the guinea pig orbitofrontal cortex by selective serotonin reuptake inhibitors: relevance to treatment of obsessive-compulsive disorder. *Neuropsychopharmacology* 1995;13:117–127.

260. Caccia S, Cappi M, Fracasso C, et al. Influence of dose and route of administration on the kinetics of fluoxetine and its metabolite norfluoxetine in the rat. *Psychopharmacology* 1990;100:509–514.

261. Dean B, Pereira A, Pavey G, et al. Repeated antidepressant drug treatment, time of death and frequency of handling do not affect [^{3}H]paroxetine binding in rat cortex. *Psychiatry Res* 1997;73:173-179.

262. Durand M, Berton O, Aguerre S, et al. Effects of repeated fluoxetine on anxiety-related behaviors, central serotoninergic systems, and the corticotropic axis in SHR and WKY rats. *Neuropharmacology* 1999;38:893–907.

263. Berton O, Durand M, Aguerre S, et al. Behavioral, neuroendocrine and serotoninergic consequences of single social defeat and repeated fluoxetine pretreatment in the Lewis rat strain. *Neuroscience* 1999;92:327–341.

264. Rutter JJ, Gundlah C, Auerbach SB. Increase in extracellular serotonin produced by uptake inhibitors is enhanced after chronic treatment with fluoxetine. *Neurosci Lett* 1994;171:183–186.

265. Le Poul E, Boni C, Hanoun N, et al. Differential adaptation of brain 5-HT$_{1A}$ and 5-HT$_{1B}$ receptors and 5-HT transporter in rats treated chronically with fluoxetine. *Neuropharmacology* 2000;39:110–122.

266. Lesch KP, Aulakh CS, Wolozin BL, et al. Regional brain expression of serotonin transporter mRNA and its regulation by reuptake inhibiting antidepressants. *Mol Brain Res* 1993;17:31–35.

267. Lopez JF, Chalmers DT, Vazquez DM, et al. Serotonin transporter mRNA in rat brain is regulated by classical antidepressants. *Biol Psychiatry* 1994;35:287–290.

268. Burnet PW, Michelson D, Smith MA, et al. The effect of chronic imipramine administration on the densities of 5-HT$_{1A}$ and 5-HT$_2$ receptors and the abundancies of 5-HT receptor and transporter mRNA in the cortex, hippocampus and dorsal raphe of three strains of rat. *Brain Res* 1994;638:311–324.

269. Spurlock G, Buckland P, O'Donovan M, et al. Lack of effect of antidepressant drugs on the levels of mRNAs encoding serotoninergic receptors, synthetic enzymes and 5-HT transporter. *Neuropharmacology* 1994;33:433–440.

270. Neumaier JF, Root DC, Hamblin MW. Chronic fluoxetine reduces serotonin transporter mRNA and 5-HT$_{1B}$ mRNA in a sequential manner in the rat dorsal raphe nucleus. *Neuropsychopharmacology* 1996;15:515–522.

271. Koed K, Linnet K. The serotonin transporter messenger RNA level in rat brain is not regulated by antidepressants. *Biol Psychiatry* 1997;42:1177–1180.

272. Benmansour S, Cecchi M, Morilak DA, et al. Effects of chronic antidepressant (AD) treatment on serotonin (SERT) and norepinephrine (NET) transporters. *Soc Neurosci Abstr* 2000;26:400(abst).

273. Tejani-Butt SM. [^{3}H]Nisoxetine: a radioligand for quantitation of norepinephrine uptake sites by autoradiography or by homogenate binding. *J Pharmacol Exp Ther* 1992;260:427–436.

274. Tejani-Butt SM, Brunswick DJ, Frazer A. [^{3}H]Nisoxetine: a new radioligand for norepinephrine uptake sites in brain. *Eur J Pharmacol* 1990;191:239–243.

275. Bauer ME, Tejani-Butt SM. Effects of repeated administration of desipramine or electroconvulsive shock on norepinephrine uptake sites measured by [^{3}H]nisoxetine autoradiography. *Brain Res* 1992;582:208–214.

276. Zhu M, Ordway GA. Down-regulation of norepinephrine transporters on PC12 cells by transporter inhibitors. *J Neurochem* 1997;68:134–141.

277. Zavosh A, Schaefer J, Ferrel A, et al. Desipramine treatment decreases ^{3}H-nisoxetine binding and norepinephrine transporter mRNA in SK-N-SHSY5Y cells. *Brain Res Bull* 1999;49:291–295.

278. Szot P, Ashliegh EA, Kohen R, et al. Norepinephrine transporter mRNA is elevated in the locus coeruleus following short- and long-term desipramine treatment. *Brain Res* 1993;618:308–312.

279. Daws LC, Toney GM, Gerhardt GA, et al. *In vivo* chronoamperometric measures of extracellular serotonin clearance in rat dorsal hippocampus: contribution of serotonin and norepinephrine transporters. *J Pharmacol Exp Ther* 1998;286:967–976.

280. Faux MC, Scott JD. More on target with protein phosphorylation: conferring specificity by location. *Trends Biochem* 1996;21:312–315.

281. Mann CD, Bich Vu T, Hrdina PD. Protein kinase C in rat brain cortex and hippocampus: effect of repeated administration of fluoxetine and desipramine. *Br J Pharmacol* 1995;115:595–600.

282. Rahimian R, Hrdina PD. Possible role of protein kinase C in regulation of 5-hydroxytryptamine 2A receptors in rat brain. *Can J Physiol Pharmacol* 1995;73:1686–1691.

283. Ramamoorthy S, Giovanetti E, Qian Y, et al. Phosphorylation and regulation of antidepressant-sensitive serotonin transporters. *J Biol Chem* 1998;273:2458–2466.

284. Perez J, Tinelli D, Bianchi E, et al. cAMP binding proteins in the rat cerebral cortex after administration of selective 5-HT and NE reuptake blockers with antidepressant activity. *Neuropsychopharmacology* 1991;4:57–64.

285. Popoli M. Synaptotagmin is endogenously phosphorylated by Ca^{2+}/calmodulin protein kinase II in synaptic vesicles. *FEBS Lett* 1993;317:85–88.

286. Nestler EJ, Terwilliger RZ, Duman RS. Chronic antidepressant administration alters the subcellular distribution of cyclic AMP-dependent protein kinase in rat frontal cortex. *J Neurochem* 1989;53:1644–1647.

287. Mori S, Garbini S, Caivano M., et al. Time-course changes in rat cerebral cortex subcellular distribution of the cyclic-AMP binding after treatment with selective serotonin reuptake inhibitors. *Int J Neuropsychopharmacol* 1998;1:3–10.

288. Miyamoto S, Asakura M, Sasuga Y, et al. Effects of long-term treatment with desipramine on microtubule proteins in rat cerebral cortex. *Eur J Pharmacol* 1997;333:279–287.

289. Racagni G, Brunello N, Tinelli D, et al. New biochemical hypotheses on the mechanism of action of antidepressant drugs: cAMP-dependent phosphorylation system. *Pharmacopsychiatry* 1992;25:51–55.

290. Schwartz JP, Costa E. Protein kinase translocation following ß-adrenergic receptor activation in C6 glioma cells. *J Biol Chem* 1980;255:2943–2948.

291. Shelton RC, Manier DH, Sulser F. Cyclic AMP-dependent protein kinase A (PKA) activity in human fibroblasts from normal subjects and from patients with major depression. *Am J Psychiatry* 1996;153:1037–1042.

292. Manier DH, Shelton RC, Ellis T, et al. Human fibroblasts as a relevant model to study signal transduction in affective disorders. *J Affect Disord* 2000;61:51–58.

293. Nichols RA, Sibra TS, Czernik AJ, et al. Calcium/calmodulin-dependent protein kinase II increases glutamate and noradrenaline release from synaptosomes. *Nature* 1990;343:647–651.

294. Nibuya M, Nestler EJ, Duman RS. Chronic anti-depressant administration increases the expression of cAMP response element-binding protein (CREB) in rat hippocampus. *J Neurosci* 1996;16:2365–2372.

295. Duman RS, Heninger GR, Nestler EJ. A molecular and cellular theory of depression. *Arch Gen Psychiatry* 1997;54:597–606.

296. Thome J, Sakai N, Shin K-H, et al. cAMP response element-mediated gene transcription is upregulated by chronic antidepressant treatment. *J Neurosci* 2000;20:4030–4036.

297. Nibuya M, Morinobu S, Duman RS. Regulation of BDNF and trkB mRNA in rat brain by chronic electroconvulsive seizure and antidepressant drug treatments. *J Neurosci* 1995;15:7539–7547.

298. Condorelli DF, Dell'Albani P, Mudo G, et al. Expression of neurotrophins and their receptors in primary astroglial cultures: induction by cAMP-elevating agents. *J Neurochem* 1994;63:509–516.

299. Duman RS, Vaidya VA, Nibuya M, et al. Stress, antidepressant treatments, and neurotrophic factors: molecular and cellular mechanisms. *Neuroscientist* 1995;1:351–360.

300. Deleted in press.

301. Prange AJ Jr. The pharmacology and biochemistry of depression. *Dis Nerv Syst* 1964;25:217–221.

302. Schildkraut JJ. The catecholamine hypothesis of affective disorders: a review of supporting evidence. *Am J Psychiatry* 1965;122:509–522.

303. Sapolsky RM, Krey LC, McEwen BS. Prolonged glucocorticoid exposure reduces hippocampal neuron number: implications for aging. *J Neurosci* 1985;5:1222–1227.

304. Sapolsky RM. The physiological relevance of glucocorticoid endangerment of the hippocampus. *Ann NY Acad Sci* 1994;746:294–304; discussion 304–307.

305. Sheline YI, Wang P, Gado MH, et al. Hippocampal atrophy in recurrent major depression. *Proc Natl Acad Sci USA* 1996;93:3908–3913.

306. Drevets WC, Price JL, Simpson JR, et al. Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* 1997;386:824–827.

307. Bremner JD, Narayan M, Anderson ER, et al. Hippocampal volume reduction in major depression. *Am J Psychiatry* 2000;157:115–118.

308. Ongur D, Drevets WC, Price JL, et al. Glial reduction in the subgenual prefrontal cortex in mood disorders. *Proc Natl Acad Sci USA* 1998;95:13290–13295.

309. Thoenen H. Neurotrophins and neuronal plasticity. *Science* 1995;270:593–598.

310. Mamounas LA, Blue ME, Siuciak JA, et al. BDNF promotes the survival and sprouting of serotoninergic axons in the rat brain. *J Neurosci* 1995;15:7929–7939.

311. Sklair-Tavron L, Nestler EJ. Opposing effects of morphine and the neurotrophins NT-3, NT-4, and BDNF on locus coeruleus neurons *in vitro. Brain Res* 1995;702:117–125.

312. Malberg JE, Eisch AJ, Nestler EJ, et al. Chronic antidepressant administration increases granule cell neurogenesis in the hippocampus of the adult male rat. *Soc Neurosci Abstr* 1999;25:1029(abst).

80

Hormonal and Gender Influences on mood Regulation

David R. Rubinow

Peter J. Schmidt

Catherine A. Roca

David R. Rubinow: National Institutes of Health, Bethesda, Maryland.

Peter J. Schmidt: Behavioral Endocrinology Branch, National Institute of Mental Health, Bethesda, Maryland.

Catherine A. Roca: Behavioral Endocrinology Branch, National Institute of Mental Health, Bethesda, Maryland.

HISTORY AND MECHANISMS

Part of "80 - Hormonal and Gender Influences on mood Regulation "

Within the past 20 years, the putative role of gender and gonadal steroids in mood regulation has been transformed from the staple of stereotype to a critical locus of research in clinical neuroscience. This transformation reflects the impact of explosive advances in molecular endocrinology and basic neuroscience, both of which suggest the myriad and dramatic neuroregulatory effects of gonadal steroids. Although direct isomorphs between basic mechanisms and clinical observations are for the most part absent, our burgeoning knowledge of the cellular and central nervous system effects of gonadal steroids is offering new models for understanding the relevance of gender and gonadal steroids in mood regulation. In this chapter, we review some of the major findings in reproductive neuroscience, emphasize the context dependency of many of these findings, and suggest that similar contextual effects underlie the inability to demonstrate uniform effects of gender or gonadal steroids on mood and behavior.

Reproductive hormones have played a central historical role in the development of our understanding of the effects of hormones on brain and behavior. More than 2,000 years ago, Aristotle, in his biological treatise *Historia Animalium*, observed that castration of immature male birds prevents the development of characteristic male singing and sexual behavior (1). One hundred fifty years ago, Berthold (2) successfully transplanted testes in castrated roosters and reversed their hypogonadal symptoms, demonstrating that reproductive organs possess factors that can dramatically alter physiology and behavior. These observations culminated in the claims by nineteenth century organotherapists (e.g., Brown-Sequard) that the administration of ovarian or testicular extracts could treat a variety of mood disorders in humans, ranging from depression to the anergy of senescence (3 ,4). In the 1920s and 1930s, the active gonadal substances—testosterone, estradiol, and progesterone—were isolated and characterized.

In 1962, Jensen and Jacobsen (5) provided evidence that the effects of estradiol are mediated through a specific intracellular hormone-binding protein, the estrogen receptor (a concept originally proposed by Langley in 1905), and by 1966, the estrogen receptor became the first hormone receptor to be isolated and identified (6). An elegant scheme for the cellular effects of hormones was subsequently elaborated. As lipophilic factors, steroid hormones would diffuse into cells, where they would bind the intracytoplasmic receptor (in contrast to the membrane-bound receptors of neurotransmitters and peptide hormones); the receptor protein would then be phosphorylated to cause dissociation of a heat shock protein and uncapping of the DNA binding domain of the receptor, which would result in binding of the receptor (frequently after dimerization) to a response element on the DNA, transcription of messenger RNA (mRNA), and finally translation of the mRNA into proteins in the cell cytoplasm. Because an enormous array of proteins relevant to neural transmission (e.g., neurotransmitter synthetic and metabolic enzymes, neural peptides, receptor proteins, signal transduction proteins) were observed to be regulated by gonadal steroids, this “genomic” mechanism promised to explain at a cellular level many of the effects of reproductive steroids observed at the level of the organism. In the past 15 years, the elegant simplicity of this genomic mechanism has given way to a model for the actions of reproductive steroids that is more comprehensive, powerful, and complex. Examination of this complexity promotes both an appreciation for the rich neuroregulatory potential of reproductive steroids and a means for understanding the diverse and wide ranging behavioral responses to alterations in reproductive steroid levels.

First, as the mechanics of transcription were elucidated, it became clear that activated steroid receptors influence transcription not as solitary agents but in combination with other intracellular proteins (7). These protein–protein interactions were such that an activated receptor might enhance, reduce, initiate, or terminate transcription of a particular gene solely as a function of the specific proteins with which it interacted (and the ability of these proteins to enhance or hinder the recruitment of the general transcription factor apparatus). The expression of these proteins—co-regulators (co-activators or co-repressors)—proved to be tissue-specific, and so suggested a means by which a hormone receptor modulator (e.g., tamoxifen) could act like an (estrogen) agonist in some tissues (e.g., bone) and like an (estrogen) antagonist in others (e.g., breast) (8 ,9). Another group of intracellular proteins, the co-integrators, provided a means by which classic hormone receptors could bind to and regulate sites other than hormone response elements [e.g., estrogen receptor (ER) or glucocorticoid receptor (GR) binding cyclic adenosine monophosphate (cAMP) response element binding (CREB) protein and, subsequently, the activator protein 1 (AP-1) binding site)] (10), and competition for co-integrator or other transcriptional regulatory proteins was demonstrated as a mechanism by which even ligand-free hormone receptors could influence (e.g., squelch or interfere with) the transcriptional efficacy of other activated hormone receptors (11). Thus, both the intracellular hormone receptor environment and the extracellular hormone environment might dictate the response to hormone receptor activation.

Second, the hormone receptors were found to exist in different forms. For example, isoforms of the progesterone receptor, PR_A and PR_B (the latter of which contains a 164-amino acid N-terminal extension), have different distributions and biological actions (12). As another example, two separate forms of the estrogen receptor, ER_α and ER_β, are encoded on different chromosomes (6 and 14, respectively) and have different patterns of distribution in the brain, different affinity patterns for certain ligands, and a range of different actions (including those created by ER heterodimers) (13 ,14 and 15). Further, a variant of ER_β, ER_β, is expressed in the brain, where it can form heterodimers with the ER_α or ER_β receptors (16) and inhibit their transcriptional actions (17).

Third, a variety of substances (e.g., nerve growth factor, insulin) are capable of activating a steroid receptor even in the absence of ligand (18 ,19). This crosstalk is exemplified by the ability of dopamine to induce lordosis by activating the PR (20 ,21).

Fourth, the relatively slow “genomic” effects of gonadal steroids have been expanded in two dimensions: time, with a variety of rapid (seconds to minutes) effects observed, and targets, which now include ion channels and a variety of second-messenger systems. For example, estradiol increases the firing of neurons in the cerebral cortex and hippocampus (CA1) (22) and decreases firing in medial preoptic neurons (23). The activity of membrane receptors like the glutamate and γ-aminobutyric acid (GABA) receptors is acutely modulated by gonadal steroids (estradiol and the 5-α reduced metabolite of progesterone, allopregnanolone, respectively) (24 ,25). Estradiol binds to and modulates the maxi-K potassium channel (26), increases cAMP levels (27), activates membrane G proteins ($G_{\alpha q}$, $G_{\alpha s}$) (28), inhibits L-type calcium channels (via nonclassic receptor) (29), and immediately activates the mitogen-activated protein kinase (MAPK) pathway (albeit in a receptor-mediated fashion) (30). The effects observed are tissue- and even cell-specific (e.g., estradiol increases MAPK in neurons but decreases it in astrocytes (31) (Zhang et al., *unpublished data*). The increase in the number of described mechanisms by which gonadal steroids can affect cell function has paralleled the rapid growth in their observed effects. Consequently, with each of these newly identified actions (which are usually, but sometimes inaccurately, called *nongenomic*), one needs to examine multiple factors before inferring the mechanism of action: (a) the duration required to see the effect, (b) the impact on the effect of inhibitors of transcription and protein synthesis, (c) the presence (or absence) of intracellular hormone receptors, (d) the stereospecificity of ligand binding (to see if effects are mediated through a classic receptor), (e) the effect of hormone receptor blockers, and (f) the ability of the ligand to initiate the action from the cell membrane (i.e., when entry into the cell is blocked). This last requirement acknowledges the presence on the membrane of binding sites for gonadal steroids that appear to be physiologically relevant (32).

Fifth, gonadal steroids regulate cell survival. Neuroprotective effects of estradiol have been described in neurons grown in serum-free media or those exposed to glutamate, amyloid-β, hydrogen peroxide, or glucose deprivation (22). Some of these effects appear to lack stereospecificity (i.e., are not classic receptor-mediated) and may be attributable to the antioxidant properties of estradiol (33 ,34), although data from one report are consistent with a receptor-mediated effect (35). Gonadal steroids may also modulate cell survival through effects on cell survival proteins (e.g., Bcl-2, Bax), MAPK, or even amyloid precursor protein metabolism (31 ,36 ,37) (Zhang et al., *unpublished data*).

Sixth, some actions of gonadal steroids on brain appear to depend on context and developmental stage. ToranAllerand (38) has shown that estrogen displays reciprocal interactions with growth factors and their receptors (e.g., p51 and trkA, neurotrophins) in such a way as to regulate, throughout development, the response to estrogen stimulation; estrogen stimulates its own receptor early in development, inhibits it during adulthood, and stimulates it again

in the context of brain injury. Additionally, we have demonstrated that the ability to modulate serotonin receptor subtype and GABA receptor subunit transcription in rat brain with exogenous administration of gonadal steroids or gonadal steroid receptor blockade largely depends on the developmental stage (e.g., last prenatal week vs. fourth postnatal week) during which the intervention occurs (39 ; Zhang et al., *unpublished data*).

Finally, the effects of gonadal steroids do not occur in isolation, but rather in exquisite interaction with the environment. Juraska (40), for example, demonstrated that the rearing environment (enriched vs. impoverished) dramatically influences sex differences in dendritic branching in the rat cortex and hippocampus. Further, the size of the spinal nucleus of the bulbocavernosus and the degree of adult male sexual behavior in rats is in part regulated by the amount of anogenital licking they receive as pups from their mothers, an activity that is elicited from the dames by the androgen the pups secrete in their urine (41).

SEXUAL DIMORPHISMS IN BRAIN STRUCTURE AND FUNCTION

Part of "80 - Hormonal and Gender Influences on mood Regulation "

In a highly influential article from the laboratory of C. W. Young, Phoenix et al. (42) demonstrated that exposure of the prenatal female guinea pig to androgens leads to defeminization of reproductive behavior in adulthood and increased sensitivity to androgen-induced male mating behavior. The authors interpreted these results as demonstrating an organizational effect of perinatal steroids on structure and subsequent behavioral function of the brain. These organizational effects, permanent changes in the brain consequent to exposure to gonadal steroids during small, critical windows of development, were contrasted with activational effects, which were impermanent effects that required continued exposure to gonadal steroids. This dichotomy was not uniformly accepted (43); Beach and Holz (44), who had demonstrated the effects on adult reproductive behavior of perinatal steroid manipulation and had interpreted them as derived from changes in gonadal morphology rather than brain function, referred to organizational effects as *imaginary brain mechanisms* (45). Nonetheless, by the late 1960s and early 1970s, the evidence for sexual dimorphisms in the brain that were organized by perinatal steroids was fairly compelling. Pfaff (46) demonstrated dimorphisms in rat brain in both gross and cellular morphology, with the dimorphisms altered by perinatal castration. Nottebohm and Arnold (47) showed that male song birds, who, in contrast to females, have the capacity to sing, had song control nuclei that were five to six times larger than comparable structures in the female. In two classic articles, Raisman and Field (48 ,49) demonstrated dimorphisms in neural connections; females showed a greater proportion of spine synapses in the preoptic area (48), and this dimorphism could be altered by perinatal steroid manipulation (49). Gorski et al. (50) identified in mammals a sexually dimorphic nucleus of the preoptic area that was two to six times larger in males than in females, and Rainbow et al. (51) demonstrated sexual differences in the response to gonadal steroids, with PR induction by estrogen seen more robustly in females than in males.

In subsequent years, sexual dimorphisms have been identified at all levels of the neuraxis and include differences in the following: nuclear volume; neuron number, size, density, morphology, and gene expression; neuritic outgrowth and arborization; synapse formation; glial number, morphology, and gene expression; and capacity for certain physiologic (e.g., cyclic gonadotropin secretion) and behavioral (e.g., song) activities (see refs. 52 , 53 for review). For example, we observed that in comparison with male astrocytes, perinatal cortical astrocytes from female rats have more activated MAPK at baseline and are more sensitive to estradiol suppression of both MAPK and cell proliferation (Zhang et al., *unpublished data*). Additionally, we observed dramatic dimorphisms in the developmental pattern and amount of expression of the cell survival/death proteins Bcl-2 and Bax (Zhang et al., *unpublished data*). The wide scope of these sex differences has become less mystifying as gonadal steroids have been seen to regulate virtually all stages of brain development, from neurogenesis to neural migration, differentiation, synaptogenesis, survival, and death (53). Nonetheless, despite the elegance of sexually dimorphic brain organization as an explanation for dimorphic behaviors, the complexity of the process underlying the development of sexual dimorphisms has assumed daunting proportions. First, it is often difficult to interpret the meaning of the dimorphisms. For example, lesions of the sexually dimorphic nucleus of the preoptic area (SDN-POA) do not compromise male copulatory behavior (despite the role of the preoptic area in reproductive behavior) (54), and DeVries and Boyle (55) have suggested that sexual dimorphisms may in some cases mediate the same behavior (e.g., the same parental behavior in prairie voles is mediated by differences in vasopressin). Second, lack of parallelism across genders complicates ascription of dimorphisms to the presence or absence of a particular steroid hormone. For example, song behavior (usually seen only in males) develops in female zebra finches if they are administered androgen or estradiol perinatally, but males deprived of androgen perinatally show no disruption of song behavior as adults (56 ,57). Third, some sexual dimorphisms appear to be organized and are independent of subsequent steroid exposure (58); others are activated (i.e., are dependent on subsequent steroid exposure) but not organized (i.e., they are not permanently influenced by perinatal steroid manipulation) (59); and still others are both organized and activated (e.g., the perinatally androgenized female zebra finch requires androgen as an

adult to express song behavior) (55 ,60). Further, Reisert and Pilgrim (61) have evidence suggesting that dimorphisms in the course of development of mesencephalic and diencephalic neurons are under genetic control (i.e., they are determined well before the appearance of any differences in gonadal steroid levels), like the genetically determined pouch or scrotum in marsupials (62). Fourth, the activational–organizational dichotomy is far more fluid and plasticity is much greater than the concept of critical periods allows. In contrast to the female zebra finch (who shows no male song behavior if androgenized during adulthood only), the female canary receiving androgen during adulthood exhibits male song behavior and shows masculine morphologic changes in the vocal control nuclei, including marked dendritic branching (63 ,64). Not only is the timing of hormonal administration (and species of the animal) important in determining outcome, but the manner of administration may also dictate the response. For example, Sodersten (65) demonstrated that one can induce typical female behavior in gonadectomized adult male rats by pulsatile but not continuous administration of estradiol followed by progesterone. Complexities notwithstanding, gonadal steroids appear capable of programming gonadal steroid sensitive circuitry in the brain, behavioral capacities, and differential response to the same physiologic stimulus.

SEXUAL DIMORPHISMS IN HUMANS

Part of "80 - Hormonal and Gender Influences on mood Regulation "

Given the relative lack of access to the brain in human studies in comparison with similar investigations in animals, the existence of gender-related differences has provided a major source of inference about the role of gonadal steroids in brain function and behavior. Reported gender dimorphisms in psychiatry include the following: prevalence, phenomenology (including characteristic symptoms, age at onset, susceptibility to recurrence, stress responsivity), and treatment characteristics. Specific examples of such dimorphisms are presented below.

Depression

Studies consistently demonstrate a twofold increased prevalence of depression in women in comparison with men (66 ,67 ,68 and 69), and this increased prevalence has been observed in a variety of countries (68). A twofold to threefold increased prevalence of dysthymia and a threefold increase in seasonal affective disorder (70) in women have also been noted (71). Although bipolar illness is equally prevalent in men and women (66 ,72 ,73 ; see ref. 74 for review), prepubertal depression prevalence rates are not higher in girls (75 ,76), possibly reflecting ascertainment bias (depressed boys may be more likely to come to the attention of health care providers) or the possibility that prepubertal major depression is premonitory of bipolar illness (77). With some exceptions, fewer differences in age at onset (68 ,69 ,78 ,79 ,80 and 81 ; but see also refs. 82 ,83 ,84 and 85), type of symptoms, severity, and likelihood of chronicity and recurrence (68 ,69 ,78 ,82 ,86 ,87 and 88 ; but see also refs. 89 ,90 ,91 ,92 and 93) are seen between men and women. Women are more likely to present with anxiety, atypical symptoms, or somatic symptoms (70 ,78 ,78 ,82 ,91 ,93 ,94 and 95) and are more likely to report symptoms, particularly in self-ratings (70 ,78 ,95). They are also more likely to report antecedent stressful events (96 ,97) and manifest a more robust effect of stress on the likelihood depression developing during adolescence (98). Women display increased comorbidities of anxiety and eating disorders (83 ,99 ,100 and 101), thyroid disease (102 ,103), and migraine headaches (104), and a lower lifetime prevalence of substance abuse and dependence (82 ,83). Reported differences in treatment response characteristics in women in comparison with men include a poor response to tricyclics (105 ,106 ,107 and 108), particularly in younger women (106); a superior response to selective serotonin reuptake inhibitors or monoamine oxidase inhibitors (109 ,110); and a greater likelihood of response to triiodothyronine (T_3) augmentation (103). The extent to which these differences reflect gender-related differences in pharmacokinetics (111 ,112 ,113 ,114 ,115 ,116 and 117) remains to be determined. Finally, although the prevalence of bipolar disorder is comparable in men and women, rapid cycling is more likely to develop in women (74), and they may be more susceptible to antidepressant-induced rapid cycling (118).

Schizophrenia

Although prevalence rates of schizophrenia appear comparable in men and women, a variety of differences in phenomenology, course, and treatment response characteristics have been identified (111 ,119 ,120).

Most but not all (121) studies suggest that the onset of schizophrenia is later in women and that they have better premorbid function (20–40 vs. 17–30; 122–126). Although symptom patterns during acute psychosis do not appear to differ substantially by gender, a variety of studies suggest that deficit or negative symptoms occur more frequently in men (123 ,127 ,128 and 129), perhaps consistent with the increased incidence in men of abnormalities of brain structure (primarily the corpus callosum) (127 ,130 ,131 ,132 and 133) and cognitive impairment (134). The course of illness in women tends to be more benign, with higher levels of psychosocial function (124), less time in a psychotic state (135), fewer days of hospitalization and fewer readmissions (126 ,136 ,137), less substance abuse (126 ,135 ,137), and less violence (aggressive episodes and completed suicide) (123 ,138). Finally, consistent with a more benign course (at least until middle age), women are reported to respond to neuroleptics more quickly, more extensively, and at lower doses (123 ,139 ,140 ,141 ,142 ,143 ,144 and 145).

The extent to which this last observation reflects gender dimorphisms in pharmacokinetics (e.g., drug absorption, storage, distribution volume, clearance) versus pharmacodynamics awaits determination (112).

Physiologic Dimorphisms

The epidemiologic observations previously described are increasingly complemented by demonstrations of sexual dimorphisms in brain structure and physiology in humans. Structural and functional brain imaging studies, for example, have shown the following: (a) differences in functional organization of the brain, with brain activation response to rhyming task lateralized in men but not in women (146); (b) gender-specific decreases in regional brain volume (caudate in males and globus pallidus, putamen in females) during development (147); (c) increased neuronal density in the temporal cortex in women (148); (d) greater interhemispheric coordinated activation of brain regions in women (149); (e) larger-volume hypothalamic nucleus (INAH3) in men (150); (f) differences in both resting blood flow and the activation pattern accompanying self-induced mood change (151); (g) decreased serotonin (5-HT_2) binding in the frontal, parietal, temporal, and singular cortices in women (152); (h) differences in whole-brain serotonin synthesis (interpreted as decreased in women but possibly increased if corrected for plasma levels of free tryptophan (153); (i) greater and more symmetric cerebral blood flow in women (154 ,155 ,156 ,157 and 158); (j) greater asymmetry in the planum temporale in men (159); and (k) higher rates of brain glucose metabolism (19%) in women (160 ,161). The potential relevance of gonadal steroids in some of these differences has also been demonstrated with the same technologies. For example, Berman et al. (162) demonstrated that the normal pattern of cognitive task-activated cerebral blood flow is eliminated by induced hypogonadism and restored by replacement with estradiol or progesterone. These findings were supported by Shaywitz et al. (163), who demonstrated estrogen enhancement of cognitive task-stimulated brain regional activation on function magnetic resonance imaging (fMRI) in postmenopausal women. Additionally, Wong et al. (164) demonstrated in a small number of subjects that dopamine receptor density in the caudate (measured by positron emission tomography) varies as a function of the menstrual cycle (lower in the follicular phase). The contribution of these and other effects of gonadal steroids to observed gender dimorphisms must, obviously, await further determination.

MOOD DISORDERS RELATED TO REPRODUCTIVE ENDOCRINE FUNCTION

Part of "80 - Hormonal and Gender Influences on mood Regulation "

Given the complexity of the factors that affect gender throughout development, it is very difficult to infer the degree to which differential exposure to gonadal steroids determines gender-related behavioral differences. A better opportunity to determine the behavioral relevance of fluctuations in gonadal steroids is provided by mood disorders that appear linked to changes in levels of reproductive steroids. In the following section, we review the role of gonadal steroids in the precipitation and treatment of mood disorders by focusing on three disorders: premenstrual syndrome (PMS), postpartum depression (PPD), and perimenopausal depression.

Premenstrual Syndrome

Although Frank is credited with the first description of “premenstrual tension” in 1931, reports of mood and behavioral disturbances confined to the luteal phase of the menstrual cycle appeared earlier, in the medical literature of the nineteenth century. For example, in 1847, Dr. Ernst G. Von Feuchtersleben stated that “the menses in sensitive women is almost always attended by mental uneasiness, irritability or sadness” (165). Two years later, the organ transplantation studies of Berthold (2) demonstrated that the reproductive organs possess factors that can markedly alter physiology and behavior, an observation that culminated in the late nineteenth century in the practice by organotherapists of administering ground-up animal glands and organs to treat a wide array of diseases and ailments (4). In this context, the isolation and characterization of ovarian steroids in the early twentieth century led to the inevitable assumption that premenstrual tension is caused by an excess or deficiency of estrogen or progesterone. In the ensuing years, each new discovery in endocrinology gave rise to a new theory of PMS if the new endocrine factor could in any way be shown to influence or be influenced by the menstrual cycle or potentially mediate any of the myriad symptoms attributed to PMS. Despite the obvious appeal of hormonal excess or deficiency as an explanation for PMS, consistent demonstration of any hormonal abnormality was lacking. A major source of study inconsistency was identified in the 1980s (166)—namely, that samples of women with PMS were selected (diagnosed) with highly unreliable techniques (i.e., unconfirmed history). Without prospective demonstration of luteal phase-restricted symptom expression, samples selected were certain to include a large number of false-positives and so make it impossible to apply the data to the population with PMS (167). This requirement for prospective confirmation of luteal phase symptomatology was ultimately incorporated into diagnostic criteria for PMS (National Institute of Mental Health, *unpublished data*, 1983) and late luteal phase depressive disorder/premenstrual depressive disorder (168). Although the improved diagnostic methods used since the mid-1980s have ensured the comparability of samples selected for study, subsequent data, if anything, have provided fairly convincing evidence against

hormonal excess or deficiency as etiologically relevant in PMS.

Individual studies have identified diagnostic group-related differences in levels of reproductive hormones, but the most consistent and compelling data support the absence of such differences. We observed no diagnosis-related differences in plasma levels, areas under the curve, or patterns of hormone secretion for estradiol, progesterone, follicle-stimulating hormone (FSH), or luteinizing hormone (LH) (169), findings consistent with those of Backstrom et al. (170) in a comparison of patients with high and low degrees of cyclic mood change.

In recent additions to the conflicting literature, Wang et al. (171) observed increased estradiol and decreased progesterone levels in women with PMS, Redei and Freeman (172) reported nonsignificant increases in both estradiol and progesterone, and Facchinetti et al. (173) found no differences between subjects and controls in integrated progesterone levels. Results of studies of androgen levels have been similarly inconsistent, demonstrating both normal and decreased testosterone levels (174 ,175 and 176) and elevated and decreased free testosterone levels (175 ,176).

Recent speculations about the etiology of PMS have focused on putative abnormal neurosteroid levels. Observations central to these speculations include the following: (a) the GABA receptor (the presumed mediator of anxiolysis) is positively modulated by the 5-α and 5-ß reduced metabolites of progesterone (allopregnanolone and pregnanolone, respectively) (25); (b) withdrawal of progesterone in rats produces anxiety and insensitivity to benzodiazepines secondary to withdrawal of allopregnanolone, with consequent induction of $GABA_A$ α_4 subunit levels and inhibition of GABA currents (177 ,178); (c) decreased plasma allopregnanolone levels are seen in major depressive disorder and in depression associated with alcohol withdrawal, with increased levels seen in plasma and cerebrospinal fluid following successful antidepressant treatment (179 ,180 ,181 and 182); (d) allopregnanolone has anxiolytic effects in several animal models of anxiety (183 ,184 and 185) and may be involved in the stress response (186); (e) antidepressants may promote the reductive activity of one of the neurosteroid synthetic enzymes (3-α-hydroxysteroid oxidoreductase) and thus favor the formation of allopregnanolone (187); (f) patients with PMS show differences in pregnanolone-modulated saccadic eye velocity and sedation in the luteal phase in comparison with controls (188) (although the reported differences seem attributable to a saccadic eye velocity response to vehicle in those with PMS and a blunted sedation response in the follicular phase in controls); (g) patients with severe PMS show blunted saccadic eye velocity and sedation responses to $GABA_A$-receptor agonists—pregnanolone (188) or midazolam (189)—in comparison with patients with mild PMS. Although one investigator observed decreased serum allopregnanolone levels in women with PMS in comparison with controls on menstrual cycle day 26 (190), other studies showed no diagnosis-related differences in allopregnanolone or pregnanolone (191 ,192) nor any difference in allopregnanolone levels in women with PMS before and after successful treatment with citalopram (193). Wang et al. (192) did find that if two cycles differed in area under the curve of a hormone by more than 10%, the cycle with the lower levels of allopregnanolone and higher levels of estradiol, pregnanolone, and pregnenolone sulfate was accompanied by more severe symptoms.

In sum, then, no consistent or convincing evidence is available that PMS is characterized by abnormal circulating plasma levels of gonadal steroids or gonadotropins or by hypothalamic–pituitary–ovarian axis dysfunction. Several studies do, however, suggest that levels of estradiol, progesterone, or neurosteroids (e.g., pregnenolone sulfate) may be correlated with symptom severity in women with PMS (171 ,194 ,195).

If one treats PMS with any of a number of therapies, one is never certain that the response seen is causally related to the pharmacologic (contrasted with the nonspecific) properties of the intervention employed. Therefore, we attempted to dissociate the symptoms of PMS from the menstrual cycle phase by targeting the menstrual cycle phase rather than the symptoms (196). We administered a progesterone receptor blocker (RU 486) with or without human chorionic gonadotropin (hCG) to women with PMS during the early luteal to midluteal phase. Within 2 days of administration, RU 486 caused menses (by blocking the endometrial progesterone receptors) and luteolysis and advanced the onset of the follicular phase of the next cycle. Addition of hCG does not alter the RU 486-induced menses but “rescues” or preserves the corpus luteum and permits a luteal phase of normal length. Consequently, after women experienced an RU 486-induced menses, they did not know whether they were in the follicular phase of the next cycle (RU 486 alone) or in the preserved luteal phase of the first cycle (RU 486 plus hCG). Women in both groups experienced typical PMS symptoms despite the fact that the women receiving RU 486 were now symptomatic in the context of an experimentally advanced follicular phase. Hence, the endocrinology of the midluteal to late luteal phase is irrelevant to the symptoms of PMS, as this phase can be eliminated without influencing the appearance of PMS symptoms. This suggested two possibilities: mood symptoms in women with PMS were entrained to the menstrual cycle but not caused by it, or mood symptoms might be triggered in the luteal phase by reproductive endocrine events occurring earlier in the menstrual cycle, a possibility that was examined in a second study.

The gonadotropin-releasing hormone (GnRH) agonist leuprolide acetate (Lupron) was administered for 3 months in a double-blinded, placebo-controlled, parallel-design study to 20 women with PMS. Women receiving Lupron, but not those receiving placebo, demonstrated a significant decrease in symptom severity and cyclicity, consistent with

several earlier demonstrations of the efficacy in PMS of medical or surgical oophorectomy (197 ,198). The ovulatory cycle, therefore, appears to be necessary for the expression of PMS (199).

To determine whether gonadal steroids were the factors that when removed resulted in the elimination of PMS, we added back estradiol and progesterone separately to women who continued to take Lupron and for whom Lupron alone successfully eliminated symptoms of PMS. Both estradiol and progesterone were associated with a return of symptoms typical of PMS. Symptoms were precipitated within 7 to 10 days and largely remitted by the end of the 4-week phase of addback. It does appear, therefore, that gonadal steroids can trigger symptoms of PMS, an observation that at first glance appears discordant with the lack of differences in gonadal steroid levels between women with PMS and controls. The reconciliation of these observations is found in the second part of the aforementioned study, in which a comparison group of women with confirmed absence of PMS received the same protocol of Lupron and hormone addback. The control women showed no perturbation of mood during Lupron-induced hypogonadism or during hormone addback with either progesterone or estradiol, despite achieving hormone levels comparable with those seen in the women with PMS. Women with PMS, therefore, are differentially sensitive to gonadal steroids such that they experience mood destabilization with levels or changes in gonadal steroids that are absolutely without effect on mood in women lacking a history of PMS. Gonadal steroids, then, are necessary but not sufficient for PMS. They can trigger PMS, but only in women, who, for undetermined reasons, are otherwise vulnerable to experience mood state destabilization (199). In other words, PMS represents an abnormal response to normal hormone levels.

Postpartum Depression

The literature examining the possible role of hormone abnormalities in postpartum is more exiguous than that for PMS. This literature, however, may be similarly distilled; the evidence for a reproductive hormone abnormality in PPD is scant (200 ,201 ,202 ,203 and 204 , but see also ref. 205). Nonetheless, it is difficult to regard as irrelevant the enormous hormonal excursions occurring during the puerperium (with precipitous drops of estradiol and progesterone from levels of up to 15,000 pg/mL and 150 ng/mL, respectively, to hypogonadal levels in just 1 to 3 days). Analagous to our observations with PMS, it is possible that women with and those without PPD differ in sensitivity to puerperal hormone changes, not in the degree to which they occur. To test this hypothesis, we created a scaled-down model of the puerperium in which women received high-dose estradiol and progesterone for 2 months (superimposed on Lupron-induced gonadal suppression to permit comparability and stability of levels achieved), followed by a blinded, precipitous withdrawal of gonadal steroids and a consequent Lupron-induced hypogonadal state. This protocol was performed in two groups: euthymic women with a history of PPD occurring no more recently than 1 year before the study (PPD+) and controls lacking a history of depression (PPD-). In the first 2 weeks following withdrawal, the women with a history of PPD experienced a significant increase in measures of depression relative to baseline, with several subjects experiencing an increase in symptoms during the last few weeks of addback. No similar symptoms were experienced by the women lacking a history of PPD. Both the levels of hormones achieved and the change from peak to withdrawal-induced hypogonadism were comparable in the two groups. It appears, therefore, that like women with PMS, women with a history of PPD experience mood state destabilization in association with changes in levels of gonadal steroids that are without effect on mood in women lacking a history of PPD. The hormonal changes can trigger the mood state change, but only in a context of increased susceptibility to affective dysregulation.

Context

The differential sensitivity to gonadal steroids seen in women with a history of PMS or PPD emphasizes that the response to a biological signal cannot be inferred absent an understanding of the context in which the signal occurs. This context includes current physiologic and external environments, prior experience, past history of exposure to the stimulus, and genetic makeup. With the imminent mapping of the human genome, this last contextual determinant becomes of great practical interest as a potential explanation for the differential response to steroids. Data already exist from both animal and human studies in support of this hypothesis. Spearow et al. (206) demonstrated greater than 16-fold differences in sensitivity to estradiol (reproductive disruption) across six different mouse strains, with genotype accounting for more of the variation than the dose of estradiol. Similarly, strain/genetic (and task-dependent) differences in behavioral sensitivity to allopregnanolone were observed by Finn et al. (207). Huizenga et al. (208) demonstrated not only intraperson stability of baseline cortisol and feedback sensitivity (to dexamethasone), which suggests a genetic influence (209), but also a higher sensitivity to exogenously administered glucocorticoid (dexamethasone) in association with a polymorphism in exon 2 of the glucocorticoid receptor. Association studies suggest a progressively increased rate and severity of prostate cancer as the number of cytosine–adenine–guanosine (CAG) trinucleotide repeats in exon 1 of the androgen receptor decreases (210). This observation is accompanied by the recent observation that androgen receptors with decreased CAG repeats demonstrate increased transcriptional efficiency (211). Steroid receptor polymorphisms, then, may alter the steroid signaling pathway in such a way as to produce or

contribute to a different behavioral/phenotypic response to a hormone signal. As appealing as this explanation is for the differential sensitivity observed in PMS and PPD, the demonstrations in animal studies that perinatal steroid manipulations alter the organization of gonadal steroid-sensitive circuitry (42) and gonadal steroid-activated gene expression (212) caution us that gene–environment interactions may yield markedly different phenotypic expressions of the same genotype.

Hormones as Therapeutic Agents

An emerging area of interest is the use of gonadal steroids in the treatment of PPD (213) and the perimenopause. As in PPD, the evidence for a reproductive hormonal abnormality in perimenopausal depression is vanishingly small (214 ,215 and 216 ; Schmidt et al., *unpublished data*). Some (217 ,218 ,219 and 220) but not all (Schmidt et al., *unpublished data*) studies have observed lower plasma LH levels in postmenopausal depressed women, but no consistent group-related differences in gonadal steroids have been demonstrated. Similarly, despite claims for the antidepressant efficacy of estrogen dating back to the nineteenth century (4 ,221), reports of the effect of estradiol on mood in perimenopausal and postmenopausal women (222 ,223 ,224 ,225 and 226) have been inconsistent (227 ,228 and 229) and have been compromised by the failure to diagnose depression (as opposed to depressive symptoms, which have different causes and treatment response characteristics), the failure (with one exception; see ref. 226) to consider remediation of hot flushes as a confound in assessment of psychotropic efficacy, and the failure to assess efficacy in perimenopausal (vs. postmenopausal) women, a potentially important distinction identified by Montgomery et al. (222). These problems were addressed in a recent study that demonstrated the antidepressant efficacy of estradiol in perimenopausal women with major and minor depression (230 ,231). The antidepressant effects were further shown in the subsample of women with no hot flushes, so that the possibility that remediation of hot flush-induced sleep disturbance might indirectly improve mood was eliminated. A subsequent study has similarly demonstrated the psychotropic efficacy of estradiol in perimenopausal depression (C. Soares et al., *unpublished data*). These observations converge with *in vitro* and epidemiologic evidence for neuroprotective effects of estradiol in suggesting that gonadal steroids (and adrenal androgens) may enter the neuropsychiatric therapeutic armamentarium, either as primary or adjunctive agents. While not permitting an inference about the etiology of reproductive endocrine-related mood disorders, the psychotropic effects of hormones may help dissect neural pathways of relevance to the regulation of affect. Attempts to define the mechanisms underlying both the psychotropic effects of gonadal steroids and the differential response to endogenous gonadal steroids should help advance our efforts to illuminate the neurobiology of mood and mood disorders.

REFERENCES

1. Dorfman RI, Shipley RA. Androgens: biochemistry, physiology, and clinical significance. New York: John Wiley and Sons, 1956:9.

2. Berthold A. Transplantation der Hoden. *Arch Anat Physiol Wiss Med* 1849;16:42–46.

3. Brown-Séquard CE. The effects produced on man by subcutaneous injections of a liquid obtained from the testicles of animals. *Lancet* 1889;2:105–107.

4. Easterbrook CC. Organo-therapeutics in mental diseases. *Br Med J* 1900;2:813–823.

5. Jensen EV, Jacobson HI. Basic guides to the mechanism of estrogen action. *Recent Prog Horm Res* 1962;18:387–414.

6. Toft D, Gorski J. A receptor molecule for estrogens: isolation from the rat uterus and preliminary characterization. *Proc Natl Acad Sci USA* 1966;55:1574–1581.

7. Halachmi S, Marden E, Martin G, et al. Estrogen receptor-associated proteins: possible mediators of hormone-induced transcription. *Science* 1994;264:1455–1458.

8. Smith CL, Nauaz Z, O'Malley BW. Co-activator and co-repressor regulation of the agonist/antagonist activity of the mixed antiestrogen 4-hydroxytamoxifen. *Mol Endocrinol* 1997;11:657–666.

9. Jackson TA, Richer JK, Bain DL, et al. The partial agonist activity of antagonist-occupied steroid receptors is controlled by a novel hinge domain-binding co-activator L7/SPA and the co-repressors N-COR or SMRT. *Mol Endocrinol* 1997;11:693–705.

10. Uht RM, Anderson CM, Webb P, et al. Transcriptional activities of estrogen and glucocorticoid receptors are functionally integrated at the AP-1 response element. *Endocrinology* 1997;138:2900–2908.

11. Meyer M-E, Gronemeyer H, Turcott B, et al. Steroid hormone receptors compete for factors that mediate their enhancer function. *Cell* 1989;57:433–442.

12. Chalbos D, Galtier F. Differential effect of forms A and B of human progesterone receptor on estradiol-dependent transcription. *J Biol Chem* 1994;269:23007–23012.

13. Kuiper GGJM, Shughrue PJ, Merchenthaler I, et al. The estrogen receptor ß subtype: a novel mediator of estrogen action in neuroendocrine systems. *Front Neuroendocrinol* 1998;19:253–286.

14. Paech K, Webb P, Kuiper GGJM, et al. Differential ligand activation of estrogen receptors ERα and ERß at AP1 sites. *Science* 1997;277:1508–1510.

15. Shughrue PJ, Lane MV, Merchenthaler I. Comparative distribution of estrogen receptor-α and -ß mRNA in the rat central nervous system. *J Comp Neurol* 1997;388:507–525.

16. Moore JT, McKee DD, Slentz-Kesler K, et al. Cloning and characterization of human estrogen receptor ß isoforms. *Biochem Biophys Res Commun* 1998;247:75–78.

17. Maruyama K, Endoh H, Sasaki-Iwaoka H, et al. A novel isoform of rat estrogen receptor ß with 18 amino acid insertion in the ligand binding domain as a putative dominant negative regulator of estrogen action. *Biochem Biophys Res Commun* 1998;246:142–147.

18. Ignar-Trowbridge DM, Nelson KG, Bidwell MC, et al. Coupling of dual signaling pathways: epidermal growth factor action involves the estrogen receptor. *Proc Natl Acad Sci USA* 1992;89:4658–4662.

19. Aronica SM, Katzenellenbogen BS. Stimulation of estrogen receptor-mediated transcription and alteration in the phosphorylation state of the rat uterine estrogen receptor by estrogen, cyclic adenosine monophosphate, and insulin-like growth factor-1. *Mol Endocrinol* 1993;7:743–752.

20. Mani SK, Blaustein JD, Allen JMC, et al. Inhibition of rat sexual behavior by antisense oligonucleotides to the progesterone receptor. *Endocrinology* 1994;135:1409–1414.

21. Power RF, Mani SK, Codina J, et al. Dopaminergic and ligand-independent activation of steroid hormone receptors. *Science* 1991;254:1636–1639.

22. McEwen BS, Alves SE. Estrogen actions in the central nervous system. *Endocr Rev* 1999;20:279–307.

23. McEwen BS. Steroid hormones: effect on brain development and function. *Horm Res Suppl* 1992;3:1–10.

24. Wong M, Moss RL. Long-term and short-term electrophysiological effects of estrogen on the synaptic properties of hippocampal CA1 neurons. *J Neurosci* 1992;12:3217–3225.

25. Majewska MD, Harrison NL, Schwartz RD, et al. Steroid hormone metabolites are barbiturate-like modulators of the GABA receptor. *Science* 1986;232:1004–1007.

26. Valverde MA, Rojas P, Amigo J, et al. Acute activation of maxi-K channels (hSlo) by estradiol binding to the ß subunit. *Science* 1999;285:1929–1931.

27. Aronica SM, Kraus WL, Katzenellenbogen BS. Estrogen action via the cAMP signaling pathway: stimulation of adenylate cyclase and cAMP-regulated gene transcription. *Proc Natl Acad Sci USA* 1994;91:8517–8521.

28. Ravindra R, Aronstam RS. Progesterone, testosterone and estradiol-17ß inhibit gonadotropin-releasing hormone, stimulation of G protein GTPase activity in plasma membranes from rat anterior pituitary lobe. *Acta Endocrinol (Copenh)* 1992;126:345–349.

29. Mermelstein PG, Becker JB, Surmeier DJ. Estradiol reduces calcium currents in rat neostriatal neurons via a membrane receptor. *J Neurosci* 1996;16:595–604.

30. Migliaccio A, Piccolo D, Castoria G, et al. Activation of the Src/p21ras/Erk pathway by progesterone receptor via cross-talk with estrogen receptor. *EMBO J* 1998;17:2008–2018.

31. Watters JJ, Campbell JS, Cunningham MJ, et al. Rapid membrane effects of steroids in neuroblastoma cells: effects of estrogen on mitogen-activated protein kinase signaling cascade and c-*fos* immediate early gene transcription. *Endocrinology* 1997;138:4030–4033.

32. Brubaker KD, Gay CV. Evidence for plasma membrane-mediated effects of estrogen. *Calcif Tissue Int* 1999;64:459–462.

33. Behl C, Skutella T, Lezoualc'h F, et al. Neuroprotection against oxidative stress by estrogens: structure–activity relationship. *Mol Pharmacol* 1997;51:535–541.

34. Mooradian AD. Antioxidant properties of steroids. *J Steroid Biochem Mol Biol* 1993;45:509–511.

35. Singer CA, Rogers KL, Strickland TM, et al. Estrogen protects primary cortical neurons from glutamate toxicity. *Neurosci Lett* 1996;212:13–16.

36. Garcia-Segura LM, Cardona-Gomez P, Naftolin F, et al. Estradiol upregulates Bcl-2 expression in adult brain neurons. *Neuroendocrinology* 1998;9:593–597.

37. Gouras GK, Xu H, Gross RS, et al. Testosterone reduces neuronal secretion of Alzheimer's beta-amyloid peptides. *Proc Natl Acad Sci U S A* 2000;97:1202–1205.

38. Toran-Allerand CD. Developmental interactions of estrogens with the neurotrophins and their receptors. In: Micevych P, Hammer RP, eds. *Neurobiological effects of sex steroid hormones*. Cambridge, UK: Cambridge University Press, 1994:391–411.

39. Zhang L, Ma W, Barker JL, et al. Sex differences in expression of serotonin receptors (subtypes 1A and 2A) in rat brain: a possible role of testosterone. *Neuroscience* 1999;94:251–259.

40. Juraska JM. The structure of the rat cerebral cortex: effects of gender and the environment. In: Kolb B, Tees RC, eds. *The cerebral cortex of the rat*. Cambridge, MA: MIT Press, 1990:483–505.

41. Moore CL, Dou H, Juraska JM. Maternal stimulation affects the number of motor neurons in a sexually dimorphic nucleus of the lumbar spinal cord. *Brain Res* 1992;572:52–56.

42. Phoenix CH, Goy RW, Gerall AA, et al. Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig. *Endocrinology* 1959;65:369–382.

43. Baum MJ. Frank Beach's research on the sexual differentiation of behavior and his struggle with the "organizational" hypothesis. *Neurosci Biobehav Rev* 1990;14:201–206.

44. Beach FA, Holz AM. Mating behavior in male rats castrated at various ages and injected with androgen. *J Exp Zool* 1946;101:91–142.

45. Beach FA. Hormonal factors controlling the differentiation, development, and display of copulatory behavior in the ramstergig and related species. In: Tobach E, Aronson LR, Shaw E, eds. *The biopsychology of development*. New York: Academic Press, 1971:249–296.

46. Pfaff DW. Morphological changes in the brains of adult male rats after neonatal castration. *J Endocrinol* 1966;36:415–416.

47. Nottebohm F, Arnold AP. Sexual dimorphism in vocal control areas of the songbird brain. *Science* 1976;194:211–213.

48. Raisman G, Field PM. Sexual dimorphism in the preoptic area of the rat. *Science* 1971;173:731–733.

49. Raisman G, Field PM. Sexual dimorphism in the neuropil of the preoptic area of the rat and its dependence on neonatal androgen. *Brain Res* 1973;54:1–29.

50. Gorski RA, Gordon JH, Shryne JE, et al. Evidence for a morphological sex difference within the medial preoptic area of the rat brain. *Brain Res* 1978;148:333–346.

51. Rainbow TC, Parsons B, McEwen BS. Sex differences in rat brain oestrogen and progestin receptors. *Nature* 1982;300:648–649.

52. Gorski RA. Sexual differentiation of the endocrine brain and its control. In: Motta M, ed. *Brain endocrinology*. New York: Raven Press, 1991:71–104.

53. Pilgrim C, Hutchison JB. Developmental regulation of sex differences in the brain: can the role of gonadal steroids be redefined? *Neuroscience* 1994;60:843–855.

54. Arendash GW, Gorski RA. Effects of discrete lesions of the sexually dimorphic nucleus of the preoptic area or other medial preoptic regions on the sexual behavior of male rats. *Brain Res Bull* 1983;10:147–150.

55. De Vries GJ, Boyle PA. Double duty for sex differences in the brain. *Behav Brain Res* 1998;92:205–213.

56. Breedlove SM. Sexual dimorphism in the vertebrate nervous system. *J Neurosci* 1992;12:4133–4142.

57. Gurney ME. Behavioral correlates of sexual differentiation in the zebra finch song system. *Brain Res* 1982;231:153–172.

58. Enriquez P, Cales JM, Sanchez-Santed F, et al. Effects of early postnatal gonadal steroids on extinction of a continuously food-rewarded running response. *Physiol Behav* 1991;49:57–61.

59. Goldman S, Nottebohm F. Neuronal production, migration, and differentiation in a vocal control nucleus of the adult female canary brain. *Proc Natl Acad Sci USA* 1983;21:1185–1201.

60. Schlinger BA, Arnold AP. Androgen effects on the development of the zebra finch song system. *Brain Res* 1991;561:99–105.

61. Reisert L, Pilgrim C. Sexual differentiation of monoaminergic neurons—genetic or epigenetic? *Trends Neurosci* 1991;14:468–473.

62. Renfree MB, Harry JL, Shaw G. The marsupial male: a role model for sexual development. *Philos Trans R Soc Lond B Biol Sci* 1995;350:243–251.

63. Nottebohm F. Testosterone triggers growth of brain vocal control nuclei in adult female canaries. *Brain Res* 1980;189:429–436.

64. DeVoogd T, Nottebohm F. Gonadal hormones induce dendritic growth in the adult avian brain. *Science* 1981;214:202–204.

65. Sodersten P. Sexual differentiation: do males differ from females in behavioral sensitivity to gonadal hormones? In: De Vries GJ, De Bruin JPC, Uylings HBM, et al., eds. *Prog Brain Res* 1984;61: 257–270.

66. Robins LN, Regier DA, eds. *Psychiatric disorders in America: the Epidemiologic Catchment Area Study*. New York: The Free Press, 1991.

67. Weissman MM, Klerman GL. Sex differences in the epidemiology of depression. *Arch Gen Psychiatry* 1977;34:98–111.

68. Weissman MM, Bland R, Joyce PR, et al. Sex differences in rates of depression: cross-national perspectives. *J Affect Disord* 1993;29:77–84.

69. Kessler RC, McGonagle KA, Swartz M, et al. Sex and depression in the National Comorbidity Survey I: lifetime prevalence, chronicity and recurrence. *J Affect Disord* 1993;29:85–96.

70. Leibenluft E, Hardin TA, Rosenthal NE. Gender differences in seasonal affective disorder. *Depression* 1995;3:13–19.

71. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington, DC: American Psychiatric Association, 1994.

72. Weissman MM, Klerman GL. Sex differences and the epidemiology of depression. In: Guggenheim FG, Nadelson CC, eds. *Major psychiatric disorders: overviews and selected readings*. New York: Elsevier Science, 1982:95–114.

73. Weissman MM, Klerman GL. Gender and depression. *Trends Neurosci* 1985;8:416–420.

74. Leibenluft E. Women with bipolar illness: clinical and research issues. *Am J Psychiatry* 1996;153:163–173.

75. Anderson JC, Williams S, McGee R, et al. DSM-III disorders in preadolescent children: prevalence in a large sample from the general population. *Arch Gen Psychiatry* 1987;44:69–77.

76. McGee R, Feehan M, Williams S, et al. DSM-III disorders from age 11 to age 15 years. *J Am Acad Child Adolesc Psychiatry* 1992;31:50–59.

77. Leibenluft E. Gender differences in major depressive disorder and bipolar disorder. *CNS Spectrums* 1999;4:25–33.

78. Frank E, Carpenter LL, Kupfer DJ. Sex differences in recurrent depression: are there any that are significant? *Am J Psychiatry* 1988;145:41–45.

79. Thase ME, Reynolds CF, Frank E, et al. Do depressed men and women respond similarly to cognitive-behavior therapy. *Am J Psychiatry* 1994;151:500–505.

80. Burke KC, Burke JD, Regier DA, et al. Age at onset of selected mental disorders in five community populations. *Arch Gen Psychiatry* 1990;47:511–518.

81. Winokur G, Tsuang MT, Crowe RR. The Iowa 500: affective disorder in relatives of manic and depressed patients. *Am J Psychiatry* 1982;139:209–212.

82. Kornstein SG, Schatzberg AF, Yonkers KA, et al. Gender differences in presentation of chronic major depression. *Psychopharmacol Bull* 1995;31:711–718.

83. Fava M, Abraham M, Alpert J, et al. Gender differences in Axis I comorbidity among depressed outpatients. *J Affect Disord* 1996;38:129–133.

84. Spicer CC, Hare EH, Slater E. Neurotic and psychotic forms of depressive illness: evidence from age-incidence in a national sample. *Br J Psychiatry* 1973;123:535–541.

85. Nolen-Hoeksema S. Sex differences in unipolar depression: evidence and theory. *Psychol Bull* 1987;101:259–282.

86. Kessler RC, McGonagle KA, Nelson CB, et al. Sex and depression in the National Comorbidity Survey II: cohort effects. *J Affect Disord* 1994;30:15–26.

87. Simpson HB, Nee JC, Endicott J. First-episode major depression. *Arch Gen Psychiatry* 1997;54:633–639.

88. Zlotnick C, Shea MT, Pilkonis PA, et al. Gender, type of treatment, dysfunctional attitudes, social support, life events, and depressive symptoms over naturalistic follow-up. *Am J Psychiatry* 1996;153:1021–1027.

89. Srikant CB, Patel YC. Somatostatin receptors. *Adv Exp Med Biol* 1985;188:291–304.

90. Aneshensel CS. The natural history of depressive symptoms. *Res Commun Ment Health* 1985;5:45–74.

91. Ernst C, Angst J. The Zurich Study. XII. Sex differences in depression. Evidence from longitudinal epidemiological data. *Eur Arch Psychiatry Clin Neurosci* 1992;241:222–230.

92. Keitner GI, Ryan CE, Miller IW, et al. Twelve-month outcome of patients with major depression and comorbid psychiatric or medical illness (compound depression). *Am J Psychiatry* 1991;148:345–350.

93. Winokur G, Coryell W, Keller M, et al. A prospective follow-up of patients with bipolar and primary unipolar affective disorder. *Arch Gen Psychiatry* 1993;50:457–465.

94. Young MA, Fogg LF, Scheftner WA, et al. Sex differences in the lifetime prevalence of depression: does varying the diagnostic criteria reduce the female/male ratio? *J Affect Disord* 1990;18:187–192.

95. Angst J, Dobler-Mikola A. Do the diagnostic criteria determine the sex ratio in depression? *J Affect Disord* 1984;7:189–198.

96. Bebbington PE, Brugha T, MacCarthy B, et al. The Camberwell Collaborative Depression Study, I: depressed probands—adversity and the form of depression. *Br J Psychiatry* 1988;152:754–765.

97. Karp JF, Frank E. Combination therapy and the depressed woman. *Depression* 1995;3:91–98.

98. Silberg J, Pickles A, Rutter M, et al. The influence of genetic factors and life stress on depression among adolescent girls. *Arch Gen Psychiatry* 1999;56:225–232.

99. Blazer DG, Kessler RC, McGonagle KA, et al. The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. *Am J Psychiatry* 1994;151:979–986.

100. Regier DA, Burke JD, Burke KC. Comorbidity of affective and anxiety disorders in the NIMH Epidemiologic Catchment Area Program. In: Maser JD, Cloninger CR, eds. *Comorbidity of mood and anxiety disorders*. Washington, DC: American Psychiatric Press, 1990:113–122.

101. Judd LL. When anxiety disorders are comorbid with major depression: social and clinical burden. Abstract presented at the 147th annual meeting of the American Psychiatric Association, 1994.

102. Reus VI. Behavioral aspects of thyroid disease in women. *Psychiatr Clin North Am* 1989;12:153–165.

103. Whybrow PC. Sex differences in thyroid axis function: relevance to affective disorder and its treatment. *Depression* 1995;3:33–42.

104. Moldin SO, Scheftner WA, Rice JP, et al. Association between major depressive disorder and physical illness. *Psychol Med* 1993;23:755–761.

105. Old Age Depression Interest Group. How long should the elderly take antidepressants? A double-blind placebo-controlled study of continuation/prophylaxis therapy with dothiepin. *Br J Psychiatry* 1993;162:175–182.

106. Raskin A. Age–sex differences in response to antidepressant drugs. *J Nerv Ment Dis* 1974;159:120–130.

107. Glassman AH, Perel JM, Shostak M, et al. Clinical implications of imipramine plasma levels for depressive illness. *Arch Gen Psychiatry* 1977;34:197–204.

108. Coppen A, Whybrow PC, Noguera R, et al. The comparative antidepressant value of L-tryptophan and imipramine with and without attempted potentiation by liothyronine. *Arch Gen Psychiatry* 1972;26:234–241.

109. Davidson J, Pelton S. Forms of atypical depression and their response to antidepressant drugs. *Psychiatry Res* 1986;17:87–95.

110. Steiner M, Wheadon DE, Kreider MS. Antidepressant response to paroxetine by gender. Abstract presented at the 146th annual meeting of the American Psychiatric Association, 1993.

111. Dawkins K, Potter WZ. Gender differences in pharmacokinetics and pharmacodynamics of psychotropics: focus on women. *Psychopharmacol Bull* 1991;27:417–426.

112. Yonkers KA, Kando JC, Cole JO, et al. Gender differences in pharmacokinetics and pharmacodynamics of psychotropic medication. *Am J Psychiatry* 1992;149:587–595.

113. Moody JP, Tait AC, Todrick A. Plasma levels of imipramine and desmethylimipramine during therapy. *Br J Psychiatry* 1967;113:183–193.

114. Preskorn SH, Mac DS. Plasma levels of amitriptyline: effects of age and sex. *J Clin Psychiatry* 1985;46:276–277.

115. Gex-Fabry M, Balant-Gorgia AE, Balant LP, et al. Clomipramine metabolism: model-based analysis of variability factors from drug-monitoring data. *Clin Pharmacokinet* 1990;19:241–255.

116. Greenblatt DJ, Friedman H, Burstein ES, et al. Trazodone kinetics: effects of age, gender, and obesity. *Clin Pharmacol Ther* 1987;42:193–200.

117. Warrington SJ. Clinical implications of the pharmacology of sertraline. *Int Clin Psychopharmacol* 1991;6:11–21.

118. Altshuler LL, Post RM, Leverich GS, et al. Antidepressantinduced mania and cycle acceleration: a controversy revisited. *Am J Psychiatry* 1995;152:1130–1138.

119. Lepine JP, Chignon JM, Teherani M. Suicide attempts in patients with panic disorder. *Arch Gen Psychiatry* 1993;50:144–149.

120. Dawkins K. Gender differences in psychiatry: epidemiology and drug response. *CNS Drugs* 1995;3:393–407.

121. Leboyer M, Filteau MJ, Jay M, et al. No gender effect on age at onset in familial schizophrenia? *Am J Psychiatry* 1992;149:1409.

122. Carpenter WT, Buchanan RW. Schizophrenia. *N Engl J Med* 1994;330:681–690.

123. Szymanski S, Lieberman JA, Alvir JM, et al. Gender differences in onset of illness, treatment response, course, and biologic indexes in first-episode schizophrenic patients. *Am J Psychiatry* 1995;152:698–703.

124. Seeman MV. Gender and the onset of schizophrenia: neurohumoral influences. *Psychiatr J Univ Ottawa* 1981;6:136–138.

125. Loranger AW. Sex differences in age of onset of schizophrenia. *Arch Gen Psychiatry* 1984;41:157–161.

126. Goldstein J. Gender differences in the course of schizophrenia. *Am J Psychiatry* 1988;145:684–689.

127. Lewine RRJ, Gulley LR, Risch SC, et al. Sexual dimorphism, brain morphology, and schizophrenia. *Schizophr Bull* 1990;16:195–203.

128. Cowell PE, Kostianovsky DJ, Gur RC, et al. Sex differences in neuroanatomical and clinical correlations in schizophrenia. *Am J Psychiatry* 1996;153:799–805.

129. Goldstein JM, Santangelo SL, Simpson JC, et al. The role of gender in identifying subtypes of schizophrenia—a latent class analytic approach. *Schizophr Bull* 1990;16:263–275.

130. Andreasen NC, Ehrhardt JC, Swayze VW, et al. Magnetic resonance imaging of the brain in schizophrenia—the pathophysiologic significance of structural abnormalities. *Arch Gen Psychiatry* 1990;47:35–44.

131. Raine A, Harrison GN, Reynolds GP, et al. Structural and functional characteristics of the corpus callosum in schizophrenics, psychiatric controls, and normal controls—a magnetic resonance imaging and neuropsychological evaluation. *Arch Gen Psychiatry* 1990;47:1060–1064.

132. Nasrallah HA, Schwarzkopf SB, Olson SC, et al. Gender differences in schizophrenia on MRI brain scans. *Schizophr Bull* 1990;16:205–209.

133. Woodruff PWR, Pearlson GD, Geer MJ, et al. A computerized magnetic resonance imaging study of corpus callosum morphology in schizophrenia. *Psychol Med* 1993;23:45–56.

134. Aylward E, Walker E, Bettes B. Intelligence in schizophrenia: metaanalysis of the research. *Schizophr Bull* 1984;10:430–459.

135. McGlashan TH, Bardenstein KK. Gender differences in affective, schizoaffective, and schizophrenic disorders. *Schizophr Bull* 1990;16:319–329.

136. Angermeyer MC, Kuhn L, Goldstein JM. Gender and the course of schizophrenia: differences in treated outcomes. *Schizophr Bull* 1990;16:293–307.

137. Bardenstein KK, McGlashan T. Gender differences in affective, schizoaffective and schizophrenic disorders. A review. *Schizophr Res* 1990;3:159–172.

138. Breier A, Astrachan BM. Characteristics of schizophrenic patients who commit suicide. *Am J Psychiatry* 1984;141:206–209.

139. Seeman MV. Interaction of sex, age and neuroleptic dose. *Compr Psychiatry* 1983;24:125–128.

140. Chouinard G, Annable L, Steinberg S. A controlled clinical trial of fluspirilene, a long-acting injectable neuroleptic, in schizophrenic patients with acute exacerbation. *J Clin Psychopharmacol* 1986;6:21–26.

141. Chouinard G, Annable L. Pimozide in the treatment of newly admitted schizophrenic patients. *Psychopharmacology* 1982;76:13–19.

142. Seeman MV. Current outcome in schizophrenia: women vs. men. *Acta Psychiatr Scand* 1986;73:609–617.

143. Goldstein MJ, Rodnick EH, Evans JR, et al. Drug and family therapy in the aftercare of acute schizophrenics. *Arch Gen Psychiatry* 1978;35:1169–1177.

144. Andia AM, Zisook S, Heaton RK, et al. Gender differences in schizophrenia. *J Nerv Ment Dis* 1995;183:522–528.

145. Seeman MV. Neuroleptic prescription for men and women. *Soc Pharmacol* 1989;3:219–236.

146. Shaywitz BA, Shaywitz SE, Pugh KR, et al. Sex differences in the functional organization of the brain for language. *Nature* 1995;373:607–609.

147. Giedd JN, Blumenthal J, Jeffries NO, et al. Brain development during childhood and adolescence: a longitudinal MRI study. *Nature* 1999;2:861–863.

148. Witelson SA. Neural sexual mosaicism: sexual differentiation of the human temporoparietal region for functional asymmetry. *Psychoneuroendocrinology* 1991;16:131–153.

149. Azari NP, Pettigrew KD, Pietrini P, et al. Sex differences in patterns of hemispheric cerebral metabolism—a multiple regression discriminant analysis of positron emission tomographic data. *Int J Neurosci* 1995;81:1–20.

150. Allen L, Hines M, Shryne J, et al. Two sexually dimorphic cell groups in the human brain. *J Neurosci* 1989;9:497–506.

151. George MS, Ketter TA, Parekh PI, et al. Gender differences in regional cerebral blood flow during transient self-induced sadness or happiness. *Biol Psychiatry* 1996;40:859–871.

152. Biver F, Lotstra F, Monclus M, et al. Sex difference in 5-HT_2 receptor in the living human brain. *Neurosci Lett* 1996;204:25–28.

153. Nishizawa S, Benkelfat C, Young SN, et al. Differences between males and females in rates of serotonin synthesis in human brain. *Proc Natl Acad Sci USA* 1997;94:5308–5313.

154. Rodriguez G, Warkentin S, Risberg J, et al. Sex differences in regional cerebral blood flow. *J Cereb Blood Flow Metab* 1988;8:783–789.

155. Gur RC, Gur RE, Obrist WD, et al. Sex and handedness differences in cerebral blood flow during rest and during cognitive activity. *Science* 1982;217:659–661.

156. Gur RC, Gur RE, Obrist WD, et al. Age and regional cerebral blood flow at rest and during cognitive activity. *Arch Gen Psychiatry* 1987;44:617–621.

157. Shaw T, Meyer JS, Mortel K, et al. Effects of normal aging, sex, and risk factors for stroke on regional cerebral blood flow (rCBF) in normal volunteers. In: Gotoh F, Nagai H, Tazaki Y, eds. *Cerebral blood flow and metabolism*. Copenhagen: Munksgaard, 1979.

158. Esposito G, Van Horn JD, Weinberger DR, et al. Gender differences in cerebral blood flow as a function of cognitive state with PET. *J Nucl Med* 1996;37:559–564.

159. Kulynych JJ, Vladar K, Jones DW, et al. Gender differences in the normal lateralization of the supratemporal cortex: MRI surface-rendering morphometry of Heschl's gyrus and the planum temporale. *Cereb Cortex* 1994;4:107–118.

160. Baxter LR Jr, Mazziotta JC, Phelps ME, et al. Cerebral glucose metabolic rates in normal human females versus normal males. *Psychiatry Res* 1987;21:237–245.

161. Andreason PJ, Zametkin AJ, Guo AC, et al. Gender-related differences in regional cerebral glucose metabolism in normal volunteers. *Psychiatry Res* 1993;51:175–183.

162. Berman KF, Schmidt PJ, Rubinow DR, et al. Modulation of cognition-specific cortical activity by gonadal steroids: a positron emission tomography study in women. *Proc Natl Acad Sci USA* 1997;94:8836–8841.

163. Shaywitz SE, Shaywitz BA, Pugh KR, et al. Effect of estrogen on brain activation patterns in postmenopausal women during working memory tasks. *JAMA* 1999;281:1197–1202.

164. Wong DF, Broussolle EP, Wand G, et al. *In vivo* measurement of dopamine receptors in human brain by positron emission tomography: age and sex differences. *Ann NY Acad Sci* 1988;515:203–214.

164a. Frank RT. The hormonal causes of premenstrual tension. *Archives of Neurology and Psychiatry* 1931;26:1053–1057.

165. von Feuchtersleben E. *The principles of medical psychology*. London: Sydenham Society, 1847.

166. Rubinow DR, Roy-Byrne PP. Premenstrual syndromes: overview from a methodologic perspective. *Am J Psychiatry* 1984;141:163–172.

167. Rubinow DR, Roy-Byrne PP, Hoban MC, et al. Prospective assessment of menstrually related mood disorders. *Am J Psychiatry* 1984;141:684–686.

168. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, third edition, revised. Washington, DC: American Psychiatric Association, 1987.

169. Rubinow DR, Hoban MC, Grover GN, et al. Changes in plasma hormones across the menstrual cycle in patients with menstrually related mood disorder and in control subjects. *Am J Obstet Gynecol* 1988;158:5–11.

170. Backstrom T, Sanders D, Leask R, et al. Mood, sexuality, hormones, and the menstrual cycle: II. Hormone levels and their relationship to the premenstrual syndrome. *Psychosom Med* 1983;45:503–507.

171. Wang M, Seippel L, Purdy RH, et al. Relationship between symptom severity and steroid variation in women with premenstrual syndrome: study on serum pregnenolone, pregnenolone sulfate, 5α-pregnane-3,20-dione and 3α-hydroxy-5α-pregnan-20-one. *J Clin Endocrinol Metab* 1996;81:1076–1082.

172. Redei E, Freeman EW. Daily plasma estradiol and progesterone levels over the menstrual cycle and their relation to premenstrual symptoms. *Psychoneuroendocrinology* 1995;20:259–267.

173. Facchinetti F, Genazzani AD, Martignoni E, et al. Neuroendocrine changes in luteal function in patients with premenstrual syndrome. *J Clin Endocrinol Metab* 1993;76:1123–1127.

174. Backstrom T, Aakvaag A. Plasma prolactin and testosterone during the luteal phase in women with premenstrual tension syndrome. *Psychoneuroendocrinology* 1981;6:245–251.

175. Eriksson E, Sundblad C, Lisjo P, et al. Serum levels of androgens are higher in women with premenstrual irritability and dysphoria than in controls. *Psychoneuroendocrinology* 1992;17:195–204.

176. Bloch M, Schmidt PJ, Su T-P, et al. Pituitary–adrenal hormones and testosterone across the menstrual cycle in women with premenstrual syndrome and controls. *Biol Psychiatry* 1998;43:897–903.

177. Smith SS, Gong QH, Hsu F-C, et al. $GABA_A$ receptor alpha-4 subunit suppression prevents withdrawal properties of an endogenous steroid. *Nature* 1998;392:926–930.

178. Smith SS, Gong QH, Li X, et al. Withdrawal from 3α-OH-5α-pregnan-20-one using a pseudopregnancy model alters the kinetics of hippocampal $GABA_A$-gated current and increases the $GABA_A$-receptor α_4 subunit in association with increased anxiety. *J Neurosci* 1998;18:5275–5284.

179. Ströhle A, Romeo E, Hermann B, et al. Concentrations of 3α-reduced neuroactive steroids and their precursors in plasma of patients with major depression and after clinical recovery. *Biol Psychiatry* 1999;45:274–277.

180. Romeo E, Brancati A, de Lorenzo A, et al. Marked decrease of plasma neuroactive steroids during alcohol withdrawal. *Clin Neuropharmacol* 1996;19:366–369.

181. Romeo E, Strohle A, Spalletta G, et al. Effects of antidepressant treatment on neuroactive steroids in major depression. *Am J Psychiatry* 1998;155:910–913.

182. Uzunova V, Sheline Y, Davis JM, et al. Increase in the cerebrospinal fluid content of neurosteroids in patients with unipolar major depression who are receiving fluoxetine or fluvoxamine. *Proc Natl Acad Sci USA* 1998;95:3239–3244.

183. Bitran D, Purdy RH, Kellogg CK. Anxiolytic effect of progesterone is associated with increases in cortical allopregnanolone and GABAa-receptor function. *Pharmacol Biochem Behav* 1993;45:423–428.

184. Bitran D, Hilvers RJ, Kellogg CK. Anxiolytic effects of 3α-hydroxy-5α[β]-pregnan-20-one: endogenous metabolites of progesterone that are active at the $GABA_A$ receptor. *Brain Res* 1991;561:157–161.

185. Wieland S, Lan NC, Mirasedeghi S, et al. Anxiolytic activity of the progesterone metabolite 5α-pregnan-3α-ol-one. *Brain Res* 1991;565:263–268.

186. Purdy RH, Morrow AL, Moore PH Jr, et al. Stress-induced elevations of gamma-aminobutyric acid type A receptor-active steroids in the rat brain. *Proc Natl Acad Sci USA* 1991;88:4553–4557.

187. Griffin LD, Mellon SH. Selective serotonin reuptake inhibitors directly alter activity of neurosteroidogenic enzymes. *Proc Natl Acad Sci U S A* 1999;96:13512–13517.

188. Sundstrom I, Andersson A, Nyberg S, et al. Patients with premenstrual syndrome have a different sensitivity to a neuroactive steroid during the menstrual cycle compared to control subjects. *Neuroendocrinology* 1998;67:126–138.

189. Sundstrom I, Nyberg S, Backstrom T. Patients with premenstrual syndrome have reduced sensitivity to midazolam compared to control subjects. *Neuropsychopharmacology* 1997;17:370–381.

190. Rapkin AJ, Morgan M, Goldman L, et al. Progesterone metabolite allopregnanolone in women with premenstrual syndrome. *Obstet Gynecol* 1997;90:709–714.

191. Schmidt PJ, Purdy RH, Moore PH Jr, et al. Circulating levels of anxiolytic steroids in the luteal phase in women with premenstrual syndrome and in control subjects. *J Clin Endocrinol Metab* 1994;79:1256–1260.

192. Wang G-J, Volkow ND, Overall J, et al. Reproducibility of regional brain metabolic responses to lorazepam. *J Nucl Med* 1996;37:1609–1613.

193. Sundstrom I, Backstrom T. Citalopram increases pregnanolone sensitivity in patients with premenstrual syndrome: an open trial. *Psychoneuroendocrinology* 1998;23:73–88.

194. Schechter D, Strasser TJ, Endicott J, et al. Role of ovarian steroids in modulating mood in premenstrual syndrome. In: *Proceedings of the 51st annual meeting of the Society of Biological Psychiatry*, 1996:646(abst).

195. Halbreich U, Endicott J, Goldstein S, et al. Premenstrual changes and changes in gonadal hormones. *Acta Psychiatr Scand* 1986;74:576–586.

196. Schmidt PJ, Nieman LK, Grover GN, et al. Lack of effect of induced menses on symptoms in women with premenstrual syndrome. *N Engl J Med* 1991;324:1174–1179.

197. Muse KN, Cetel NS, Futterman LA, et al. The premenstrual syndrome: effects of "medical ovariectomy." *N Engl J Med* 1984;311:1345–1349.

198. Casson P, Hahn PM, VanVugt DA, et al. Lasting response to ovariectomy in severe intractable premenstrual syndrome. *Am J Obstet Gynecol* 1990;162:99–105.

199. Schmidt PJ, Nieman LK, Danaceau MA, et al. Differential behavioral effects of gonadal steroids in women with and in those without premenstrual syndrome. *N Engl J Med* 1998;338:209–216.

200. Hendrick V, Altshuler LL, Suri R. Hormonal changes in the postpartum and implications for postpartum depression. *Psychosomatics* 1998;39:93–101.

201. Harris B, Lovett L, Newcombe RG, et al. Maternity blues and major endocrine changes: Cardiff puerperal mood and hormone study II. *BMJ* 1994;308:949–953.

202. Harris B. Biological and hormonal aspects of postpartum depressed mood: working towards strategies for prophylaxis and treatment. *Br J Psychiatry* 1994;164:288–292.

203. Wieck A. Endocrine aspects of postnatal mental disorders. *Baillieres Clin Obstet Gynaecol* 1989;3:857–877.

204. Harris B, Lovett L, Smith J, et al. Cardiff puerperal mood and hormone study. III. Postnatal depression at 5 to 6 weeks postpartum, and its hormonal correlates across the peripartum period. *Br J Psychiatry* 1996;168:739–744.

205. Buckwalter JG, Stanczyk FZ, McCleary CA, et al. Pregnancy, the postpartum, and steroid hormones: effects on cognition and mood. *Psychoneuroendocrinology* 1999;24:69–84.

206. Spearow JL, Doemeny P, Sera R, et al. Genetic variation in susceptibility to endocrine disruption by estrogen in mice. *Science* 1999;285:1259–1261.

207. Finn DA, Roberts AJ, Lotrich F, et al. Genetic differences in behavioral sensitivity to a neuroactive steroid. *J Pharmacol Exp Ther* 1997;280:820–828.

208. Huizenga NATM, Koper JW, De Lange P, et al. A polymorphism in the glucocorticoid receptor gene may be associated with an increased sensitivity to glucocorticoids *in vivo*. *J Clin Endocrinol Metab* 1998;83:144–151.

209. Huizenga NATM, Koper JW, De Lange P, et al. Interperson variability but intraperson stability of baseline plasma cortisol concentrations, and its relation to feedback sensitivity of the hypothalamo–pituitary–adrenal axis to a low dose of dexamethasone in elderly individuals. *J Clin Endocrinol Metab* 1998;83:47–54.

210. Giovannucci E, Stampfer MJ, Krithivas K, et al. The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc Natl Acad Sci USA* 1997;94:3320–3323.

211. Beilin J, Zajac JD. Function of the human androgen receptor varies according to CAG repeat number within the normal range. In: *Proceedings of the 81st annual meeting of the Endocrine Society*, 1999:500(abst).

212. Salo LK, Makela SI, Stancel GM, et al. Neonatal exposure to diethylstilbestrol permanently alters the basal and 17 beta-estradiol induced expression of c-*fos* proto-oncogene in mouse urethroprostatic complex. *Mol Cell Endocrinol* 1997;126:133–141.

213. Gregoire AJP, Kumar R, Everitt B, et al. Transdermal oestrogen for treatment of severe postnatal depression. *Lancet* 1996;347:930–933.

214. Saletu B, Brandstatter N, Metka M, et al. Hormonal, syndromal and EEG mapping studies in menopausal syndrome patients with and without depression as compared with controls. *Maturitas* 1996;23:91–105.

215. Cawood EHH, Bancroft J. Steroid hormones, the menopause, sexuality and well-being of women. *Psychol Med* 1996;26:925–936.

216. Barrett-Connor E, von Muhlen D, Laughlin GA, et al. Endogenous levels of dehydroepiandrosterone sulfate, but not other sex hormones, are associated with depressed mood in older women: the Rancho Bernardo study. *J Am Geriatr Soc* 1999;47:685–691.

217. Brambilla F, Maggioni M, Ferrari E, et al. Tonic and dynamic gonadotropin secretion in depressive and normothymic phases of affective disorders. *Psychiatry Res* 1990;32:229–239.

218. Amsterdam JD, Winokur A, Lucki I, et al. Neuroendocrine regulation in depressed postmenopausal women and healthy subjects. *Acta Psychiatr Scand* 1983;67:43–49.

219. Altman N, Sachar EJ, Gruen PH, et al. Reduced plasma LH concentration in postmenopausal depressed women. *Psychosom Med* 1975;37:274–276.

220. Guicheney P, Léger D, Barrat J, et al. Platelet serotonin content and plasma tryptophan in peri- and postmenopausal women: variations with plasma oestrogen levels and depressive symptoms. *Eur J Clin Invest* 1988;18:297–304.

221. Werner AA, Johns GA, Hoctor EF, et al. Involutional melancholia: probable etiology and treatment. *JAMA* 1934;103:13–16.

222. Montgomery JC, Brincat M, Tapp A, et al. Effect of oestrogen and testosterone implants on psychological disorders in the climacteric. *Lancet* 1987;1:297–299.

223. Saletu B, Brandstatter N, Metka M, et al. Double-blind, placebo-controlled, hormonal, syndromal and EEG mapping studies with transdermal oestradiol therapy in menopausal depression. *Psychopharmacology* 1995;122:321–329.

224. Sherwin BB. Affective changes with estrogen and androgen replacement therapy in surgically menopausal women. *J Affect Disord* 1988;14:177–187.

225. Klaiber EL, Broverman DM, Vogel W, et al. Estrogen therapy for severe persistent depressions in women. *Arch Gen Psychiatry* 1979;36:550–554.

226. Ditkoff EC, Crary WG, Cristo M, et al. Estrogen improves psychological function in asymptomatic postmenopausal women. *Obstet Gynecol* 1991;78:991–995.

227. Coope J. Is oestrogen therapy effective in the treatment of menopausal depression? *J R Coll Gen Pract* 1981;31:134–140.

228. Campbell S. Double-blind psychometric studies on the effects of natural estrogens on post-menopausal women. In: Campbell S, ed. *The management of the menopause and postmenopausal years*. Lancaster, UK: MTP Press, 1976:149–158.

229. George GCW, Utian WH, Beaumont PJV, et al. Effect of exogenous oestrogens on minor psychiatric symptoms in postmenopausal women. *S Afr Med J* 1973;47:2387–2388.

230. Schmidt PJ, Nieman L, Danaceau MA, et al. Estrogen replacement in perimenopause-related depression: a preliminary report. *Am J Obstet Gynecol* 2000;183:414–420.

231. Wolkowitz OM, Reus VI, Keebler A, et al. Double-blind treatment of major depression with dehydroepiandrosterone. *Am J Psychiatry* 1999;156:646–649.

81

Depression and the Medically Ill

Robert G. Robinson

K. Ranga Rama Krishnan

Robert G. Robinson: Department of Psychiatry, University of Iowa, Iowa City, Iowa.

K. Ranga Rama Krishnan: Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, North Carolina.

The lifetime prevalence of depression in the general U.S. population has been reported to be as high as 17% (1). The rates of depression in healthy persons are significantly lower than those among the medically ill. The rates of depression among medically ill patients range as high as 20% to 40%. Depression in the presence of medical illness, in comparison with depression in the absence of medical illness, is associated with more severely impaired physical or cognitive function. Untreated, depression can persist for many months and may complicate recovery from the medical illness.

Each year, approximately $44 billion is spent in the treatment of depression (2). Although depression associated with medical illness has been shown to increase mortality (3), the benefits of treating depression on medical morbidity and mortality have not yet been established. In this chapter, we review the relationship between depression and medical illness, with cerebrovascular and cardiovascular disease used as a prototype of medical illness.

PREVALENCE OF DEPRESSION AFTER MYOCARDIAL INFARCTION OR STROKE

Part of "81 - Depression and the Medically Ill "

The prevalence of major depression in patients after myocardial infarction (MI) has been estimated to be about 20% (3 ,4 ,5 ,6 ,7 and 8). Depressive symptoms following an acute MI have been reported in 60% of patients (9 and 10). Even in patients with only angiographically proven coronary artery disease, the prevalence of depression is approximately 18% (11).

The prevalence of depression after stroke has been studied in numerous countries of the world (12 ,13 and 14). These studies have found a mean prevalence of major depression of 20% among hospitalized and outpatient victims of stroke, and a prevalence of 13% has been found in community surveys (14). The mean prevalence of minor depression, defined as a subsyndromal form of major depression or the symptoms of dysthymia without the 2-year duration, has been found to be 19% in hospital and 10% in community samples (15) (Fig. 81.1).

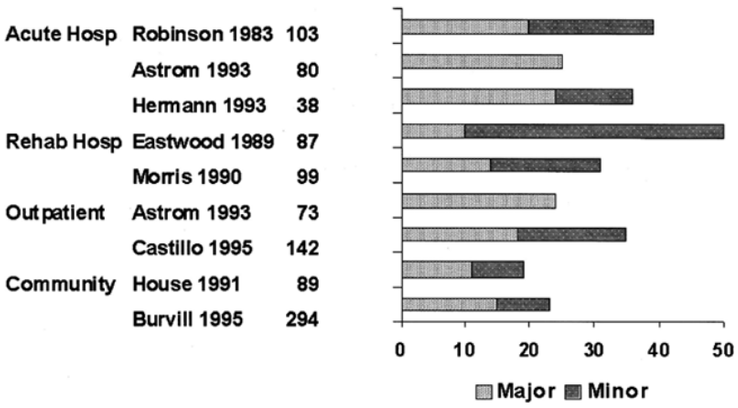


FIGURE 81.1. The percentage of patients found to be depressed after stroke based on the setting in which they were evaluated. Note that hospitalized patients and those in outpatient clinics generally show higher rates of depression than those studied in the community. This finding probably reflects the fact that the strokes of patients who seek medical service are more severe than those of patients in community surveys, in which persons with very mild or no impairment are studied. These studies represent findings from the United States, Canada, England, Sweden, Germany, and Australia.

DIAGNOSIS OF DEPRESSION IN PATIENTS WITH MYOCARDIAL INFARCTION OR STROKE

Part of "81 - Depression and the Medically Ill "

The problem of diagnosing depression in patients with medical illness has been a focus of research and a source of controversy among consultation and liaison psychiatrists for many years. Cohen-Cole and Stoudemire (16) reported that four approaches have been used to deal with this problem in the medically ill. In the “inclusive approach,” depressive diagnostic symptoms are counted whether or not they are related to physical illness. In the “etiologic approach” of Rifkin et al. (17), a symptom is counted only if the diagnostician feels it is not caused by the physical illness. Rapp and Varna (18) use the “substitutive approach” of Endicott (19), in which the psychological symptoms of depression are substituted for the vegetative symptoms, which tend to be nonspecific in a physically ill population, and the “exclusive approach” of Bukberg et al. (20), in which symptoms are removed from the diagnostic criteria if they are not found to be more frequent in depressed than in nondepressed patients.

The utility of these methods in the diagnosis of post-stroke depression was examined in a study that included 142 patients with acute stroke who were reexamined at 3, 6, 12, and 24 months after their stroke. Of the 142 patients, 60 (42%) reported a depressed mood during their acute hospitalization, and the remaining 82 patients denied a depressed mood. No significant differences in background characteristics were found between the depressed and nondepressed groups except that the depressed group was significantly younger (p =.006) and the frequency of personal

history of a psychiatric disorder was significantly higher in the depressed group (p = .04). Throughout the 2 years of follow-up, the depressed patients reported a significantly higher frequency of both vegetative and psychological symptoms of depression than did the patients who were not depressed. The vegetative symptoms that were examined included anxiety, anxious foreboding, morning depression, weight loss, delayed sleep, subjective anergia, early morning awakening, and loss of libido (21). The only vegetative symptoms that were not more frequent in the depressed than in the nondepressed group were weight loss and early morning awakening at the initial evaluation; weight loss, delayed sleep, and early morning awakening at 3 months; weight loss and early morning awakening at 6 months; weight loss, early morning awakening, anxious foreboding, and loss of libido at 1 year; and weight loss and loss of libido at 2 years. Most psychological symptoms were more frequent in the depressed patients throughout the 2-year period. The psychological symptoms assessed included worrying, brooding, loss of interest, hopelessness, suicidal plans, social withdrawal, self-deprecation, lack of self-confidence, simple ideas of reference, guilty ideas of reference, pathologic guilt, and irritability. The only psychological symptoms that were not significantly more frequent in the depressed than in the nondepressed group were suicidal plans, simple ideas of reference, and pathologic guilt at 3 months; pathologic guilt at 6 months; pathologic guilt, suicidal plans, guilty ideas of reference, and irritability at 1 year; and pathologic guilt and self-deprecation at 2 years.

The effect of using each of the proposed alternative diagnostic methods for post-stroke depression based on DSM-IV criteria was examined (22). Symptoms were assessed with the inclusive approach (i.e., a symptom was included if acknowledged, even if it was suspected that the symptom was related to the physical illness). Thus, the initial diagnoses were based on inclusive criteria during the in-hospital evaluation. When this approach was used, 26 patients (18%) met the DSM-IV diagnostic criteria for major depression. The DSM-IV diagnostic criteria were then modified by imposing a requirement for five or more specific symptoms (weight loss and early morning awakening were excluded as criteria for major depression because they were not significantly more frequent in the depressed than in the nondepressed patients). Of the 27 patients with major depression, three were excluded in comparison with the diagnosis based on inclusive criteria. The diagnoses based on unmodified symptoms had a specificity of 98% and a sensitivity of 100% in comparison with use of the “exclusive” criteria as the “gold standard.”

Next, the DSM-IV criteria were modified to examine the substitutive approach (i.e., all vegetative symptoms were eliminated and the presence of four psychological symptoms plus depressed mood was required for the diagnosis of major depression). When this approach was used, none of the original 27 patients in whom major depression had been diagnosed with the inclusive approach was excluded.

At the 3-month follow-up, use of the exclusive approach, which requires only specific symptoms (i.e., weight loss, insomnia, and suicidal ideation were eliminated), resulted in the exclusion of 1 of 12 patients (16%) with major depression. When a diagnosis based on specific symptoms was used as the gold standard, the unmodified DSM-IV criteria and inclusive approach had a sensitivity of 100% and a specificity of 97%. If the substitutive approach, which requires depression plus four psychological symptoms, had been used, none of the 12 patients would have been excluded.

At the 6-month follow-up, when the exclusive approach (i.e., weight loss and insomnia were excluded) was used, 3 of 15 patients no longer met the criteria for major depression. When specific symptoms were used as the gold standard, the unmodified inclusive DSM-IV criteria had a sensitivity of 100% and a specificity of 95%. If the substitutive approach had been used, none of the 15 patients with major depression would have been excluded. At the 1-year follow-up, when the exclusive approach (i.e., weight loss, difficulty concentrating, and suicidal ideation were excluded) was used, 3 of 7 patients no longer met the diagnostic criteria, and the unmodified inclusive DSM-IV criteria had a sensitivity of 100% and a specificity of 95%. With use of the substitutive approach, none of the 7 patients was excluded. At the 2-year follow-up, when the exclusive approach (i.e., weight loss was excluded) was used, 2 of 16 patients with major depression were excluded. With use of the unmodified inclusive DSM-IV criteria, the sensitivity was 100% and the specificity was 96% in comparison with the exclusive approach. The substitutive approach excluded none of the 16 patients.

Although Kathol et al. (23) concluded that the substitutive approach is the best one given our current knowledge, the inclusive approach previously described for patients with stroke during the first 2 years of follow-up had a sensitivity of 100% and specificity of more than 95% in comparison with the exclusive approach. Weight loss was the only symptom that was not significantly more frequent among depressed than among nondepressed patients during the entire 2-year period. Given that the unmodified DSM-IV criteria consistently showed a sensitivity of 100% and a specificity that ranged from 95% to 98% in comparison with "exclusive" criteria that counted only specific symptoms, one could reasonably conclude that the use of unmodified DSM-IV criteria and the inclusive approach is the most rational way to diagnose depression in patients with vascular disease.

DEPRESSION AS A RISK FACTOR FOR CARDIAC DISEASE

Part of "81 - Depression and the Medically Ill "

It has been suggested that depression may play a role in the development of cardiac disease (24 ,25 ,26 ,27 ,28 and 29). For example, data from a study of 2,832 U.S. adults found that persons with symptoms of depression are at 50% greater risk of fatal ischemic heart disease (relative risk, 1.5; 95% confidence interval, 1.0 to 2.3) than are persons without depression (24). A notable exception, however, is a study of more than 2,500 persons followed for 15 years in which an association between depression and increased cardiovascular mortality or ischemic heart disease was not observed (30). The relationship of depression to the development of stroke has not been studied systematically. Although a significant number of patients in whom acute stroke develops have a preexisting depressive disorder (31), a causal relationship of depression and stroke has not been established.

IMPACT OF DEPRESSION IN MYOCARDIAL INFARCTION

Part of "81 - Depression and the Medically Ill "

A higher cardiovascular mortality among patients with major depression has been reported by numerous investigators (24 ,26 ,27 ,32 ,33 ,34 ,35 ,36 ,37 ,38 ,39 ,40 ,41 ,42 and 43). Large, controlled studies have confirmed an increased cardiovascular risk in persons in whom depression develops following an ischemic event (3 ,7 ,26 ,44 ,45 and 46). Frasure-Smith and colleagues (3) studied 222 patients who had experienced an acute MI. Patients were screened with the Beck Depression Inventory and interviewed after their MI with a modified version of the National Institute of Mental Health Diagnostic Interview Schedule to determine if criteria for depression were met. Even after corrections were made for relevant variables that determine outcome, including previous MI and Killip class, major depression was a significant independent predictor of adverse outcome (adjusted hazard ratio, 4.29; 95% confidence interval, 3.14 to 5.44; $p = .013$). The same results were found for patients with modest depressive symptoms (i.e., Beck Depression Inventory > 10). During a subsequent 18-month follow-up, the risk persisted. The underlying pathophysiology and the effects of antidepressant treatment on mortality are the focus of ongoing clinical investigations.

An increased risk for mortality among patients with depression following stroke has been reported in two studies utilizing different patient populations (47 and 48). One study examined the 10-year follow-up of 91 of 103 patients evaluated following acute stroke. Among the 48 patients who had died, it was found that those with major or minor depression while in the hospital were 3.4 times more likely to have died (confidence interval, 1.4 to 8.4; $p = .007$) than were those who were not depressed while in the hospital, even when other variables (e.g., lesion volume) related to mortality were controlled (47 ,48) (Fig. 81.2).

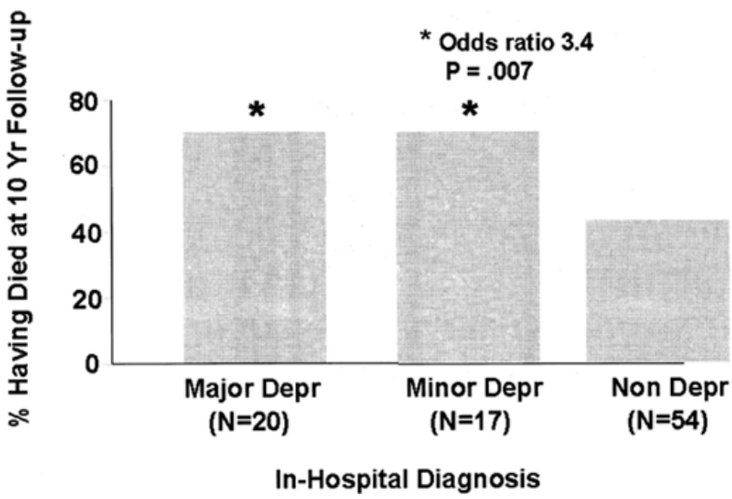


FIGURE 81.2. Survival curves during 10 years for 37 patients with major or minor depression at the time of in-hospital post-stroke evaluation and for 54 patients without in-hospital depression. By 10-year follow-up, 14 of 20 patients with major depression and 12 of 17 with minor depression had died, in comparison with only 22 of 54 nondepressed patients. (Data extracted from Morris PLP, Robinson RG, Andrezejewski P, et al. Association of depression with 10 year post-stroke mortality. *Am J Psychiatry* 1993;150:124–129, with permission.)

A 15-month follow-up of 84 Australian patients who were initially examined in a rehabilitation hospital found that 23% of the patients with major depression at initial evaluation had died, in comparison with 10% of those with minor depression and 2% of those without depression. Patients with major or minor depression were 8.1 times more likely to have died during the 15-month follow-up (confidence interval, 0.9 to 72.9; $p = .06$) than were nondepressed patients (47 ,48). However, Astrom et al. (49) found that in 21 of 80 patients who had initially been examined within the first 2 weeks following an acute stroke and had died during 3 years of follow-up, mortality was associated

with older age, disorientation during the hospitalization, greater impairment in activities of daily living, and a greater degree of cortical atrophy, but not with depression. Despite this negative finding, numerous studies have found that depression associated with stroke or MI increases the risk for death during the first 3 to 5 years after the vascular illness.

Concurrent depression also increases the risk for poor outcome among patients with MI, unstable angina, and heart failure. Unpublished data from Duke University suggest a twofold increase in mortality in depressed patients with heart failure. Besides the increase in mortality, a substantial effect on morbidity has been noted. Post-MI patients who are depressed take longer to return to work than those without depression (10 ,50 ,51). The most parsimonious explanation is that patients who are depressed following an MI are more likely to drop out of cardiac rehabilitation and exercise programs than are patients who are not depressed (52). For example, depressed smokers are 40% less likely to stop smoking than nondepressed smokers (relative risk, 0.6) (53), and depressed patients with coronary artery disease are less likely to comply with low-dose aspirin therapy than are nondepressed patients (54). These findings would suggest that all depressed patients with coronary artery disease should be treated, but data indicating the efficacy and safety of treatments for depression associated with heart disease are very limited.

Numerous studies have documented the adverse effect of depression on physical recovery from stroke (55 ,56 and 57). In a study of 25 patients with major or minor depression after stroke and some impairment in their activities of daily living versus a comparable group of 38 nondepressed patients, the degree of recovery in activities of daily living was significantly greater in the nondepressed than in the depressed patients at 2-year follow-up (p <.01), even when other factors that influence outcome (e.g., baseline deficits, early intervention, specialized stroke and rehabilitation care, nature and size of the lesion) were controlled (57). This delayed recovery was evident as early as 3 to 6 months following stroke (57 ,58) (Fig. 81.3).

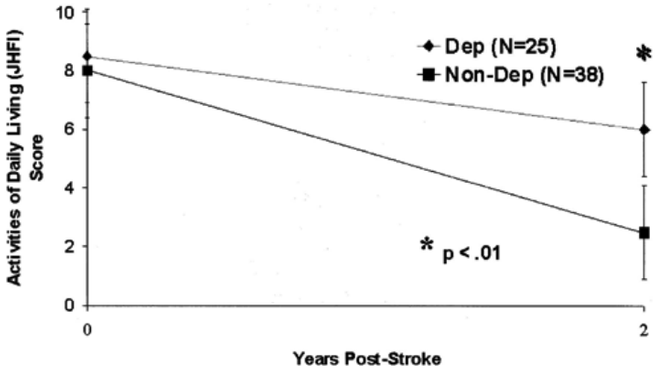


FIGURE 81.3. Changes in activities-of-daily-living scores for depressed (major or minor) patients and nondepressed patients at the time of in-hospital evaluation and 2 years later. Depressed patients show less recovery than nondepressed patients. (Reprinted from Parikh RM, Robinson RG, Lipsey JR, et al. The impact of post-stroke depression on recovery in activities of daily living over two-year follow-up. *Arch Neurol* 1990;47:785–789, with permission.)

In addition to the adverse effects of depression on activities of daily living, numerous studies have demonstrated the adverse effects of major depression after stroke on cognitive function (59 ,60 ,61 and 62). In a study of 275 patients with acute stroke, the mean Mini-Mental State Examination (MMSE) score for patients with major depression (n = 56) was 20.0 ± 6.2; for those with minor depression (n = 49), it was 22.9 ± 6.3; and for those without depression (n = 170), it was 23.3 ± 5.3 (p = .001) (63) (Fig. 81.4). To control for the possibility that the location of lesions, which has been correlated with affect during acute stroke, might have influenced these findings, patients with and without major depression were matched for size and location of stroke lesion (64). Patients with major depression had significantly lower MMSE scores than did their lesion-matched counterparts. Furthermore, on a battery of detailed neuropsychological tests, patients with major depression after stroke showed significantly greater impairment in orientation, language, visual-spatial skills, and executive motor and frontal lobe tasks than did nondepressed patients with lesions in similar locations (59). These findings indicate that major depression following stroke leads to a dementia of depression. Furthermore, this adverse effect of major depression on cognitive function lasts for the first year following stroke (63).

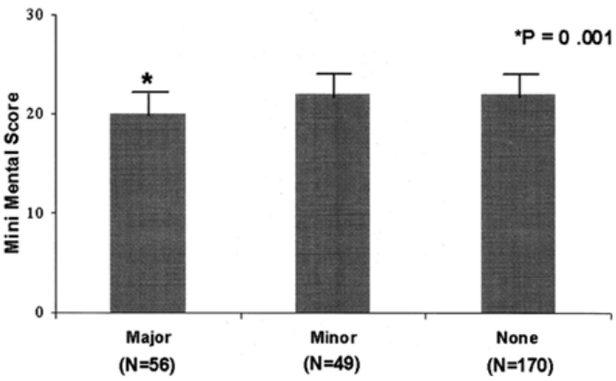


FIGURE 81.4. Mini-Mental State Examination scores in patients with major and minor depression and patients without depression.

The Enhancing Recovery in Coronary Heart Disease (ENRICHD) trial, sponsored by the National Heart, Lung, and Blood Institute, is currently under way (65). Psychosocial interventions in post-MI patients with depression or a high level of perceived social isolation are being evaluated in this multicenter mortality trial of 3,000 patients. Results should be available by the end of 2001 and should provide

answers to the question of whether psychosocial treatments should be initiated in patients with post-MI depression.

TREATMENT WITH ANTIDEPRESSANTS

Part of "81 - Depression and the Medically Ill "

The evidence for antidepressant use is also limited. Tricyclic antidepressants are known to cause adverse cardiovascular effects, including orthostatic hypotension and slowed intraventricular conduction (66 ,67); therefore, it would not be prudent to use these agents in a population at risk.

Pilot studies suggest that the selective serotonin reuptake inhibitors are safe and effective in persons with ischemic heart disease and depression (68 ,69). The only study of an antidepressant in post-MI patients is that of Shapiro et al. (70). In this study, sertraline (Zoloft) was well tolerated, and no unexpected cardiac effects were noted. Large, randomized, controlled trials are necessary to assess the effects of long-term antidepressant treatment on morbidity and mortality. A multicenter study of sertraline is under way in post-MI patients with major depression (SADHART).

Currently, at least four double-blinded, placebo-controlled studies have examined the efficacy of antidepressant medication in the treatment of post-stroke depression (71 ,72 ,73 and 74). In the first study, reported in 1984, 11 patients given nortriptyline showed a significantly greater improvement on the Hamilton Depression Scale (HAM-D), the Zung Self-Rating Depression Scale, and the profile of depressive symptoms assessed by the Present State Examination than did 14 placebo-treated controls (71). It is worth noting that three of the original 14 patients treated with nortriptyline dropped out of the study. Two patients became delirious, and one had a sudden syncopal episode of unknown cause.

In a controlled study by Reding et al. (72), seven patients with abnormal results on the dexamethasone suppression test and post-stroke depression were treated with trazodone for 5 weeks; these patients showed a significantly greater improvement in activities of daily living as measured by the Barthel Activities of Daily Living Scale than did nine patients with positive Problem Solving Therapies (PSTs) who were treated with placebo.

Andersen et al. (73) assessed the efficacy and tolerability of the selective serotonin reuptake inhibitor citalopram in a controlled study of 66 patients with stroke. HAM-D and Melancholia Scale scores were significantly better after 3 and 6 weeks of treatment in the 33 patients given citalopram (under age 65, 20-mg dose; over age 65, 10-mg dose) than in the 33 patients given placebo.

The most recent of the four studies compared nortriptyline (*n* = 16) with fluoxetine (*n* = 23) and placebo (*n* = 17) (74). About half of the patients had major depression, and the other half had minor depression based on DSM-IV diagnostic criteria elicited by the semistructured Present State Examination. The response rate (defined as a reduction of > 50% in the HAM-D score and no longer meeting criteria for depression) of patients treated with nortriptyline who completed the trial was significantly greater (77%) than the response rate of the patients treated with placebo (31%) or fluoxetine (14%). The HAM-D scores of nortriptyline-treated patients were significantly lower after 12 weeks of treatment than the scores of the patients treated with fluoxetine or placebo, which were not significantly different from each other (Fig. 81.5). The dropout rate of the patients treated with fluoxetine in doses that were increased from 10 to 40 mg during 12 weeks (10-mg increase every 3 weeks) was significantly higher (9 of 23) than those of the other two groups (3 of 16 and 4 of 17).

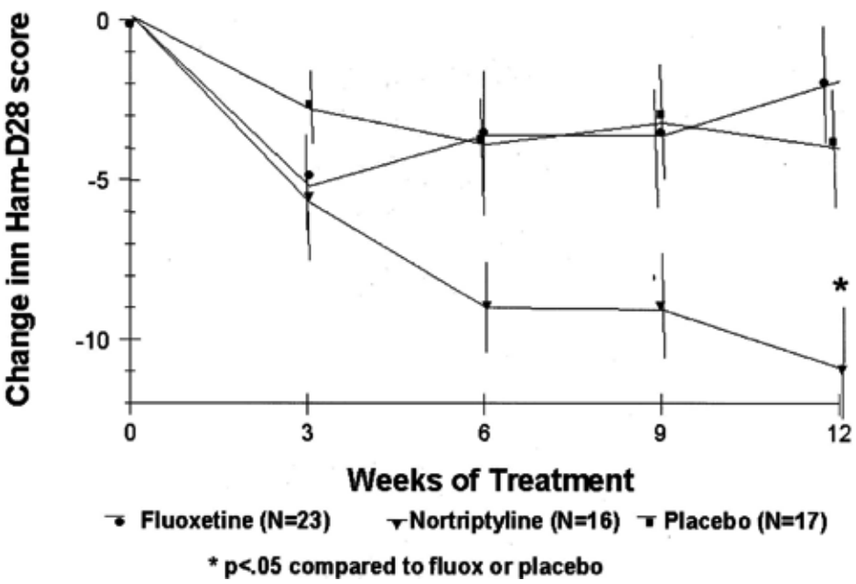


FIGURE 81.5. Change in Hamilton Depression Scale (28 items) scores during 12 weeks of treatment for all patients entered in the study (i.e., intention-to-treat analysis). *Significant group-by-time interaction (F = 3.45, df = 8, 212, p = .0035) and *post hoc* significantly greater change in patients treated with nortriptyline than in those treated with fluoxetine or placebo at 12 weeks.

CONCLUSION

Part of "81 - Depression and the Medically Ill "

Identifying depression in the medically ill population is difficult. Often, the symptoms that are used to identify depression are confused with the underlying symptoms of the medical illness. One approach is to exclude symptoms (e.g., fatigue) that may arise from the medical condition. Another approach is to replace some of the vegetative symptoms, and a third approach is to include all symptoms. The inclusive approach is favored because of its ease of use and practical applicability, and it appears to be sufficient. It is also more sensitive and relates well to functional impairment (75).

In summary, evidence is increasing that depression in patients with vascular disease can be successfully treated with pharmacologic agents. The effects of antidepressant treatment on the risk for exacerbating medical illness and on improving physical, cognitive, and quality-of-life outcomes remain to be determined. The findings of ongoing pharmacologic treatment studies will be viewed with interest.

REFERENCES

1. Kessler RC, McGonagle KA, Zaho S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. *Arch Gen Psychiatry* 1994;51:8–19.

2. Greenberg PE, Stiglin LE, Finkelstein SN, et al. The economic burden of depression in 1990. *J Clin Psychiatry* 1993;54:405–418.

3. Frasure-Smith N, Lesperance F, Talajic M. Depression following myocardial infarction: impact on 6-month survival. JAMA 1993;270:1819–1825.

4. Carney RM, Freedland KE, Jaffe AS. Insomnia and depression prior to myocardial infarction. *Psychosom Med* 1990;52:603–609.

5. Forrester AW, Lipsey JR, Teitelbaum ML, et al. Depression following myocardial infarction. *Int J Psychiatry Med* 1992;22:33–46.

6. Hance M, Carney RM, Freedland KE, et al. Depression in patients with coronary heart disease: a 12-month follow-up. *Gen Hosp Psychiatry* 1996;18:61–65.

7. Ladwig KH, Lehmacher W, Roth R, et al. Factors which provoke post-infarction depression: results from the Post-Infarction Late Potential study (PILP). *J Psychosom Res* 1992;36:723–729.

8. Schleifer SJ, Macari-Hinson MM, Coyle DA, et al. Platelet monoamine oxidase activity in elderly depression outpatients. *Biol Psychiatry* 1986;21:1360–1364.

9. Cassem NH, Hackett TP. Psychiatric consultation in a coronary care unit. *Ann Intern Med* 1971;75:9–14.

10. Cay EL, Vetter N, Philip AE, et al. Psychological status during recovery from an acute heart attack. *J Psychosom Res* 1972;16:425–435.

11. Carney RM, Rich MW, Tevelde AJ, et al. Major depressive disorder in coronary artery disease. *Am J Cardiol* 1987;60:1273–1275.

12. Eastwood MR, Rifat SL, Nobbs H, et al. Mood disorder following cerebrovascular accident. *Br J Psychiatry* 1989;154:195–200.

13. Robinson RG, Starr LB, Kubos KL, et al. A two-year longitudinal study of post-stroke mood disorders: findings during the initial evaluation. *Stroke* 1983;14:736–744.

14. Burvill PW, Johnson GA, Jamrozik KD, et al. Prevalence of depression after stroke: the Perth Community Stroke Study. *Br J Psychiatry* 1995;166:320–327.

15. Robinson RG, Schultz SK, Paradiso S. Treatment of poststroke psychiatric disorders. In: Nelson JC, ed. *Geriatric psychopharmacology*. New York: Marcel Dekker Inc, 1998:161–185.

16. Cohen-Cole SA, Stoudemire A. Major depression and physical illness: special considerations in diagnosis and biologic treatment. *Psychiatr Clin North Am* 1987;10:1–17.

17. Rifkin A, Reardon G, Siris S, et al. Trimipramine in physical illness with depression. *J Clin Psychiatry* 1985;46:4–8.

18. Rapp SR, Vrana S. Substituting nonsomatic for somatic symptoms in the diagnosis of depression in elderly male medical patients. *Am J Psychiatry* 1989;146:1197–1200.

19. Endicott J. Measurement of depression in patients with cancer. *Cancer* 1984;53[Suppl]:2243–2248.

20. Bukberg J, Penman D, Holland JC. Depression in hospitalized cancer patients. *Psychosom Med* 1984;46:199–212.

21. Davidson J, Turnbull CD. Diagnostic significance of vegetative symptoms in depression. *Br J Psychiatry* 1986;148:442–446.

22. Robinson RG. *The clinical neuropsychiatry of stroke*. Cambridge, UK: Cambridge University Press, 1998.

23. Kathol RG, Noyes R, Williams J, et al. Diagnosing depression in patients with medical illness. *Psychosomatics* 1990;31:434–440.

24. Anda R, Williamson D, Jones D, et al. Depressed affect, hopelessness, and the risk of ischemic heart disease in a cohort of US adults. *Epidemiology* 1993;4:285–294.

25. Aromaa A, Raitasalo R, Reunanen A, et al. Depression and cardiovascular diseases. *Acta Psychiatr Scand Suppl* 1994;377:77–82.

26. Barefoot JC, Schroll M. Symptoms of depression, acute myocardial infarction, and total mortality in a community sample. *Circulation* 1996;78:613–617.

27. Everson SA, Goldberg DE, Kaplan GA, et al. Hopelessness and risk of mortality and incidence of myocardial infarction and cancer. *Psychosom Med* 1996;58:113–121.

28. Ford DE, Mead LA, Chang PP, et al. Depression is a risk factor for coronary artery disease in men: the precursors study. *Arch Intern Med* 1998;158:1422–1426.

29. Pratt LA, Ford DE, Crum RM, et al. Depression, psychotropic medication, and risk of myocardial infarction: prospective data from the Baltimore ECA follow-up. *Circulation* 1996;94:3123–3129.

30. Vogt T, Pope C, Mullooly J, et al. Mental health status as a predictor of morbidity and mortality: a 15-year follow-up of members of a health maintenance organization. *Am J Public Health* 1994;84:227–231.

31. Morris PLP, Robinson RG, Raphael B. Prevalence and course of depressive disorders in hospitalized stroke patients. *Int J Psychiatry Med* 1990;20:349–364.

32. Avery D, Winokur G. Mortality in depressed patients treated with electroconvulsive therapy and antidepressants. *Arch Gen Psychiatry* 1976;33:1029–1037.

33. Black DW, Warrack G, Winokur G. Excess mortality among psychiatric patients: the Iowa record-linkage study. *JAMA* 1985;253:58–61.

34. Bruce ML, Leaf PJ, Rozal GPM, et al. Psychiatric status and 9-year mortality data in the New Haven Epidemiological Catchment Area Study. *Am J Psychiatry* 1994;151:716–721.

35. Carney RM, Rich MW, Freedland KE, et al. Major depressive disorder predicts cardiac events in patients with coronary artery disease. *Psychosom Med* 1988;50:627–633.

36. Maltzberg B. Mortality among patients with involutional melancholia. *Am J Psychiatry* 1937;93:1231–1238.

37. Murphy JM, Monson RR, Olivier DC, et al. Affective disorders and mortality: a general population study. *Arch Gen Psychiatry* 1987:473–480.

38. Weeke A, Vaeth M. Excess mortality of bipolar and unipolar manic-depressive patients. *J Affect Disord* 1986;11:227–234.

39. Norton B, Whalley LJ. Mortality of a lithium-treated population. *Br J Psychiatry* 1984;145:277–282.

40. Rabins PV, Harvis K, Koven S. High fatality rates of late-life depression associated with cardiovascular disease. *J Affect Disord* 1985;9:165–167.

41. Sharma R, Markar HR. Mortality in affective disorder. *J Affect Disord* 1994;31:91–96.

42. Tsuang MT, Woolson RF, Fleming JA. Premature deaths in schizophrenia and affective disorders: analysis of survival curves and variables affecting the shortened survival. *Arch Gen Psychiatry* 1980;37:979–983.

43. Vestergaard P, Aagaard J. Five-year mortality in lithium-treated manic-depressive patients. *J Affect Disord* 1991;21:33–38.

44. Ahern DK, Gorkin L, Anderson JL, et al. Biobehavioral variables and mortality or cardiac arrest in the Cardiac Arrhythmia Pilot Study (CAPS). *Am J Cardiol* 1990;66:59–62.

45. Frasure-Smith N, Lesperance F, Talajic M. Depression and 18-month prognosis after myocardial infarction. *Circulation* 1995;91:999–1005.

46. Kennedy GJ, Hofer MA, Cohen D, et al. Significance of depression and cognitive impairment in patients undergoing programmed stimulation of cardiac arrhythmias. *Psychosom Med* 1987;49:410–421.

47. Morris PL, Robinson RG, Samuels J. Depression, introversion and mortality following stroke. *Aust N Z J Psychiatry* 1993;27:443–449.

48. Morris PU, Robinson RG, Raphael B. Emotional lability following stroke. *Aust N Z J* Psychiatry 1993;27:601–605.

49. Astrom M, Adolfsson R, Asplund K. Major depression in stroke patients: a 3-year longitudinal study. *Stroke* 1993;24:976–982.

50. Lloyd GG, Cawley RH. Distress or illness? A study of psychological symptoms after myocardial infarction. *Br J Psychiatry* 1983;142:120–125.

51. Mayou R, Foster A, Williamson B. Psychosocial adjustment in patients 1 year after myocardial infarction. *J Psychosom Res* 1978;22:447–453.

52. Blumenthal JA, Williams RS, Wallace AG, et al. Physiological and psychological variables predict compliance to prescribed exercise therapy in patients recovering from myocardial infarction. *Psychosom Med* 1982;44:519–529.

53. Anda RF, Williamson DF, Escobedo LG, et al. Depression and the dynamics of smoking: a national perspective. *JAMA* 1990;264:1541–1545.

54. Carney RM, Freedland KE, Eisen SA, et al. Major depression and medication adherence in elderly patients with coronary artery disease. *Health Psychol* 1995;14:88–90.

55. Kotilla M, Waltimo O, Niemim L, et al. The profile of recovery from stroke in factors influencing outcome. *Stroke* 1984;15:1039–1044.

56. Sinyor D, Jacques P, Kaloupek DG, et al. Post-stroke depression and lesion location: an attempted replication. *Brain* 1986;109:539–546.

57. Parikh RM, Robinson RG, Lipsey JR, et al. The impact of post-stroke depression on recovery in activities of daily living over two-year follow-up. *Arch Neurol* 1990;47:785–789.

58. Shimoda K, Robinson RG. The relationship between social impairment and recovery from stroke. *Psychiatry* 1998;61:101–111.

59. Bolla-Wilson K, Robinson RG, Starkstein SE, et al. Lateralization of dementia of depression in stroke patients. *Am J Psychiatry* 1989;146:627–634.

60. House A, Dennis M, Warlow C, et al. The relationship between intellectual impairment and mood disorder in the first year after stroke. *Psychol Med* 1990;20:805–814.

61. Andersen G, Vestergaard K, Riis JO, et al. Dementia of depression or depression of dementia in stroke? *Acta Psychiatr Scand* 1996;94:272–278.

62. Kauhanen M, Korpelainen JT, Hiltunen P, et al. Poststroke depression correlates with cognitive impairment and neurological deficits. *Stroke* 1999;30:1875–1880.

63. Downhill JE, Jr, Robinson RG. Longitudinal assessment of depression and cognitive impairment following stroke. *J Nerv Ment Dis* 1994;182:425–431.

64. Starkstein SE, Robinson RG, Price TR. Comparison of patients with and without post-stroke major depression matched for size and location of lesion. *Arch Gen Psychiatry* 1988;45:247–252.

65. Blumenthal JA, O'Connor C, Hinderliter A, et al. Psychosocial factors and coronary disease: a national multicenter clinical trial (ENRICHD) with a North Carolina focus. *N C Med J* 1997;58:440–444.

66. Glassman AH. Cardiovascular effects of antidepressant drugs: updated. *J Clin Psychiatry* 1998;59:13–18.

67. Glassman AH, Roose SP, Bigger JT Jr. The safety of tricyclic antidepressants in cardiac patients: risk-benefit reconsidered. *JAMA* 1993;269:2673–2675.

68. Roose SP, Glassman AH, Attia E, et al. Cardiovascular effects of fluoxetine in depressed patients with heart disease. *Am J Psychiatry* 1998;155:660–665.

69. Roose SP, Laghrissi-Thode F, Kennedy JS, et al. Comparison of paroxetine and nortriptyline in depressed patients with ischemic heart disease. *JAMA* 1998;279:287–291.

70. Shapiro PA, Lesperance F, Frasure-Smith N. An open-label preliminary trial of sertraline for treatment of major depression after acute myocardial infarction (the SADHAT Trial). *Am Heart J* 1999;137:1110–1106.

71. Lipsey JR, Robinson RG, Pearlson GD, et al. Nortriptyline treatment of post-stroke depression: a double-blind study. *Lancet* 1984;1:297–300.

72. Reding MJ, Orto LA, Winter SW, et al. Antidepressant therapy after stroke: a double-blind trial. *Arch Neurol* 1986;43:763–765.

73. Andersen G, Vestergaard K, Riis J. Citalopram for post-stroke pathological crying. *Lancet* 1993;342:837–839.

74. Robinson RG, Schultz SK, Castillo C, et al. Nortriptyline versus fluoxetine in the treatment of depression and in short-term recovery after stroke: a placebo-controlled, double-blind study. *Am J Psychiatry* 2000;157:351–359.

75. Koenig HG, George LK, Peterson BL, et al. Depression in medically ill hospitalized older adults: prevalence, characteristics, and course of symptoms according to six diagnostic schemes. *Am J Psychiatry* 1997;154:1376–1383.

Section IX
Alzheimer and other dementias

Kenneth L. Davis

Alzheimer and other dementias - Introduction

As average life expectancies continue to rise dramatically, the aging population is at risk, now more than ever, for Alzheimer disease and other dementias characterized by impaired memory and other cognitive disabilities. In this section, age-associated alterations in cognition, brain structure, and neurochemistry of stroke, Alzheimer disease, and neuropsychiatric manifestations of HIV-1 and AIDS are addressed.

Mohs and Haroutunian summarize the issues involved in early diagnosis of Alzheimer disease. This is an area that has received intense interest, given the hope that future therapies may some day alter the course of the disease. Clearly, in such a circumstance, making a diagnosis at the earliest possible time becomes critical. Ideally, diagnosis in the premorbid state will someday be possible, a possibility further suggested in Chapter 86 by Small, who reviews positron emission tomography techniques that, when combined with the apolipoprotein E status of patients with potential Alzheimer disease, may be a very powerful tool for exceptionally early diagnosis. Yet another aspect of both these chapters, and one that is drawing increasing attention, is the question of the utility of current diagnostic criteria for Alzheimer disease and such potentially related conditions as mild cognitive impairment. Increasingly, these distinctions are becoming blurred.

Parvathy and Buxbaum provide a detailed overview of the molecular and pathologic changes in Alzheimer disease. The work they review increasingly points toward the centrality of amyloid in the pathophysiology of Alzheimer disease and makes the development of transgenic animals that overexpress ß-amyloid secondary to the insertion of various human mutations that have been linked to Alzheimer disease in humans so crucial. Chapter 85 , by Nixon, is a companion piece to Chapter 83 . Whereas Nixon focuses on cellular events, Parvathy and Buxbaum focus on genetic and molecular issues. Of particular interest in Chapter 85 is the discussion of the possible role of lysosomal enzymes in the cellular damage of Alzheimer disease. Chapter 84 , by Duff, is a concise summary of developments in transgenics, as well as the challenges the field faces in using these models, particularly from the perspective of relative contributions of tau and amyloid to the pathogenesis of Alzheimer disease. As these models continue to improve and to demonstrate their congruence with the Alzheimer phenotype, they will prove a cornerstone drug discovery in Alzheimer disease. In my chapter, Chapter 87 , I try to draw on these advances in cellular and molecular biology to discuss the exciting opportunities for the therapeutics of Alzheimer disease. In addition, I summarize the status of current treatments.

Raskind and Barnes review informative studies of psychopharmacologic management of noncognitive behavioral problems in Alzheimer disease, including depressive signs and symptoms, psychotic symptoms, and disruptive agitated behavior. It is becoming increasingly clear that such behaviors can be even more problematic than the cognitive disturbances. Regrettably, this has been a very difficulty therapeutic area, although one that is now receiving a good deal of attention. The data they review can guide the clinician in making some very difficult choices among a broad spectrum of agents that have been employed to ameliorate the host of behavioral symptoms that patients with Alzheimer disease can display.

Manca, Davies, and Burns discuss the implications of the demographic trend toward an aging population and the economic impact of neuropsychiatric disease. As is appropriate, considerable concern is raised about the magnitude of the economic implications of this disease. Given the need to justify a new therapeutic agent on its cost effectiveness,

this kind of discussion is increasingly becoming a part of drug development.

Since the publication of the *Fourth Generation of Progress*, the recognition of Lewy body dementia has increased substantially. To many clinicians, this is a diagnostic entity that was previously incorrectly diagnosed as either Alzheimer disease with some parkinsonian features or Parkinson disease with dementia. In a wonderfully lucid chapter (Chapter 91), McKeith et al. distinguish dementia with Lewy bodies from Alzheimer disease and provide a detailed account of the clinical features discovered in the 1990s.

An entire new term for a set of diseases, tauopathies, was coined in the last few years, and the group that is largely responsible for characterizing the molecular and cellular pathology of these conditions has contributed a key chapter to this section. In Chapter 94 , Higuchi, Trojanowski, and Lee address tau-positive filamentous lesions in neurodegenerative disease.

One of the most active areas of central nervous system therapeutics has been in developing drugs to decrease the cellular disease that follows stroke. Many drugs have shown promise in what seem valid animal models, but, as the chapter by Small, Morley, and Buchan points out, have not been efficacious in the clinic. Nevertheless, as this chapter details, this is a particularly rich area of experimental therapeutics and one of the best examples of the ways in which fundamental advances in neuroscience can drive rational drug development. This theme is made all the more apparent when Chapter 92 , by Graham and Hickey, is read alongside Chapter 93 , because the former so elegantly summarizes the mechanisms that exacerbate neuronal death resulting from hypoxia and hypoglycemia.

Highly active antiretroviral therapy (HAART) has revolutionized the treatment of AIDS and has had a major impact on the neuropsychiatric manifestations of HIV infection. Evans and Mason address both the neurocognitive functioning and the psychiatric manifestations of HIV-1 infection, as well as its treatment in the HAART era.

Taken together, the chapters in this section are an impressive compendium of advances in understanding and treating some of the worst diseases faced by humans. That so much has been learned in so short a time is truly remarkable, but even more remarkable is the advances that will undoubtedly occur in the future, advances whose foundations are eloquently elaborated in the following pages.

82

Alzheimer Disease: From Earliest Symptoms to End Stage

Richard C. Mohs

Vahram Haroutunian

Richard C. Mohs and Vahram Haroutunian: Research Service, Mount Sinai School of Medicine, Veterans Affairs Medical Center, Bronx, New York.

INTRODUCTION TO THE NATURAL HISTORY

Part of "82 - Alzheimer Disease: From Earliest Symptoms to End Stage "

Alzheimer disease (AD) is a progressive, degenerative brain disease that is the most common cause of dementia in elderly persons. Clinically, patients with AD have impairments in memory, language, praxis, and other cognitive functions that develop very gradually but progress relentlessly. Longitudinal studies leading to autopsy have shown that the most common neuropathologic findings in elderly patients with these symptoms are neuritic plaques (NPs) and neurofibrillary tangles (NFTs). Modern diagnostic criteria for AD recognize that AD is both a clinical and pathologic entity. By definition, all patients with AD must have dementia, a progressive loss of memory, and at least one other cognitive function that is sufficiently severe to interfere with daily functioning. To differentiate AD from other acquired neuropsychiatric conditions associated with cognitive impairment, the clinical diagnosis of AD is only made when no other conditions could account for the patient's progressive cognitive impairment. Patients who meet the clinical criteria for AD are very likely to have the characteristic neuropathologic features as well. Sometimes, however, the clinical diagnosis is not confirmed at autopsy, and hence the most widely used criteria for AD reserve the term *definite AD* for those patients in whom both clinical and neuropathologic data support the diagnosis of AD.

Although progressive cognitive impairment is the core or defining characteristic of AD clinically, patients with this disease have other symptoms as well. Many patients also have other neuropsychiatric symptoms including agitation, psychosis, depressed mood, and personality change. These other symptoms are not necessary for a diagnosis of AD and tend to be quite variable both within a given patient and from one patient to another. When they are present, these symptoms can be a major problem for caregivers of AD patients, and behavioral problems have been linked to an increased need for health services including nursing home care.

The first definite symptoms of AD are often quite mild and are difficult to differentiate from the mild memory loss that is a frequent consequence of normal or usual aging. Inevitably, however, the degenerative changes of AD become sufficiently severe so the patient has difficulty with daily functioning. The functional change can be observed first in the performance of cognitively demanding tasks such as handling money, remembering appointments, following directions, and using appliances. As the disease progresses and the patient's cognitive abilities deteriorate, the patient has difficulty in more functional domains including the basic activities of daily living such as feeding, toileting, dressing, and personal hygiene. In the later stages of AD, patients are often unable to remember even very simple things, have great difficulty talking and understanding language, and may be confined to bed or to a chair. The average life expectancy of a patient with AD after the initial diagnosis is approximately 10 years, but with a great deal of variability around that mean.

In the sections that follow, we review studies of the development of the neurobiological changes responsible for AD. Later, we briefly review the epidemiology of AD and review in greater detail the development of the disease clinically. In each section, we emphasize the need to understand AD from a longitudinal perspective because both the underlying neurobiology and clinical presentation of the illness vary substantially across the course of illness. Because of intense recent interest in understanding the very early development of AD to develop preventive therapies, our presentation emphasizes recent findings on the earliest manifestations of disease.

NEUROBIOLOGICAL STUDIES ACROSS THE SEVERITY SPECTRUM

Part of "82 - Alzheimer Disease: From Earliest Symptoms to End Stage "

It has already been noted that the definitive diagnosis of AD depends on neuropathologic changes that characterize

the disease. No single neuropathologic lesion is in itself adequate for the diagnosis of AD; rather, the neuropathologic diagnosis of AD is based on the presence of multiple AD-related lesions, the density of these lesions relative to the age of the subject, and the absence of lesions characteristic of other neuropathologic diseases. The absolute weight of the brain is decreased relative to normal controls, but this decrease is generally less than 10% relative to age-matched controls, and it is neither diagnostic of nor specific to AD (1). Gross examination of the brain in AD also reveals significant apparent atrophy, widening of the sulci, and erosion of the gyri, but these changes also reflect advanced age with significant overlap between AD and normal elderly controls. However, the atrophy of the cortex is associated with significant reductions in the numbers of neurons (2 ,3). For example, Terry et al. reported 40% to 46% losses of large neurons in the frontal and temporal cortices of specimens derived from patients with AD (4). Similarly, Gomez-Isla and colleagues (5 ,6), using unbiased stereologic sampling techniques, reported approximately 50% losses in neurons of the superior temporal sulcus with even more pronounced losses in specific cortical laminae. These neuronal losses were observed not only in brain specimens from patients with severe dementia, but also in specimens derived from patients with relatively mild or questionable dementia. The magnitude of neuronal loss increases systematically with increasing dementia severity and increasing disease duration. Neuronal degeneration is not restricted to the cortex, but it is also reflected in neuronal losses in subcortical nuclei such as the nucleus basalis of Meynert (7) (the cells of origin of the cholinergic input to the cerebral cortex), the locus ceruleus, and raphe aminergic nuclei (8 ,9). Neuronal loss in these subcortical structures, especially in the nucleus basalis of Meynert (10), has also been found to correlate significantly with dementia severity and cognitive deficits.

Neuronal loss and degeneration are accompanied by significant decreases in markers of synaptic density. Although synaptic markers such as synaptophysin are reduced significantly in the cerebral cortex, especially the frontal and parietal cortices and in the hippocampus, with increasing age (11), further losses are encountered in AD, whether assessed by immunohistochemical techniques or by direct assessment of synaptic specializations and profiles (3 ,12 ,13 and 14). The loss of synaptophysin immunoreactivity in the frontal and parietal cortices, and in the hippocampus, is among the strongest correlates of dementia severity (10 ,15 ,16). These losses and correlations with cognitive function are not only evident at the immunohistochemical level, but they have also been observed with quantitative enzyme-linked immunosorbent assay techniques (10). This loss of synaptic markers is not merely a reflection of the degeneration of the cortical neurons noted earlier, but it also reflects the loss of presynaptic terminals and neuropeptide- and neurotransmitter-containing vesicles.

Various specific lesions have been found to be associated with AD. The most prominent are NPs and NFTs. However, these lesions are not exclusively associated with AD. Age-related accumulations of NPs and diffuse plaques have been noted in elderly persons who are otherwise normal. Similarly, NFTs have been found in the brains of nondemented elderly persons and in association with non–AD-like neurodegenerative diseases (17). Although NP and NFT lesions can be present in diseases other than AD, other markers and clinical phenotypes can be used to distinguish among them, and the presence of NPs and NFTs in the absence of other confounding neuropathologic lesions provides the basis for the diagnosis of AD.

NPs are extracellular deposits of varying sizes with an amyloid ß-peptide core (Aß) and neuritic inclusions. Aß is a 40 to 43 amino acid long peptide that is generated from a larger peptide (Alzheimer amyloid precursor protein or APP) by two cleavage events (18). The cleavage mechanisms that lead to the production of Aß from APP are under intense investigation. ß-Amyloid cleavage enzyme was isolated and cloned and proposed as the enzyme responsible for cleavage at the N-terminus (19 ,20 and 21). Cleavage at the C-terminus is attributed to an as yet unidentified enzyme termed γ-secretase. Some evidence suggests that presenilin 1 may be the γ-secretase (22), but this hypothesis is still under investigation. That Aß deposition plays a critical role in the pathogenesis of AD was recognized with the accumulation of evidence showing that mutations in the APP gene, as well as mutations in the gene encoding for presenilin 1 and 2, were invariably associated with AD (23 ,24). Studies in transgenic mice demonstrated that the introduction of these mutations leads to the development deposition of Aß plaques and learning and memory deficits in some mutants (25 ,26 and 27).

Despite the clear evidence implicating NP deposition in the pathogenesis of AD, few studies have addressed the relationship of NP deposition with the symptoms of AD (dementia) during the early phases of the disease. Ascertainment of the relationship between specific pathologic lesions and symptoms of AD has been difficult, because most studies have focused on the neuropathology of AD at the terminal stages of the disease, when dementia has been fully developed and neuropathologic lesions have been profuse. Studies have suggested that increases in the densities of neocortical NPs occur very early during the course of cognitive deterioration (6 ,10 ,28 ,29 ,30 ,31 and 32), and they may be among the initial pathologic events in the development of AD (31). In some of these studies, brain specimens were grouped according to the severity of dementia before death according to the Clinical Dementia Rating (CDR) scale. The density of NPs and Aß

immunoreactivity were then quantified in different brain regions. These studies showed that increases in NP density and quantitatively measured Aß immunoreactivity are evident even in those patients who die at the earliest stages of dementia, when dementia severity is very mild or even questionable. The density of NPs and Aß immunoreactivity then increase systematically as a function of increasing dementia severity. In one study (29), elevated levels of Aß-42 were detected in multiple neocortical regions before NFTs and significant immunoreactivity to abnormal tau (see later) could be demonstrated in the same cortical regions. Increases in NP density and Aß immunoreactivity were observed in cases of mild dementia before the density of neuropathologic lesions was high enough for the patients to meet the threshold criteria for the definitive diagnosis of AD.

NFTs constitute the second hallmark of AD neuropathology. Immunohistochemical and biochemical studies have shown that NFTs consist of paired helical filaments that are abnormal aggregates of abnormally folded (33 ,34) or phosphorylated (35 ,36) forms of the microtubule-associated protein tau. The progressive involvement and distribution of NFTs to different brain regions have been used to stage the neuropathologic severity of AD (37 ,38). These studies have suggested that the first signs of NFT are found within the entorhinal cortex, followed by the hippocampus, and the eventual involvement of virtually all regions of the isocortex. There is also clear evidence that the density and severity of NFTs increase as a function of increasing disease duration (5 ,6). Thus, both the density of NFTs in any given brain region and the regions of the brain affected increase with increasing disease duration. In a study identical to that described earlier, study subjects were grouped on the basis of the severity of dementia before death, and the density of NFT-bearing neurons in different brain regions was quantified as a function of dementia severity (39). The density of NFTs in all brain regions increased as a function of increasing dementia severity. However, moderate NFT involvement was documented in the entorhinal cortex of elderly patients with no clinical evidence of dementia (see also ref. 40). Neocortical NFTs were abundant only in patients with moderate to severe dementia, and NFT density increased as a function of increasing dementia severity. Although neocortical NFTs were present in patients with moderate dementia, NFT abundance was low or absent in patients with mild or questionable dementia. Similar findings have been reported in other studies (10 ,11 ,40). These studies suggest that NFTs are most abundant in the entorhinal cortex, where they can be observed in nondemented elderly subjects as well as in patients with AD. NFTs involve neocortical structures later in the course of the AD and are associated with significant dementia. As dementia severity increases, so does the density of neocortical NFTs. This correlation of NFT density with dementia severity is not restricted to the neocortex and to the hippocampus, but it also applies to the subcortical nuclei, such as the forebrain cholinergic nucleus basalis of Meynert (15). Thus, NFTs are a significant neuropathologic feature of AD and contribute to the progression of dementia.

In addition to the neuropathologic lesions associated with AD, significant deficits in neurochemical functions and indices have been observed (41). Chief among these neurochemical deficits are deficits in neocortical indices of cholinergic function and decreases in the concentrations of several neuropeptides such as somatostatin and corticotropin-releasing hormone (42). Deficits in several other neurochemicals and neurotransmitters such as norepinephrine and serotonin have also been reported, but their alterations are not as profound and do not appear to be as consistently observed (41 ,43 ,44). Deficits in the activity of cholinergic marker enzymes (choline acetyltransferase and acetylcholinesterase) were among the first to be reported in AD (45 ,46). Deficits in cortical cholinergic marker enzymes have been among the most consistently replicated neurochemical findings in AD. Some studies have reported that compensatory mechanisms interact with cholinergic enzyme deficits and lead to an up-regulation of high-affinity choline transport (47). Irrespective of compensatory mechanisms that may be engaged, the loss of cortical cholinergic enzyme activity is associated with severe degeneration of cholinergic neurons in the basal forebrain including the neocortically projecting neurons of the nucleus basalis of Meynert (7). The discovery of these profound forebrain cholinergic system deficits and the growing understanding of the role of the forebrain cholinergic system in learning and memory (48) were pivotal to the development of the current therapeutic strategies in AD (49), which focus on the restoration of these cholinergic deficits by inhibiting the activity of the acetylcholine catabolic enzyme, acetylcholinesterase.

As with neuropathologic studies of AD, most postmortem studies assessing cholinergic markers in AD were derived from patients with end-stage dementia. Those few studies in which brain biopsies were obtained and cholinergic markers were assessed ante mortem were generally restricted to patients who have a very early onset of dementia or to patients with relatively advanced dementia (50 ,51 and 52). Thus, although deficits in the activity of cholinergic marker enzymes have been shown to correlate significantly with dementia severity (43 ,53), the question whether these profound deficits in cholinergic markers found in patients with end-stage dementia extend to patients with much earlier disease remained unanswered until recently. Using the same strategy as that described for studying neuropathologic changes in early dementia, Davis and colleagues assigned patients to groups on the basis of their cognitive status at the time of death according to the CDR scale (54). After stratification of subjects to different dementia conditions, the activity of cholinergic marker enzymes was assessed in multiple neocortical regions that encompassed representative regions within the frontal, temporal, parietal, and occipital cortices. Cholinergic marker enzyme activity was profoundly diminished in patients with end-stage, severe dementia, but neither the activity of acetylcholinesterase nor the activity of choline acetyltransferase was reduced in subjects with mild and moderate dementia. Although it can be argued that adaptive changes compensate for cholinergic

deficits in mild dementia (47), the most parsimonious interpretation of these results is that cholinergic deficits are characteristic of relatively advanced dementia and contribute relatively less to the early phases of cognitive impairment in AD.

Deficits in selected neuropeptides have also been consistently reported in AD (41). The levels of somatostatin (SLI) and corticotropin-releasing factor (CRF) are the most consistently affected (55 ,56 and 57). Deficits in these neuropeptides are often found to be as profound as those observed for the cholinergic marker enzymes and are specific in that not all neuropeptides are diminished in AD cortex (58 ,59). The CRF deficits are accompanied by the up-regulation of CRF receptors (60), whereas SLI receptors density is either unchanged or down-regulated (56). Evidence of the relationship of the concentration of these neuropeptides to the severity of dementia has been sparse. Some insight into CRF concentrations in earlier stages of AD has been gained from negative correlations between CRF levels and duration of illness (61), as well as from studies of cerebrospinal fluid (62). Correlations between the severity of Alzheimer dementia and cerebrospinal fluid CRF have been found, suggesting that the CRF deficiency may be a relatively early marker of AD. This relationship has not been observed consistently (63 ,64), however. Postmortem studies of CRF and SLI concentrations in the cortices of subjects stratified to groups on the basis of their cognitive status at the time of death have suggested that although the concentrations of both neuropeptides are significantly and severely diminished in patients with severe or terminal dementia, only the levels of CRF are significantly altered in patients with mild to moderate dementia (65).

In the past few years, many epidemiologic studies have addressed the possible protective effect of antiinflammatory drug use with regard to AD (66 ,67). At a molecular level, it is apparent that an inflammatory response accompanies the neuropathologic features of AD (66 ,67 ,68 and 69). There is clear evidence of an acute-phase response with up-regulation of inflammatory cytokines such as interleukin 1 (IL-1) and IL-6 and tumor necrosis factor-α, accompanied by an increase in acute-phase proteins such as α1-antichymotrypsin and α_2-macroglobulin (66). The complement system is active in the AD brain (70 , 71), with generation of the lytic membrane attack complex and presumably with release of anaphylatoxins. Up-regulation of cyclooxygenase 2 in AD neurons (69) suggests that inflammatory lipids may also be involved in the pathogenesis of the disease. It has been hypothesized that inflammatory responses can be autotoxic to neurons and may exacerbate the fundamental pathology of AD (72). The epidemiologic studies with antiinflammatory agents and a prospective study suggesting some slowing of disease progression after indomethacin treatment support a role for inflammatory processes in AD progression (66). Although studies seeking direct evidence of the role of inflammatory processes in the progression of AD and dementia have been initiated only recently (73), one study has examined cytokine gene expression during AD progression. In this study, cytokine gene expression (IL-6 mRNA) in the hippocampus was found to increase as dementia severity progressed from moderate (CDR 2) to severe (CDR 5). Neither the epidemiologic studies nor the neurobiological studies directly address the cause and effect relationship among AD, dementia progression, and inflammatory responses within the brain. These studies do suggest, however, that even if inflammatory responses are not a critical feature of the etiology of AD, they may nevertheless play an important role in mediating the development and progression of dementia.

The results of the studies summarized earlier provide only a very general and global review of the tens of thousands of published reports relevant to the pathogenesis of AD. The results of the more recent studies, especially those that relate to the progression of the disease and dementia, have shed some new light on the pathophysiologic mechanisms involved at the onset of dementia and its progression during the disease process. These relatively recent findings have suggested that, in contrast to some earlier views, the deposition of amyloid plaques is integral to the onset and progression of dementia, and, at least in some brain regions such as the cerebral cortex, they may precede the involvement of some of the other prominent deficits and lesions (e.g., NFTs and cholinergic and neuropeptidergic deficits) characteristic of later stages of the disease. These results have also emphasized that many different lesions contribute to AD neuropathology, and each lesion (NP, NFT, neuronal loss, synaptic loss, cholinergic deficit, neuropeptide deficit, inflammatory response, and countless others) contributes significantly to the dementia symptoms of AD. Conversely, these studies have shown that AD is not characterized by random or general neural system failures, but rather that the pathologic features of AD appear to follow a course of progressive involvement of different neuronal systems, the characteristics of which are only now beginning to be elucidated.

CLINICAL AND NEUROPSYCHOLOGIC STUDIES

Part of "82 - Alzheimer Disease: From Earliest Symptoms to End Stage "

Epidemiologic Studies of Persons at Risk

Precise estimates of the proportion of dementia cases that are attributable to AD are difficult to obtain because few population-based studies obtain autopsy data that would enable a definitive diagnosis of AD. In most large-scale autopsy series, AD lesions are the primary neuropathologic finding in more than 50% of all dementia cases (74). Studies using clinical criteria also find that AD accounts for more than 50% of all dementias, with mixed AD plus vascular dementia and AD plus parkinsonism accounting for significant proportions of the remaining cases (75). Pure vascular

dementia and Lewy body dementia are also found with some regularity, but they are both far less common than AD. Because of the high prevalence of AD among dementia cases, the epidemiology of dementia in old age is largely the epidemiology of AD. Because AD is so much more common than other types of old-age dementia, some clinical guidelines have argued that AD should be treated as a diagnosis of inclusion rather than one of exclusion (76); that is, an older person with dementia should be diagnosed with AD unless there is substantial clinical evidence supporting another cause of the dementia.

The prevalence of AD rises dramatically with age, and age is the most potent risk factor for AD. Less than 1% of new cases of AD are found in persons younger than age 65 years (77), and the prevalence of AD rises steadily after that. By age 90, approximately 35% of persons remaining alive will have AD (77). Men and women are equally vulnerable to AD, but because women live longer than men on average, there are more women than men with AD. Studies looking at different ethnic and cultural groups have found that AD is common in elderly persons from all ethnic and socioeconomic backgrounds, but there may be some Asian ethnic groups who are less vulnerable to AD (78). Environmental risk factors for AD have been difficult to identify, but there is some evidence that persons with higher educational attainment are less likely to develop AD in old age (79). Neurobiological mechanisms that may account for the protective effect of education have not been elucidated, but it is possible that persons with more education have a greater reserve of brain capacity that enables those persons to remain cognitively intact for longer periods of time during the early stages of AD.

Certain genetic factors have been identified that contribute to the development of AD. Specific genetic mutations that cause AD have been identified in the gene coding for the amyloid precursor protein, in the presenilin 1 gene, and in the presenilin 2 gene (24 ,80). Persons who inherit one of these mutations develop AD when they are quite young, often as early as age 40 to 50 years. In families carrying one of these mutations, the inheritance of AD follows the classic pattern of autosomal dominant inheritance, with 50% of each generation developing the disease. Investigations of these mutations are very important because of the information they provide about possible pathophysiologic mechanisms leading to the development of plaques, tangles, cell loss, and dementia. From a population standpoint, however, these genetically determined cases of AD are of less interest because they constitute a small fraction of all cases observed clinically. Most estimates are that less than 2% of all AD cases result from specific genetic mutations (80).

Family (81) and population (82) studies have demonstrated that persons who carry the ε4 form of the apolipoprotein E (Apo E) gene *(APOE)* have a greater likelihood of developing AD than do persons who carry only the ε3 and the ε2 forms. Apo E is a cholesterol-transporting protein that is coded by a gene on chromosome 14. The gene has three allelic forms, ε3, which is by far the most common, and two rarer forms, ε2 and ε4. Persons carrying the ε4 form are at increased risk of developing AD, particularly between the ages of 65 and 75 years. The mechanism by which *APOE* genotype influences the risk of AD is currently under investigation. From a clinical standpoint, *APOE* can be useful for identifying persons at increased risk of developing AD. It is not useful in routine diagnostic evaluations, however, because many patients who develop AD do not carry the ε4 allele, and some who do carry the high-risk form of *APOE* do not develop AD (82). There is extensive research to identify other common genes that influence the likelihood of developing AD, but none have been identified that consistently associate with the risk of AD as *APOE* genotype does.

Predictive Neuropsychological Deficits

AD is a progressive disease with insidious onset in which the underlying neurodegenerative changes probably begin years before clinical symptoms are obvious. Studies of populations at risk of developing AD have been conducted to determine whether there are changes in cognitive function that can be detected with neuropsychological tests before patients meet clinical criteria for the diagnosis of AD. For these studies, persons who are cognitively normal but who are at increased risk of developing AD, usually because of old age, are followed longitudinally with a structured battery of neuropsychological tests. After a period of 1 to 5 years, the baseline performance of patients who have subsequently been diagnosed with AD is compared with the remainder of the population that has remained free of dementia. Several studies using this model have demonstrated consistently that impairment in memory is significantly worse at baseline in those persons who subsequently are diagnosed with AD (83 ,84 and 85). In most instances, the memory tests most impaired before diagnosis are those measuring delayed recall, that is, recall of newly learned information but after a delay of several minutes during which the subject must perform other cognitive tasks. A deficit in the rate of new learning for verbal material (e.g., a list of words) has also be found to predict subsequent dementia in some studies (85). Language function, particularly difficulty with naming, has also been found to differentiate those persons who subsequently develop dementia from others who remain free of dementia (83). Occasionally, other cognitive tasks such as those placing great demands on executive function and working memory show deficits before the onset of dementia, but memory impairment is uniformly the most pronounced deficit (84).

Evidence indicates that some of the predictive power of poor performance on neuropsychological tests results from the fact that memory deficits are, in part, a subclinical surrogate identifying those at increased risk because of old age or presence of an *APOE*ε4 genotype. Because studies have

clarified that *APOE* genotype may confer additional risk of AD primarily within a certain age range (86), it is likely that *APOE* genotype and neuropsychological test performance are independent predictors of dementia in most instances. Analyses of data from these data on neuropsychological antecedents of dementia have consistently shown that the memory and other deficits cannot be accounted for simply by considering age as a predictor. Rather, it appears that deficits in memory and, to a lesser extent, language and executive function are predictors of subsequent dementia across a broad range of ages and for all *APOE* genotypes.

Longitudinal Studies

Numerous longitudinal studies have documented the progression of cognitive, behavioral, and functional changes throughout the course of illness. As expected, given the studies of populations at risk for AD described earlier, studies of very mild AD have documented that memory impairment is the earliest and most prominent feature of the illness (87). As a consequence, memory measures, particularly those employing a measure of delayed recall memory, are now frequently used to identify persons thought to be in the very earliest stages of AD or who may be at high risk of developing AD. As the disease progresses, deficits in both expressive and receptive language and deficits in praxis and visuospatial ability become quite pronounced. Longitudinal studies have also documented that cognitive deterioration in AD is relentlessly progressive, with little evidence of improvement (88).

Some standard assessment tools have been developed to measure the cognitive deficits in AD in a semiquantitative fashion. Among the most commonly used assessment tools are the Mini-Mental State Examination (89), the Blessed Test of Information, Memory, and Concentration (90), and the Alzheimer's Disease Assessment Scale (91). Each of these instruments includes brief tests to assess dysfunction in cognitive domains typically impaired in AD, particularly memory, language, orientation, and praxis. The Mini-Mental State Examination and the Blessed test are quite brief and are often used as screening instruments in research and clinical practice. The Alzheimer's Disease Assessment Scale was developed as a tool for use in clinical trials and is almost always used as one of the primary efficacy measures in clinical trials of antidementia drugs (49 ,92). Longitudinal studies with each of these instruments have been performed. Those studies demonstrate that the measured rate of cognitive decline in AD is quite consistent from study to study and across different populations (88 ,93). In addition, the rate of cognitive decline in AD is curvilinear with time, such that deterioration is quite slow at the start of the illness, is faster during the middle years of illness, and is again slow when patients reach the near terminal phase of the illness. This relationship of rate of deterioration with stage of illness has important implications for clinical trials of agents that are expected to slow the rate of cognitive deterioration (88 ,94).

Factors that may be associated with differences in the rate of cognitive deterioration have been investigated extensively. Apart from the relationship of rate with stage of disease described earlier, no other factors have been found to affect the rate of deterioration consistently. Age, age of disease onset, gender, ethnicity, and *APOE* genotype have all been examined as possible predictors, and none has consistently been shown to affect the rate of decline. Once patients develop the disease, cognitive function deteriorates relentlessly and at approximately the same rate regardless of these variables (88 ,93 ,94).

Behavioral disturbances have also been investigated longitudinally, and it is clear that symptoms such as psychosis, agitation, and depressed mood can be very disturbing both to the patient and to caregivers. Because of the importance of these symptoms in patient management, new tools have been developed in an effort to provide reliable and valid assessment of their severity. Commonly used tools include the Neuropsychiatric Inventory (95), the BEHAVE-AD (96), and the noncognitive subscale of the Alzheimer's Disease Assessment Scale (91). In contrast to the cognitive deficits of AD, however, these behavioral disturbances are quite variable from one patient to another and over time in individual patients (95 ,97). These disturbances are episodic phenomena that wax and wane over the course of AD, with little evidence of progression. Most trials of potential new treatments for AD now include some assessment of these symptoms, and the overall effectiveness of treatments for AD is at least partly determined by the extent to which they improve behavioral symptoms.

By definition, all patients with AD have some impairment in their ability to perform daily activities (98). The assessment of functional ability in patients with dementia includes both an assessment of the basic activities of daily living such as feeding, toileting, dressing, and grooming and an assessment of more cognitively demanding, instrumental activities of daily living such as handling money, using the telephone, performing household chores, and using appliances (99 ,100). The definition of basic activities of daily living is quite consistent from study to study, but there is much less consensus on the kinds of activities that must be surveyed in any assessment of instrumental activities of daily living. Longitudinal data are consistent, however, in demonstrating that impairments in instrumental activities of daily living appear very early in the course of AD, whereas impairments in basic activities of daily often do not appear until patients are quite cognitively impaired (100 ,101). Thus, any comprehensive assessment of functional status in AD must include both basic and instrumental activities of daily living.

The longitudinal progression of functional impairment is relentless, and functional abilities, once lost, are rarely

regained (100 ,101). In this regard, functional impairment follows a progression similar to that of cognitive decline. Correlational studies invariably find a close relationship of functional impairment with the degree of cognitive impairment (101). There is some evidence that the severity of behavioral symptoms, including psychosis and agitation, is associated with some excess disability over that resulting from cognitive impairment, but the contribution of behavioral pathology to functional impairment is relatively small (102). Overall, the trajectory of functional impairment in AD follows the progressive downhill course determined by cognitive loss.

CONCLUDING COMMENT

Part of "82 - Alzheimer Disease: From Earliest Symptoms to End Stage "

Development of effective treatments for AD will be facilitated by a detailed understanding of the neurobiological mechanisms underlying the disease at each of its stages. Testing of treatments for AD requires an understanding of the clinical manifestations and consequences of the disease at each of its stages. Studies investigating neurobiological and clinical changes of the disease over its entire course are providing the tools necessary to develop effective interventions for the prevention and treatment of AD.

DISCLAIMER

Part of "82 - Alzheimer Disease: From Earliest Symptoms to End Stage "

Dr. Mohs serves as a research consultant to the following companies: Pfizer, Eli Lilly, Janssen, Orth-Biotech, and Forrest.

REFERENCES

1. Terry RD. Structural changes in senile dementia of the Alzheimer type. In: Amaducci L, Davison AN, Antuono P, eds. *Aging of the brain and dementia*. New York: Raven, 1980:23–32.

2. Ball M. Neuronal loss, neurofibrillary tangles and granulovacuolar degeneration in the hippocampus with ageing and dementia. *Acta Neuropathol (Berl)* 1977;37:111–118.

3. Masliah E, Terry R, Alford M, et al. Cortical and subcortical patterns of synaptophysin-like immunoreactivity in Alzheimer's disease. *Am J Pathol* 1991;138:235–246.

4. Terry RD, Peck A, DeTeresa R, et al. Some morphometric aspects of the brain in senile dementia of the Alzheimer type. *Ann Neurol* 1981;10:184–192.

5. Gomez-Isla T, Hollister R, West H, et al. Neuronal loss correlates with but exceeds neurofibrillary tangles in Alzheimer's disease. *Ann Neurol* 1997;41:17–24.

6. Gomez-Isla T, Price JL, McKeel DW Jr, et al. Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. *J Neurosci* 1996;16:4491–4500.

7. Whitehouse PJ, Price DL, Struble RG, et al. Alzheimer's disease and senile dementia: loss of neurons in the basal forebrain. *Science* 1982;215:1237–1239.

8. Chan-Palay V, Asan E. Alterations in catecholamine neurons of the locus coeruleus in senile dementia of the Alzheimer type and in Parkinson's disease with and without dementia and depression. *J Comp Neurol* 1989;287:373–392.

9. Zweig RM, Ross CA, Hedreen JC, et al. The neuropathology of aminergic nuclei in Alzheimer's disease. *Ann Neurol* 1988;24:233–242.

10. Dickson DW, Crystal HA, Bevona C, et al. Correlations of synaptic and pathological markers with cognition of the elderly. *Neurobiol Aging* 1995;16:285–298.

11. Nagy Z, Esiri MM, Jobst KA, et al. Relative roles of plaques and tangles in the dementia of Alzheimer's disease: correlations using three sets of neuropathological criteria. *Dementia* 1995;6:21–31.

12. DeKosky ST, Scheff SW, Styren SD. Structural correlates of cognition in dementia: quantification and assessment of synapse change. *Neurodegeneration* 1996;5:417–421.

13. Lassmann H, Fischer P, Jellinger K. Synaptic pathology of Alzheimer's disease. *Ann NY Acad Sci* 1993;695:59–64.

14. Scheff SW, DeKosky ST, Price DA. Quantitative assessment of cortical synaptic density in Alzheimer's disease. *Neurobiol Aging* 1990;11:29–37.

15. Samuel W, Terry RD, DeTeresa R, et al. Clinical correlates of cortical and nucleus basalis pathology in Alzheimer dementia. *Arch Neurol* 1994;51:772–778.

16. Masliah E, Miller A, Terry RD. The synaptic organization of the neocortex in Alzheimer's disease. *Med Hypotheses* 1993;41:334–340.

17. Wisniewski K, Jervis GA, Moretz RC, et al. Alzheimer neurofibrillary tangles in diseases other than senile and presenile dementia. *Ann Neurol* 1979;5:288–294.

18. Selkoe DJ. Normal and abnormal biology of the B-amyloid precursor protein. *Annu Rev Neurosci* 1994;17:489–517.

19. Vassar R, Bennett BD, Babu-Khan S, et al. ß-Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* 1999;286:735–741.

20. Sinha S, Anderson JP, Barbour R, et al. Purification and cloning of amyloid precursor protein beta-secretase from human brain. *Nature* 1999;402:537–540.

21. Yan R, Bienkowski MJ, Shuck ME, et al. Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity. *Nature* 1999;402:533–537.

22. Wolfe MS, Xia W, Ostaszewski BL, et al. Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity. *Nature* 1999;398:513–517.

23. Hutton M, Perez-tur J, Hardy J. Genetics of Alzheimer's disease. *Essays Biochem* 1998;33:117–131.

24. Selkoe DJ. Alzheimer's disease: genotypes, phenotype, and treatment. *Science* 1997;275:630–631.

25. Price DL, Sisodia SS. Mutant genes in familial Alzheimer's disease and transgenic models. *Annu Rev Neurosci* 1998;21:479–505.

26. Erlinge D, Edvinsson L, Brunkwall J, et al. Human neuropeptide Y Y1 receptor antisense oligodeoxynucleotide specifically inhibits neuropeptide Y-evoked vasoconstriction. *Eur J Pharmacol* 1993;240:77–80.

27. Hsiao K, Chapman P, Nilsen S, et al. Correlative memory deficits, Aß elevation, and amyloid plaques in transgenic mice. *Science* 1996;274:99–102.

28. Haroutunian V, Perl DP, Purohit DP, et al. Regional distribution of neuritic plaques in nondemented elderly and cases of very mild Alzheimer's disease. *Arch Neurol* 1998;55:1185–1191.

29. Näslund J, Haroutunian V, Mohs R, et al. Elevated amyloid ß-peptides in brain: correlation with cognitive decline. *JAMA* 2000;283:1571–1577.

30. Berg L, McKeel DW Jr, Miller JP, et al. Clinicopathologic studies in cognitively healthy aging and Alzheimer's disease: relation of histologic markers to dementia severity, age, sex, and apolipoprotein E genotype. *Arch Neurol* 1998;55:326–335.

31. Morris JC, Storandt M, McKeel DW Jr, et al. Cerebral amyloid deposition and diffuse plaques in "normal" aging: evidence for presymptomatic and very mild Alzheimer's disease. *Neurology* 1996;46:707–719.

32. Morris JC, McKeel DW Jr, Storandt M, et al. Very mild Alzheimer's disease: informant-based clinical, psychometric, and pathologic distinction from normal aging. *Neurology* 1991;41:469–478.

33. Jicha GA, Berenfeld B, Davies P. Sequence requirements for formation of conformational variants of tau similar to those found in Alzheimer's disease. *J Neurosci Res* 1999;55:713–723.

34. Jicha GA, Lane E, Vincent I, et al. A conformation- and phosphorylation-dependent antibody recognizing the paired helical filaments of Alzheimer's disease. *J Neurochem* 1997;69:2087–2095.

35. Holzer M, Holzapfel HP, Zedlick D, et al. Abnormally phosphorylated tau protein in Alzheimer's disease: heterogeneity of individual regional distribution and relationship to clinical severity. *Neuroscience* 1994;63:499–516.

36. Mandelkow EM, Mandelkow E. Tau as a marker for Alzheimer's disease. *Trends Biochem Sci* 1993;18:480–483.

37. Braak H, Braak E. Evolution of the neuropathology of Alzheimer's disease. *Acta Neurol Scand Suppl* 1996;165:3–12.

38. Braak E, Braak H, Mandelkow EM. A sequence of cytoskeleton changes related to the formation of neurofibrillary tangles and neuropil threads. *Acta Neuropathol (Berl)* 1994;87:554–567.

39. Haroutunian V, Purohit DP, Perl DP, et al. Neurofibrillary tangles in nondemented elderly and very mild Alzhimer's disease. *Arch Neurol* 1999;57:713–718.

40. Zubenko G, Henderson R, Stiffler J, et al. Association of the *APOE*ε4 allele with clinical subtypes of late depression. *Biol Psychiatry* 1996;40:1008–1016.

41. Gsell W, Strein I, Riederer P. The neurochemistry of Alzheimer type, vascular type and mixed type dementias compared. *J Neural Transm Suppl* 1996;47:73–101.

42. Nemeroff CB, Bissette G, Slotkin TA, et al. Recent advances in the neurochemical pathology of Alzheimer's disease: studies of neuropeptides, cholinergic function and Alzheimer's disease–associated protein. *Ann NY Acad Sci* 1991;640:193–196.

43. Bierer LM, Haroutunian V, Gabriel S, et al. Neurochemical correlates of dementia severity in Alzheimer's disease: relative importance of the cholinergic deficits. *J Neurochem* 1995;64:749–760.

44. Perry EK, Tomlinson BE, Blessed G, et al. Neuropathological and biochemical observations on the noradrenergic system in Alzheimer's disease. *J Neurol Sci* 1981;51:279–337.

45. Davies P, Maloney AJF. Selective loss of central cholinergic neurons in Alzheimer's disease. *Lancet* 1976;2:1403.

46. Perry EK, Perry RH, Blessed G, et al. Necropsy evidence of central cholinergic deficits in senile dementia. *Lancet* 1977;1:189.

47. Bissette G, Seidler FJ, Nemeroff CB, et al. High affinity choline transporter status in Alzheimer's disease tissue from rapid autopsy. *Ann NY Acad Sci* 1996;777:197–204.

48. Haroutunian V, Santucci AC. Pharmacological animal models of dementia. In: Charney DS, Nestler EJ, Bunney BS, eds. *Neurobiological foundations of mental illness*. New York: Oxford University Press, 1999:669–678.

49. Davis KL, Thal LT, Gamzu ER, et al. A double-blind, placebo-controlled multicenter study of tacrine for Alzheimer's disease. *N Engl J Med* 1992;327:1253–1259.

50. DeKosky ST, Harbaugh RE, Schmitt FA, et al. Cortical biopsy in Alzheimer's disease: diagnostic accuracy and neurochemical, neuropathological, and cognitive correlations. Intraventricular Bethanecol Study Group. *Ann Neurol* 1992;32:625–632.

51. Bowen DM, Benton JS, Spillane JA, et al. Choline acetyltransferase activity and histopathology of frontal neocortex from biopsies of demented patients. *J Neurol Sci* 1982;57:191–202.

52. Martin EM, Wilson RS, Penn RS, et al. Cortical biopsy results in Alzheimer's disease: correlation with cognitive deficits. *Neurology* 1987;37:1201–1204.

53. Perry EK, Blessed G, Tomlinson BE, et al. Neurochemical activities in human temporal lobe related to aging and Alzheimer-type changes. *Neurobiol Aging* 1981;2:251–256.

54. Davis KL, Mohs RC, Marin DB, et al. Cholinergic markers are not decreased in early Alzheimer's disease. *JAMA* 1999;281:1401–1406.

55. Davies P, Katzman R, Terry RD. Reduced somatostatin-like immunoreactivity in cerebral cortex from cases of Alzheimer's disease Alzheimer's senile dementia. *Nature* 1980;288:279–280.

56. Bissette G. Neuropeptides and Alzheimer's disease pathology. In: Beckwith BE, Saria A, Chronwall BM, et al., eds. *Neuropeptides in development and aging*. New York: New York Academy of Sciences, 1997:17–29.

57. De Souza EB, Whitehouse PJ, Price DL, et al. Abnormalities in corticotropin-releasing hormone (CRH) in Alzheimer's disease and other human disorders. *Ann NY Acad Sci* 1987;512:237–247.

58. Gabriel SM, Bierer LM, Haroutunian V, et al. Widespread deficits in somatostatin but not neuropeptide-Y consentrations in Alzheimer's disease cerebral cortex. *Neurosci Lett* 1993;155:116–120.

59. Perry RH, Perry EK, Smith CJ, et al. Cortical neuropathological and neurochemical substrates of Alzheimer's and Parkinson's diseases. *J Neural Transm Suppl* 1987;24:131–136.

60. De Souza EB, Whitehouse PJ, Kuhar MJ, et al. Reciprocal changes in corticotropin-releasing factor (CRF)–like immunoreactivity and CRF receptors in cerebral cortex of Alzheimer's disease. *Nature* 1986;319:593–595.

61. Bissette G, Cook L, Smith W, et al. Regional neuropeptide pathology in Alzheimer's disease: corticotropin-releasing factor and somatostatin. *J Alzheimer Dis* 1998;1:1–15.

62. Pomara N, Singh RR, Deptula D, et al. CSF corticotropin-releasing factor (CRF) in Alzheimer's disease: its relationship to severity of dementia and monoamine metabolites. *Biol Psychiatry* 1989;26:500–504.

63. Molchan SE, Hill JL, Martinez RA, et al. CSF somatostatin in Alzheimer's disease and major depression: relationship to hypothalamic-pituitary-adrenal axis and clinical measures. *Psychoneuroendocrinology* 1993;18:509–519.

64. Mouradian MM, Farah JM Jr, Mohr E, et al. Spinal fluid CRF reduction in Alzheimer's disease. *Neuropeptides* 1986;8:393–400.

65. Davis KL, Mohs RC, Marin DB, et al. Neuropeptide abnormalities in patients with early Alzheimer's disease. *Arch Gen Psychiatry* 1999;56:981–987.

66. McGeer PL, McGeer EG. Mechanisms of cell death in Alzheimer disease: immunopathology. *J Neural Transm Suppl* 1998;54:159–166.

67. Aisen PS. Inflammation and Alzheimer's disease: mechanisms and therapeutic strategies. *Gerontology* 1997;43:143–149.

68. Rogers J. Inflammation as a pathogenic mechanism in Alzheimer's disease. *Arzneimittelforschung* 1995;45:439–442.

69. Pasinetti GM, Aisen PS. Cyclooxygenase-2 expression is increased in frontal cortex of Alzheimer's disease brain. *Neuroscience* 1998;87:319–324.

70. Johnson SA, Lampert-Etchells M, Pasinetti GM, et al. Complement mRNA in the mammalian brain: responses to Alzheimer's disease and experimental brain lesioning. *Neurobiol Aging* 1992;13:641–648.

71. Pasinetti GM. Inflammatory mechanisms in neurodegeneration and Alzheimer's disease: the role of the complement system. *Neurobiol Aging* 1996;17:707–716.

72. McGeer PL, McGeer EG. The inflammatory response system of brain: implications for therapy of Alzheimer and other neurodegenerative diseases. *Brain Res Brain Res Rev* 1995;21:195–218.

73. Luterman JD, Haroutunian V, Yemul S, et al. Cytokine gene expression as a function of the clinical progression of Alzheimer disease dementia. *Arch Neurol* 2000;57:1153–1160.

74. Gearing J, Mirra SS, Hedreen J, et al. CERAD. X. Neuropathology confirmation of the clinical diagnosis of Alzheimer's disease. *Neurology* 1995;45:461–466.

75. Thal LJ, Grundman M, Klauber MR. Dementia: characteristics of a referral population and factors associated with progression. *Neurology* 1988;38:1083–1090.

76. Small GW, Rabins PV, Barry PP, et al. Diagnosis and treatment of Alzheimer's disease and related disorders. *JAMA* 1997;278:1363–1371.

77. Jorm AF, Jolley D. The incidence of dementia: a meta-analysis. *Neurology* 1998;51:728–733.

78. Ho LX, Keller DM. Prevalence of AD among whites: a summary by levels of severity. *Neurology* 2000;55:198–204.

79. Stern Y, Gurland B, Tatamichi TK, et al. Influence of education and occupation on the incidence of Alzheimer's disease. *JAMA* 1994;271:1004–1010.

80. Saunders AM, Strittmatter WJ, Schmechel D, et al. Association of apolipoprotein E allele 4 with late-onset familial Alzheimer disease. *Neurology* 1993;43:1467–1472.

81. Meyers RH, Schaefer EJ, Wilson PWF, et al. Apolipoprotein E ε4 association with dementia in a population-based study: the Framingham Study. *Neurology* 1996;46:673–677.

82. Jacobs DM, Sano M, Dooneief G, et al. Neuropsychological detection and characterization of preclinical Alzheimer's disease. *Neurology* 1995;45:957–962.

83. Masur DM, Sliwinski M, Lipton RB, et al. Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons. *Neurology* 1994;44:1427–1432.

84. Petersen RC, Smith GE, Ivnik RJ, et al. Memory function in very early Alzheimer's disease. *Neurology* 1994;44:867–872.

85. Petersen RC, Smith GE, Ivnik RJ, et al. Apolipoprotein E status as a predictor of the development of Alzheimer's disease in memory-impaired individuals. *JAMA* 1995;273:1274–1278.

86. Blacker D, Haines JL, Rodes L, et al. ApoE-4 and age at onset of Alzheimer's disease. *Neurology* 1997;48:139–147.

87. Welsh KA, Butter N, Hughes J, et al. Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. *Arch Neurol* 1991;48:278–281.

88. Morris JC, Edland S, Clark C, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). IV. Rates of cognitive change in the longitudinal assessment of probable Alzheimer's disease. *Neurology* 1993;43:2457–2465.

89. Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–198.

90. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br J Psychiatry* 1968;114:797–811.

91. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry* 1984;141:1356–1364.

92. Rogers SL, Doody RS, Mohs RC, et al. Donepezil improves cognition and global function in Alzheimer's disease: a 15-week, double-blind, placebo-contolled study. *Arch Intern Med* 1998;158:1021–1031.

93. Stern RG, Mohs RC, Bierer LM, et al. Deterioration on the Blessed test in Alzheimer's disease: longitudinal data and their implications for clinical trials and identification of subgroups. *Psychiatry Res* 1992;42:101–110.

94. Stern RG, Mohs RC, Davidson M, et al. A longitudinal study of Alzheimer's disease: measurement, rate and predictors of cognitive deterioration. *Am J Psychiatry* 1994;151:390–396.

95. Cummings JL, Mega M, Gray K, et al. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology* 1994;44:2308–2314.

96. Mohs RC, Ferris SH. Measuring response to treatment in Alzheimer's disease: what constitutes meaningful change? *Int J Geriatr Psychopharmacol* 1998;1:S7–S14.

97. Marin DB, Green CR, Schmeidler J, et al. Noncognitive disturbances in Alzheimer's disease: frequency, longitudinal course and relationship to cognitive symptoms. *J Am Geriatr Soc* 1997;45:1331–1338.

98. McKhann G, Drachman DA, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939–944.

99. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;9:179–186.

100. Galasko D, Bennett D, Sano M, et al. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. *Alzheimer Dis Assoc Disord* 1997;11:S33–S39.

101. Green CR, Mohs RC, Schmeidler J, et al. Functional decline in Alzheimer's disease: a longitudinal study. *J Am Geriatr Soc* 1993;41:654–661.

102. Green CR, Marin DB, Mohs RC, et al. The impact of behavioral impairment on functional ability in Alzheimer's disease. *Int J Geriatr Psychiatry* 1999;14:307–316.

83

Molecular Genetics of Alzheimer Disease

S. Parvathy

Joseph D. Buxbaum

S. Parvathy and Joseph D. Buxbaum: *Department of Psychiatry, Mount Sinai School of Medicine, New York, New York.*

MOLECULAR PATHOLOGY OF ALZHEIMER DISEASE

Part of "83 - Molecular Genetics of Alzheimer Disease "

Pathologic Changes

Alzheimer disease (AD) is characterized histopathologically by the intraneuronal accumulation of paired helical filaments (PHFs) composed of abnormal tau proteins and extracellular deposits of an amyloid peptide (Aß) in plaques (1). AD plaques are round, spheric structures, 15 to 20 µM in diameter, consisting of a peripheral rim of abnormal neuronal processes and glial cells surrounding a core of deposited material. Several associated proteins have been identified in the plaques including heparan sulfate proteoglycans (2), apolipoprotein E (Apo E), and α-antichymotrypsin (3), as well as metal ions (4).

Alzheimer Amyloid Precursor Protein

Aß is proteolytically derived from a larger integral membrane protein, the amyloid precursor protein (APP). Because there are mutations in APP (discussed in detail later), that lead to AD, the molecular and cell biology of APP will be discussed here at some length. APP is a type I integral membrane glycoprotein containing the Aß region, which includes 28 amino acids of the ectodomain and 12 to 14 amino acids of the adjacent transmembrane domain (1 ,5). APP is a member of a family that also includes the amyloid precursor–like proteins 1 and 2 (APLP1 and APLP2). The *APP* gene is localized on chromosome 21 at 21q21.2 (6 ,7), and it is encoded by 18 exons, of which exons 16 and 17 encode the Aß peptide domain (8). Three major splice variants of APP have been identified containing the Aß sequence, that is, the APP695, APP751, and the APP770 isoform, of which APP695 is the major isoform found in neurons (6 ,7 ,9). The two longer forms (APP751 and APP770) contain a 56-amino-acids domain with homology to Kunitz family of serine protease inhibitors (KPI) (10).

APP Processing

APP can be processed by at least three secretases: α-, ß-, and γ-secretases. The site of cleavage of each of these enzymes is shown in Fig. 83.1 . In the nonamyloidogenic pathway, α-secretase cleaves the amyloid precursor protein within the Aß domain. The cleavage within the Aß domain prevents deposition of the intact amyloidogenic peptide. α-Secretase activity generates a soluble N-terminal fragment of APP known as sAPPα, and its C-terminal counterpart of approximately 10 to 11 kd remains embedded in the membrane. The site of cleavage targeted by α-secretase has been identified at the Lys16-Leu17 bond of the Aß sequence corresponding to Lys687-Leu688 peptidyl bond of APP770 (11). The 10- to 11-kd C-terminal product may undergo an additional cleavage by a protease γ-secretase activity. This process leads to the formation of p3 and its complementary product p7 (Fig. 83.1).

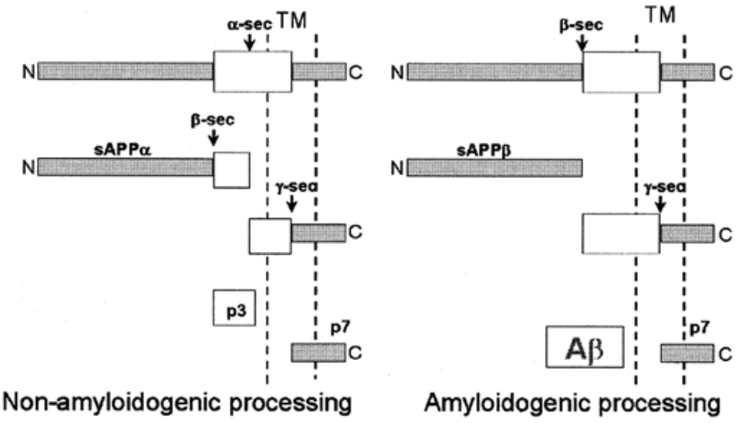


FIGURE 83.1. Processing of APP by the secretases.

The protease termed ß-secretase initiates Aß generation by cleaving APP after methionine 671 (using APP770 numbering), thus creating an approximately 12-kd membrane-retained C-terminal fragment having residue 1 (aspartate) of Aß at its N-terminus (12). This can result in the secretion of a truncated soluble APP molecule, called sAPPß, into the medium (13). The 12-kd fragment may then undergo γ-secretase cleavage within the hydrophobic transmembrane domain at either valine 710, alanine 712, or threonine 713 to release the 40, 42, or 43 residue Aß peptides (13). The varying C-terminal of Aß may be a feature of crucial importance because Aß peptides display distinct physical properties and, in particular, exhibit aggregation behavior that can vary according to their length (14).

Buxbaum et al. showed by using fibroblasts with a disrupted *TACE* (tumor necrosis factor α-converting enzyme)

gene that two classes of α-secretase exist (16), one involved in the basal secretion and the other involved in regulated secretion (16). These investigators demonstrated that TACE, a member of the ADAM family (a disintegrin and metalloprotease family) of proteases, plays a central role in regulated α-secretase cleavage of APP. The existence of two classes of α-secretases is supported by the finding that the potency of inhibition by different hydroxamic acid–based inhibitors is different between basal α-secretase activity and TACE (17). The mammalian kuzbanian orthologue mKuz (ADAM 10) has α-secretase activity and is involved in the basal release of sAPPα (18). Four groups have now identified a candidate for ß-secretase (BACE), also known as Asp-2 (19 ,20 ,21 and 22). BACE has been shown to be an aspartyl protease that is activated from the proenzyme form in cell lines. Finally, presenilin 1 (PS1) appears to facilitate a proteolytic activity that cleaves the integral membrane domain of APP by γ-secretase (23). It is possible that presenilins are γ-secretase or they facilitate γ-secretase activity through some other mechanism (discussed in detail later).

Aß Peptide

As indicated earlier, the chief peptide component of the AD plaque, Aß, has been isolated as a peptide of 40 to 43 amino acids in length (1). Brain Aß has both soluble and insoluble species with aggregation states from monomer to higher-molecular-weight oligomers. Soluble brain Aß is predominately a random coil and an α-helical folded peptide. Insoluble Aß is ß-sheeted and forms either fibrillar or amorphous deposits. These Aß fibrillar aggregates are thought to act as a nidus for subsequent deposits of other proteins, including α-antichymotrypsin, components of the complement cascade, and Apo E, and Apo J (2).

Aß production and release are normal physiologic events. Aß peptides are normally present in the media of APP-expressing cultured cells and in human and rodent cerebrospinal fluid (13 ,24 ,25 and 26). Some reports have showed evidence that there could be two distinct pools of intracellularly generated Aß: a pool that is eventually secreted and a pool that is destined to remain within the cell (27 ,28). The relative importance of intracellular versus extracellular Aß has not yet been determined.

Aß can stimulate inflammatory responses from microglia, can inhibit neuritic outgrowth, and can activate protein phosphorylation, and it is neurotoxic (29). In addition, oxidation of Aß can promote aggregation by peptide cross-linking (30), a potentially important factor in light of increasing evidence of the involvement of oxidative stress in AD. Exposure of neurons, cell lines, or endothelial cells to high concentration of aggregated Aß causes cell death, although it has been suggested that low concentrations of Aß are neurotrophic (31). Altogether, it is thought that the deleterious effect of Aß can cause the pathogenic features that lead to AD.

Functions of APP

The functions of APP and related members of the APP superfamily remain to be fully clarified (32), although APP has been shown to have a wide range of biological properties. The extracellular domain of APP is capable of binding to a range of proteoglycan molecules, and this allows it to function as a regulator of cell–cell or cell–matrix interactions, cell growth, and synaptic plasticity. APP itself may function as a cell-surface receptor. Secreted forms of APP liberated from the cells by the action of secretases may regulate Ca^{2+}, thereby having a neuroprotective effect (33). In addition, sAPPα has been shown to stimulate a protein kinase–mediated signal transduction cascade in cultured cells (34). APP also contains heparin-binding sites and metal ion-binding sites. Finally, APP isoforms containing the KPI domain can act as potent inhibitor of serine proteases (9). The physiologic consequences of the enhanced secretion of either sAPPα or Aß have not been defined.

The intracellular C-terminal domain of APP has been shown to bind to several intracellular proteins, including X11 (35), Fe65 (36), disabled protein (Dab) (37), G protein Go (38), and BP1 (39). The phosphotyrosine interaction domains of X11, Fe65, and Dab bind to the YENPTY motif found in the C-terminus of APP. These proteins can function as adapter proteins enabling APP to bind to other proteins.

Regulation of APP Processing

Several lines of evidence suggest that protein phosphorylation can modulate APP processing in various cell lines. This was shown by a series of studies that consistently reported on the effect of several agents known to activate protein kinase C (PKC) directly. Thus, cell treatment with phorbol esters such as phorbol 12-myristate 13-acetate (PMA) or phorbol 12,13-dibutyrate (PDBu) leads to significant increases in the production of sAPPα (40 ,41). Concomitant

with increased sAPPα secretion, PKC activation lowers the production of Aß (42 ,43 and 44).

The hypothesis that APP processing could be regulated by phosphorylation-dependent processes was reinforced by a study showing that okadaic acid, an inhibitor of phosphatases 1 and 2A, potentiated the PDBu-induced increase of sAPPα (42). The same studies showed that Aß formation was also under the control of phosphatase activity (42). It was demonstrated that drugs that increase cytoplasmic Ca^{2+} levels stimulated sAPPα release in the extracellular medium and increased Aß formation (45 ,46). Therefore it can be said APP processing can be regulated by PKC and by intracellular Ca^{2+} concentrations. The exact mechanism by which PKC and Ca^{2+} modulate APP processing is not clear.

The density of cholinergic receptors is affected in patients with AD, an observation that prompted some authors to examine whether the stimulation of muscarinic receptors could, through activation of the phospholipase C/PKC cascade, ultimately modulate APP processing. Release of sAPPα is found to be stimulated after treatment with various agonists of muscarinic receptors in cells overexpressing M1 and M3 receptor subtypes (47 ,48).

Several other agents have been known to activate APP processing. These include nerve growth factor, epidermal growth factor, tacrine (anticholinesterase drug), estrogen, and electrical depolarization (49). All the foregoing observations emphasize the complexity of the regulatory mechanisms of APP processing.

Localization of APP Processing

By using a mixture of protease inhibitors and by approaches such as metabolic labeling, pulse chase experiments, and cell treatments with various agents known to affect a characteristic step of the intracellular trafficking, De Strooper et al. established that α-secretase activity occurred in a late compartment of the constitutional secretory pathway (50 ,51). These data were supported by Sambamurti and co-workers (52), who demonstrated using [^{35}S] labeling and temperature block that α-secretase activity occurs in the *trans*-Golgi network or in a late *trans*-Golgi compartment just after sulfate incorporation (53).

APP can escape intracellular cleavage by α-secretase and reach the cell surface as a full-length mature product, as evidenced by labeling of APP after biotinylation (54) or radioiodination (55) of cell-surface membranes. Such an approach allows recovering secreted biotinylated or iodinated sAPPα in the cell medium (55 ,56). This finding indicates that α-secretase activity can also occur at the plasma membrane level in several cell systems (see also ref. 57).

Full-length APP that arrives to the plasma membrane can be endocytosed and recycled (58), and Aß appears to be generated both in the endocytic and the exocytic pathway. The cellular sites of Aß production have been thoroughly investigated in cell culture systems. Initial studies indicated that endosomal/lysosomal processing of APP leads to the production of fragments that contain the APP C-terminus and entire Aß region and hence are potentially amyloidogenic (56 ,59). Despite the initial excitement generated by the discovery of these Aß-containing fragments, several lines of evidence suggest that the lysosomal degradation of APP is unlikely to contribute to the production of Aß. However, agents that interfere with pH gradients (i.e., ammonium chloride and chloroquine) inhibit the production of Aß (25), a finding suggesting that Aß may be generated in acidic compartments (i.e., endosomes or late Golgi).

Furthermore, cells that express APP with various deletions in the cytoplasmic tail release lower levels of soluble Aß (60), a finding suggesting that the internalization of APP from the cell surface and subsequent recycling to the plasma membrane may be responsible for the generation of Aß. Two lines of evidence are consistent with this idea: secreted Aß can be generated from [^{125}I]surface-labeled APP (55), and the surface APP "tagged" with either monoclonal antibodies or biotin can be recycled to the plasma membrane after endocytosis (58). These studies offered a model wherein ß-secretase cleavage occurs within endocytic compartments, and γ-secretase cleavage of the residual approximately 100 amino acid membrane-bound fragment within the APP transmembrane domain occurs virtually simultaneously with the formation and release of Aß.

Although the majority of Aß appears to be generated in the endosomal recycling pathway, a few secreted Aß species are also generated in a secretory pathway (27 ,61). Aß secretion could be totally prevented by brefeldin A (24 ,62), and by monensin (24), two agents that potently block intracellular trafficking. This finding suggests that Aß may be generated during the transport of APP through the Golgi apparatus *en route* to the plasma membrane.

Taking together all the available data, it is possible to formulate a model for APP processing. APP polypeptide undergoes complex post-translational modifications, including sulfation, phosphorylation, and both N- and O-linked glycosylation (44 ,63). These modifications occur during the trafficking of the protein through the secretory pathway. Thus, APP is cotranslationally translocated into the endoplasmic reticulum by its signal peptide and then matured during passage through the Golgi by acquiring sulfate, phosphate, and sugar groups. At this stage, some of the APP may have already been processed by ß- and γ-secretases. Then, a percentage of mature molecules is transported to the plasma membrane by secretory vesicles (64). At or near the cell surface, some APP molecules undergo proteolysis by the protease designated α-secretase. Alternatively, uncleaved surface APP molecules can undergo endocytosis through clathrin-coated vesicles, apparently mediated by a YENPTY signal sequence in the distal cytoplasmic tail (65), after which the full-length precursor is trafficked to late endosomes and lysosomes for apparent

degradation (56 ,59), or it is rapidly recycled within early endosomes to the cell surface (66). The latter pathway has been shown to be principal site for the two proteolytic cleavages that generate the Aß peptide (55). It appears that, at least in cultured cells, only a few of all biosynthesized APP molecules undergo either the α-secretase or the ß-secretase fate; many full-length precursor molecules remain inserted into internal membranes, particularly in the Golgi.

Tangles

Neurofibrillary tangles are another important histologic feature of AD. They consist of PHFs in a double helix (diameter, 20 nm) found in the cytoplasm of neurons, particularly of the pyramidal cells of the cerebral cortex and hippocampus (67). PHFs are composed principally of phosphorylated tau, a low-molecular-weight microtubule-associated protein (68).

Tau is a family of six proteins derived by alternative mRNA splicing from a single gene located on chromosome 17. These molecular isoforms of tau differ in whether they contain three or four tubulin-binding domains of 31 or 32 amino acids each near the C-terminal end and no, one, or two inserts of 29 amino acids each at N-terminal end of the molecule.

Tau in AD brain, especially in PHFs, is abnormally hyperphosphorylated and glycosylated. At the later stages of tangle formation, the tau is increasingly ubiquinated. In a normal neuron, biological function depends on an intact microtubule network through which much of the axoplasmic transport is supported. The AD abnormally phosphorylated tau (AD P-tau) competes with tubulin in binding to normal tau, MAP1, and MAP2 and inhibits their microtubule assembly-promoting activities. The disruption of the microtubule network probably compromises the axonal transport and starts retrograde degeneration of the affected neurons. The neuronal cytoskeleton in AD is progressively disrupted and is replaced by bundles of PHFs, leading to the formation of neurofibrillary tangles (67).

To date, 21 phosphorylation sites in the AD abnormally phosphorylated tau have been identified (69). Among the several protein kinases that have been implicated in the phosphorylation are glycogen synthase kinase-3 (70), neuronal Cdc-like protein kinase (71), mitogen-activated protein kinase (72), Ca^{2+}/calmodulin-dependent protein kinase II (73), casein kinase I (74), and cyclic adenosine monophosphate–dependent protein kinase (75). AD P-tau can be dephosphorylated by protein phosphatases PP-2B, PP-2A, and PP-1, but not by PP-2C. Tau phosphatase activity is decreased by approximately 30% in AD brain (76); thus, the decrease in phosphatase activities may contribute to the abnormal hyperphosphorylation of tau in AD resulting from the inhibition of dephosphorylation reaction. To date, no mutations in tau have been implicated in AD; however, mutations in tau are associated with frontotemporal dementia (77).

MUTATIONS IN APP IN EARLY-ONSET AD

Part of "83 - Molecular Genetics of Alzheimer Disease "

Genes may be related to disease in two ways: through mutations that by themselves are sufficient to cause the disease (i.e., deterministic mutations), or alternatively, through gene variations (polymorphisms) that may increase disease risk without being sufficient (or necessary) by themselves to cause the disorder. This latter group is referred to as susceptibility genes. Although it is currently thought that most cases of AD occur sporadically, autosomal dominant transmission has been identified in families with early-onset AD, defined as beginning before the age of 65 years. These cases are relatively rare; worldwide, only several hundred families are currently known to carry deterministic mutations (78). Extensive research carried out since the early 1980s has isolated certain genes that, when mutated, cause AD, notably *APP* on chromosome 21 (79 ,80), the *PS1* gene on chromosome 14 (81), and the *PS2* gene on chromosome 1 (81). Mutations in these genes lead to early-onset AD and explain only a small proportion of total AD cases. Furthermore, trisomy 21 (Down syndrome or DS) increases the risk of AD, perhaps because of the tripled genetic dosage of APP. In addition, some susceptibility genes are currently being studied, of which polymorphisms of the *APOE* gene have received the most attention. The presence of the *APOE*-4 allele has been identified as a genetic risk factor for sporadic AD and familial AD (FAD) of late onset. All these genetic causes of AD are discussed in detail later.

AD in Down Syndrome

In considering the molecular pathology of DS, a matter of critical interest is that virtually all patients with DS who survive beyond 35 years of age develop neuropathologic changes that closely resemble AD (82). Thus, there is the abnormal accumulation of Aß in the brains of both patients with AD and those with DS, followed by cognitive decline. Patients with DS are thought to express high levels of APP because of an extra copy of chromosome 21. The most straightforward explanation for dementia in DS is the presence of three instead of two copies of APP in patients with DS (83). Other genes that are potentially overexpressed in DS are located within a segment of chromosome 21, termed the Down locus. Genes that are contained within the Down locus include APP, superoxide dismutase I, S100-ß (a calcium-binding protein), and BACE-2 (a homologue of BACE).

The occurrence of +1 frameshifted proteins, such as APP+1 and ubiquitin-B+1 (UBB+1) has been linked to the onset of AD in patients with DS. Frameshifts are caused by dinucleotide deletions in GAGAG motifs in mRNA and are now thought to be the result of unfaithful transcription

of normal DNA by a novel process termed *molecular misreading*. The aberrant mRNAs are translated in the +1 reading frame as "+1 proteins," that is, proteins with a wild-type N-terminus and frameshifted and often truncated C-terminus. It has been shown that expression of APP+1 protein (84 ,85) and UBB+1 protein is found in all patients with DS (85). In patients with DS and AD, the mRNA decay pathway may be impaired, and therefore +1mRNA is translated into +1 protein. Conversely, UBB+1 may have a role in directly interfering with the ubiquitin/proteasome system and may lead to an inefficient protein breakdown through the proteasomal pathway.

APP Mutations and Their Effect on Aß Formation

Several different pathogenic mutations have been found in exons 16 and 17 of the *APP* gene to date (Table 83.1). These mutations are missense mutations. The early-onset AD mutations (according to APP770 numbering), APP715, 716, 717, and APP670/671, are located outside the Aß amyloid sequence, with APP715-717 (Val715Met, Ile716Val, and Val717Ile/Gly/Phe) close to the C-terminal γ-secretase cleavage site (79) within the transmembrane domain and APP670/671 (Lys670Asn/Met671Leu) at the N-terminal ß-secretase cleavage site within the extracellular part of APP (90). In contrast, the APP692 (Ala692Gly) mutation is located inside the Aß amyloid sequence next to the α-secretase cleavage site (89). Other sequence variations in the Aß region are thought not to be pathogenic.

TABLE 83.1. APP MUTATIONS THAT CAUSE AD

Mutation	Age of Onset (y)	References
V717I (London)	55 (50–60)	Goate et al., 1991 (79)
V717F	47 (42–57)	Murrell et al., 1991 (86)
V717L	38	Murrell et al., 2000 (87)
V717G	55 (45–62)	Chartier-Harlin et al., 1991 (88)
A692G (Flemish)	40–60 but variable	Hendriks et al., 1992 (89)
K/M670/1N/L (Swedish)	52 (44–59)	Mullan et al., 1992 (90)
I716V (Florida)	55	Eckman et al., 1997 (91)
V715M (French)	52 (40–60)	Ancolio et al., 1999 (92)

The localization of mutations led to the hypothesis that these mutations could influence the activity of the respective secretases, resulting in the aberrant processing of APP (95). Indeed, mutations at codons 716 and 717 lead to a selective increase in the production of Aß peptides ending at residue 42/43 (91 ,96 ,97 ,98 and 99). The Lys670Asn/Met671Leu mutation, conversely, appears to augment the production of both Aß40 and Aß42/43 (100), whereas the Ala692Gly mutation has a more complicated effect on APP processing by causing impaired α-secretase cleavage, increased heterogeneity of secreted Aß species, and increased hydrophobicity of the Aß (98). The Ala692Gly mutation also has clinical features in some cases similar to those of cerebral hemorrhage with amyloidosis of the Dutch type (HCHWA-D) (93), and in other cases more similar to AD. Recently, another mutation in APP (E693G) has been identified as the Arctic APP mutation that enhances Aß protofibril formation (94).

MUTATIONS IN THE PRESENILINS IN EARLY-ONSET AD

Part of "83 - Molecular Genetics of Alzheimer Disease "

The homologous membrane proteins presenilin 1 (PS1) and presenilin 2 (PS2) were identified in 1995 as the genes responsible for a substantial fraction of early onset, autosomal dominant AD (81 ,101). The most common causes of autosomal dominant FAD are mutations in the *PS1* gene on chromosome 14 (81). These account for 30% to 50% of all early-onset cases (102), and they are the primary cause of AD with onset before the age of 55 years. To date, more than 50 *PS1* mutations and two *PS2* mutations have been reported (Table 83.2 and Table 83.3) in FAD. All mutations in the presenilins (PS) are missense mutations, except for the mutation of a splice acceptor site resulting in the deletion of exon 9. The lack of mutations leading to loss of gene expression or frameshifts suggests that the disease phenotype results from a gain of function. The hydrophobic regions of PS1 and PS2 are almost completely identical and are highly conserved among species.

Mutation	Age of Onset (y)	References
V82L	55	Campion et al., 1995 (103)
V96F		Kamino et al., 1996 (104)
Y115H	37	Campion et al., 1996 (105)
Y115C		Hardy, 1997 (106)
P117L		Wisniewski et al., 1998 (107)
E120D	48	St. George-Hyslop, 1998 (78)
E120K		Hardy, 1997 (106)
K123E		Yasuda et al., 1999 (108)
N135D		Hardy, 1997 (106)
M139T	49	Campion et al., 1996 (105)
M139V	40	Alzheimer's Disease Collaborative Group, 1995 (109)
M139K		Dumanchin et al., 1998 (110)
I143F		Hardy, 1997 (106)
I143T	35	Cruts et al., 1995 (111)
M146L	45	Sherrington et al., 1995 (81)
M146V	38	Alzheimer's Disease Collaborative Group, 1995 (109)
M146I	40	St. George-Hyslop, 1998 (78)
H163R	50	Sherrington et al., 1995 (81)
H163Y	47	Alzheimer's Disease Collaborative Group, 1995 (109)
L171P	40	Ramirez-Duenas et al., 1998 (112)
G209V		Kamino et al., 1996 (104)
I213T		Kamino et al., 1996 (104)
A231T	52	Campion et al., 1996 (105)
A231V		Hardy, 1997 (106)
M233T	35	Kwok et al., 1997 (113)
L235P	32	Campion et al., 1996 (105)
A246E	55	Sherrington et al., 1995 (81)
L250S		Hardy, 1997 (106)
A260V	40	Ikeda et al., 1996 (114)
L262F	50	Forsell et al., 1997 (115)
C263R	47	St. George-Hyslop, 1998 (78)
P264L	45	Campion et al., 1995 (103)
P267S	35	Alzheimer's Disease Collaborative Group, 1995 (109)
R269G		Hardy, 1997 (106)
R269H		Hardy, 1997 (106)
E273A		Kamimura et al., 1998 (116)
R278T		Kwok et al., 1997 (113)
E280A	47	Alzheimer's Disease Collaborative Group, 1995 (109)
E280G	42	Alzheimer's Disease Collaborative Group, 1995 (109)
L282R		Alduda et al., 1998 (117)
A285V	50	Ikeda et al., 1996 (114)
L286V	50	Sherrington et al., 1995 (81)
291-319 deletion		Crook et al., 1998 (118); Perez-Tur et al., 1996 (119)
E317G		Hardy, 1997 (106)
G384A	35	Cruts et al., 1995 (111)
L392V	25–40	Campion et al., 1995 (103)
C410Y	48	Alzheimer's Disease Collaborative Group, 1995 (109)
A426P		Hardy, 1997 (106)
P436S		Hardy, 1997 (106)

TABLE 83.2. MISSENSE MUTATIONS IN THE PRESENILIN 1 GENE

Mutation	Age of Onset (y)	References
A141I (Volga Germans)	50–65	Rogaev et al., 1995 (120)
M239V (Florence)	Onset variable	Rogaev et al., 1995 (120)

TABLE 83.3. MISSENSE MUTATIONS IN THE PRESENILIN 2 GENE

Structure, Localization, and Post-Translational Modification of Presenilins

Structure

Hydropathy analysis of PS1 and PS2, using the indices of Kyte and Doolittle, indicates the presence of a hydrophilic N-terminal followed by ten hydrophobic regions (HR) of at least 15 amino acids in length, which could potentially span the membrane (81 ,101 ,120). Most of these segments are connected by small hydrophilic loops, except one longer stretch of mostly hydrophilic residues between HR7 and HR8 called the "large loop."

Multiple studies have been aimed to determine the number of transmembrane domains as well as the orientation of the N- and C-terminus of PS. These approaches included staining with antibodies after selective permeabilization of cellular membranes, construction of chimeric proteins with protease cleavage sites, and the use of ß-galactosidase or glycosylation tags (121 ,122 ,123 ,124 ,125 ,126 and 127). All the proposed models agree that the first six hydrophobic regions represent transmembrane domains, whereas the predictions of the total number of regions varies between six and eight. In all these models except one (126), the N-terminal and the C-terminal

domains were shown to protrude into the cytoplasm. Two studies using ß-galactosidase fusion proteins support the model of eight transmembrane domains (TM) with N-terminal and C-terminal in the cytosol (124 ,125). In this model, HR7 and HR10 do not pass through but are associated with the membrane. Alternative topologies have also been suggested.

Post-Translational Modification

Presenilins are neither glycosylated nor modified by sulfation, acylation, or the addition of glycosaminoglycans (121), but both proteins are phosphorylated on serine residues (128 ,129). The most prominent post-translational modification of both PS1 and PS2 is proteolytic cleavage (130 ,131). PS1 is rapidly cleaved into a 27- to 28-kd N-terminal fragment (NTF) and an 18- to 20-kd C-terminal fragment (CTF), and PS2 is cleaved into two polypeptides of 34 and 20 kd, respectively.

Epitope mapping studies (131) and radiosequencing analysis (130) revealed that PS1 endoproteolysis occurs in the cytoplasmic loop domain, within a domain in which several of the identified FAD-linked PS1 mutations occur. The N-terminal of the CTF is heterogeneous, with the two predominant species beginning at amino acids 293 and 299 (130), encoded by exon 9. These findings are consistent with the demonstration that the FAD-linked PS1ΔE9 variant, which lacks exon 9 encoded sequences (amino acids 290 to 319), fails to be cleaved (131). Several of the PS mutations are clustered around this region.

The PS holoproteins are unstable, with half-lives about 1.5 hours (130 ,132), and their degradation is apparently mediated in part through the proteosome (133 ,134). In contrast, PS fragments produced through normal endoproteolysis of wild type and FAD-mutant presenilins are quite stable (half-life of approximately 24 hours) (130 ,132), a finding consistent with the hypothesis that the heterodimeric complexes represent the biologically active form of the protein. Presenilin molecules that are not incorporated into the complex are rapidly degraded by several proteases, including the caspases and calpain-like enzymes (134 ,135). It appears that endoproteolysis of the presenilins is not needed for activation of their putative activities but may be required to convert unstable presenilins into stable complexes (134).

Various C-terminally truncated and chimeric PS polypeptides were used to characterize the interaction between NTF and CTF. It was observed that transgene-derived human PS1 NTF expressed in mouse N2a cells neither assembled with the endogenous CTF nor inhibited the cleavage or accumulation of the endogenous mouse PS1. Furthermore, in cells coexpressing PS1 and PS2, PS1-and PS2-derived fragments did not form mixed assemblies. In contrast, cells expressing a chimeric PS1/PS2 polypeptide formed PS1 NTF.PS2 CTF assemblies. These studies provide strong evidence that intramolecular associations between PS domains precede endoproteolytic processing (136).

In addition to the regulated endoproteolytic processing cleavage by the yet hypothetical presenilinase, presenilins also undergo additional cleavage, termed *alternative cleavage*, within the hydrophilic loop domain (133). Full-length PS1, as well as PS1- or PS2 derived CTFs, can be cleaved by caspases in transfected cells and cells induced to undergo apoptosis. Several members of the caspase family of proteases, including caspases 1, 3, 6, 7, 8, and 11, are capable of cleaving PS1 and PS2 *in vitro* (138).

Localization

Endogenous presenilins have a relatively limited subcellular distribution; they are found in the early compartments of biosynthetic pathway. Presenilin proteins have been localized to the endoplasmic reticulum (ER) and the Golgi subcellular compartments (137). Confocal and electron microscopy, combined with subcellular fractionation experiments, show that presenilins in neurons reside in the smooth and rough ER, the ER Golgi intermediate compartments, and, to a limited extent, in the *cis*-Golgi, but not beyond (139). The finding of overexpressed presenilin proteins within Golgi compartments should, however, be interpreted with caution, because evidence indicates that membrane proteins with many transmembrane domains can accumulate in structures called aggresomes, structures that reflect cell stress (140). Studies provide convincing evidence that some mammalian PS1 can be found at the cell surface, where it can be biotinylated (141).

Biological Functions of Presenilins

Interaction with APP

There is strong evidence that presenilins are able to interact directly with APP. Complex formation between APP and presenilins has been demonstrated by coimmunoprecipitation of both proteins in cells either transfected or with endogenous proteins as well as with the yeast two-hybrid system (142 ,143). Thinakaran and colleagues, in contrast to these other researchers, did not observe physiologic complexes between PS1 and PS2 derivatives with APP (144).

Studies with progressive deletion of presenilin showed that the hydrophilic N-terminal of PS2 (1-87) is sufficient for the interaction with APP (127). Two different domains of APP appear to be involved in the APP-PS interaction. The last 100 C-terminal residues of APP (C100) encompassing Aß and the TM region are able to interact with PS1 and PS2 (142 ,143). However, deletion of the cytoplasmic C-terminus domain does not abrogate PS1 binding (143). In addition, two APP constructs representing physiologically secreted forms of APP (sAPPα and sAPPß) were shown to coprecipitate with PS2 in transfected COS cells (127). Taken together, these result suggest presenilin binds to the Aß/transmembrane region and at least one additional interaction domain N-terminal of Aß. Full maturation of APP does not seem to be required for the interaction, because the APP form detected in precipitated complexes is mostly immature.

Role in APP Processing

Pathogenic mutations in PS modify APP processing, thereby leading to an augmentation of Aß42 secretion. Patients with AD who carry PS1 or PS2 mutations have significant increase of plasma Aß42 levels (145) together with deposition of Aß42 in the brain (146 ,147). In fibroblasts from such patients, the APP metabolism is shifted toward an increase of Aß42 production. Similarly, the presence of mutated PS1 increases Aß42 in transfected cells (148 ,149 ,150 and 151), as well as in transgenic mice (148 ,149 and 150). How the mutant PS influences the production of Aß42 peptides remains uncertain, but these PS mutations appear to cause aberrant gain, rather than loss, of function. In neurons of PS1-knockout mice, secretion of Aß is drastically reduced, leading to the accumulation of α- and ß-cleaved C-terminus stubs of APP (23 ,152). This gives evidence that PS1 is obligatory for proteolysis of APP at the γ-secretase cleavage site.

Wolf et al. mutated the two aspartates located in analogous positions near the middle of TM6 and TM7 to alanines (153). Either mutation, when expressed in various mammalian cell types, prevented both the normal endoproteolysis of PS1 within the hydrophobic region of the TM6-TM7 loop and markedly inhibited the γ-secretase cleavage of the 99-residue C-terminal fragment of APP (C99), thereby lowering levels of Aß40 and Aß42. Conservative substitution of aspartate by glutamate still abrogated the γ-secretase cleavage of APP, a finding indicating a specific requirement for the two TM aspartates. These results are consistent with one of two mechanisms: a role for presenilin as a unique cofactor for γ-secretase that could play a role in protein trafficking or a role as a functional γ-secretase, making it an unprecedented intramembranous aspartyl protease. There is evidence for and against both possibilities.

A major concern with the hypothesis that presenilins are proteases is their subcellular localization. Presenilin proteins

have been localized to early transport compartments, whereas abundant γ-secretase activity is restricted to late transport compartments and the endosomal pathway (55 ,61). The same holds true for the release of the Notch intracellular domain (see later), which occurs after ligand binding by Notch at the cell surface (154 ,155). In contrast, the cellular localization of presenilin proteins in ER and early Golgi overlaps to some degree with the intracellular site of generation of the highly amyloidogenic Aβ42. An additional concern is that the presenilin sequences have no homology to any of the proteases identified so far. Confirmation that PS is γ-secretase will require reconstitution of the γ-secretase/presenilinase activities in artificial lipid bilayers using appropriate substrates and cellular factors. Partial characterization of detergent-solubilized γ-secretase activity shows that γ-secretase activity is catalyzed by PS1-containing macromolecular complex (156).

The alternate hypothesis holds that PS1 influences endoproteolysis of APP and Notch 1 indirectly, by altering trafficking and cocompartmentalization of each substrate and elusive protease. It was demonstrated that CTF derived from the APP homologue, APLP1, also accumulates in PS-/- neurons (152). Because APP and APLP1 transmembrane domains have very limited homology, it may be more difficult to envision that PS1 plays a role as a specific γ-secretase involved in the cleavage of APP, Notch (see later), and APLP1. Rather, a broader role for PS1 in directing membrane-bound CTFs derived from APP family members or other transmembrane proteins to appropriate cleavage or degradation compartments may be considered. It is possible that the presenilins are critical cofactors for γ-secretases, analogous to the sterol-cleavage activating protein involved in the extramembranous processing of the sterol regulatory element binding protein, (157), which is a transcription factor essential for cholesterol biosynthesis.

Notch Signaling Pathways

The Notch gene family encodes large, single transmembrane proteins in the plasma membrane. Notch function is involved in various signaling pathways, and Notch is crucial for many steps in the development of an organ and organism. A similar pathway is used in *Caenorhabditis elegans* at multiple steps in development, including singling out precursor cells involved in vulva differentiation (158). For this purpose, two cells that are initially functionally identical crosstalk through the activity of the lin-12/Notch gene product. The consequence of this interaction is that one cell will specialize in producing vulva cells, and the other will specialize in generating uterine cells. A reduction on lin-12/Notch function causes an egg-laying defect that results from failure in vulva induction.

A functional link between APP and Notch receptor was first suggested when several groups reported that disruption of PS1 not only prevented APP processing by γ-secretase, but also prevented the cleavage of the Notch C-terminus in the membrane. This result immediately suggested that either presenilins are directly involved in cleaving both Notch and APP or mediate both cleavages in a more indirect way. Processing of Notch resembles in some aspects the processing of APP. Notch is processed by a furin-mediated cleavage during its passage through the Golgi system. The resultant two fragments remain in the same protein complex and localize in the cellular membrane to form the functional receptor. The binding of the ligand to the receptor stimulates the cleavage of one of the subunits at a specific extracellular site close to the membrane. A subsequent intramembranous cleavage liberates an intracellular fragment that translocates into the nucleus. This peptide forms an active transcription complex, which activates transcription of Notch target genes. The last of these proteolytic cleavage steps of Notch resembles γ-secretase cleavage of APP.

Complex formation between Notch and PS has been observed (141). There are also data showing that presenilins are functionally implicated in the Notch signaling pathway. The phenotype observed in PS1 knockout animal models (159) consists of a severe impairment of the development of the axial skeleton. The origin of these skeleton abnormalities lies in the impairment of the segmentation of the somites. Interestingly, Notch-1 (160) knockout animals suffered from similar abnormal skeleton deformations, a finding consistent with interaction of presenilins with Notch signaling pathway. De Strooper et al. used a Semiliki forest virus system to express truncated fragments of Notch-1 in primary neuronal cultures of PS1 knockout animals to identify the steps in Notch signaling depending on PS1(161). The authors provide evidence strongly suggesting that PS1 is crucial for the final processing of Notch-1 that generates the intracellular fragments, which subsequently will translocate to the nucleus and activate the expression of Notch-1 target genes.

The genetic studies in *C. elegans* and *Drosophila* offer a powerful approach to study presenilin function. Sel-12, a nematode homologue of presenilin, was identified by screening for suppressors of lin-12 (*C. elegans* homologue of Notch) gain of function mutation (162). Sel-12 is able to facilitate the signaling of transmembrane receptors of the lin-12/Notch family, and human presenilins have been shown to complement for Sel-12 function effectively (163). The egg-laying deficiency in *C. elegans* caused by the null mutations of Sel-12, the worm's homologue of mammalian presenilin, can be rescued by wild-type, but not mutated, human presenilin (163 ,164). However, presenilin cleavage does not seem to be essential for functional activity, because the PS1 with the Δ-exon 9 mutation is still able at least partially to rescue the egg-laying defective phenotype of *C. elegans* Sel-12 mutants (164).

Proteins Interacting with PS

Presenilins have been found to interact directly with a variety of proteins. Proteins interacting with presenilins include members of the catenin family (165 ,166 and 167). Catenins have at least two different functions in the cell. First, they are components of cell–cell adhesive junctions interacting with the cytoskeletal anchors of cadherin adhesion molecules. Second, there is compelling evidence that ß-catenin is a key effector in the Wingless/Wnt signaling cascade. Wingless and its vertebrate counterpart Wnt signaling direct many crucial developmental decisions in *Drosophila* and vertebrates. ß-Catenin has been shown to become destabilized in case of PS1 mutations and PS1 deficiency (168). ß- and ð-Catenin interact with the large loop of PS1 (165 ,166). The yeast two-hybrid system was used for the identification of a neuronal calcium-binding protein termed calsenilin (15), which binds to the C-terminus of PS2. Calsenilin was shown to interact with both PS1 and PS2 in cultured cells and could link presenilin function to pathways regulating intracellular calcium levels.

Several other proteins have been identified that interact with presenilins including the cytoskeletal proteins filamin and filamin homologue (168), µ-calpain (169), Rab11, a small guanosine triphosphatase belonging to the p21 ras-related superfamily (170), G-protein Go (171), and glycogen synthase kinase-3b (172).

Apoptosis and Cell Death

There is increasing evidence of causal involvement of presenilins in apoptosis. ALG3, a 103-residue C-terminal fragment of PS2, was isolated in death trap assay as rescuing a T-cell hybridoma from T-cell receptor and Fas-induced apoptosis (173). In PC12 cells, the down-regulation of PS2 by antisense RNA protects the cells from glutamate toxicity. Similar effects of ALG3 overexpression and PS2 down-regulation suggest that this C-terminal fragment of PS2 acts as a dominant negative form of PS2. Expression of mutant PS1 (L286V) in PC12 cells enhanced apoptosis on trophic factor withdrawal or Aß toxicity (174). The alternative caspase cleavage in the C-terminal fragment of PS1 has been shown to abrogate the binding of PS1 to ß-catenin (167) and could therefore modulate the apoptotic outcome.

Finally, the FAD mutation M146V was inserted by homologous recombination in the mouse genome (knock-in) to drive the expression at normal levels of a mutated mouse PS1 protein (175). The knock-in mutation was shown to increase ER calcium mobilization and superoxide and mitochondrial reactive oxygen species production leading to caspase activation (175).

Evidence shows that mutant PS1 also renders cells less efficient to respond to stress conditions in ER. Mutations in PS1 may increase vulnerability to ER stress by altering the unfolded protein response (UPR) signaling pathway that is responsible for ensuring the proper folding of newly synthesized proteins (176 ,177).

ROLE OF APOLIPOPROTEIN E ISOFORMS IN LATE-ONSET AD

Part of "83 - Molecular Genetics of Alzheimer Disease "

In addition to the deterministic genetic mutations found in APP and presenilins, genetic factors modify the risk of developing AD. The *APOE* gene on chromosome 19 is considered as an important risk factor for the development of late-onset AD. Apo E is a 34-kd component of various lipoproteins, including chylomicrons, very-low-density lipoproteins (VLDLs), and a subset of high-density lipoproteins (HDLs) (178). These lipoproteins regulate plasma lipid transport and clearance by acting as ligand for lipoprotein receptors such as LDL-R and low-density receptor-related protein (LRP) (179 ,180). Apo E has been implicated in the transport of cholesterol and phospholipids for the repair, growth, and maintenance of membranes that occur during development or after injury (178).

Apo E is polymorphic and is encoded by three alleles (*APOE2,3,4*) that differ in two amino acid positions. The most common isoform, E3, has a Cys residue at position 112 and an Arg at position 158. The two variants contain either two Cys residues (E2) or two Arg residues (E4) at these positions. In general, it seems that E4 allele increases the risk of developing AD by approximately threefold, and that the E2 allele decreases the risk (181). The presence of one or two E4 alleles is associated with earlier onset of disease and an enhanced amyloid burden in brain, but it has little effect on the rate of progression of dementia (182). Thus, homozygous E4/E4 subjects have an earlier onset (mean age less than 70 years) than heterozygous E4 subjects (mean age of onset for E2/E3 is more than 90 years) (183).

The most obvious hypothesis is that *APOE* polymorphisms may influence the production, distribution, or clearance of Aß. This hypothesis is supported by observations that the subjects with one or more *APOE4 alleles have a higher amyloid burden than do subjects with no APOE4* alleles (184). Second, there is evidence that both Apo E and Aß may be cleared through the LRP receptor, and Apo E4 and Aß peptide may compete for clearance through the LRP receptor (179). Third, transgenic mice that overexpress APP develop a significantly lower number of Aß deposits when they are bred to an *APOE* knockout background (185). These findings strongly support a role of Apo E in the aggregation or clearance of Aß in the brain. The *APOE* genotype influences the onset of AD in patients with DS and in those with APP mutations but not in families with presenilin mutations (185 ,186 and 187).

OTHER GENETIC RISK FACTORS IN AD

Part of "83 - Molecular Genetics of Alzheimer Disease "

In addition to the *APOE* gene, which has been confirmed as a strong risk factor in various studies, polymorphisms in

several other genes have been described to increase susceptibility for AD. Most of these genetic polymorphisms are still subject to discussion because they either need to be confirmed in larger studies or show only weak influence toward the risk of developing AD.

A family-based study revealed an association between late-onset AD and the presence of an exon 2 splice acceptor deletion in the α_2-macroglobulin *(A2M)* gene on the short arm of chromosome 12 (188). Significantly, A2M binds to a variety of proteins, including proteases (189 ,190) and Aß peptide (191). In addition, A2M is also present in senile plaques and can attenuate Aß fibrillogenesis and neurotoxicity (191). Moreover, A2M may, through LRP-mediated endocytosis, allow the internalization and subsequent lysosomal degradation of Aß (192).

Other candidate AD genes that have been reported include α_1-antichymotrypsin (193), bleomycin hydrolase (194), and a gene on chromosome 3q25-26 (195). The candidacy for these genes as AD loci awaits further testing and confirmatory studies in greater numbers of AD samples.

Among the several risk factors, age is one of the most important elements that has to be considered with respect to sporadic Alzheimer-type dementia. Various cellular and molecular changes take place in the brain during normal aging, among which changes in glucose and energy metabolism are of pivotal significance (196). Reduced acetylcholine synthesis, formation of advanced glycation end products, membrane instability, and reduced energy availability are some of the other changes that occur with age. The decrease in the pool of available energy may lower or even arrest the transport of newly synthesized membrane proteins such as APP, thereby facilitating Aß synthesis intracellularly. Therefore, it can be hypothesized that these changes somehow contribute to the formation of Aß peptide and hyperphosphorylated tau. In a general context, age may be considered as a risk factor for neuronal damage and thus for age-related brain disorders such as sporadic dementia of the Alzheimer type.

AD has been reported to be in higher prevalence among old women compared with men. There is evidence suggesting that estrogen may have a protective role against AD, perhaps through its action as a trophic factor for cholinergic neurons (197), a modulator for the expression of Apo E in the brain (198 ,199), an antioxidant compound decreasing the neuronal damage caused by oxidative stress (200 ,201), or a promoter of physiologic nonamyloidogenic processing of the APP decreasing the production of Aß (202). Therefore, in older postmenopausal women, the lack of circulation of estrogen may contribute to developing AD.

CONCLUSIONS

Part of "83 - Molecular Genetics of Alzheimer Disease "

There are clearly very significant genetic contributions to AD. Practically, Apo E is the most significant genetic factor in AD because of the very high prevalence of the *APOE4* allele, even though *APOE* is only a risk factor for AD.

The mutations in *APP* and *PS* account for only a small percentage of AD cases, but when present, these are deterministic mutations leading to an aggressive early-onset form of the disease. These mutations in *APP* and *PS*, rare though they are, give crucial insight into the molecular process underlying all forms of AD. Thus, the *APP* mutations clearly underscore critical role of APP in disease initiation. The *PS* mutations implicate Aß, and particularly Aß42, in disease. A similar, crucial role for Aß in sporadic AD is supported by postmortem and tau studies. Polymorphism in *APOE* and *A2M* may also have their effects by interacting with APP and Aß. It therefore seems likely that treatments that modulate Aß formation or Aß clearance may be of benefit in AD.

If Aß changes initiate disease, it is likely that changes in tau are the actual cause of neuronal dysfunction and cognitive decline. Therefore, abrogating deleterious effects of Aß on tau phosphorylation may represent another viable therapeutic approach to AD.

REFERENCES

1. Glenner GG, Wong CW. Alzheimer’s disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem Biophys Res Commun* 1984;120:885–890.

2. Snow AD, Mar H, Nochlin D, et al. The presence of heparan sulfate proteoglycans in the neuritic plaques and congophilic angiopathy in Alzheimer’s disease. *Am J Pathol* 1988;133:456–463.

3. Abraham CR, Selkoe DJ, Potter H. Immunochemical identification of the serine protease inhibitor alpha 1-antichymotrypsin in the brain amyloid deposits of Alzheimer’s disease. *Cell* 1988;52:487–501.

4. Atwood CS, Moir RD, Huang X, et al. Dramatic aggregation of Alzheimer abeta by Cu(II) is induced by conditions representing physiological acidosis. *J Biol Chem* 1998;273:12817–12826.

5. Masters CL, Simms G, Weinman NA, et al. Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proc Natl Acad Sci USA* 1985;82:4245–4249.

6. Kang J, Lemaire HG, Unterbeck A, et al. The precursor of Alzheimer’s disease amyloid A4 protein resembles a cell-surface receptor. *Nature* 1987;325:733–736.

7. Tanzi RE, Gusella JF, Watkins PC, et al. Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. *Science* 1987;235:880–884.

8. Lemaire HG, Salbaum JM, Multhaup G, et al. The PreA4(695) precursor protein of Alzheimer’s disease A4 amyloid is encoded by 16 exons. *Nucleic Acids Res* 1989;17:517–522.

9. Kitaguchi N, Takahashi Y, Tokushima Y, et al. Novel precursor of Alzheimer’s disease amyloid protein shows protease inhibitory activity. *Nature* 1988;331:530–532.

10. Ponte P, Gonzalez-DeWhitt P, Schilling J, et al. A new A4 amyloid mRNA contains a domain homologous to serine proteinase inhibitors. *Nature* 1988;331:525–527.

11. Esch FS, Keim PS, Beattie EC, et al. Cleavage of amyloid beta peptide during constitutive processing of its precursor. *Science* 1990;248:1122–1124.

12. Citron M, Teplow DB, Selkoe DJ. Generation of amyloid beta protein from its precursor is sequence specific. *Neuron* 1995;14:661–670.

13. Seubert P, Vigo-Pelfrey C, Esch F, et al. Isolation and quantification of soluble Alzheimer's beta-peptide from biological fluids. *Nature* 1992;359:325–327.

14. Burdick D, Soreghan B, Kwon M, et al. Assembly and aggregation properties of synthetic Alzheimer's A4/beta amyloid peptide analogs. *J Biol Chem* 1992;267:546–554.

15. Buxbaum JD, Choi EK, Luo Y, et al. Calsenilin: a calcium-binding protein that interacts with the presenilins and regulates the levels of a presenilin fragment. *Nat Med* 1998;4:1177–1181.

16. Buxbaum JD, Liu KN, Luo Y, et al. Evidence that tumor necrosis factor alpha converting enzyme is involved in regulated alpha-secretase cleavage of the Alzheimer amyloid protein precursor. *J Biol Chem* 1998;273:27765–27767.

17. Parvathy S, Hussain I, Karran EH, et al. Cleavage of Alzheimer's amyloid precursor protein by alpha-secretase occurs at the surface of neuronal cells. *Biochemistry* 1999;38:9728–9734.

18. Lammich S, Kojro E, Postina R, et al. Constitutive and regulated alpha-secretase cleavage of Alzheimer's amyloid precursor protein by a disintegrin metalloprotease. *Proc Natl Acad Sci USA* 1999;96:3922–3927.

19. Hussain I, Powell D, Howlett DR, et al. Identification of a novel aspartic protease (Asp 2) as beta-secretase. *Mol Cell Neurosci* 1999;14:419–427.

20. Sinha S, Anderson JP, Barbour R, et al. Purification and cloning of amyloid precursor protein beta-secretase from human brain. *Nature* 1999;402:537–540.

21. Vassar R, Bennett BD, Babu-Khan S, et al. Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* 1999;286:735–741.

22. Yan R, Bienkowski MJ, Shuck ME, et al. Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity. *Nature* 1999;402:533–537.

23. De Strooper B, Saftig P, Craessaerts K, et al. Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein [see Comments]. *Nature* 1998;391:387–390.

24. Busciglio J, Gabuzda DH, Matsudaira P, et al. Generation of beta-amyloid in the secretory pathway in neuronal and nonneuronal cells. *Proc Natl Acad Sci USA* 1993;90:2092–2096.

25. Haass C, Schlossmacher MG, Hung AY, et al. Amyloid beta-peptide is produced by cultured cells during normal metabolism [see Comments]. *Nature* 1992;359:322–325.

26. Shoji M, Golde TE, Ghiso J, et al. Production of the Alzheimer amyloid beta protein by normal proteolytic processing. *Science* 1992;258:126–129.

27. Cook DG, Forman MS, Sung JC, et al. Alzheimer's A beta(1–42) is generated in the endoplasmic reticulum/intermediate compartment of NT2N cells. *Nat Med* 1997;3:1021–1023.

28. Skovronsky DM, Doms RW, Lee VM. Detection of a novel intraneuronal pool of insoluble amyloid beta protein that accumulates with time in culture. *J Cell Biol* 1998;141:1031–1039.

29. Mattson MP. Cellular actions of beta-amyloid precursor protein and its soluble and fibrillogenic derivatives. *Physiol Rev* 1997;77:1081–1132.

30. Dyrks T, Dyrks E, Hartmann T, et al. Amyloidogenicity of beta A4 and beta A4-bearing amyloid protein precursor fragments by metal-catalyzed oxidation. *J Biol Chem* 1992;267:18210–18217.

31. Yankner BA, Duffy LK, Kirschner DA. Neurotrophic and neurotoxic effects of amyloid beta protein: reversal by tachykinin neuropeptides. *Science* 1990;250:279–282.

32. Storey E, Cappai R. The amyloid precursor protein of Alzheimer's disease and the Abeta peptide. *Neuropathol Appl Neurobiol* 1999;25:81–97.

33. Mattson MP, Cheng B, Culwell AR, et al. Evidence for excitoprotective and intraneuronal calcium-regulating roles for secreted forms of the beta-amyloid precursor protein. *Neuron* 1993;10:243–254.

34. Greenberg SM, Koo EH, Selkoe DJ, et al. Secreted beta-amyloid precursor protein stimulates mitogen-activated protein kinase and enhances tau phosphorylation. *Proc Natl Acad Sci USA* 1994;91:7104–7108.

35. Borg JP, Ooi J, Levy E, et al. The phosphotyrosine interaction domains of X11 and FE65 bind to distinct sites on the YENPTY motif of amyloid precursor protein. *Mol Cell Biol* 1996;16:6229–6241.

36. Fiore F, Zambrano N, Minopoli G, et al. The regions of the Fe65 protein homologous to the phosphotyrosine interaction/phosphotyrosine binding domain of Shc bind the intracellular domain of the Alzheimer's amyloid precursor protein. *J Biol Chem* 1995;270:30853–30856.

37. Howell BW, Lanier LM, Frank R, et al. The disabled 1 phosphotyrosine-binding domain binds to the internalization signals of transmembrane glycoproteins and to phospholipids. *Mol Cell Biol* 1999;19:5179–5188.

38. Nishimoto I, Okamoto T, Matsuura Y, et al. Alzheimer amyloid protein precursor complexes with brain GTP-binding protein G(o) [see Comments]. *Nature* 1993;362:75–79.

39. Chow N, Korenberg JR, Chen XN, et al. APP-BP1, a novel protein that binds to the carboxyl-terminal region of the amyloid precursor protein. *J Biol Chem* 1996;271:11339–11346.

40. Caporaso GL, Gandy SE, Buxbaum JD, et al. Protein phosphorylation regulates secretion of Alzheimer beta/A4 amyloid precursor protein. *Proc Natl Acad Sci USA* 1992;89:3055–3059.

41. Dyrks T, Monning U, Beyreuther K, et al. Amyloid precursor protein secretion and beta A4 amyloid generation are not mutually exclusive. *FEBS Lett* 1994;349:210–214.

42. Buxbaum JD, Koo EH, Greengard P. Protein phosphorylation inhibits production of Alzheimer amyloid beta/A4 peptide. *Proc Natl Acad Sci USA* 1993;90:9195–9198.

43. Gabuzda D, Busciglio J, Yankner BA. Inhibition of beta-amyloid production by activation of protein kinase C. *J Neurochem* 1993;61:2326–2329.

44. Hung AY, Selkoe DJ. Selective ectodomain phosphorylation and regulated cleavage of beta-amyloid precursor protein. *Embo J* 1994;13:534–542.

45. Buxbaum JD, Ruefli AA, Parker CA, et al. Calcium regulates processing of the Alzheimer amyloid protein precursor in a protein kinase C–independent manner. *Proc Natl Acad Sci USA* 1994;91:4489–4493.

46. Querfurth HW, Selkoe DJ. Calcium ionophore increases amyloid beta peptide production by cultured cells. *Biochemistry* 1994;33:4550–4561.

47. Nitsch RM, Slack BE, Wurtman RJ, et al. Release of Alzheimer amyloid precursor derivatives stimulated by activation of muscarinic acetylcholine receptors. *Science* 1992;258:304–307.

48. Buxbaum JD, Oishi M, Chen HI, et al. Cholinergic agonists and interleukin 1 regulate processing and secretion of the Alzheimer beta/A4 amyloid protein precursor. *Proc Natl Acad Sci USA* 1992;89:10075–10078.

49. Checler F. Processing of the beta-amyloid precursor protein and its regulation in Alzheimer's disease. *J Neurochem* 1995;65:1431–1444.

50. De Strooper B, Van Leuven F, Van den Berghe H. Alpha 2-macroglobulin and other proteinase inhibitors do not interfere with the secretion of amyloid precursor protein in mouse neuroblastoma cells. *FEBS Lett* 1992;308:50–53.

51. De Strooper B, Umans L, Van Leuven F, et al. Study of the synthesis and secretion of normal and artificial mutants of murine amyloid precursor protein (APP): cleavage of APP occurs in a late compartment of the default secretion pathway. *J Cell Biol* 1993;121:295–304.

52. Sambamurti K, Refolo LM, Shioi J, et al. The Alzheimer's amyloid precursor is cleaved intracellularly in the trans-Golgi network or in a post-Golgi compartment. *Ann NY Acad Sci* 1992;674:118–128.

53. Sambamurti K, Shioi J, Anderson JP, et al. Evidence for intracellular cleavage of the Alzheimer's amyloid precursor in PC12 cells. *J Neurosci Res* 1992;33:319–329.

54. Nordstedt C, Caporaso GL, Thyberg J, et al. Identification of the Alzheimer beta/A4 amyloid precursor protein in clathrin-coated vesicles purified from PC12 cells. *J Biol Chem* 1993;268:608–612.

55. Koo EH, Squazzo SL. Evidence that production and release of amyloid beta-protein involves the endocytic pathway. *J Biol Chem* 1994;269:17386–17389.

56. Haass C, Koo EH, Mellon A, et al. Targeting of cell-surface beta-amyloid precursor protein to lysosomes: alternative processing into amyloid-bearing fragments. *Nature* 1992;357:500–503.

57. Sisodia SS. Beta-amyloid precursor protein cleavage by a membrane-bound protease. *Proc Natl Acad Sci USA* 1992;89:6075–6079.

58. Koo EH, Squazzo SL, Selkoe DJ, et al. Trafficking of cell-surface amyloid beta-protein precursor. I. Secretion, endocytosis and recycling as detected by labeled monoclonal antibody. *J Cell Sci* 1996;109:991–998.

59. Golde TE, Estus S, Younkin LH, et al. Processing of the amyloid protein precursor to potentially amyloidogenic derivatives. *Science* 1992;255:728–730.

60. Selkoe DJ, Yamazaki T, Citron M, et al. The role of APP processing and trafficking pathways in the formation of amyloid beta-protein. *Ann NY Acad Sci* 1996;777:57–64.

61. Hartmann T, Bieger SC, Bruhl B, et al. Distinct sites of intracellular production for Alzheimer's disease A beta40/42 amyloid peptides. *Nat Med* 1997;3:1016–1020.

62. Haass C, Selkoe DJ. Cellular processing of beta-amyloid precursor protein and the genesis of amyloid beta-peptide. *Cell* 1993;75:1039–1042.

63. Weidemann A, Konig G, Bunke D, et al. Identification, biogenesis, and localization of precursors of Alzheimer's disease A4 amyloid protein. *Cell* 1989;57:115–126.

64. Selkoe DJ. Amyloid beta-protein precursor: new clues to the genesis of Alzheimer's disease. *Curr Opin Neurobiol* 1994;4:708–716.

65. Lai A, Sisodia SS, Trowbridge IS. Characterization of sorting signals in the beta-amyloid precursor protein cytoplasmic domain. *J Biol Chem* 1995;270:3565–3573.

66. Yamazaki T, Koo EH, Selkoe DJ. Trafficking of cell-surface amyloid beta-protein precursor. II. Endocytosis, recycling and lysosomal targeting detected by immunolocalization. *J Cell Sci* 1996;109:999–1008.

67. Kidd M. Alzheimer's disease: an electron microscopical study. *Brain* 1964;87:307–320.

68. Grundke-Iqbal I, Iqbal K, Tung YC, et al. Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc Natl Acad Sci USA* 1986;83:4913–4917.

69. Morishima-Kawashima M, Hasegawa M, Takio K, et al. Hyperphosphorylation of tau in PHF. *Neurobiol Aging* 1995;16:365–371; discussion 371–380.

70. Ishiguro K, Shiratsuchi A, Sato, S, et al. Glycogen synthase kinase 3 beta is identical to tau protein kinase I generating several epitopes of paired helical filaments. *FEBS Lett* 1993;325:167–172.

71. Paudel HK, Zwiers H, Wang JH. Phosphorylase kinase phosphorylates the calmodulin-binding regulatory regions of neuronal tissue-specific proteins B-50 (GAP-43) and neurogranin. *J Biol Chem* 1993;268:6207–6213.

72. Drewes G, Lichtenberg-Kraag B, Doring F, et al. Mitogen activated protein (MAP) kinase transforms tau protein into an Alzheimer-like state. *Embo J* 1992;11:2131–2138.

73. Sironi JJ, Yen SH, Gondal JA, et al. Ser-262 in human recombinant tau protein is a markedly more favorable site for phosphorylation by CaMKII than PKA or PhK. *FEBS Lett* 1998;436:471–475.

74. Singh TJ, Grundke-Iqbal I, Iqbal K. Phosphorylation of tau protein by casein kinase-1 converts it to an abnormal Alzheimer-like state. *J Neurochem* 1995;64:1420–1423.

75. Jicha GA, Weaver C, Lane E, et al. cAMP-dependent protein kinase phosphorylations on tau in Alzheimer's disease. *J Neurosci* 1999;19:7486–7494.

76. Gong CX, Shaikh S, Wang JZ, et al. Phosphatase activity toward abnormally phosphorylated tau: decrease in Alzheimer disease brain. *J Neurochem* 1995;65:732–738.

77. Goedert M, Crowther RA, Spillantini MG. Tau mutations cause frontotemporal dementias. *Neuron* 1998;21:955–958.

78. St. George-Hyslop S. Role of genetics in test of genotype, status, and disease progression in early-onset Alzheimer's disease. *Neurobiol Aging* 1998;19:133–137.

79. Goate A, Chartier-Harlin MC, Mullan M, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 1991;349:704–706.

80. Tanzi RE, St George-Hyslop PH, Haines JL, et al. The genetic defect in familial Alzheimer's disease is not tightly linked to the amyloid beta-protein gene. *Nature* 1987;329:156–157.

81. Sherrington R, Rogaev EI, Liang Y, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 1995;375:754–760.

82. Mann DM. The pathological association between Down syndrome and Alzheimer disease. *Mech Ageing Dev* 1988;43:99–136.

83. Beyreuther K, Pollwein P, Multhaup G, et al. Regulation and expression of the Alzheimer's beta/A4 amyloid protein precursor in health, disease, and Down's syndrome. *Ann NY Acad Sci* 1993;695:91–102.

84. Hol EM, Neubauer A, de Kleijn DP, et al. Dinucleotide deletions in neuronal transcripts: a novel type of mutation in non-familial Alzheimer's disease and Down syndrome patients. *Prog Brain Res* 1998;117:379–395.

85. van Leeuwen FW, de Kleijn DP, van den Hurk HH, et al. Frameshift mutants of beta amyloid precursor protein and ubiquitin-B in Alzheimer's and Down patients [see Comments]. *Science* 1998;279:242–247.

86. Murrell J, Farlow M, Ghetti B, et al. A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease. *Science* 1991;254:97–99.

87. Murrell J, Hake A, Quaid K, et al. Early-onset alzheimer disease caused by a new mutation (V717L) in the amyloid precursor protein gene. *Arch Neurol* 2000;57:885–887.

88. Chartier-Harlin MC, Crawford F, Houlden H, et al. Early-onset Alzheimer's disease caused by mutations at codon 717 of the beta-amyloid precursor protein gene. *Nature* 1991;353:844–846.

89. Hendriks L, van Duijn CM, Cras P, et al. Presenile dementia and cerebral haemorrhage linked to a mutation at codon 692 of the beta-amyloid precursor protein gene. *Nat Genet* 1992;1:218–221.

90. Mullan M, Crawford F, Axelman K, et al. A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. *Nat Genet* 1992;1:345–347.

91. Eckman CB, Mehta ND, Crook R, et al. A new pathogenic mutation in the APP gene (I716V) increases the relative proportion of A beta 42(43). *Hum Mol Genet* 1997;6:2087–2089.

92. Ancolio K, Dumanchin C, Barelli H, et al. Unusual phenotypic alteration of beta amyloid precursor protein (betaAPP) maturation by a new Val-715 → Met betaAPP-770 mutation responsible for probable early-onset Alzheimer's disease. *Proc Natl Acad Sci USA* 1999;96:4119–4124.

93. Levy E, Carman MD, Fernandez-Madrid IJ, et al. Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. *Science* 1990;248:1124–1126.

94. Nilsberth C, Westlind-Danielsson A, Eckman CB, et al. The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced Abeta protofibril formation. *Nat Neurosci* 2001;4:887–893.

95. Hardy J. The Alzheimer family of diseases: many etiologies, one pathogenesis? [Comment]. *Proc Natl Acad Sci USA* 1997;94:2095–2097.

96. Cai XD, Golde TE, Younkin SG. Release of excess amyloid beta protein from a mutant amyloid beta protein precursor. *Science* 1993;259:514–516.

97. Citron M, Oltersdorf T, Haass C, et al. Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production. *Nature* 1992;360:672–674.

98. Haass C, Hung AY, Selkoe DJ, et al. Mutations associated with a locus for familial Alzheimer's disease result in alternative processing of amyloid beta-protein precursor. *J Biol Chem* 1994;269:17741–17748.

99. Suzuki N, Cheung TT, Cai XD, et al. An increased percentage of long amyloid beta protein secreted by familial amyloid beta protein precursor (beta APP717) mutants. *Science* 1994;264:1336–1340.

100. Citron M, Vigo-Pelfrey C, Teplow DB, et al. Excessive production of amyloid beta-protein by peripheral cells of symptomatic and presymptomatic patients carrying the Swedish familial Alzheimer disease mutation. *Proc Natl Acad Sci USA* 1994;91:11993–11997.

101. Levy-Lahad E, Wijsman EM, Nemens E, et al. A familial Alzheimer's disease locus on chromosome 1. *Science* 1995;269:970–973.

102. Cruts M, Hendriks L, Van Broeckhoven C. The presenilin genes: a new gene family involved in Alzheimer disease pathology. *Hum Mol Genet* 1996;5:1449–1455.

103. Campion D, Flaman JM, Brice A, et al. Mutations of the presenilin I gene in families with early-onset Alzheimer's disease. *Hum Mol Genet* 1995;4:2373–2377.

104. Kamino K, Sato S, Sakaki Y, et al. Three different mutations of presenilin 1 gene in early-onset Alzheimer's disease families. *Neurosci Lett* 1996;208:195–198.

105. Campion D, Brice A, Dumanchin C, et al. A novel presenilin 1 mutation resulting in familial Alzheimer's disease with an onset age of 29 years. *Neuroreport* 1996;7:1582–1584.

106. Hardy J. Amyloid, the presenilins and Alzheimer's disease [see Comments]. *Trends Neurosci* 1997;20:154–159.

107. Wisniewski T, Dowjat WK, Buxbaum JD, et al. A novel Polish presenilin-1 mutation (P117L) is associated with familial Alzheimer's disease and leads to death as early as the age of 28 years. *Neuroreport* 1998;9:217–221.

108. Yasuda M, Maeda K, Hashimoto M, et al. A pedigree with a novel presenilin 1 mutation at a residue that is not conserved in presenilin 2. *Arch Neurol* 1999;56:65–69.

109. Alzheimer's Disease Collaborative Group. The structure of the presenilin 1 (S182) gene and identification of six novel mutations in early onset AD families. *Nat Genet* 1995;11:219–222.

110. Dumanchin C, Brice A, Campion D, et al. *De novo* presenilin 1 mutations are rare in clinically sporadic, early onset Alzheimer's disease cases: French Alzheimer's Disease Study Group. *J Med Genet* 1998;35:672–673.

111. Cruts M, Backhovens H, Wang SY, et al. Molecular genetic analysis of familial early-onset Alzheimer's disease linked to chromosome 14q24.3. *Hum Mol Genet* 1995;4:2363–2371.

112. Ramirez-Duenas MG, Rogaeva EA, Leal CA, et al. A novel Leu171Pro mutation in presenilin-1 gene in a Mexican family with early onset Alzheimer disease. *Ann Genet* 1998;41:149–153.

113. Kwok JB, Taddei K, Hallupp M, et al. Two novel (M233T and R278T) presenilin-1 mutations in early-onset Alzheimer's disease pedigrees and preliminary evidence for association of presenilin-1 mutations with a novel phenotype. *Neuroreport* 1997;8:1537–1542.

114. Ikeda M, Sharma V, Sumi SM, et al. The clinical phenotype of two missense mutations in the presenilin I gene in Japanese patients. *Ann Neurol* 1996;40:912–917.

115. Forsell C, Froelich S, Axelman K, et al. A novel pathogenic mutation (Leu262Phe) found in the presenilin 1 gene in early-onset Alzheimer's disease. *Neurosci Lett* 1997;234:3–6.

116. Kamimura K, Tanahashi H, Yamanaka H, et al. Familial Alzheimer's disease genes in Japanese. *J Neurol Sci* 1998;160:76–81.

117. Aldudo J, Bullido MJ, Arbizu T, et al. Identification of a novel mutation (Leu282Arg) of the human presenilin 1 gene in Alzheimer's disease. *Neurosci Lett* 1998;240:174–176.

118. Crook R, Verkkoniemi A, Perez-Tur J, et al. A variant of Alzheimer's disease with spastic paraparesis and unusual plaques due to deletion of exon 9 of presenilin 1. *Nat Med* 1998;4:452–455.

119. Perez-Tur J, Croxton R, Wright K, et al. A further presenilin 1 mutation in the exon 8 cluster in familial Alzheimer's disease. *Neurodegeneration* 1996;5:207–212.

120. Rogaev EI, Sherrington R, Rogaeva EA, et al. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature* 1995;376:775–778.

121. De Strooper B, Beullens M, Contreras B, et al. Phosphorylation, subcellular localization, and membrane orientation of the Alzheimer's disease-associated presenilins. *J Biol Chem* 1997;272:3590–3598.

122. Doan A, Thinakaran G, Borchelt DR, et al. Protein topology of presenilin 1. *Neuron* 1996;17:1023–1030.

123. Lehmann S, Chiesa R, Harris DA. Evidence for a six-transmembrane domain structure of presenilin 1. *J Biol Chem* 1997;272:12047–12051.

124. Li X, Greenwald I. Membrane topology of the *C. elegans* SEL-12 presenilin. *Neuron* 1996;17:1015–1021.

125. Li X, Greenwald I. Additional evidence for an eight-transmembrane-domain topology for *Caenorhabditis elegans* and human presenilins. *Proc Natl Acad Sci USA* 1998;95:7109–7114.

126. Dewji NN, Singer SJ. The seven-transmembrane spanning topography of the Alzheimer disease-related presenilin proteins in the plasma membranes of cultured cells. *Proc Natl Acad Sci USA* 1997;94:14025–14030.

127. Pradier L, Carpentier N, Delalonde L, et al. Mapping the APP/presenilin (PS) binding domains: the hydrophilic N-terminus of PS2 is sufficient for interaction with APP and can displace APP/PS1 interaction. *Neurobiol Dis* 1999;6:43–55.

128. Seeger M, Nordstedt C, Petanceska S, et al. Evidence for phosphorylation and oligomeric assembly of presenilin 1. *Proc Natl Acad Sci USA* 1997;94:5090–5094.

129. Walter J, Grunberg J, Capell A, et al. Proteolytic processing of the Alzheimer disease-associated presenilin-1 generates an *in vivo* substrate for protein kinase C. *Proc Natl Acad Sci USA* 1997;94:5349–5354.

130. Podlisny MB, Citron M, Amarante P, et al. Presenilin proteins undergo heterogeneous endoproteolysis between Thr291 and Ala299 and occur as stable N- and C-terminal fragments in normal and Alzheimer brain tissue. *Neurobiol Dis* 1997;3:325–337.

131. Thinakaran G, Borchelt DR, Lee MK, et al. Endoproteolysis of presenilin 1 and accumulation of processed derivatives *in vivo. Neuron* 1996;17:181–190.

132. Ratovitski T, Slunt HH, Thinakaran G, et al. Endoproteolytic processing and stabilization of wild-type and mutant presenilin. *J Biol Chem* 1997;272:24536–24541.

133. Kim TW, Pettingell WH, Hallmark OG, et al. Endoproteolytic cleavage and proteasomal degradation of presenilin 2 in transfected cells. *J Biol Chem* 1997;272:11006–11010.

134. Steiner H, Capell A, Pesold B, et al. Expression of Alzheimer's disease-associated presenilin-1 is controlled by proteolytic degradation and complex formation. *J Biol Chem* 1998;273:32322–32331.

135. Kim TW, Pettingell WH, Jung YK, et al. Alternative cleavage of Alzheimer-associated presenilins during apoptosis by a caspase-3 family protease. *Science* 1997;277:373–376.

136. Saura CA, Tomita T, Davenport F, et al. Evidence that intramolecular associations between presenilin domains are obligatory for endoproteolytic processing. *J Biol Chem* 1999;274:13818–13823.

137. Kovacs DM, Fausett HJ, Page KJ, et al. Alzheimer-associated presenilins 1 and 2: neuronal expression in brain and localization to intracellular membranes in mammalian cells. *Nat Med* 1996;2:224–229.

138. Thinakaran G. The role of presenilins in Alzheimer's disease. *J Clin Invest* 1999;104:1321–1327.

139. Annaert WG, Levesque L, Craessaerts K, et al. Presenilin 1 controls gamma-secretase processing of amyloid precursor protein in pre-golgi compartments of hippocampal neurons. *J Cell Biol* 1999;147:277–294.

140. Johnston JA, Ward CL, Kopito RR. Aggresomes: a cellular response to misfolded proteins. *J Cell Biol* 1998;143:1883–1898.

141. Ray WJ, Yao M, Mumm J, et al. Cell surface presenilin-1 participates in the gamma-secretase–like proteolysis of Notch. *J Biol Chem* 1999;274:36801–36807.

142. Waragai M, Imafuku I, Takeuchi S, et al. Presenilin 1 binds to amyloid precursor protein directly. *Biochem Biophys Res Commun* 1997;239:480–482.

143. Xia W, Zhang J, Perez R, et al. Interaction between amyloid precursor protein and presenilins in mammalian cells: implications for the pathogenesis of Alzheimer disease. *Proc Natl Acad Sci USA* 1997;94:8208–8213.

144. Thinakaran G, Regard JB, Bouton CM, et al. Stable association of presenilin derivatives and absence of presenilin interactions with APP. *Neurobiol Dis* 1998;4:438–453.

145. Scheuner D, Eckman C, Jensen M, et al. Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased *in vivo* by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. *Nat Med* 1996;2:864–870.

146. Iwatsubo T. Abeta42, presenilins, and Alzheimer's disease. *Neurobiol Aging* 1998;19[Suppl 1]:S11–S13.

147. Lemere CA, Lopera F, Kosik KS, et al. The E280A presenilin 1 Alzheimer mutation produces increased A beta 42 deposition and severe cerebellar pathology. *Nat Med* 1996;2:1146–1150.

148. Borchelt DR, Thinakaran G, Eckman CB, et al. Familial Alzheimer's disease-linked presenilin 1 variants elevate Abeta1-42/1–40 ratio *in vitro* and *in vivo. Neuron* 1996;17:1005–1013.

149. Citron M, Eckman CB, Diehl TS, et al. Additive effects of PS1 and APP mutations on secretion of the 42-residue amyloid beta-protein. *Neurobiol Dis* 1998;5:107–116.

150. Duff K, Eckman C, Zehr C, et al. Increased amyloid-beta42(43) in brains of mice expressing mutant presenilin 1. *Nature* 1996;383:710–713.

151. Tomita T, Maruyama K, Saido TC, et al. The presenilin 2 mutation (N141I) linked to familial Alzheimer disease (Volga German families) increases the secretion of amyloid beta protein ending at the 42nd (or 43rd) residue. *Proc Natl Acad Sci USA* 1997;94:2025–2030.

152. Naruse S, Thinakaran G, Luo JJ, et al. Effects of PS1 deficiency on membrane protein trafficking in neurons. *Neuron* 1998;21:1213–1221.

153. Wolfe MS, Xia W, Ostaszewski BL, et al. Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity. *Nature* 1999;398:513–517.

154. Kopan R, Turner DL. The Notch pathway: democracy and aristocracy in the selection of cell fate. *Curr Opin Neurobiol* 1996;6:594–601.

155. Schroeter EH, Kisslinger JA, Kopan R. Notch-1 signalling requires ligand-induced proteolytic release of intracellular domain [see Comments]. *Nature* 1998;393:382–386.

156. Li YM, Lai MT, Xu M, et al. Presenilin 1 is linked with γ-secretase activity in the detergent solubilized state. *Proc Nat Acad Sci USA* 2000;97:6138–6143.

157. Brown MS, Goldstein JL. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* 1997;89:331–340.

158. Kimble J, Simpson P. The LIN-12/Notch signaling pathway and its regulation. *Annu Rev Cell Dev Biol* 1997;13:333–361.

159. Shen J, Bronson RT, Chen DF, et al. Skeletal and CNS defects in presenilin-1-deficient mice. *Cell* 1997;89:629–639.

160. Conlon RA, Reaume AG, Rossant J. Notch1 is required for the coordinate segmentation of somites. *Development* 1995;121:1533–1545.

161. De Strooper B, Annaert W, Cupers P, et al. A presenilin-1–dependent gamma-secretase-like protease mediates release of Notch intracellular domain. *Nature* 1999;398:518–522.

162. Levitan D, Greenwald I. Facilitation of lin-12–mediated signalling by sel-12, a *Caenorhabditis elegans* S182 Alzheimer's disease gene. *Nature* 1995;377:351–354.

163. Levitan D, Doyle TG, Brousseau D, et al. Assessment of normal and mutant human presenilin function in *Caenorhabditis elegans. Proc Natl Acad Sci USA* 1996;93:14940–14944.

164. Baumeister R, Leimer U, Zweckbronner I, et al. Human presenilin-1, but not familial Alzheimer's disease (FAD) mutants, facilitate *Caenorhabditis elegans* Notch signalling independently of proteolytic processing. *Genes Funct* 1997;1:149–159.

165. Levesque G, Yu G, Nishimura M, et al. Presenilins interact with armadillo proteins including neural-specific plakophilin-related protein and beta-catenin. *J Neurochem* 1999;72:999–1008.

166. Murayama M, Tanaka S, Palacino J, et al. Direct association of presenilin-1 with beta-catenin. *FEBS Lett* 1998;433:73–77.

167. Tesco G, Kim TW, Diehlmann A, et al. Abrogation of the presenilin 1/beta-catenin interaction and preservation of the heterodimeric presenilin 1 complex following caspase activation. *J Biol Chem* 1998;273:33909–33914.

168. Zhang W, Han SW, McKeel DW, et al. Interaction of presenilins with the filamin family of actin-binding proteins. *J Neurosci* 1998;18:914–922.

169. Shinozaki K, Maruyama K, Kume H, et al. The presenilin 2 loop domain interacts with the mu-calpain C-terminal region. *Int J Mol Med* 1998;1:797–799.

170. Dumanchin C, Czech C, Campion D, et al. Presenilins interact with Rab11, a small GTPase involved in the regulation of vesicular transport. *Hum Mol Genet* 1999;8:1263–1269.

171. Smine A, Xu X, Nishiyama K, et al. Regulation of brain G-protein go by Alzheimer's disease gene presenilin-1. *J Biol Chem* 1998;273:16281–16288.

172. Takashima A, Murayama M, Murayama O, et al. Presenilin 1 associates with glycogen synthase kinase-3beta and its substrate tau. *Proc Natl Acad Sci USA* 1998;95:9637–9641.

173. Vito P, Wolozin B, Ganjei J K, et al. Requirement of the familial Alzheimer's disease gene PS2 for apoptosis: opposing effect of ALG-3. *J Biol Chem* 1996;271:31025–31028.

174. Guo Q, Furukawa K, Sopher BL, et al. Alzheimer's PS-1 mutation perturbs calcium homeostasis and sensitizes PC12 cells to death induced by amyloid beta-peptide. *Neuroreport* 1996;8:379–383.

175. Guo Q, Sebastian L, Sopher BL, et al. Increased vulnerability of hippocampal neurons from presenilin-1 mutant knock-in mice to amyloid beta-peptide toxicity: central roles of superoxide production and caspase activation. *J Neurochem* 1999;72:1019–1029.

176. Katayama T, Imaizumi K, Sato N, et al. Presenilin-1 mutations downregulate the signalling pathway of the unfolded-protein response. *Nat Cell Biol* 1999;1:479–485.

177. Niwa M, Sidrauski C, Kaufman RJ, et al. A role for presenilin-1 in nuclear accumulation of Ire1 fragments and induction of the mammalian unfolded protein response. *Cell* 1999;99:691–702.

178. Mahley RW. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science* 1988;240:622–630.

179. Kounnas MZ, Moir RD, Rebeck GW, et al. LDL receptor-related protein, a multifunctional ApoE receptor, binds secreted beta-amyloid precursor protein and mediates its degradation. *Cell* 1995;82:331–340.

180. Krieger M, Herz J. Structures and functions of multiligand lipoprotein receptors: macrophage scavenger receptors and LDL receptor-related protein (LRP). *Annu Rev Biochem* 1994;63:601–637.

181. Corde, EH, Saunders AM, Strittmatter WJ, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families [see Comments]. *Science* 1993;261:921–923.

182. Gomez-Isla T, West HL, Rebeck GW, et al. Clinical and pathological correlates of apolipoprotein E epsilon 4 in Alzheimer's disease. *Ann Neurol* 1996;39:62–70.

183. Corder EH, Saunders AM, Risch NJ, et al. Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat Genet* 1994;7:180–184.

184. Schmechel DE, Saunders AM, Strittmatter WJ, et al. Increased amyloid beta-peptide deposition in cerebral cortex as a consequence of apolipoprotein E genotype in late-onset Alzheimer disease. *Proc Natl Acad Sci USA* 1993;90:9649–9653.

185. Bales KR, Verina T, Cummins DJ, et al. Apolipoprotein E is essential for amyloid deposition in the APP(V717F) transgenic mouse model of Alzheimer's disease. *Proc Natl Acad Sci USA* 1999;96:15233–15238.

186. Lendon CL, Martinez A, Behrens IM, et al. E280A PS-1 mutation causes Alzheimer's disease but age of onset is not modified by ApoE alleles. *Hum Mutat* 1997;10:186–195.

187. Van Broeckhoven C, Backhovens H, Cruts M, et al. APOE genotype does not modulate age of onset in families with chromosome 14 encoded Alzheimer's disease. *Neurosci Lett* 1994;169:179–180.

188. Blacker D, Wilcox MA, Laird NM, et al. Alpha-2 macroglobulin is genetically associated with Alzheimer disease. *Nat Genet* 1998;19:357–360.

189. Borth W. Alpha 2-macroglobulin, a multifunctional binding protein with targeting characteristics. *FASEB J* 1992;6:3345–3353.

190. Rebeck GW, Harr SD, Strickland DK, et al. Multiple, diverse senile plaque-associated proteins are ligands of an apolipoprotein E receptor, the alpha 2-macroglobulin receptor/low- density-lipoprotein receptor-related protein. *Ann Neurol* 1995;37:211–217.

191. Du Y, Ni B, Glinn M, et al. alpha2-Macroglobulin as a beta-amyloid peptide-binding plasma protein. *J Neurochem* 1997;69:299–305.

192. Narita M, Holtzman DM, Schwartz AL, et al. Alpha2-macroglobulin complexes with and mediates the endocytosis of beta-amyloid peptide via cell surface low-density lipoprotein receptor-related protein. *J Neurochem* 1997;69:1904–1911.

193. Kamboh MI, Sanghera DK, Ferrell RE, et al. APOE*4-associated Alzheimer's disease risk is modified by alpha 1-antichymotrypsin polymorphism [published erratum appears in *Nat Genet* 1995;11:104]. *Nat Genet* 1995;10:486–488.

194. Montoya SE, Aston CE, DeKosky ST, et al. Bleomycin hydrolase is associated with risk of sporadic Alzheimer's disease [Letter] [published erratum appears in *Nat Genet* 1998;19:404]. *Nat Genet* 1998;18:211–212.

195. Price DL, Tanzi RE, Borchelt DR, et al. Alzheimer's disease: genetic studies and transgenic models. *Annu Rev Genet* 1998;32:461–493.

196. Hoyer S. Risk factors for Alzheimer's disease during aging. Impacts of glucose/energy metabolism. *J Neural Transm Suppl* 1998;54:187–194.

197. Honjo H, Ogino Y, Naitoh K, et al. *In vivo* effects by estrone sufate on the central nervous system: senile dementia. *J Steroid Biochem* 1989;41:633–635.

198. Sacks FM, Walsh BW. Sex hormones and lipoprotein metabolism. *Curr Opin Lipidol* 1994;5:236–240.

199. Srivastava R, Bhasin N, Srivastava N. ApoE gene expression in various tissues of mouse adn regulation by estrogen. *Biochem Mol Biol Int* 1996;38:91–101.

200. Behl C, Skutella T, Lezoualc'h F, et al. Neuroprotection against oxidative stress by estrogens: structure-activity relationship. *Mol Pharmacol* 1997;51:535–541.

201. Behl C, Widmann M, Trapp T, et al. 17-Beta estradiol protects neurons from oxidative stress-induced cell death *in vitro*. *Biochem Biophys Res Commun* 1995;216:473–482.

202. Jaffe AB, Toran-Allerand CD, Greengard P, et al. Estrogen regulates metabolism of Alzheimer amyloid beta precursor protein. *J Biol Chem* 1994;269:13065–13068.

84

Transgenic Mouse Models of Alzheimer Disease

Karen Duff

Karen Duff: Nathan Kline Institute, Orangeburg, New York.

GENERAL PRINCIPLES

Part of "84 - Transgenic Mouse Models of Alzheimer Disease "

Transgenic models of Alzheimer disease (AD) continue to gain credibility as more features of the human disease are shown to be represented in the mice. Tau pathology and extensive cell loss are still not seen, however. Despite these shortcomings, the mice are excellent models of amyloidosis, and this field of study has been highly informative in advancing our understanding of *in vivo* responses to amyloid insult and the mechanism by which genetic lesions may cause AD. Because many investigators believe that amyloid (or its precursors) is critical to initiation of the disease, the mice are now being used to test therapeutic agents that may have utility in patients with AD. In addition, new models that address the issue of tau pathogenesis have been created that may help to explain the relative contribution of tau and amyloid to the pathogenesis of AD.

AD is a progressive neurodegenerative disease. Most cases of AD occur sporadically, but familial forms of the disease have been most widely studied because of the insight they give into disease origin. Genetic causes of the disease are heterogeneous and include mutations or variants in several genes including the amyloid precursor protein *(APP)* gene, the presenilins *(PS)*, and apolipoprotein E *(APOE)* (reviewed in ref. 1). The disease phenotype is remarkably consistent and includes the accumulation of ß-amyloid (Aß) and its deposition into senile plaques, the formation of tau-containing tangles, reactive gliosis, inflammation and an immune response, neurodegeneration, cholinergic deficit, and cognitive impairment.

Report of the first transgenic mouse to develop a robust AD-related phenotype was published in 1995 by the Exemplar/Athena Neuroscience group (2). This line (known as PDAPP) overexpresses mutant *APP* at high enough levels to generate sufficient Aß for extracellular deposits (plaques) to form in relevant regions of the brain. In 1996, a second line (Tg2576), created by Karen Hsiao and colleagues, also made sufficient amyloid for deposits to form, and, in addition, this mouse showed age-related cognitive impairment (3). Subsequently, other cDNA mice (4 ,5) and mice overexpressing genomic constructs (6) have also been shown to form amyloid in old age. Several research groups have created transgenic mice that overexpress mutant presenilin (7 ,8 and 9), but these mice do not show amyloid deposition, most likely because they have insufficient levels of the Aß peptide.

Most of the current work on the mice focuses on cellular response to amyloid accumulation and its relevance to AD. Recently, the PDAPP line of mice was used to test the feasibility of modulating amyloid levels by immunization with Aß (10). The results of this experiment suggested that amyloid modulation is indeed possible and that some of the secondary effects of amyloidosis (gliosis and neuritic changes) can be prevented. This work opens up a new direction in amyloid research and may well have significant impact on the development of human therapies.

RECENT ADVANCES IN PHENOTYPE ASSESSMENT IN TRANSGENIC MODELS OF AD

Part of "84 - Transgenic Mouse Models of Alzheimer Disease "

Amyloidosis

Several studies aimed to modulate the amyloid phenotype by crossing in other transgenes such as *PS1* or *TGF-ß* (transforming growth factor-ß). The studies showed that when a *PS1* mutant mouse was crossed to an *APP* overexpressing mouse, the levels of Aß42 (43) were increased in the double transgenic mice, and this elevation had a profound influence on the age at which amyloid deposition could first be detected (6 ,11 ,12). In one cross, the age at which amyloid deposits were first identified was reduced from 9 to 12 months to 10 to 12 weeks (13) which is the earliest age at which amyloid has been reported. When *TGF-ß* cDNA mice were crossed into an *APP* overexpressing line, amyloid deposition was again accelerated, and deposition was far more prominent in the vasculature (14). Apart from showing

that these pathways interact, one outcome of the crossed-mouse studies is to enhance, and in some cases to modify, the phenotype, thus providing us with better models.

Presenilin Transgenics

In terms of the amyloid phenotype seen in AD, the most significant phenotype in the mutant *PS* transgenics continues to be the specific elevation of Aß1-42(43) (7 ,8 and 9). Mutant *PS2* transgenics have been created that also show elevation in Aß1-42(43) (15), a finding that strengthens the argument that APP and the presenilins interact, either directly, as suggested by Wolfe et al. (16), or indirectly, and *PS* mutations cause AD through APP/Aß modulation. More recently, several studies on cDNA and targeted knock-in *PS1* mice have shown that mutant *PS* mice show deficits in calcium homeostasis (17 ,18 and 19) and, in some models, impaired mitochondrial function (17).

Studies of knockout *PS1* animals have shown that PS1 plays an important role in development, because lack of the protein leads to a deficiency in somitogenesis during early embryogenesis that results in severe skeletal abnormalities and prenatal death (20 ,21). These abnormalities strongly resemble those seen in Notch knockout mice (22), and this observation, coupled with several studies *in vivo* and *in vitro*, suggests that Notch and PS1 interact in some way to affect normal cellular function that may downstream affect signaling, differentiation, and development.

PS1 has been strongly implicated in other signaling pathways because several potential components of signal transduction pathways have been identified as presenilin-interacting proteins. These include several proteins containing an armadillo repeat region, which is a 42 amino acid motif that has been identified in proteins involved in cell–cell adhesion, protein–protein interaction, and signal transduction. The best known of these is ß-catenin, and both ß- and its homologue, ð-catenin, have been shown to interact with PS1 (23). The effect of *PS1* mutations on ß-catenin stability and hence its downstream effects are controversial. In transgenic mice and human familial AD brain homogenates, for example, mutations in PS1 are linked to increased degradation of ß-catenin (24), whereas other studies have suggested no effect or increased activity for the wild-type and mutant PS1 protein (25 ,26 and 27). Quite how *PS1* mutations may lead to AD through a ß-catenin pathway is unclear, although in addition to the effect on Aß42, one action maybe on the action of GSK-ß and hence tau phosphorylation (for a review, see Alzheimer forum panel discussion at www.alzforum.org/members/forums/journals/catenin/index.html).

Neurodegeneration

Human AD brain shows extensive neurodegeneration, both in cholinergic neurons of the nucleus basalis (28) and in noncholinergic neurons throughout the cortex and hippocampus. Studies showed that fibrillar Aß peptides are toxic to neurons in culture (29), and the overproduction of human Aß in the brains of transgenic mice was therefore expected to cause extensive neurodegeneration. Several of the best characterized mouse models were examined for overt cell loss (30 ,31 and 32), but PDAPP, Tg2576, and the Tg2576/PS1 cross-mouse did not show significant cell loss, even though the amyloid burden exceeded 30%. One model was reported to show significant cell loss, but only in the hippocampus (32). Although overt cell loss is not seen in mice such as Tg2576/PS1, neurites in close proximity to amyloid deposits are severely dystrophic, and cellular disturbance is rife, as shown by extensive lysosomal activation (unpublished data). In addition, magnetic resonance imaging (MRI) revealed differences in the volume of structures such as the lateral ventricles and the corpus callosum between mice with and without amyloid, a finding that may reflect loss of neuropil, or shrinkage of cells rather than cell death *per se* (unpublished data). Advances in MRI in the study of human AD brain (33) suggest that it will soon be possible to image plaques directly *in vivo*, although contrasting agents may be required for mouse imaging because resolution is poor. Although the field is in its infancy, the application of functional and structural MRI to the analysis of models is predicted to have significant impact, especially because longitudinal analyses can be performed on the same mouse, including the effects of drug treatments.

Cholinergic Deficits

Although modulation of the cholinergic system has been a therapeutic treatment for AD dementia for many years, investigations into the response of cholinergic neurons to amyloid insult has only recently been studied in the transgenic mice. Immunohistochemical analysis has shown that cholinergic markers accumulate in swollen abnormal neurites around amyloid deposits in the cortex (4), and our own studies have shown that in the early stages of deposition, neurons in the nucleus basalis of depositing mice appear normal, but their projection areas in certain regions of the cortex show a significant reduction in synapse density and size (34). Significantly, this finding correlates with reduced cholinergic neurotransmitter activity (unpublished data). Further work to assess how cholinergic neurons in all areas of the brain respond to increasing amyloid burden and age is under way because cholinesterase inhibitors are currently considered to be valid therapeutic agents for AD.

Cognitive Impairments and Neurodegeneration

Recreating human cognitive impairment is perhaps the greatest challenge facing genetic engineers working on

transgenic models of human dementia. Mice are genetically less suitable to behavioral testing than rats because performance data on normal mice from different strains are often contradictory. Despite these reservations, most of the transgenic mice that form amyloid deposits have been tested for cognitive impairment. The PDAPP mouse (35 ,36), the Tg2576 mouse (3), and the Tg2576/PS1 cross-mouse (12) have all shown a deficit in tests of hippocampal dysfunction before amyloid deposits form, a finding that strongly suggests that overt amyloid accumulation or deposition is not responsible for this early cognitive impairment. Deficits in water maze performance that correlate with increasing age and amyloid burden and with decreasing long-term potentiation have been reported for Tg2576 mice (3 ,37).

Tau Pathology

One of the major deficits in the current AD mice is the lack of tau pathology. In humans, tau pathology takes the form of intracellular tangles of an abnormally phosphorylated form of the tau protein, which associates into paired helical filaments (PHFs). Both amyloid plaques and tau tangles are pathognomic features of human AD, and their relative contribution to the disease has long been disputed. The identification of AD-causing mutations in *APP* and the presenilin genes, however, adds weight to the amyloid-based hypothesis of pathogenesis, which assumes that tau abnormalities are a secondary lesion that form in response to amyloid accumulation. To examine this link, transgenic mice with extensive amyloid burden were examined for abnormal tau pathology by immunohistochemical analysis. This work showed that amyloid deposits in transgenic mice are ringed by dystrophic neurites that are immunoreactive for markers of phosphotau epitopes such as phosphoserine 202 (4 ,38). These epitopes are phosphorylated to some degree in normal brain, but they are hyperphosphorylated in AD brain. In the mice, it is not yet clear whether the immunoreactivity around deposits reflects local hyperphosphorylation of tau or simple elevation of tau protein levels in response to neuritic damage. In the human AD brain, subsets of neurons are also immunolabeled with antibodies to both tyrosine phosphotau and the signaling protein fyn, which is an src, nonreceptor tyrosine kinase (39). Co-IP has shown that the N-terminus of tau and the SH-3 domain of fyn interact directly, a finding suggesting that tau may be involved in signal transduction pathways (40). Interestingly, fyn binds another signaling protein, FAK (41), which is itself up-regulated by Aß (42). Ongoing work includes a study of how FAK, fyn, tau, and Aß interact in transgenic animals with and without elevated amyloid.

It is clear, however, that tau pathology does not develop further in the mutant *APP* and *PS* transgenic animals, a finding that suggests that either Aß/amyloid accumulation is not detrimental to tau or differences between mouse and human brain preclude the formation of pathogenic tau, as suggested by Geula et al. (43). This suggestion has been countered by reports that mouse tau is capable of forming PHF-like structures *in vitro* (44). We have created a line of transgenic mice that overexpress all isoforms of human tau under the control of the human tau promoter (45), in an effort to “humanize” the tau environment in amyloiddepositing mice by cross-breeding the different lines. Unfortunately, even in the presence of extensive amyloid, the mice do not form neurofibrillary pathology (unpublished data), a finding suggesting that the mice are either still deficient in another human component or that Aß accumulation does not stimulate tau pathogenesis, at least in mice.

Mutations in tau have been shown to cause frontal temporal lobe dementia (FTD-17) (46 ,47 and 48), which, in some cases, appears to result from an imbalance in tau isoform ratios (47). Transgenic mice that overexpress normal human tau isoforms and recreate this imbalance show a partly unexpected phenotype in that the mice develop hind limb weakness mainly resulting from spheroid accumulations composed of tau and neurofilament form, which are particularly prevalent in the axons of motor neurons (45 ,49 ,50). Although the tau does appear to be in an abnormal conformation, the relevance of the axopathy to human FTD-17 or the tauopathies is unclear. Mutant tau transgenic mice have been created, but their phenotype has not yet been described. It is likely that the creation of a battery of tau transgenic mice will provide long-awaited resources for the study of the normal and abnormal biology of this important cell component.

ACKNOWLEDGMENT

Part of "84 - Transgenic Mouse Models of Alzheimer Disease "

This work is supported by National Institute of Health grants AG146133, AG10485, AG17585, and AG17216.

REFERENCES

1. Cruts M, Van Broeckhoven C. Molecular genetics of Alzheimer’s disease. *Ann Med* 1998;30:560-565.

2. Games D, Adams D, Alessandrini R, et al. Alzheimer-type neuropathology in transgenic mice overexpressing V717F ß-amyloid precursor protein. *Nature* 1995;373:523–527.

3. Hsiao K, Chapman P, Nilsen S, et al. Correlative memory deficits, Aß elevation and amyloid plaques in transgenic mice. *Science* 1996;274:99—102.

4. Sturchler-Pierrat C, Abramowski D, Duke M, et al. Two amyloid precursor protein transgenic mouse models with Alzheimer disease-like pathology. *Proc Natl Acad Sci USA* 1997;94:13287–13292.

5. Nalbantoglu J, Tirado-Santiago G, Lahsaini A, et al. Impaired learning and LTP in mice expressing the carboxy terminus of the Alzheimer amyloid precursor protein. *Nature* 1997;387:500–505.

6. Lamb BA, Bardel KA, Kulnane LS, et al. Amyloid production and deposition in mutant amyloid precursor protein and presenilin-1 yeast artificial chromosome transgenic mice. *Nat Neurosci* 1999;8:695–697.

7. Duff K, Eckman C, Zehr C, et al. Increased amyloid-Aß42(43) in brains of mice expressing mutant presenilin 1. *Nature* 1996;383:710—713.

8. Borchelt DR, Thinakaran G, Eckman CB, et al. Familial Alzheimer's disease-linked presenilin 1 variants elevate Abeta1-42/1-40 ratio *in vitro* and *in vivo*. *Neuron* 1996;5:1005–1013.

9. Citron M, Westaway D, Xia W, et al. Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid ß-protein in both transfected cells and transgenic mice. *Nat Med* 1997;3:67–68.

10. Schenk D, Barbour R, Dunn W, et al. Immunization with amyloid-beta attenuates Alzheimer-disease–like pathology in the PDAPP mouse. *Nature* 1999;400:173–177.

11. Borchelt DR, Ratovitski T, van Lare J, et al. Accelerated amyloid deposition in the brains of transgenic mice co-expressing mutant presenilin 1 and amyloid precursor proteins. *Neuron* 1997;4:939—945.

12. Holcomb L, Gordon MN, McGowan E, et al. Accelerated Alzheimer-type phenotype in transgenic mice carrying both mutant amyloid precursor protein and presenilin 1 transgenes. *Nat Med* 1998;1:97–100.

13. McGowan E, Sanders S, Iwatsubo T, et al. Amyloid phenotype characterization of transgenic mice over-expressing both mutant amyloid precursor protein and mutant presenilin 1 transgenes. *Neurobiol Dis* 1999;6:231–244.

14. Wyss-Coray T, Masliah E, Mallory M, et al. Amyloidogenic role of cytokine TGF-beta1 in transgenic mice and in Alzheimer's disease. *Nature* 1997;389:603–606.

15. Oyama F, Sawamura N, Kobayashi K, et al. Mutant presenilin 2 transgenic mouse: effect on an age-dependent increase of amyloid beta-protein 42 in the brain. *J Neurochem* 1998;71:313–322.

16. Wolfe MS, Xia W, Ostaszewski BL, et al. Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity. *Nature* 1999;398:513–517.

17. Guo Q, Fu W, Sopher BL, et al. Increased vulnerability of hippocampal neurons to excitotoxic necrosis in presenilin-1 mutant knock-in mice. *Nat Med* 1999;5:101–106.

18. Parent A, Linden DJ, Sisodia SS, et al. Synaptic transmission and hippocampal long-term potentiation in transgenic mice expressing FAD-linked presenilin 1. *Neurobiol Dis* 1999;6:56–62.

19. Barrow P, Empson R, Jefferys J, et al. Functional phenotype in transgenic mice expressing mutant human presenilin-1. *Neurobiol Dis* 2000;7:119–126.

20. Wong PC, Zheng H, Chen H, et al. Presenilin 1 is required for Notch1 and DII1 expression in the paraxial mesoderm. *Nature* 1997;387:288–292.

21. Shen J, Bronson RT, Chen DF, et al. Skeletal and CNS defects in presenilin-1–deficient mice. *Cell* 1997;89:629–639.

22. Conlon RA, Reaume AG, Rossant J. Notch1 is required for the coordinate segmentation of somites. *Development* 1995;121:1533–1545.

23. Zhou J, Liyanage U, Medina M, et al. Presenilin 1 interaction in the brain with a novel member of the Armadillo family. *Neuroreport* 1997;8:2085–20890.

24. Zhang Z, Hartmann H, Do VM, et al. Destabilization of beta-catenin by mutations in presenilin-1 potentiates neuronal apoptosis. *Nature* 1998;395:698–702.

25. Murayama M, Tanaka S, Palacino J, et al. Direct association of presenilin-1 with beta-catenin. *FEBS Lett* 1998;433:73–77.

26. Kang DE, Soriano S, Frosch MP, et al. Presenilin 1 facilitates the constitutive turnover of beta-catenin: differential activity of Alzheimer's disease-linked PS1 mutants in the beta-catenin-signaling pathway. *J Neurosci* 1999;19:4229–4237.

27. Nishimura M, Yu G, Levesque G, et al. Presenilin mutations associated with Alzheimer disease cause defective intracellular trafficking of beta-catenin, a component of the presenilin protein complex. *Nat Med* 1999;5:164–169.

28. Whitehouse PJ, Price DL, Struble RG, et al. Alzheimer's disease and senile dementia: loss of neurons in the basal forebrain. *Science* 1982;215:1237–1239.

29. Lorenzo A, Yankner BA. ß-Amyloid neurotoxicity requires fibril formation and is inhibited by congo red. *Proc Natl Acad Sci USA* 1994;91:12243–12247.

30. Irizarry MC, Soriano F, McNamara M, et al. Aß deposition is associated with neuropil changes, but not with overt neuronal loss in the human amyloid precursor protein V717F (PDAPP) transgenic mouse. *J Neurosci* 1997;17:7053–7059.

31. Irizarry MC, McNamara M, Fedorchak K, et al. APPSw transgenic mice develop age-related Aß deposits and neuropil abnormalities, but no neuronal loss in CA1. *J Neuropathol Exp Neurol* 1997;56:965–973.

32. Calhoun ME, Wiederhold KH, Abramowski D, et al. Neuron loss in APP transgenic mice. *Nature* 1998;395:755–756.

33. Benveniste H, Einstein G, Kim KR, et al. Detection of neuritic plaques in Alzheimer's disease by magnetic resonance microscopy. *Proc Natl Acad Sci USA* 1999;96:14079–14084.

34. Wong TP, Debeir T, Duff K, et al. Reorganization of cholinergic terminals in the cerebral cortex and hippocampus in transgenic mice carrying mutated presenilin-1 and amyloid precursor protein transgenes. *J Neurosci* 1999;19:2706–2716.

35. Dodart J, Meziane H, Mathis C, et al. Memory and learning impairment precede amyloid deposition in the V717F PDAPP transgenic mouse [Abstract 636.5]. Society for Neurosciences, New Orleans, 1997.

36. Justice A, Motter R. Behavioral characterisation of PDAPP transgenic Alzheimer mice [Abstract 636.6]. Society for Neurosciences, New Orleans, 1997.

37. Chapman PF, White GL, Jones MW, et al. Impaired synaptic plasticity and learning in aged amyloid precursor protein transgenic mice. *Nat Neurosci* 1999;2:271–276.

38. Chen KS, Masliah E, Grajeda H, et al. Neurodegenerative Alzheimer-like pathology in PDAPP 717V→F transgenic mice. *Prog Brain Res* 1998;117:327–334.

39. Shirazi SK, Wood JG. The protein tyrosine kinase, fyn, in Alzheimer's disease pathology. *Neuroreport* 1993;4:435–437.

40. Lee G, Newman ST, Gard DL, et al. Tau interacts with src-family non-receptor tyrosine kinases. *J Cell Sci* 1998;111:3167–3177.

41. Cobb BS, Schaller MD, Leu TH, et al. Stable association of pp60src and pp59fyn with the focal adhesion-associated protein tyrosine kinase, pp125FAK. *Mol Cell Biol* 1994;14:147–155.

42. Zhang C, Qiu HE, Krafft GA, et al. A beta peptide enhances focal adhesion kinase/Fyn association in a rat CNS nerve cell line. *Neurosci Lett* 1996;211:187–190.

43. Geula C, Wu CK, Saroff D, et al. Aging renders the brain vulnerable to amyloid ß-protein neurotoxicity. *Nat Med* 1998;4:827–828.

44. Kampers T, Pangalos M, Geerts H, et al. Assembly of paired helical filaments from mouse tau: implications for the neurofibrillary pathology in transgenic mouse models for Alzheimer's disease. *FEBS Lett* 1999;451:39–44.

45. Duff K, Knight H, Refolo LM, et al. Characterization of pathology in transgenic mice over-expressing human genomic and cDNA tau transgenes. *Neurobiol Dis* 2000;7:87–98.

46. Spillantini MG, Murrell JR, Goedert M, et al. Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proc Natl Acad Sci USA* 1998;95:7737–41.

47. Hutton M, Lendon CL, Rizzu P, et al. Association of missense and 5′-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* 1998;393:702–705.

48. Goedert M, Spillantini MG, Crowther RA, et al. Tau gene mutation in familial progressive subcortical gliosis. *Nat Med* 1999;5:454–457.

49. Ishihara T, Hong M, Zhang B, et al. Age-dependent emergence and progression of a tauopathy in transgenic mice overexpressing the shortest human tau isoform. *Neuron* 1999;24:751–762.

50. Spittaels K, Van den Haute C, Van Dorpe J, et al. Prominent axonopathy in the brain and spinal cord of transgenic mice overexpressing four-repeat human tau protein. *Am J Pathol* 1999;155:2153–2165.

85

Cell and Molecular Neuropathology of Alzheimer Disease

Ralph A. Nixon

Ralph A. Nixon: Nathan Kline Institute, Orangeburg, New York.

Although forms of dementia arising late in life had been identified by Kraepelin and his colleagues in the 1800s, it was not until 1907 that Alois Alzheimer identified the presenile form of dementia with unique neuropathologic features that now bears his name. Alzheimer described a 51-year-old woman who presented with personality changes and soon developed progressively worsening memory loss, disorientation to time, and language disturbances but who had relatively normal neurologic function. Mental deterioration progressed, and she died 4½ years later. On autopsy, her brain showed clear evidence of cerebral atrophy. Under the microscope, Alzheimer discovered that many cortical neurons contained argyrophilic fibrous structures—neurofibrillary tangles (NFTs)—now known to be mainly composed of abnormal filamentous forms of the microtubule-associated protein, tau. Coexisting in the same affected brain regions were extracellular plaquelike lesions. These neuritic or “senile” plaques were later discovered to contain ß-amyloid, a fibrillar form of the Aß peptide, as their signature constituent. In 1910, Kraepelin proposed that this neuropathologic picture was pathognomic of a new presenile dementing disease and introduced the eponym *Alzheimer disease* (AD).

Nearly a century later, genetic analyses have shown AD to be the common neuropathologic outcome of different primary etiologic factors (1). Roughly one-third of patients with AD have a familial predisposition, with at least one other affected first-degree relative. In families with early-onset AD arising before the age of 65 years, a type accounting for 2% to 10% of all AD cases, the transmission pattern is consistent with an autosomal dominant disorder with age-dependent penetrance. To date, early-onset familial AD (FAD) has been linked to mutations of three different genes: the amyloid precursor protein *(APP)* gene and the presenilin 1 *(PS1)* and presenilin 2 *(PS2)* genes. Together, these mutations account for nearly half of the families with autosomal dominant early-onset AD (2). In FAD, a single gene defect, interacting with the brain aging process, causes the disease (reviewed in Chapter 83). However, in the other 90% of cases, designated *sporadic AD*, the emergence of disease is influenced by environmental factors as well as by multiple genes with either neuroprotective or disease-facilitating effects. Although all forms of AD, by definition, share common neuropathologic features, the metabolic antecedents of this pathology in sporadic forms of the disease are poorly understood. Much is known about the genetic forms of AD, yet issues as fundamental as the anatomic and cellular substrate of cognitive decline, the toxic cascades mediating cell degeneration, and the roles of Aß, ß-amyloid, and tau in the neurodegenerative process are just now being resolved. In this chapter, a discussion of Alzheimer neuropathology provides the starting point for understanding how the diagnosis of AD is made and how clinical symptoms arise and progress. Later sections address the neuroanatomic and cellular basis for dementia and the molecular events that lead to neuropathologic lesions and, ultimately, to the death of neurons. A consideration of current hypotheses on the evolution of cellular pathology identifies common features that help to reconcile the differing views of disease pathogenesis. New studies are beginning to shed light on underlying mechanisms in prevalent sporadic forms of AD, and their review complements a brief discussion of pathogenesis in the rare familial forms here and a more detailed consideration in Chapter 83 .

DIAGNOSTIC FEATURES OF ALZHEIMER DISEASE

Part of "85 - Cell and Molecular Neuropathology of Alzheimer Disease "

AD remains a diagnosis based on its histopathologic features. By the time AD can be clinically diagnosed by DSM-IV criteria, neuropathologic features are already quite extensive in some regions of the brain (3 ,4). Patients with Down

syndrome, for example, develop probable AD by neuropathologic criteria years before clinical dementia can be detected (5). According to the most widely accepted neuropathologic criteria (those of the Consortium to Establish a Registry for Alzheimer's Disease or CERAD), a diagnosis of definite AD can be made in a demented patient when neuritic plaques reach a requisite number, adjusted for age, in the most severely affected regions of the neocortex (6) (e.g., superior and middle temporal gyrus, middle frontal gyrus, inferior parietal lobule, hippocampal/entorhinal cortex, and midbrain) in the absence of other neuropathologic lesions likely to cause dementia. More stringent research criteria require semiquantitative estimates of both neuritic plaques and NFTs (7 ,8 and 9). In the most recent guidelines developed by the National Institute of Health and the Reagan Institute (9), topographic staging of NFT accumulation (7) is incorporated into the criteria and significantly increases diagnostic precision (10).

Neuritic plaques are complex spheric lesions of varying sizes, usually many times larger than a single neuron. These typically contain an extracellular core of ß-amyloid surrounded by dystrophic dendrites and axons, loosely organized fibrils of ß-amyloid, and many other proteins and protein fragments derived from degenerating cells or liberated from neurons, reactive astrocytes, and phagocytic cells (11). ß-Amyloid is a fibrillar form of the ß-amyloid peptide (Aß), a 40- to 43-amino acid peptide derived from the normal processing of a larger ubiquitous membrane glycoprotein, the ß-APP (12). As it forms fibrils, Aß assumes a ß-sheet conformation recognizable by probes such as thioflavin S. This conformation distinguishes ß-amyloid from the diffuse nonfibrillar deposits of Aß that appear several or more years before the neuritic plaques. Diffuse plaques may be widespread throughout the brains of elderly persons with no measurable cognitive impairment. By contrast, neuritic plaques are primarily confined to the neocortex, hippocampus, and amygdala (13).

Although the relationship of diffuse plaques with the development of neuritic plaques is not yet established, studies in transgenic mouse models of cerebral ß-amyloidosis have suggested that neuritic plaques originate from diffuse Aß plaques through a "maturation" process. One possibility to explain this process is based on findings that Aß peptide aggregates bind and activate the complement protein C1 and, hence, the classic complement pathway. This, in turn, sets off a local non–immune-mediated, chronic inflammatory response involving microglial activation and stimulation of the acute-phase response (14). This proposed cascade explains the presence within neuritic plaques of additional cell types (microglia, reactive astrocytes) and many proteins, including glial-derived acute-phase proteins, such as α-antichymotrypsin, apolipoprotein E (Apo E), and serum amyloid P, which may, in some cases, accelerate fibrilization of Aß. Activated microglia release potentially toxic products, including proinflammatory cytokines (e.g., interleukin-1 and interleukin-6), reactive oxygen species, and proteases, all of which contribute to the local development of dystrophic neurites (15 ,16). According to a second possibility, not incompatible with the first, primary damage to neurites initiates local Aß overproduction, neurite degenerationregeneration responses, and secondary inflammatory reactions involved in removing cellular debris (17 ,18). Indeed, ultrastructural studies show that dystrophic neurites contain increased levels of both APP and organellar machinery needed for Aß generation. Moreover, neurons injured by various toxic factors produce more Aß (19 ,20 and 21).

The NFTs first seen by Alzheimer are skeins of twisted abnormal filaments, whose presence in neurons reflects a global disorganization of the neuronal cytoskeleton (22). The abnormal filaments, which assume a paired helical structure, hence the name *paired helical filaments* (PHFs), are composed of tau protein. Although its full repertoire of functions is still unclear, tau is known to bind to microtubules and to stabilize their polymeric structure, thereby facilitating the microtubule's function in axonal transport and structural support (23). Over a half-dozen protein kinases regulate the function of tau, including its affinity for microtubules. The observation that the tau in PHFs is hyperphosphorylated has suggested that altered phosphorylation is important for the development of these lesions (24). Other modifications of tau, such as proteolysis and glycation, are also considered to be important for PHF formation and for the resistance of PHF to degradation and removal (25). The importance of tau-related pathology to AD pathogenesis is strongly suggested by the identification of tau mutations in 20% of patients with frontotemporal dementia and in nearly half the patients with frontotemporal dementia who have an affected first-degree relative. In addition, other dementing disorders previously linked to chromosome 17 and characterized by NFT formation in specific neuronal populations are now being found to involve tau mutations or polymorphisms, including progressive supranuclear palsy, corticobasal degeneration, and Pick disease (23).

PHF coexists in tangles together with fragments of various cytoskeletal proteins. Notably, abnormally phosphorylated neurofilaments may accumulate as the earliest cytoskeletal alteration associated with dystrophic neurite formation (26). Neurites (axons and dendrites) containing abnormal organized cytoskeletal elements are referred to as *neuropil threads* and are a feature of the dystrophic neurites abundant within neuritic plaques. The appearance of neuropil threads before NFTs develop reflects the slow centrifugal progression of cellular compromise from the synaptic endings toward the neuronal perikaryon over several or more years and indicates that synapse function becomes impaired well before the neuron dies.

Although they are not part of current diagnostic criteria for AD, other characteristic pathologic responses of neurons begin before the first traces of ß-amyloid are deposited. One of these responses involves endocytosis, the process by which

the cell internalizes materials that are extracellular or on the cell's surface. Endocytosis allows continuous sampling of the external environment, a process that is important for the uptake of nutrients and for cellular responses to toxic foreign agents. Constant remodeling of plasma membrane receptor topography by endocytosis also allows cells to control how they respond to external signaling molecules. Very early in AD, neuronal early endosomes, which are a major site of Aß peptide formation, display prominent morphologic and biochemical alterations reflecting increased activity of the endocytic pathway and appear to be highly specific for AD (27). Individual early endosomes in pyramidal neurons of the Alzheimer brain may enlarge as much as 32-fold compared with the normal average endosomal volume (28). These changes coincide with the initial rise in brain production of Aß peptide, which precedes ß-amyloid deposition, and they are detectable even before birth in patients with Down syndrome, a population that invariably develops in AD after the age of 40 years.

Cell stress or injury to neurons also manifests early as robust activation of the lysosomal system, a major cellular degradative pathway. This activation, which involves the proliferation of lysosomes and increased expression of a dozen or more lysosomal hydrolases (29 ,30), progressively intensifies as neurons become more compromised. Because lysosomal activity controls cell size, up-regulation of this system in AD is likely to be the molecular basis for neuronal shrinkage (31), and it is also believed to contribute to neuritic dystrophy (32 ,33) and neuronal cell death (34). At end stages of neuronal injury, gradual dysfunction of the lysosomal system is implied by the prominent accumulation within lysosomes of undigested, oxidized proteins and lipids in the form of autofluorescent lipofuscin and ceroid. Although lysosomal system activation is not entirely disease specific, its magnitude in AD is much greater than in other diseases and may reflect a specialized neurodegenerative response characteristic of AD and a subset of related disorders (34).

CELLULAR SUBSTRATES OF DEMENTIA

Part of "85 - Cell and Molecular Neuropathology of Alzheimer Disease "

That AD-related lesions may already be well developed in persons who have apparently normal cognitive function has fueled the debate about the actual neuropathologic substrate of clinical dementia. Attention is increasingly being paid to the roles of less overt pathologic features, such as rising intracellular Aß levels, synapse loss, or subtle forms of tau pathology in dendrites or axons. For example, although ß-amyloid or "plaque burden" often correlates poorly with cognitive impairment (35), levels of soluble Aß peptide may correlate better (36 ,36a), a concept supporting the possibility that accelerated formation of Aß peptide inside the neuron interferes with neuronal function or reflects dysfunction that is present before ß-amyloid fibrils form extracellularly (27). Synapse loss is not revealed by routine neuropathologic analysis but may reach 50% of synapses in affected neocortical areas by late stages of the disease (37 ,38). A high correlation with cognitive decline has suggested that the loss of synapses may be the cellular basis of dementia, although this hypothesis needs further confirmation. Neurofibrillary change in neurons (NFTs and neuropil threads) is closely associated with synaptic disease, and because it is more easily traced histologically, it is the index most frequently used to correlate structural disease progression to cognitive symptoms.

The progressive appearance of NFTs follows a consistent cytoarchitectonic pattern that parallels the severity of clinical dementia and neuronal cell loss more closely than does the evolution of senile plaques (39 ,40 and 41). The transentorhinal region, particularly layers II and IV of the entorhinal cortex, usually shows the first lesions in AD. By the time the mildest stage of cognitive impairment is detectable, the entorhinal cortex may have one-third fewer neurons than normal. Extrapolations from the rate of subsequent fallout of this cell population suggest that neuronal loss may be initiated up to 7 to 10 years before detectable cognitive symptoms.

The highly predictable development of neurofibrillary degeneration in the entorhinal cortex and its progressive extension into the hippocampus, neocortex, and later into various subcortical structures were the basis for a pathologic staging system developed by Braak and Braak (7 ,39), which uses cross-sectional data on NFT distribution to distinguish six stages of disease evolution. Inclusion of this grading system into current CERAD criteria for neuropathologic diagnosis of AD significantly increases diagnostic specificity and reduces false-negative diagnoses (10). According to this scheme, disease begins in stage 1 with the involvement of only a few transentorhinal projection cells, and it progresses in stage 2 to involve many entorhinal neurons, particularly those in layer II. In stages 3 and 4, neurofibrillary degeneration remains restricted to limbic regions, but it now begins to invade the hippocampal formation. The two principal targets of the CA1/subicular projection are the accessory nucleus of the amygdala and layer IV of the entorhinal cortex, which provide the principal output from the hippocampus to cortical and subcortical regions. The progression of changes to these targets, which gradually isolates the hippocampus from other brain structures, is associated with impaired cognitive functioning and subtle changes in personality in some persons. By stage 5, cognitive deficits have become broader, and the clinical diagnosis of Alzheimer-type dementia can usually be made (10). At this point, NFTs have increasingly appeared in projection neurons within layers II, III, and V of higher-order association cortices (42 ,43), beginning with the temporal lobe, which is more severely affected than parietal and frontal association cortices. Large cortical projection neurons in layers III and V display the most prominent cytoskeletal alterations, and

these cells may be lost to a greater extent than smaller neurons. This pattern of cell loss reflects the special vulnerability of feed-forward and feedback circuitry linking the hemispheres with each other and with the cortex of the limbic lobe and subcortical structures. The basis for the selective vulnerability is poorly understood, but it has been conjectured that it is related, in part, to unique features of the cytoskeleton in these neurons and particularly to the abundance of neurofilament proteins and their relatively low phosphorylation state (44).

Although the cerebral cortex is the primary target in AD, degeneration of subcortical structures may also contribute to memory impairment and their behavioral disturbances (45 ,46). The nucleus basalis of Meynert provides the major cholinergic input to the cortex and is important for memory, but the variability and timing of cholinergic changes suggest that they may not be the key factors in early cognitive impairment (35). The amygdala receives prominent projections from cortical areas and subcortical areas; degeneration therein is particularly relevant to disease-related impairments in motivated and emotional behavior. Extensive cell loss in the noradrenergic locus ceruleus, which richly innervates the cortex, has been associated with depressive symptoms. Changes in the serotonergic raphe nuclei and involvement of hypothalamic nuclei, including the suprachiasmatic nucleus, may explain commonly observed impairments of sleep and circadian rhythm in AD. Although dopaminergic neurons of the ventral tegmentum are severely depleted, cell loss is only moderate in the substantia nigra, as reflected by the absence of Lewy body pathology and associated extrapyramidal symptoms. The well-documented reductions in levels of various neurotransmitters and their receptors (47) are almost certainly a secondary consequence of the loss or functional deafferentation of these subcortical projection neurons.

INITIATION OF CELLULAR PATHOLOGY

Part of "85 - Cell and Molecular Neuropathology of Alzheimer Disease "

Familial Alzheimer Disease

The identification of genes that, when mutated, cause early-onset autosomal dominant FAD have provided strong clues to the cellular pathogenesis of AD (see Chapter 83).

Pathogenetic Mechanisms in Sporadic Alzheimer Disease

However, more than 90% of all cases of AD are not caused by single gene mutations, and, in these cases, the origin is less well understood. Although the factors that accelerate ß-amyloidogenesis in sporadic AD are not established, clues are emerging from studies of genes that influence the risk of developing late-onset AD. Topping the list of factors that increase AD risk is inheritance of the ε4 allele of the gene encoding Apo E, a protein that transports cholesterol and certain phospholipids into cells (48 ,49). The virtual absence of other key plasma apolipoproteins such as Apo A1, C1, and B in the brain emphasizes the critical role of Apo E in this tissue. In addition to its role in lipid transport, Apo E has antioxidant and growth-promoting properties on cells (50 ,51), and it interacts with Aß, thereby influencing its endocytosis and clearance (52), its ability to aggregate, and its neurotoxicity (53 ,54)—effects that may all be relevant to AD pathogenesis. The three isoforms of Apo E, ε2, ε3, and ε4, vary by only a single amino acid substitution, but they differ markedly in their binding affinities for low-density lipoprotein (LDL) receptors and other proteins. The most common allele, ε3, occurs in the general population with a frequency of 75%, whereas ε2 and ε4 occur with frequencies of 10% and 15%, respectively (1). Inheritance of a single ε4 allele increases the risk of AD threefold, whereas homozygosity for ε4 is associated with an eightfold increase in risk (49). The Apo E ε4 allele appears to lower the age of onset in persons in their sixties and seventies rather than influence the duration and severity of the disease. There is evidence that the ε2 allele confers some protection from the development of late-onset AD and Down syndrome (5).

Additional genes that may influence the risk of developing late-onset disease are being identified at an increasing pace (49 ,55 ,56). Among the possible positive risk factors are common population polymorphisms of the *APOE* promoter (57) and genes encoding the LDL-receptor–related protein (LRP-1) (58 ,59 ,60 and 61), α_2-macroglobulin (62), FE65 (63 ,64), very LDL (VLDL)-R receptor (65), the lysosomal protease cathepsin D (66), the lysosomal cysteine protease inhibitor cystatin C (67), bleomycin hydrolase (68), and interleukin-1 (1A and B) (69 ,70). In light of the evidence that neuronal endocytosis is altered at the very earliest stages of AD, many of these genetic modifiers of AD risk encode proteins that depend on endocytosis for their function. Apo E, its receptor on neurons (LRP), another LRP ligand (α_2-macroglobulin), and the VLDL receptor all are molecules that traffic through early endosomes as they bring cholesterol or other ligands into the cell. Inheritance of the *APOE* ε4 allele accentuates endocytic abnormalities in AD (27). Cathepsin D is a major protease of the endosomal-lysosomal pathway previously implicated by other neuropathologic and biochemical data in AD pathogenesis (34). Similarly, cystatin C mutations, which cause the Icelandic form of hemorrhagic cerebral amyloid angiopathy (71), are key regulators of proteases within the lysosomal system. The endocytic pathway is also responsible for the internalization and initial processing of APP at the cell surface. Fe65 binds to the internalization domain of APP and modulates its processing to Aß (72). Early endosomes are also a principal site of Aß generation in normal cells and mediate the cellular uptake of Aε and APPs (73 ,74).

The convergence of such a diverse group of etiologically important molecules at a known major site of Aß production

has suggested that altered early endosome function may contribute to ß-amyloidogenesis in sporadic AD (28). Supporting this hypothesis are studies showing the principal ß-secretase in cells resides largely in endosomes (75) and that cathepsin D, a protease with ß-secretase activity (76 ,77 ,78 and 79), and other “lysosomal” proteases that influence Aß formation become more abundant in neuronal early endosomes when the lysosomal system becomes activated in AD. This latter effect reflects not only the markedly increased expression of these proteases but also their enhanced targeting to early endosomes by the cation-dependent mannose-6-phosphate receptor (MPR-46), which is also more highly expressed in AD brain (28 ,80). When these conditions are recreated experimentally in cells by modestly overexpressing MPR-46, Aß generation is substantially increased (80). Thus, proteases that normally may not even be involved in Aß formation could, under pathologic conditions, become abnormally routed to cellular compartments where they promote Aß generation. This is one mechanism that explains how ß-amyloidogenesis may be accelerated in sporadic AD in the absence of a causative gene mutation.

EVOLUTION OF CELLULAR PATHOBIOLOGY

Part of "85 - Cell and Molecular Neuropathology of Alzheimer Disease "

The genetic heterogeneity of AD suggests that the disease may be initiated through distinct cellular cascades, which then converge on the final common pathways responsible for ß-amyloidogenesis, neurofibrillary pathology, and, ultimately, neuronal cell death. Secondary and tertiary responses of the brain to the presence of these neuropathologic lesions may further compromise neuronal function, making it difficult to establish what is cause or effect. Current hypotheses on the cellular pathobiology of AD emphasize different aspects of this complex multifactorial process, and, not surprisingly, these “different” views overlap considerably. To illustrate this, three perspectives on cellular pathogenesis are discussed in the following paragraphs; these emphasize metabolic decline, defective cell repair, or Aß toxicity as the driving pathophysiologic mechanism in AD.

From the *metabolic decline* perspective, cellular oxidative stress leading to neurodegeneration is a final common pathway of metabolic insults originating from different sources. According to this view, metabolic disturbances are both a cause and a consequence of ß-amyloidogenesis. The onslaught on metabolic function begins with effects of normal aging and specific genetic factors. For example, aging-related cerebral hypoperfusion leading to reduced brain glucose and oxygen utilization impairs energy production at the mitochondrial level and promotes the production of free radicals (81). The detection of regional hypometabolism in AD patients with mild cognitive impairment suggests that such hypometabolism may not be simply a result of neurodegeneration but may also precede it. Moreover, cerebral ischemia, coronary artery disease, *APP* mutations, and some inherited mitochondrial DNA mutations may increase AD risk, in part, by creating additional oxidative stress through these same pathways (82). Oxidative damage from these and other sources leads to mitochondrial membrane depolarization and increased levels of mitochondrial reactive oxygen species (82a). The resultant oxidative damage to proteins and membranes activates degradative pathways, notably the lysosomal system (31), and in doing so up-regulates cathepsins and other proteases (83), which have been implicated in mechanisms of cell death, Aß production, and cytoskeletal protein modification (34). Free radicals subsequently impair the function of glucose and glutamate transporters and damage ion-channel adenosine triphosphatases (sodium-calcium pumps), thereby reducing the ability of cells to buffer calcium (84) and rendering them more vulnerable to excitotoxins (85).

Calcium homeostasis is further altered by glutamate and other excitotoxins that stimulate receptor-mediated influx of calcium or, in FAD, by mutations of presenilin that lead to the release of intracellular calcium stores (86 ,87). Elevated cellular calcium activates major signaling cascades including stress-related protein kinases acting on tau and the cytoskeleton (24 ,88). Tau hyperphosphorylation decreases its binding to microtubules and promotes loss of microtubule stability and impaired axonal transport (89). NFT formation may compound this effect on transport by imposing physical obstructions to the movement of vital organelles to the axon and synapse. Calcium-activated neutral protease (calpain) systems, which are highly activated in AD brain (90 ,91), contribute to the truncation and breakdown of cytoskeletal proteins including tau, alter the activity of the protein kinase C cascade, cdk5, and other signaling pathways, and participate in the mechanisms underlying apoptotic and necrotic cell death (92 ,93). Ultimately, in certain cells, mitochondrial damage leads to the release of cytochrome C, which activates caspases that mediate apoptosis. FAD-linked *PS* and *APP* mutations increase the vulnerability of cultured neurons to apoptosis, presumably through one or more of the metabolic pathways discussed above (94 ,95 and 96).

Complementary to the foregoing metabolic decline perspective is a *cell repair* hypothesis, which emphasizes a putative failure by the brain to repair the cumulative neuronal damage arising from normal aging processes, ischemic and environmental insults, and genetic factors. The neurotrophic actions of APP or its mobilization during neuronal injury are most relevant here. Cells normally secrete a proteolytic derivative of APP, designated APPs, which promotes neuron growth and increases neuronal survival after certain types of injury (84). APP expression and distribution dramatically increases after neuronal injury, ischemia or oxidative stress, head injury, and exposure to toxins (97). Cerebrospinal fluid levels of APPs, however, may be reduced. Apo E also figures prominently in the processes of cell repair and regeneration by coordinating the mobilization and redistribution

of cholesterol needed for myelin and neuronal membrane synthesis (48). Functional synaptic remodeling *in vivo* is markedly compromised in mice lacking the Apo E gene (98). During regeneration, Apo E expression may increase up to 100-fold. The ε3 allele seems to be more effective as a growth-promoting or repair factor than the ε4 allele, which is linked to an increased risk of AD (99). Because Apo E is synthesized in glial cells, neurons depend heavily on receptor-mediated endocytosis to internalize Apo E–cholesterol complexes, a process that is altered at the earliest stages of AD (27). The increased levels of protease seen in neuronal early endosomes of the AD brain likely promote the degradation of internalized molecules and may prematurely abrogate their trophic or nutrient functions.

A third perspective, commonly referred to as the *Aß cascade hypothesis*, places the Aß peptide at the center of AD pathogenesis based on its neurotoxic properties in either soluble or fibrillar form. Aß deposition within senile plaques involves a balance between forces that enhance the overproduction and aggregation of Aß and countervailing forces that promote the uptake and degradation of Aß from the extracellular space. FAD-linked mutations cause varying degrees of Aß overproduction (97), but they may also favor aggregation by increasing the relative production of Aß42 or mutant Aß forms that aggregate more easily. Aß aggregation is also facilitated by additional proteins released by reactive or damaged cells (100 ,100a). In this regard, Apo E is particularly critical to Aß deposition (101 ,102 and 103). Cellular uptake and clearance of Aß involve interactions of an Apo E/Aß complex with LRP. The influence of Apo E on amyloid deposition is underscored by the observation that Apo E gene ablation abolishes amyloid deposition in transgenic mice overexpressing APP containing the London mutation (104). Microglial function also seems to be critical to Aß removal (15). In transgenic models of FAD, Aß deposition is almost completely prevented when microglia are experimentally activated by immunizing mice with Aß protein (105).

The fibrillar state of Aß is considered to be a crucial factor in Aß neurotoxicity (106 ,107); however, it is still unclear whether intraneuronal Aß accumulation or extracellular soluble Aß forms may be relevant to neurotoxicity (108 ,109). Although a sequence of events after Aß deposition has not been confirmed, it is hypothesized that Aß accumulation within diffuse plaques eventually leads to local microglial activation, cytokine release, increases in astrocyte numbers, and an inflammatory response involving the classic complement cascade, as discussed earlier (14). It has been further proposed that these glial responses and any direct neurotoxic effects of Aß initiate a cascade of biochemical and structural changes in surrounding axons, dendrites, and neuronal cell bodies in AD. Aß-initiated inflammatory and neurotoxic processes generate excessive free radicals and peroxidative injury to proteins and alterations of ionic homeostasis, particularly excessive calcium entry into neurons accompanied by the activation of calpains and kinases acting on cytoskeletal proteins (16 ,82 ,84). Aß is one of various factors that may stimulate the glycogen synthase kinase pathway, which, among other roles, is involved in both the phosphorylation of tau and still unclarified aspects of presenilin and APP processing (110). Finally, endocytic uptake of Aß 42 into cells activates and destabilizes the lysosomal system, and this further compounds the insult to this system from other sources and promotes cell death (111).

Thus, the Aß cascade hypothesis ultimately reaches the same metabolic endpoints as the metabolic decline hypothesis, but it distinguishes itself by proposing that Aß accumulation is the germinal event, rather than being a secondary, albeit important, consequence of accumulating metabolic or functional deficits within neurons. To become a comprehensive hypothesis of AD pathogenesis, the Aß cascade hypothesis still must explain the nature of the initial disturbance that causes Aß to accumulate in the 90% of AD cases that are not caused by FAD-linked mutations. Most likely, AD pathogenesis is a multifactorial process, in which Aß is necessary but not a sufficient factor.

CONCLUSIONS

Part of "85 - Cell and Molecular Neuropathology of Alzheimer Disease "

In this chapter, a discussion of Alzheimer neuropathology provides the starting point for understanding how the diagnosis of AD is made and how clinical symptoms arise and progress. Later sections address the neuroanatomic and cellular basis for dementia and the molecular events that lead to neuropathologic lesions and, ultimately, to the death of neurons. A consideration of current hypotheses on the evolution of cellular pathology identifies common features that help to reconcile the differing views of disease pathogenesis. New studies are beginning to shed light on underlying mechanisms in prevalent sporadic forms of AD, and their review complements the discussion of pathogenesis in the rare familial forms here and, in more detail, in Chapter 83 .

ACKNOWLEDGMENTS

Part of "85 - Cell and Molecular Neuropathology of Alzheimer Disease "

I am grateful to Janet Rosdil for expert assistance in the preparation of this manuscript. Studies from my laboratory were supported by grants from Leadership and Excellence in Alzheimer’s Disease (LEAD) from the National Institute on Aging and P01 AG17617-01.

DISCLAIMER

Part of "85 - Cell and Molecular Neuropathology of Alzheimer Disease "

Dr. Nixon has received research support from Johnson and Johnson (Janssen).

REFERENCES

1. Cruts M, Van Broeckhoven C. Molecular genetics of Alzheimer's disease. *Ann Med* 1998;30:560–565.

2. Kehoe P, Wavrant-De Vrieze F, Crook R, et al. A full genome scan for late onset Alzheimer's disease. *Hum Mol Genet* 1999;8:237–245.

3. Davis DG, Schmitt FA, Wekstein DR, et al. Alzheimer neuropathologic alterations in aged cognitively normal subjects. *J Neuropathol Exp Neurol* 1999;58:376–388.

4. Mufson EJ, Chen EY, Cochran EJ, et al. Entorhinal cortex beta-amyloid load in individuals with mild cognitive impairment. *Exp Neurol* 1999;158:469–490.

5. Lai F, Kammann E, Rebeck GW, et al. APOE genotype and gender effects on Alzheimer disease in 100 adults with Down syndrome. *Neurology* 1999;53:331–336.

6. Mirra SS, Heyman A, McKeel D, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). II. Standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology* 1991;41:479–486.

7. Braak H, Braak E. Aspects of cortical destruction in Alzheimer's disease. In: Hyman B, Duyckaerts C, Christen Y, eds. *Connections, cognition, and Alzheimer's disease*. Berlin: Springer-Verlag, 1997:1–16.

8. Jellinger KA, Bancher C. Proposals for re-evaluation of current autopsy criteria for the diagnosis of Alzheimer's disease. *Neurobiol Aging* 1997;18:S55–65.

9. Institute R. Consensus recommendations for the postmortem diagnosis of Alzheimer's disease: the National Institute on Aging, and Reagan Institute Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease. *Neurobiol Aging* 1997:S1–2.

10. Newell KL, Hyman BT, Growdon JH, et al. Application of the National Institute on Aging (NIA)–Reagan Institute criteria for the neuropathological diagnosis of Alzheimer disease. *J Neuropathol Exp Neurol* 1999;58:1147–1155.

11. Mann DM, Brown SM, Owen F, et al. Amyloid beta protein (A beta) deposition in dementia with Lewy bodies: predominance of A beta 42(43) and paucity of A beta 40 compared with sporadic Alzheimer's disease. *Neuropathol Appl Neurobiol* 1998;24:187–194.

12. Checler F. Processing of the beta-amyloid precursor protein and its regulation in Alzheimer's disease. *J Neurochem* 1995;65:1431–1444.

13. Mann D. *Sense and senility: the neuropathology of the aged human brain*. Austin, Texas: RG Landes, 1997.

14. Eikelenboom P, Veerhuis R. The importance of inflammatory mechanisms for the development of Alzheimer's disease. *Exp Gerontol* 1999;34:453–461.

15. Kalaria RN. Microglia and Alzheimer's disease. *Curr Opin Hematol* 1999;6:15–24.

16. Griffin WS, Sheng JG, Royston MC, et al. Glial-neuronal interactions in Alzheimer's disease: the potential role of a "cytokine cycle" in disease progression. *Brain Pathol* 1998;8:65–72.

17. Cataldo AM, Nixon RA. Enzymatically active lysosomal proteases are associated with amyloid deposits in Alzheimer brain. *Proc Natl Acad Sci USA* 1990;87:3861–3865.

18. Cataldo AM, Hamilton DJ, Nixon RA. Lysosomal abnormalities in degenerating neurons link neuronal compromise to senile plaque development in Alzheimer disease. *Brain Res* 1994;640:68–80.

19. Ohyagi Y, Yamada T, Nishioka K, et al. Selective increase in cellular A beta 42 is related to apoptosis but not necrosis. *Neuroreport* 2000;11:167–171.

20. LeBlanc A, Liu H, Goodyer C, et al. Caspase-6 role in apoptosis of human neurons, amyloidogenesis, and Alzheimer's disease. *J Biol Chem* 1999;274:23426–23436.

21. Wiedemann C, Cockcroft S. Vesicular transport: sticky fingers grab a lipid [Comment]. *Nature* 1998;394:426–427.

22. Goedert M. Tau protein and the neurofibrillary pathology of Alzheimer's disease. *Trends Neurosci* 1993;16:460–465.

23. Buee M, Rossignol M, Jauneau A, et al. The pre-symbiotic growth of arbuscular mycorrhizal fungi is induced by a branching factor partially purified from plant root exudates. *Mol Plant Microbe Interact* 2000;13:693–698.

24. Lovestone S, Davis DR, Webster MT, et al. Lithium reduces tau phosphorylation: effects in living cells and in neurons at therapeutic concentrations. *Biol Psychiatry* 1999;45:995–1003.

25. Grynspan F, Griffin WR, Cataldo A, et al. Active site-directed antibodies identify calpain II as an early-appearing and pervasive component of neurofibrillary pathology in Alzheimer's disease. *Brain Res* 1997;763:145–158.

26. Dickson TC, King CE, McCormack GH, et al. Neurochemical diversity of dystrophic neurites in the early and late stages of Alzheimer's disease. *Exp Neurol* 1999;156:100–110.

27. Cataldo A, Peterhoff C, Troncoso J, et al. Endocytic pathway abnormalities precede beta-amyloid deposition in sporadic Alzheimer's disease: differential effects of APOE genotype and presenilin mutations. *Am J Pathol* 2000;157:277–286.

28. Cataldo AM, Barnett JL, Pieroni C, et al. Increased neuronal endocytosis and protease delivery to early endosomes in sporadic Alzheimer's disease: neuropathologic evidence for a mechanism of increased beta-amyloidogenesis. *J Neurosci* 1997;17:6142–6151.

29. Cataldo AM, Paskevich PA, Kominami E, et al. Lysosomal hydrolases of different classes are abnormally distributed in brains of patients with Alzheimer disease [published erratum appears in *Proc Natl Acad Sci USA* 1992;89:2509]. *Proc Natl Acad Sci USA* 1991;88:10998–11002.

30. Cataldo AM, Barnett JL, Berman SA, et al. Gene expression and cellular content of cathepsin D in Alzheimer's disease brain: evidence for early up-regulation of the endosomal-lysosomal system. *Neuron* 1995;14:671–680.

31. Nixon RA, Cataldo AM. The endosomal-lysosomal system of neurons: new roles. *Trends Neurosci* 1995;18:489–496.

32. Bednarski E, Ribak CE, Lynch G. Suppression of cathepsins B and L causes a proliferation of lysosomes and the formation of meganeurites in hippocampus. *J Neurosci* 1997;17:4006–4021.

33. Bi X, Zhou J, Lynch G. Lysosomal protease inhibitors cause meganeurite formation and phosphorylated tau abnormalities in entorhinal-hippocampal regions vulnerable to Alzheimer's disease. *Exp Neurol* 1999;158:312–327.

34. Nixon RA, Cataldo AM, Mathews PM. The endosomal-lysosomal system of neurons in Alzheimer's disease pathogenesis: a review. *Neurochem Res* 2000;25:1161–1172.

35. Davis KL, Mohs RC, Marin D, et al. Cholinergic markers in elderly patients with early signs of Alzheimer disease [see Comments]. *JAMA* 1999;281:1401–1406.

36. Buxbaum JD, Liu KN, Luo Y, et al. Evidence that tumor necrosis factor alpha converting enzyme is involved in regulated alpha-secretase cleavage of the Alzheimer amyloid protein precursor. *J Biol Chem* 1998;273:27765–27767.

36a. Lue LF, Kuo YM, Roher AE, et al. Soluble amyloid beta peptide concentration as a predictor of synaptic change in Alzheimer disease. *Am J Pathol* 1999;155:853–862.

37. Terry RD. The pathogenesis of Alzheimer disease: an alternative to the amyloid hypothesis. *J Neuropathol Exp Neurol* 1996;55:1023–1025.

38. Terry RD, Masliah E, Salmon DP, et al. Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. *Ann Neurol* 1991;30:572–580.

39. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol* 1991;82:239–259.

40. Arnold SE, Hyman BT, Van Hoesen GW, et al. Some cytoarchitectural abnormalities of the entorhinal cortex in schizophrenia. *Arch Gen Psychiatry* 1991;48:625–632.

41. Gomez-Isla T, Price JL, McKeel DW Jr, et al. Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. *J Neurosci* 1996;16:4491–4500.

42. Lewis SE, Nixon RA. Multiple phosphorylated variants of the high molecular mass subunit of neurofilaments in axons of retinal cell neurons: characterization and evidence for their differential association with stationary and moving neurofilaments. *J Cell Biol* 1988;107:2689–2701.

43. Arnold SE, Hyman BT, Flory J, et al. The topographical and neuroanatomical distribution of neurofibrillary tangles and neuritic plaques in the cerebral cortex of patients with Alzheimer's disease. *Cereb Cortex* 1991;1:103–116.

44. Morrison BM, Hof PR, Morrison JH. Determinants of neuronal vulnerability in neurodegenerative diseases. *Ann Neurol* 1998;44:S32–44.

45. Damasio A. The anatomic basis of memory disorders. *Semin Neurol* 1984:223–225.

46. Mann DM, Iwatsubo T, Pickering-Brown SM, et al. Preferential deposition of amyloid beta protein (Abeta) in the form Abeta40 in Alzheimer's disease is associated with a gene dosage effect of the apolipoprotein E E4 allele. *Neurosci Lett* 1997;221:81–84.

47. Young AB, Penney JB Jr. Neurotransmitter receptors in Alzheimer disease. In: Terry R, Katzman R, Bick K, eds. *Alzheimer's disease*. New York: Raven, 1994:293–304.

48. Mahley RW. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science* 1988;240:622–630.

49. Roses AD. Apolipoprotein E and Alzheimer's disease: the tip of the susceptibility iceberg. *Ann N Y Acad Sci* 1998;855:738–743.

50. Gutman CR, Strittmatter WJ, Weisgraber KH, et al. Apolipoprotein E binds to and potentiates the biological activity of ciliary neurotrophic factor. *J Neurosci* 1997;17:6114–6121.

51. Laskowitz DT, Horsburgh K, Roses AD. Apolipoprotein E and the CNS response to injury. *J Cereb Blood Flow Metab* 1998;18:465–471.

52. Permanne B, Perez C, Soto C, et al. Detection of apolipoprotein E/dimeric soluble amyloid beta complexes in Alzheimer's disease brain supernatants. *Biochem Biophys Res Commun* 1997;240:715–720.

53. Jordan J, Galindo MF, Miller RJ, et al. Isoform-specific effect of apolipoprotein E on cell survival and beta-amyloid–induced toxicity in rat hippocampal pyramidal neuronal cultures. *J Neurosci* 1998;18:195–204.

54. Miyata M, Smith JD. Apolipoprotein E allele-specific antioxidant activity and effects on cytotoxicity by oxidative insults and beta-amyloid peptides. *Nat Genet* 1996;14:55–61.

55. Price DL, Sisodia SS. Mutant genes in familial Alzheimer's disease and transgenic models. *Annu Rev Neurosci* 1998;21:479–505.

56. Tanzi RE. A genetic dichotomy model for the inheritance of Alzheimer's disease and common age-related disorders. *J Clin Invest* 1999;104:1175–1179.

57. Casadei VM, Ferri C, Veglia F. APOE-491 promoter polymorphism is a risk factor for late-onset Alzheimer's disease. *Neurology* 1999;53:1888–1889.

58. Kamboh MI, Ferrell RE, DeKosky ST. Genetic association studies between Alzheimer's disease and two polymorphisms in the low density lipoprotein receptor-related protein gene. *Neurosci Lett* 1998;244:65–68.

59. King ND, O'Brian MR. Identificatin of the Irp gene in *Bradyrhizobium japonicum* and its role in regulation of delta-aminolevulinic acid uptake. *J Bacteriol* 1997;179:1828–1831.

60. Papassotiropoulos A, Bagli M, Feder O, et al. Genetic polymorphism of cathepsin D is strongly associated with the risk for developing sporadic Alzheimer's disease. *Neurosci Lett* 1999;262:171–174.

61. Tanzi RE, Kovacs DM, Kim TW, et al. The gene defects responsible for familial Alzheimer's disease. *Neurobiol Dis* 1996;3:159–168.

62. Blacker D, Wilcox MA, Laird NM, et al. Alpha-2 macroglobulin is genetically associated with Alzheimer disease [see Comments]. *Nat Genet* 1998;19:357–360.

63. Xia W, Zhang J, Perez R, et al. Interaction between amyloid precursor protein and presenilins in mammalian cells: implications for the pathogenesis of Alzheimer disease. *Proc Natl Acad Sci USA* 1997;94:8208–8213.

64. Xu X, Yang D, Wyss-Coray T, et al. Wild-type but not Alzheimer-mutant amyloid precursor protein confers resistance against p53–mediated apoptosis. *Proc Natl Acad Sci USA* 1999;96:7547–7552.

65. Yamanaka H, Kamimura K, Tanahashi H, et al. Genetic risk factors in Japanese Alzheimer's disease patients: alpha1–ACT, VLDLR, and ApoE. *Neurobiol Aging* 1998;19:S43–46.

66. Papassotiropoulos A, Bagli M, Kurz A, et al. A genetic variation of cathepsin D is a major risk factor for Alzheimer's disease. *Ann Neurol* 2000;47:399–403.

67. Crawford C, Freeman M, Schinka J, et al. The cystatin C gene as a novel genetic risk factor for late onset Alzheimer's disease. In: 7th International Conference on Alzheimer's Disease and Related Disorders, Washington, DC. *Neurobiol Aging* 2000;204.

68. Papassotiropoulos A, Bagli M, Jessen F, et al. Confirmation of the association between bleomycin hydrolase genotype and Alzheimer's disease. *Mol Psychiatry* 2000;5:213–215.

69. Grimaldi LM, Casadei VM, Ferri C, et al. Association of early-onset Alzheimer's disease with an interleukin-1alpha gene polymorphism [see Comments]. *Ann Neurol* 2000;47:361–365.

70. Nicoll JA, Mrak RE, Graham DI, et al. Association of interleukin-1 gene polymorphisms with Alzheimer's disease [see Comments]. *Ann Neurol* 2000;47:365–368.

71. Levy E, Carman MD, Fernandez-Madrid IJ, et al. Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. *Science* 1990;248:1124–1126.

72. Sabo SL, Lanier LM, Ikin AF, et al. Regulation of beta-amyloid secretion by FE65, an amyloid protein precursor-binding protein. *J Biol Chem* 1999;274:7952–7957.

73. Perez RG, Soriano S, Hayes JD, et al. Mutagenesis identifies new signals for beta-amyloid precursor protein endocytosis, turnover, and the generation of secreted fragments, including Abeta42. *J Biol Chem* 1999;274:18851–18856.

74. Cescato R, Dumermuth E, Spiess M, et al. Increased generation of alternatively cleaved beta-amyloid peptides in cells expressing mutants of the amyloid precursor protein defective in endocytosis. *J Neurochem* 2000;74:1131–1139.

75. Vassar R, Bennett BD, Babu-Khan S, et al. Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* 1999;286:735–741.

76. Gruninger-Leitch F, Berndt P, Langen H, et al. Identification of beta-secretase–like activity using a mass spectrometry–based assay system. *Nat Biotechnol* 2000;18:66–70.

77. Mackay EA, Ehrhard A, Moniatte M, et al. A possible role for cathepsins D, E, and B in the processing of beta-amyloid precursor protein in Alzheimer's disease. *Eur J Biochem* 1997;244:414–425.

78. Dreyer RN, Bausch KM, Fracasso P, et al. Processing of the pre-beta-amyloid protein by cathepsin D is enhanced by a familial Alzheimer's disease mutation. *Eur J Biochem* 1994;224:265–271.

79. Chevallier N, Vizzavona J, Marambaud P, et al. Cathepsin D displays *in vitro* beta-secretase-like specificity. *Brain Res* 1997;750:11–19.

80. Mathews PM, Guerra CB, Jiang Y, et al. Overexpression of the cation-dependent mannose 6-phosphate receptor increases abeta secretion: role for altered lysosomal hydrolase distribution in beta-amyloidogenesis. *J Biol Chem*, in press.

81. Hoyer S. Models of Alzheimer's disease: cellular and molecular aspects. *J Neural Transm Suppl* 1997;49:11–21.

82. Markesbery WR. Oxidative stress hypothesis in Alzheimer's disease. *Free Radic Biol Med* 1997;23:134–147.

82a. Keller JN, Pang Z, Geddes JW, et al. Impairment of glucose and glutamate transport and induction of mitochondrial oxidative stress and dysfunction in synaptosomes by amyloid beta peptide: role of the lipid peroxidation product 4-hydroxynonenal. *J Neurochem* 1997;69:273–289.

83. Adamec M, Vonsattel J, Nixon R. DNA strand breaks in Alzheimer's disease. *Brain Res* 1999;849:67–77.

84. Mattson MP. Cellular actions of beta-amyloid precursor protein and its soluble and fibrillogenic derivatives. *Physiol Rev* 1997;77:1081–1132.

85. Mattson MP, Barger SW, Cheng B, et al. Beta-amyloid precursor protein metabolites and loss of neuronal Ca^{2+} homeostasis in Alzheimer's disease. *Trends Neurosci* 1993;16:409–414.

86. Mattson MP. Calcium as sculptor and destroyer of neural circuitry. *Exp Gerontol* 1992;27:29–49.

87. Mattson MP, Guo Q, Furukawa K, et al. Presenilins, the endoplasmic reticulum, and neuronal apoptosis in Alzheimer's disease. *J Neurochem* 1998;70:1–14.

88. Patrick GN, Zukerberg L, Nikolic M, et al. Conversion of p35 to p25 deregulates Cdk5 activity and promotes neurodegeneration [see Comments]. *Nature* 1999;402:615–622.

89. Lee VM, Trojanowski JQ. Neurodegenerative tauopathies: human disease and transgenic mouse models. *Neuron* 1999;24:507–510.

90. Saito K, Elce JS, Hamos JE, et al. Widespread activation of calcium-activated neutral proteinase (calpain) in the brain in Alzheimer disease: a potential molecular basis for neuronal degeneration. *Proc Natl Acad Sci USA* 1993;90:2628–2632.

91. Grynspan F, Griffin WB, Mohan PS, et al. Calpains and calpastatin in SH-SY5Y neuroblastoma cells during retinoic acid-induced differentiation and neurite outgrowth: comparison with the human brain calpain system. *J Neurosci Res* 1997;48:181–191.

92. Lee MS, Kwon YT, Li M, et al. Neurotoxicity induces cleavage of p35 to p25 by calpain. *Nature* 2000;405:360–364.

93. Nixon R, Mohan P. Calpains in the pathogenesis of Alzheimer's disease. In: Wang K, Yuen W, eds. *Calpain: pharmacology and toxicology of calcium-dependent protease*. Ann Arbor, MI: Taylor & Francis, 1999:267–291.

94. Kwok JB, Li QX, Hallupp M, et al. Novel Leu723Pro amyloid precursor protein mutation increases amyloid beta42(43) peptide levels and induces apoptosis. *Ann Neurol* 2000;47:249–253.

95. Tanii H, Ankarcrona M, Flood F, et al. Alzheimer's disease presenilin-1 exon 9 deletion and L250S mutations sensitize SH-SY5Y neuroblastoma cells to hyperosmotic stress-induced apoptosis. *Neuroscience* 2000;95:593–601.

96. Weihl CC, Ghadge GD, Kennedy SG, et al. Mutant presenilin-1 induces apoptosis and downregulates Akt/PKB. *J Neurosci* 1999;19:5360–5369.

97. Hardy J. Amyloid, the presenilins and Alzheimer's disease [see Comments]. *Trends Neurosci* 1997;20:154–159.

98. Masliah E, Mallory M. Abnormal synaptic regeneration in APOE-knockout. In: Iqbal K, Mortimer J, Winblad B, eds. *Research advances in Alzheimer's and related disorders*. New York: John Wiley, 1995:405–414.

99. Buttini M, Orth M, Bellosta S, et al. Expression of human apolipoprotein E3 or E4 in the brains of Apoe-/- mice: isoform-specific effects on neurodegeneration. *J Neurosci* 1999;19:4867–4880.

100. Mann DM, Iwatsubo T, Nochlin D, et al. Amyloid (Abeta) deposition in chromosome 1-linked Alzheimer's disease: the Volga German families. *Ann Neurol* 1997;41:52–57.

100a. Ma J, Yee A, Brewer HB Jr., et al. Amyloid associated proteins α1 antichymotrypsin apolipoprotein E promotes amyloid beta protein into filaments. *Nature* 1994;372:92–94.

101. Webster S, Rogers J. Relative efficacies of amyloid beta peptide (A beta) binding proteins in A beta aggregation. *J Neurosci Res* 1996;46:58–66.

102. Holtzman DM, Bales KR, Wu S, et al. Expression of human apolipoprotein E reduces amyloid-beta deposition in a mouse model of Alzheimer's disease. *J Clin Invest* 1999;103:R15–R21.

103. Wood SJ, Chan W, Wetzel R. Seeding of A beta fibril formation is inhibited by all three isotypes of apolipoprotein E. *Biochemistry* 1996;35:12623–12628.

104. Bales KR, Verina T, Dodel, RC, et al. Lack of apolipoprotein E dramatically reduces amyloid beta-peptide deposition. *Nat Genet* 1997;17:263–264.

105. Schenk D, Barbour R, Dunn W, et al. Immunization with amyloid-beta attenuates Alzheimer-disease-like pathology in the PDAPP mouse. *Nature* 1999;400:173–177.

106. Lansbury PT Jr. Inhibition of amyloid formation: a strategy to delay the onset of Alzheimer's disease. *Curr Opin Chem Biol* 1997;1:260–267.

107. Cotman C, Pike C. Beta-amyloid and its contributions to neurodegeneration in Alzheimer disease. In: Terry R, Katzman R, Bick K, eds. *Alzheimer's disease*. New York: Raven, 1994:305–315.

108. Buxbaum JN, Tagoe CE. The genetics of the amyloidoses. *Annu Rev Med* 2000;51:543–569.

109. Lu PJ, Wulf G, Zhou XZ, et al. The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein [see Comments]. *Nature* 1999;399:784–788.

110. Imahori K, Hoshi M, Ishiguro K, et al. Possible role of tau protein kinases in pathogenesis of Alzheimer's disease. *Neurobiol Aging* 1998;19:S93–98.

111. Yang AJ, Chandswangbhuvana D, Margol L, et al. Loss of endosomal/lysosomal membrane impermeability is an early event in amyloid Abeta1–42 pathogenesis. *J Neurosci Res* 1998;52:691–698.

86

Structural and Functional Brain Imaging of Alzheimer Disease

Gary W. Small

Gary A. Small: Department of Psychiatry and Biobehavioral Sciences, Neuropsychiatric Institute, Alzheimer's Disease Centers, Center on Aging, University of California; Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, California.

Of the many laboratory measures and techniques available for understanding and quantifying biological aspects of *Alzheimer disease* (AD), imaging the structure and function of the brain is a particularly attractive approach in that it can provide highly relevant and diverse information using a variety of techniques. The application and interpretation of such information have considerable practical clinical relevance, but as these technologies and our understanding of the disease pathogenesis continue their rapid evolution, so do the potential utilities of these imaging methods in addressing timely neuropsychopharmacologic research issues.

Brain imaging techniques are often categorized as either structural or functional, based on the primary form of information provided. This classification method breaks down, however, when considering newer applications of these techniques. For example, magnetic resonance imaging (MRI) equipment is used to provide functional brain responses with functional MRI (fMRI). Moreover, both positron emission tomography (PET) and single photon emission computed tomography (SPECT) have the potential to provide visualizations of the pathognomonic structural lesions of AD, the amyloid neuritic plaques (NPs), and neurofibrillary tangles (NFTs).

The *in vivo* visualization of relevant structures and functions through brain imaging has several clinical and research applications for AD and other dementias. Recognition of dementia is particularly difficult in its early stages, when family members and physicians often incorrectly attribute the patient's symptoms to normal aging (1 ,2). Systematic studies indicate that the frequency of unrecognized memory impairment, beyond that associated with normal aging, or a dementia diagnosis can range from 50% to 90% of cases (3 ,4). A related application is the differential diagnosis of various dementia causes. The gradual onset and progressive cognitive decline of AD may be difficult to distinguish clinically from other chronic dementias, including dementia with Lewy bodies, vascular dementia, frontotemporal dementia, and late-life depression. Brain imaging techniques may sort out these various causes. The marginal diagnostic value (i.e., added specificity and sensitivity) that a brain imaging procedure provides also has applications to neuropharmacologic clinical trials. When brain imaging improves diagnostic homogeneity, drug efficacy and safety studies are likely to be more informative.

Another application of brain imaging is in the preclinical detection of AD. Neuropathologic, neuropsychological, and brain imaging data point to a form of gradual age-related cognitive decline that precedes AD (5). Use of imaging studies, particularly when coupled with data on genetic risk of AD, is an emerging strategy to identify candidates for pharmacologic interventions that delay cognitive decline progression and disease onset. A related application is the use of brain imaging data to predict and follow treatment response in patients with the full dementia syndrome of AD. Finally, imaging studies also may provide information that clarifies underlying disease mechanisms, which, in turn, may foster improved drug development. In this chapter, I review both available and developing brain imaging techniques and emphasize neuroimaging techniques and measures for presymptomatic AD detection and monitoring pharmacologic interventions.

STRUCTURAL NEUROIMAGING TECHNIQUES

Part of "86 - Structural and Functional Brain Imaging of Alzheimer Disease "

Computed Tomography

Computed tomography (CT) measures the attenuation of an x-ray beam through body tissues. A tissue's appearance will vary according to its attenuation. Bone has the highest attenuation and appears white, whereas gas has the lowest

and appears black. A ring of x-ray generators and detectors obtains images of multiple brain slices as the patient is advanced through the scanner (6). CT can differentiate bone, soft tissue, fluid, and gas with spatial resolution of less than 1 mm. Intravenous contrast medium enhances such pathologic features as bleeding, neoplasm, infection, and inflammation. Limitations of CT include its inability to differentiate gray and white matter and to visualize the posterior fossa clearly (6). Quantitative CT measures have demonstrated greater brain atrophy and ventricular dilatation in patients with AD compared with controls (7). The rate of clinical decline in AD is also related to the rate of ventricular volume change (8).

Magnetic Resonance Imaging

MRI measures the radiofrequency energy that hydrogen atoms of water molecules emit. In a static magnetic field, lower-energy nuclei align with the field, whereas higher-energy nuclei align against the field. When irradiated at a specific frequency, some lower-energy nuclei absorb energy and align against the field. The MRI scanner detects energy emitted when the radiation is discontinued and the nuclei return to their lower-energy state (9). Such energy level changes provide measures of brain structure representations. The rate that nuclei return to their low-energy state determines the type of image produced: T1-weighted images differentiate gray and white matter, and T2 images delineate white matter hyperintensities (9). Because MRI does not involve ionizing radiation, patients can have multiple scans. Spatial resolution is 1 to 2 mm, usually less than that of CT. Much of the work using MRI has focused on regional volumetric changes in patients with AD compared with controls, with an emphasis on atrophy of the hippocampus and nearby medial temporal structures (10).

FUNCTIONAL NEUROIMAGING TECHNIQUES

Part of "86 - Structural and Functional Brain Imaging of Alzheimer Disease "

Quantitative Electroencephalography

The development of computer-analyzed EEGs and the ability to examine regional differences in EEG activity have potential applications to the study of dementia (11). Quantitative EEG coherence measures the synchronization of neuronal activity at two different cortical sites. If different brain areas are simultaneously activated during a task, coherence between these areas will increase. In AD, reductions in resting state coherence occur between intrahemispheric parietal and prefrontal cortical areas, whereas in vascular dementia, reduction in coherence occurs between occipital and parietal areas (12). Changes in coherence in both the resting state and during task performance may become techniques for the differential diagnosis of dementia. Advantages of quantitative EEG are availability, low cost, and lack of radiation exposure. Disadvantages include the possibility of artifact and the fact that the measures are relatively distant from the brain. Moreover, the precise physiologic meaning of the measure is unclear. Resting state coherence in specific areas can be reduced in AD and in vascular dementia. In AD, the greatest reductions in coherence occur between intrahemispheric parietal and prefrontal cortical areas, whereas in vascular dementia, this reduction occurs between occipital and parietal areas (13).

Single Photon Emission Computed Tomography

SPECT involves administration of an inhaled or injected tracer or unstable isotope. Tracer decay leads to single photon emission, the scanner determines the site of the photon source, and a computer generates a three-dimensional image reflecting cerebral blood flow or receptor distribution (14). In comparison with PET, SPECT has lower spatial resolution, particularly for imaging deep structures. Moreover, determining the source of single photon emitters is less precise compared with determining the two photons traveling in opposite directions in PET scanning. Unlike PET, SPECT cannot demonstrate glucose metabolism.

Positron Emission Tomography

PET tracers are positron-emitting nuclides. When a positron encounters an electron, the positron is destroyed and releases photons traveling in opposite directions. A scanner records the simultaneous arrival of two different photons at different detectors (180 degrees apart) and determines the line along which the photons travel. PET images are then constructed from information received by the scanner (15). PET, like SPECT, delineates cerebral blood flow and receptor characteristics. Injection of high-affinity receptor ligands labeled with nuclides can measure receptor density and affinity. Studies of AD often use fluorodeoxyglucose (FDG) to measure cerebral glucose metabolism, which reflects synaptic activity. PET studies have demonstrated characteristic alterations in cerebral blood flow and metabolism in patients with AD that begin in the parietal cortex and spread to the temporal and prefrontal cortices. The degree of hypometabolism correlates with the severity of cognitive impairment (16). PET images can differentiate patients with AD from patients with other dementias and from cognitively intact people (17).

Although investigators have focused considerable attention on the cholinergic system in AD, numerous other neurotransmitter systems are affected, and PET has been used to study them. For example, striatal uptake of the dopamine reuptake ligand [11C]ß-CFT is decreased in AD, a finding indicating involvement of the brain dopaminergic system (18). In addition to serotonergic deficits (19), cholinergic

nicotinic and muscarinic receptors have been studied using PET radioligands (20).

Both PET and SPECT are noninvasive procedures that demonstrate neuronal activity or receptor characteristics. Advantages of PET include its better spatial resolution and the type of biological information it provides. Because of their radiochemical characteristics, positron emitters (PET tracers) can produce more ligands than photon emitters (SPECT tracers) for receptor studies. Lower scanner costs and greater availability of PET tracers have led to wider availability.

Magnetic Resonance Spectroscopy

Nuclei produce magnetic fields that modify the fields of neighboring atoms of the same molecule. Such “shielding” produces a small variation in the resonant frequency known as a chemical shift. The magnetic resonance spectrum display according to frequency demonstrates an element’s different chemical forms as characteristic peaks. These spectroscopy displays provide information on biologically important elements, thus reflecting tissue metabolite concentrations (21). Magnetic resonance spectroscopy (MRS) is noninvasive, lacks ionizing radiation exposure, and can provide quantitative regional measures of biochemical and physiologic processes. Schuff and associates used proton MRS (1H MRS) and tissue-segmented and volumetric MRI to determine whether hippocampal *N*-acetylaspartate (NAA, a neuronal marker) and volume used together provided greater discrimination between patients with AD and normal elderly persons than either measure alone (22). These investigators found that NAA reductions and volume losses made independent contributions to the discrimination of patients with AD from controls. Concentrations of myoinositol- and choline-containing compounds are higher in the occipital and parietal regions of adults with Down syndrome compared with controls (23).

Functional Magnetic Resonance Imaging

Developments in MRI techniques have allowed investigators to use the device to measure brain activity. The altered MRI signal intensity reflects local changes in blood volume or blood flow. The signal intensity of deoxygenated hemoglobin (highly paramagnetic) differs from that of oxygenated hemoglobin. During brain activity, increased blood flow brings more oxygenated blood into the capillary bed. The brain does not metabolize this excess oxygen, and this causes a greater concentration of oxygenated blood to cross over to the venous side leading to a decrease in the magnetic field gradient at the capillaries. The resultant greater magnetic field homogeneity yields a higher MRI signal intensity. Thus, brain regions receiving greater blood flow during brain activity produce a stronger MRI signal than do other regions. By comparing perfusion in activated and nonactivated states, areas of relative brain activity can be identified (24). Thus, fMRI provides measures of signal intensity that are associated with relative cerebral blood flow during memory or other cognitive tasks (25 ,26 ,27 ,28 ,29 and 30), and it has the advantages of high resolution in space and time and lack of radiation exposure. The MRI signal intensity associated with a particular task in comparison with the control condition reflects blood flow and consequently neural activity, but only indirectly (31 ,32). fMRI studies of patients with AD reveal lowered brain activity in parietal and hippocampal regions and relatively higher activity in primary cortices unaffected by the disease (33).

Diffusion Tensor Imaging

A critical aspect of the interpretation of normal and abnormal brain function is neuronal connectivity. One method that provides visualizations of projections of axonal fibers is diffusion tensor MRI (34). The technique offers quantitative information on the directionality (anisotropy) of water diffusion and thus information on local fiber orientation and integrity of white matter tracks. Diffusion tensor imaging (DTI) quantifies and visualizes diffusional anisotropy within each voxel, and computer algorithms relate DTI data to three-dimensional projections of axonal fibers. The degree of neuronal connectivity loss observed in AD is clearly a useful measure to monitor as the disease progresses, and combining DTI with other imaging modalities (e.g., PET, fMRI) may be a useful approach, which has been described in other neuropsychiatric disorders (35). A DTI study of diffusion anisotropy of pyramidal tract in ten older and ten younger adults subjects found that older persons had lower values in the cerebral peduncle, with no differences in the pons and medulla (36). A study of hippocampal water diffusion changes and temporal white matter using DTI in patients with AD and controls suggests that decreased fiber density occurs early in the temporal white matter, probably related to secondary degeneration from gray matter disease of the medial temporal lobe (37). Moreover, studies using DTI indicate mild myelin loss in patients with AD, even though white matter appears normal on MRI, and areas of periventricular hypertrophy show a definite loss of myelin and axons, including incomplete infarction (38).

IMAGING ANALYSIS TOPICS RELEVANT TO DEMENTIA RESEARCH

Part of "86 - Structural and Functional Brain Imaging of Alzheimer Disease "

Numerous variables influence the methodologic error introduced into imaging studies of dementia, including the stability and resolution of imaging systems, the reliability of image analysis, the effect size, undefined neuropathology, the stage of illness, and various confounding factors. The particular method of image analysis provides different levels of image detail and sources of error.

Imaging Registration

In early studies of SPECT and PET imaging, regions of interest (ROIs) were drawn directly on the PET images, matched to a standard atlas. This approach has high interrater reliability (39), but better anatomic definition is possible with computer software that provides image overlay programs merging structural and functional imaging data within the same subject for ROI analyses (40). Such algorithms also allow alignment of multiple PET images obtained from a single subject (41). Registration of PET images that have uniform three-dimensional resolution permits direct regional metabolic comparisons, whereas MRI and PET registration allows precise anatomic localization of those metabolic data in terms of the individual's structural anatomy.

Statistical Parametric Mapping

In statistical parametric mapping (SPM) analysis (42 ,43), images are coregistered and reoriented into a standardized coordinate system, spatially smoothed, and normalized to mean global activity. The set of pooled data are then assessed on a voxel-by-voxel basis, to identify the profile of voxels that significantly change between conditions (e.g., baseline versus follow-up scan). The probability of finding by chance any region containing its voxel of maximal significance is assessed after adjusting for multiple comparisons. It is not surprising that results vary according to analytic method. The ROI approach depends on *a priori* assumptions on size and shape of regions defined by structural criteria. If functionally relevant areas deviate from *a priori* assumptions, an area not functionally involved will dilute the statistical effect. By contrast, SPM analysis relies on pooled brain images spatially normalized into a common space; the extent that the original size and shapes of brains differ will inevitably introduce some error. Minoshima and colleagues applied an automated image analysis method, wherein metabolic reductions were standardized using three-dimensional stereotactic surface projections from FDG PET scans of patients with AD compared with controls (44). This approach has been useful in studies of asymptomatic subjects at risk of developing AD (45).

Atrophy Correction

Decreased functional imaging signal intensity in patients with AD may result from local atrophy causing partial volume effects. Approaches to correcting for cerebral atrophy and partial volume effects include a binary method, wherein cerebrospinal fluid (CSF) and brain tissue are segmented and the composite tissue images are convolved to the in-plane resolution of the PET image. The binary method ignores averaging between gray matter and white matter, and pathologic and imaging data suggest gray matter losses in AD greater than white matter losses (46). In trinary correction methods, CSF, gray matter, and white matter segmentation are included (47) (Fig. 86.1). Computer simulation studies (47) have shown close to 100% recovery of radiotracer concentration in neocortical gray matter and hippocampus, and they indicate that errors in gray matter segmentation and errors in registration of PET and MRI images result in less than 15% inaccuracy in the corrected image. Other work indicates that the neocortical deficits observed in AD reflect true metabolic reductions and are not just the result of atrophy (48).

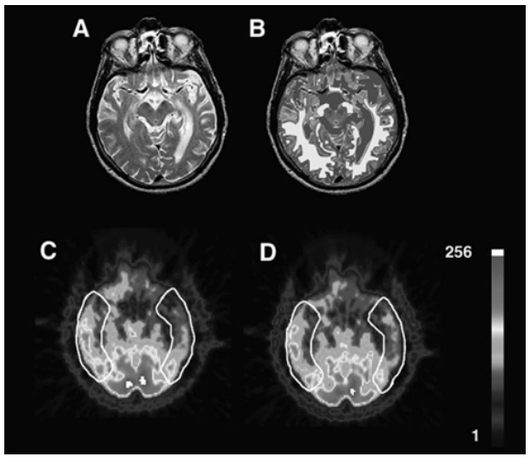


FIGURE 86.1. Partial volume correction using the trinary method. Example of the method using MRI and FDG PET images of a patient with primary progressive aphasia and left (seen on the **right** in figure image planes viewed from below the subject's head) temporal atrophy. Trinary segmentation (gray matter, white matter, CSF) was performed on the MRI image at the level of the temporal lobe. The MRI image was then registered and resliced to align with the PET image. The corrected left temporal glucose metabolic rate was higher than the uncorrected left temporal glucose metabolic rate, an expected result in light of the left temporal lobe atrophy. The corrected and uncorrected glucose metabolic raters for the right temporal lobe were nearly the same, consistent with the minimal atrophic changes in the right hemisphere. **A:** MRI scan without tissue segmentation. **B:** MRI scan with tissue segmentation (*yellow*, white matter; *gray*, gray matter; *blue*, CSF). **C:** Uncorrected PET image (*white lines* indicate temporal region of interest) showing left-to-right temporal lobe asymmetry of glucose metabolic rate. **D:** Corrected PET image showing less striking asymmetry. (Courtesy of Dr. Henry Huang, Department of Molecular and Medical Pharmacology, UCLA School of Medicine, Los Angeles.) See color version of figure.

USE OF NEUROIMAGING FOR PRESYMPTOMATIC AD DETECTION AND PHARMACOLOGIC TREATMENT MONITORING

Part of "86 - Structural and Functional Brain Imaging of Alzheimer Disease "

During the past decade, investigators have been focusing their efforts on early detection of AD at clinical stages before

the time when a physician confirms a clinical diagnosis of probable AD (49). The aim is to begin preventive pharmacologic treatments before extensive neuronal damage develops. Brain imaging has become an important tool for the development of surrogate markers that will effectively identify people with only mild cognitive losses who are likely to progress in their cognitive loss and who will eventually develop the full dementia syndrome of AD. As novel, disease-modifying agents emerge, these surrogate brain imaging markers will be critical in determining drug efficacy and will facilitate drug development in both animal models and human studies.

Several diagnostic entities have been described in efforts to characterize age-related cognitive decline better. The mildest form of age-related memory decline is known as age-associated memory impairment (AAMI) (50), characterized by self-perception of memory loss and a standardized memory test score greater than or equal to 1 standard deviation (SD) below the aged norms. In people 65 years of age or older, its estimated prevalence is 40%, afflicting approximately 16 million people in the United States (51). Only about 1% of such patients develop dementia each year. A more severe form of memory loss is mild cognitive impairment (MCI), often defined by significant memory deficits without functional impairments. People with MCI show memory impairment that is greater than or equal to 1.5 SD below aged norms on such memory tasks as delayed paragraph recall (52). Approximately 10% of people 65 years old or older suffer from MCI, and nearly 15% develop AD each year (52 ,53). Brain imaging studies of presymptomatic AD focus on both these forms of age-related memory decline.

Evidence of Presymptomatic Disease

Neuropathologic, neuroimaging, and clinical research supports the idea that the dementing process leading to AD begins years before a clinical diagnosis of probable AD can be confirmed (49). Postmortem studies of nondemented older people indicate that tangle density in healthy aging correlates with age (54), but that some persons demonstrate widely distributed neuritic and diffuse plaques throughout neocortex and limbic structures. Other studies have found that NFT density increases in some persons (55), presumably those who will eventually develop AD, very early in adult life, perhaps even by the fourth decade. The diffuse amyloid deposits in middle-aged nondemented persons are consistent with an early or “preclinical” stage of AD and suggest that the pathologic process progresses gradually, taking 20 to 30 years to proceed to the clinical manifestation of dementia (56). Other supportive evidence includes findings that linguistic ability in early life predicts cognitive decline in late life (57). High diffuse plaque density in nondemented older persons has been observed in the entorhinal cortex and inferior temporal gyrus, in association with acetylcholinesterase fiber density (58). Evidence from animal models also supports compromised hippocampal cholinergic transmission during aging (59). Studies of glucose metabolic rates using PET (45 ,60 ,61) indicate lower regional brain metabolism in middle-aged and older persons with a genetic risk (apolipoprotein E-4 [APOE-4]), lending further support to a prolonged presymptomatic AD stage.

Structural Imaging

Computed Tomography and Magnetic Resonance Imaging

Studies of early detection logically follow from initial work demonstrating the differential diagnostic utility of a brain imaging marker. For structural imaging, particularly MRI, data have emerged on the use of regional atrophy patterns for the positive diagnosis of AD and other neurodegenerative disorders. Studies without neuropathologic confirmation report the utility of medial temporal lobe atrophy, particularly hippocampal atrophy, on CT or MRI for the clinical diagnosis of AD (62). Some, but not all, quantitative MRI studies indicate that white matter hyperintensities correlate with neuropsychological functioning in both healthy elderly persons and demented patients (63 ,64). Other studies indicate loss of cerebral gray matter (46), hippocampal and parahippocampal atrophy (65), and lower left amygdala and entorhinal cortex volumes (66) in patients with AD. In differentiating AD from older normal controls, the sensitivity of various medial temporal atrophy measures ranges from 77% to 92%, with specificities ranging from 49% to 95% (67 ,68 and 69). In older patients with MCI, hippocampal atrophy predicts subsequent conversion to AD (70). Of various analytic methods, computerized volumetric techniques are most accurate, but they are currently labor intensive and are not widely available.

A modified negative-angle axial view designed to cut parallel to the anterior-posterior plane of the hippocampus has been used to assess hippocampal volume using CT or MRI (62). Such hippocampal atrophy is a sensitive and specific predictor of future AD in patients with MCI. Baseline hippocampal ratings accurately predicted decliners with an overall accuracy of 91%. Neuropathologic studies found that the sites of maximal neuronal loss for both AD and MCI are in the CA1, subiculum, and entorhinal cortex (62). Hippocampal atrophy was also found to predict future cognitive decline in older persons without cognitive impairment who were followed-up for nearly 4 years. Visual assessments of medial temporal lobe atrophy on coronal MRI sections show significant correlations between estimated and stereologically measured volumes (71). Because the latter is much more labor intensive, visual readings may be an alternative approach with greater efficiency.

The hippocampus and the temporal horn of the lateral ventricles also may serve as antemortem AD markers in

mildly impaired patients (mean Mini-Mental State Examination [MMSE] score of 24) (72). Although hippocampal atrophy may enable one to distinguish AD from normal aging, such atrophy may be nonspecific, occurring in other dementing disorders (73). MRI hippocampal atrophy measures are not as sensitive as PET glucose metabolism measures, which begin decreasing before the onset of memory decline (74). The presence of MRI white matter hyperintensities does not improve diagnostic accuracy because they occur both in AD and in healthy normal elderly persons (75 ,76).

The entorhinal cortex, a region involved in recent memory performance, is one of the earliest areas to accumulate NFTs (55). Histologic boundaries of the entorhinal cortex from patients with autopsy-confirmed AD and controls have been used to validate a method for measurement of entorhinal cortex size relying on gyral and sulcal landmarks visible on MRI (77). Such measures may be additional early AD detection markers.

Several studies have addressed the interaction between regional atrophy and *APOE* genotype. Increasing dose of *APOE*-4 allele was associated with smaller hippocampal, entorhinal cortical, and anterior temporal lobe volumes in already demented patients (78). A study of nondemented older persons found an association between *APOE*-4 dose and a larger left than right hippocampus (79). Combining medial temporal measures with other functional neuroimaging (80) or *APOE* genotyping may improve the ability of any of these measures alone to predict cognitive decline (81).

In Vivo Imaging of Amyloid Plaques and Neurofibrillary Tangles

The evidence of NP and NFT accumulation years before clinical AD diagnosis suggests that *in vivo* methods that directly image these pathognomic lesions would be useful presymptomatic detection technologies. Current methods for measuring brain amyloid, such as histochemical stains, require tissue fixation on postmortem or biopsy material. Available *in vivo* methods for measuring NPs or NFTs are indirect (e.g., CSF measures) (82). Studies that may lead to direct *in vivo* human Aß imaging include various radiolabeled probes using small organic and organometallic molecules capable of detecting differences in amyloid fibril structure or amyloid protein sequences (83). Investigators also have used chrysamine-G, a carboxylic acid analogue of congo red, an amyloid-staining histologic dye (84), serum amyloid P component, a normal plasma glycoprotein that binds to amyloid deposit fibrils (85), or monoclonal antibodies (86). Methodologic difficulties that hinder progress with these techniques include poor blood–brain barrier crossing and limited specificity and sensitivity. In addition, most approaches do not measure both NPs and NFTs.

In a breakthrough, Barrio and colleagues (87) used a hydrophobic radiofluorinated derivative of 1,1-dicyano-2-[6-(dimethylamino)naphthalen-2-yl]propene (FDDNP) (88) with PET to measure the cerebral localization and load of NFTs and SPs in patients with AD (n = 7) and controls (n = 3). The FDDNP was injected intravenously and was found to cross the blood–brain barrier readily in proportion to blood flow, as expected from highly hydrophobic compounds with high membrane permeability. Greater accumulation and slower clearance of FDDNP were observed in brain regions with high concentrations of NPs and NFTs, particularly the hippocampus, amygdala, and entorhinal cortex. The FDDNP residence time in these regions showed significant correlations with immediate and delayed memory performance measures (89), and areas of low glucose metabolism correlated with high FDDNP activity retention. The probe showed visualization of NFTs, NPs, and diffuse amyloid in AD brain specimens using *in vitro* fluorescence microscopy, which matched results using conventional stains (e.g., thioflavin S) in the same tissue specimens. Thus, FDDNP-PET imaging is a promising noninvasive approach to longitudinal evaluation of NP and NFT deposition in preclinical AD.

Magnetic Resonance Spectroscopy

Initial studies of MRS as a preclinical AD detection technique found significantly lower NAA concentrations in persons with AD and AAMI compared with controls (90). Mean inositol concentration was significantly higher in AD than in controls, whereas persons with AAMI had intermediate values. Another study focused on patients with Down syndrome because they invariably develop AD by the time they reach their thirties or forties. Concentrations of myoinositol- and choline-containing compounds found using 1H MRS were significantly higher in the occipital and parietal regions in 19 nondemented adults with Down syndrome and in 17 age- and sex-matched healthy controls (23). Moreover, older patients with Down syndrome (42 to 62 years) had higher myoinositol levels than younger subjects (28 to 39 years), a finding suggesting that this approach may be eventually useful as a preclinical AD marker.

Functional Imaging

Positron Emission Tomography

Using FDG PET, our group reported that parietal hypometabolism predicted future AD in people with questionable dementia (91), and even people with very mild age-related memory complaints have baseline PET patterns predicting cognitive decline after 3 years (92). These initial studies using PET for early AD detection emphasized family history of AD as a risk factor for future cognitive decline. A change in focus came with the discovery of the *APOE* genetic risk for AD. The first report combining PET imaging and *APOE*

genetic risk in people with a family history of AD included 12 nondemented relatives with *APOE*-4 and 19 relatives without *APOE*-4 and compared them with seven patients with probable AD (61). "At-risk" subjects had mild memory complaints, normal cognitive performance, and at least two relatives with AD. Persons with *APOE*-4 did not differ from those without *APOE*-4 in mean age at examination (56.4 versus 55.5 years) or in neuropsychologic performance. Parietal metabolism was significantly lower and left-right parietal asymmetry was higher in at-risk subjects with *APOE*-4 compared with those without *APOE*-4. Patients with dementia had significantly lower parietal metabolism than did at-risk persons with *APOE*-4.

The following year, Reiman and associates replicated these results and extended them to other brain regions (45). These investigators found hypometabolism in temporal, prefrontal, and posterior cingulate regions in a study of 11 nondemented *APOE*-4 homozygotes (4/4 genotype) and in 22 *APOE*-3 homozygotes (3/3 genotype) of similar ages to those in our own initial study (midfifties). They also applied an automated image analysis method, wherein metabolic reductions were standardized using three-dimensional stereotactic surface projections from FDG PET scans of patients with AD compared with controls (44). The results from these two studies (45 ,61) provided independent confirmation of an association between genetic risk and regional cerebral glucose hypometabolism.

Our group confirmed these two initial reports in a study that included none of the subjects participating in our previous report on *APOE* and PET (61), in a study of 65 persons in the 50- to 84-year age range (mean ± = 67.3±9.4 years), with or without a family history of AD (93). Of the 65 study subjects, 54 were nondemented (27 were *APOE*-4 carriers and 27 were subjects without *APOE*-4), and 11 were demented and were diagnosed with probable AD (49). The nondemented study subjects were aware of a gradual onset of mild memory complaints (e.g., misplacing familiar objects, difficulty in remembering names). The nondemented subjects, however, had memory performance scores within the norms for cognitively intact persons of the same age and educational level. The *APOE*-4 carriers had a small but consistent nonsignificant reduction in cognitive performance. As predicted, baseline comparisons among the three subject groups indicated the lowest metabolic rates for the AD group, intermediate rates for the nondemented *APOE*-4 carriers, and highest rates for the nondemented group without *APOE*-4 in several cortical regions, including inferior parietal, lateral temporal, and posterior cingulate (Fig. 86.2).

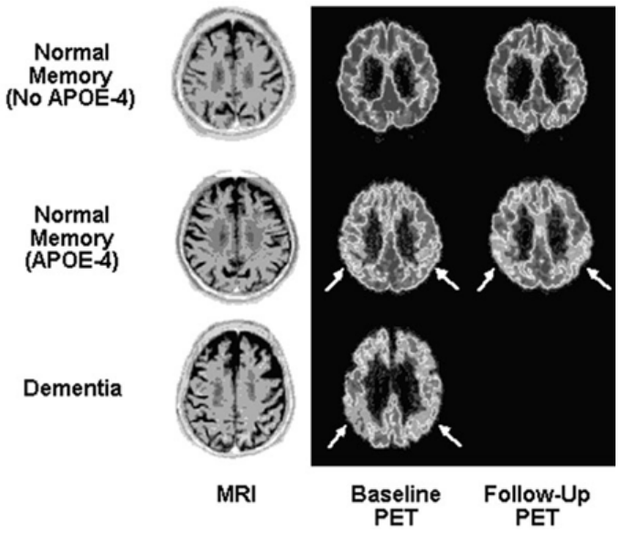


FIGURE 86.2. Examples of PET images (comparable parietal lobe levels) coregistered to each subject's baseline MRI scan for an 81-year-old nondemented woman (*APOE*-3/3 genotype; **upper images**), a 76-year-old nondemented woman (*APOE*-3/4 genotype; **middle images**), and a 79-year-old woman with AD (*APOE*-3/4 genotype; **lower images**). The last column shows 2-year follow-up scans for the nondemented women. Compared with the nondemented patient without *APOE*-4, the nondemented *APOE*-4 carrier had 18% **(right)** and 12% **(left)** lower inferior parietal cortical metabolism, whereas the demented woman's parietal cortical metabolism was 20% **(right)** and 22% **(left)** lower, as well as more widespread metabolic dysfunction resulting from disease progression. Two-year follow-up scans showed minimal parietal cortical decline for the woman without *APOE*-4, but bilateral parietal cortical decline for the nondemented woman with *APOE*-4, who also met clinical criteria for mild AD at follow-up. MRI scans were within normal limits. (From Small GW, Ercoli LM, Silverman DHS, et al. Cerebral metabolic and cognitive decline in persons at genetic risk for Alzheimer's disease. *Proc Natl Acad Sci USA* 2000;97:6037–6042, with permission.) See color version of figure.

Another FDG PET study focused on older patients with Down syndrome who were at risk of AD (94). The investigators hypothesized that an audiovisual stimulation paradigm would serve as a stress test and would reveal abnormalities in parietal and temporal cerebral glucose metabolism before dementia developed. At mental rest, younger and older patients with Down syndrome did not differ in glucose metabolic patterns. During audiovisual stimulation, however, the older patients showed significantly lower parietal and temporal metabolism. Families with familial AD linked to chromosome 14 or amyloid precursor protein (*APP*) mutations have been studied with FDG PET as well (95). In such families with early-onset AD, approximately half of relatives who live to the age at risk will develop AD. Although pedigree members with AD show typical parietal and temporal hypometabolism, asymptomatic relatives at risk of AD show a similar but less severe hypometabolic pattern.

Single Photon Emission Computed Tomography

Johnson and associates (96) used SPECT with technetium-hexamethylpropyleneamineoxime (HMPAO) to study longitudinal cerebral perfusion of patients with questionable AD (clinical dementia rating = 0.5) (97) and controls. Regional decreases in perfusion in patients whose diagnosis

converted to AD were most prominent in the hippocampal-amygdaloid complex, the anterior and posterior cingulate, and the anterior thalamus. Including *APOE* status did not influence results. A direct comparison of FDG PET and HMPAO-SPECT in their ability to differentiate AD from vascular dementia indicated higher diagnostic accuracy for PET regardless of dementia severity (98). Using ROC curves, PET diagnostic accuracy was better than SPECT for an MMSE score greater than 20 (87.2% versus 62.9%) and for an MMSE score less than or equal to 20 (100% versus 81.2%). Other studies confirmed a lower sensitivity for even high-resolution SPECT compared with PET (99). Moreover, the parietal hypoperfusion observed using SPECT in patients with AD has been observed in such other conditions as normal aging, vascular dementia, posthypoxic dementia, and sleep apnea (100).

Functional Magnetic Resonance Imaging

Two studies combined *APOE* genotyping and fMRI in persons at risk of AD. Bookheimer and associates (101) performed fMRI studies while 30 cognitively intact middle-aged and older persons (mean age, 63 years) memorized and retrieved unrelated word pairs. The 16 *APOE*-4 carriers did not differ significantly from the 14 persons without *APOE*-4 in age, prior educational achievement, or rates of AD family history. Brain activation patterns were determined during both learning and retrieval task periods and were analyzed using between-group and within-subject approaches. Memory performance was reassessed on 12 subjects after 2 years of follow-up. The *APOE*-4 carriers had significantly greater magnitude and spatial extent of MRI signal intensity during memory performance in regions affected by AD, including bilateral hippocampal and left parietal and prefrontal (Fig. 86.3). This pattern of activation was greater in the left hemisphere, consistent with the verbal nature of the task, and during the retrieval rather than the learning condition. Longitudinal data indicated that greater baseline brain activation correlated with verbal memory decline assessed 2 years later. The greater signal in persons with the *APOE*-4 genetic risk suggests that the brain may recruit additional neurons to compensate for subtle deficits. Moreover, the longitudinal data are encouraging that fMRI may be a useful approach to prediction of future cognitive decline and early AD detection.

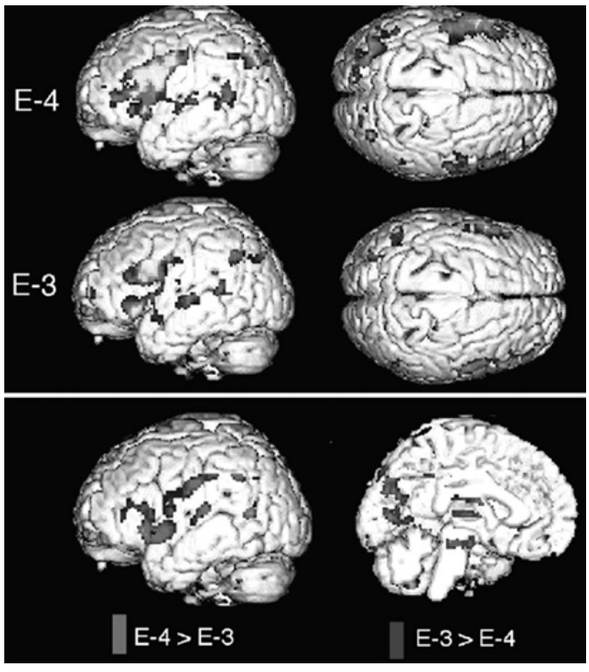


FIGURE 86.3. Statistical parametric maps of recall versus control blocks for *APOE*-4 carriers and noncarriers. Maps were standardized into a common coordinate system. Both groups showed significant MRI signal intensity increases in frontal, temporal, and parietal regions, and the *APOE*-4 group had greater extent and intensity of activation. The *APOE*-4 group showed additional activations in the left parahippocampal region, left dorsal prefrontal cortex, and other regions in the inferior and superior parietal lobes, and anterior cingulate. (From Bookheimer SY, Strojwas MH, Cohen MS, et al. Brain activation in older people at genetic risk for Alzheimer's disease. *N Engl J Med* 2000;343:450-456, with permission.) See color version of figure.

By contrast, other types of memory tasks may produce different patterns of brain activation. In another study of persons at risk for AD, visual naming and letter fluency tasks were used to activate brain areas involved in object and face recognition during fMRI scanning (102). Subjects in the high-risk group had at least one first-degree relative with AD and one *APOE*-4 allele. The low-risk group was matched for age, education, and cognitive performance. The high-risk group showed reduced activation in the middle and posterior inferotemporal regions bilaterally. Such decreased activation patterns could result from subclinical neuropathology in the inferotemporal region or in the inputs to that region.

Longitudinal Studies of Glucose Metabolism of Persons At Risk of Dementia

Both the University of California, Los Angeles (UCLA) and the University of Arizona groups have reported on longitudinal FDG PET follow-up data on nondemented persons at risk of AD. At UCLA, a total of 20 nondemented subjects (ten *APOE*-4 carriers and ten without *APOE*-4) received repeat PET and neuropsychologic testing 2 years after baseline assessment (mean±SD for follow-up was 27.9±1.7 months) (93). The ten *APOE*-4 carriers available for longitudinal study were similar to the ten noncarriers in mean±SD age (67.9±8.9 versus 69.6±8.1 years) and educational achievement (14.4±1.8 versus 16.4±2.8 years). Memory performance scores did not differ significantly

according to genetic risk either at baseline or follow-up, and the *APOE*-4 carriers and noncarriers did not differ significantly in cognitive change after 2 years.

The ROI analysis of PET scans performed after 2 years showed significant glucose metabolic decline (4%) in the left posterior cingulate region in *APOE*-4 carriers. The SPM analysis showed significant metabolic decline in the inferior parietal and lateral temporal cortices with the greatest magnitude (5%) of metabolic decline in the temporal cortex (Fig. 86.4). After correction for multiple comparisons, this decline remained significant for the *APOE*-4 group, wherein a decrease in metabolism was documented for every subject. Based on these data from only ten subjects, the estimated power of PET under the most conservative circumstances is 0.9 to detect a one-unit decline from baseline to follow-up using a one-tailed test. Such findings suggest that combining PET and AD genetic risk measures will allow investigators to use relatively small sample sizes when testing antidementia treatments in preclinical AD stages. The University of Arizona group also found that *APOE*-4 heterozygotes had significant 2-year declines in regional brain activity, the largest of which was in temporal cortex, and that these reductions were significantly greater than those in *APOE*-4 noncarriers. Their findings suggest that as few as 22 cognitively normal, middle-aged *APOE*-4 heterozygotes would be needed in each treatment arm (i.e., active drug and placebo) to test a prevention therapy over a 2-year period (103).

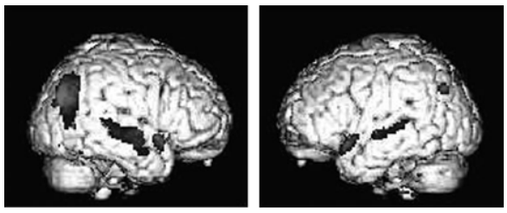


FIGURE 86.4. Regions showing the greatest metabolic decline after 2 years of longitudinal follow-up in nondemented patients with *APOE*-4 (SPM analysis) included the right lateral temporal and inferior parietal cortex (brain on the **left side** of the figure). Voxels undergoing metabolic decline ($p < .001$, before correction) are displayed in color, with peak significance ($z = 4.35$) occurring in Brodmann's area 21 of the right middle temporal gyrus. (From Small GW, Ercoli LM, Silverman DHS, et al. Cerebral metabolic and cognitive decline in persons at genetic risk for Alzheimer's disease. *Proc Natl Acad Sci USA* 2000; 343:450–456, with permission.) See color version of figure.

Clinical Trials of Presymptomatic Patients Using Neuroimaging Surrogate Markers

The longitudinal findings of significant parietal and temporal metabolic decline in asymptomatic persons at risk of AD because of age or genetic risk or both have now been confirmed at two centers in separate subject cohorts. Together, these studies indicate that combining PET imaging of glucose metabolism and genetic risk may be useful outcome markers in AD prevention trials. Functional brain imaging techniques could be used to track preclinical cognitive decline and to test candidate prevention therapies without having to perform prolonged multisite studies using incipient AD as the primary outcome measure. The consistency and extent of the metabolic decline in these well-screened populations indicate that the PET measures provide adequate power to observe such decline in relatively small subject groups. A similar but less striking metabolic decline pattern was noted in subjects without *APOE*-4 such that larger groups per treatment arm would be needed.

These observations provide an opportunity for presymptomatic treatment trials not previously available. Until now, such trials involved studies of preclinical subjects with more severe memory impairments consistent with MCI, wherein approximately 50% of subjects actually develop dementia over a 3- to 4-year period. The MCI trials have required hundreds of subjects for adequate power. These trials use a categorical variable, incipient dementia, as the primary outcome measure. The introduction of FDG PET imaging combined with *APOE*-4 genetic risk increases efficiency and reduces costs by addressing the research questions with fewer subjects. Our group is currently performing two such placebo-controlled trials, one using the cyclooxygenase-2 inhibitor celecoxib and the other using the cholinesterase inhibitor donepezil.

Cost Benefit and Cost Effectiveness

Using neuroimaging as a surrogate marker early in the disease course, even in preclinical stages, has potential cost benefits beyond the greater efficiency in preclinical trials. Because FDG PET increases diagnostic sensitivity and specificity of AD (104), the technique could improve diagnostic homogeneity in clinical trials of mild to moderate AD. Rather than treating the conventional clinical syndrome of AD, the refined phenotype would include a specific neuroimaging pattern (e.g., parietal and temporal hypometabolism). If PET can improve diagnostic accuracy, particularly in the preclinical and early disease stages, then patients would be treated earlier, with resulting improvements in their daily functioning and quality of life. When uncertain about diagnosis, clinicians generally perform costly repetitive examinations. The greater accuracy of early AD detection that neuroimaging may offer would facilitate early intervention. Offsetting the pharmacy costs would be the cost savings from avoidance of repetitive and unnecessary examinations. Following evidence from placebo-controlled studies, the assessment of economic impact would be another level of analysis driving decision makers to fund new neuroimaging technologies. Definitive diagnosis and treatment

during presymptomatic stages of AD would likely decrease both direct and indirect costs. The improved diagnostic accuracy could improve efficacy in clinical trials and could thus facilitate early optimal treatment, delay further cognitive decline, and meet patient and family expectations of the highest-quality care.

ACKNOWLEDGMENTS

Part of "86 - Structural and Functional Brain Imaging of Alzheimer Disease "

This work is supported in part by the following: the Alzheimer's Association, the Charles A. Dana Foundation, the Montgomery Street Foundation, San Francisco; the Fran and Ray Stark Foundation Fund for Alzheimer's Disease Research, Los Angeles; and National Institutes of Health grants MH52453, AG10123, and AG13308. The views expressed are mine and do not necessarily represent those of the Department of Veterans Affairs.

REFERENCES

1. Mant A, Eyland EA, Pond DC, et al. Recognition of dementia in general practice: comparison of general practitioners' opinions with assessments using the Mini-Mental State Examination and Blessed dementia rating scale. *Fam Pract* 1988;5:184–188.

2. McCartney JR, Palmateer LM. Assessment of cognitive deficit in geriatric patients: a study of physician behavior. *J Am Geriatr Soc* 1985;33:467–471.

3. Ross GW, Abbott RD, Petrovitch H, et al. Frequency and characteristics of silent dementia among elderly Japanese-American men: the Honolulu-Asia Aging Study. *JAMA* 1997;277:800–805.

4. Ryan DH. Misdiagnosis in dementia: comparisons of diagnostic error rate and range of hospital investigation according to medical specialty. *Int J Geriatr Psychiatry* 1994;9:141–147.

5. Small GW, Komo S, La Rue A, et al. Early detection of Alzheimer's disease by combining apolipoprotein E and neuroimaging. *Ann NY Acad Sci* 1996;802:70–78.

6. Gibby WA, Zimmerman RA. X-ray computed tomography. In: Mazziotta JC, Gilman S, eds. *Clinical brain imaging: principles and applications*. Philadelphia: FA Davis, 1992:2–38.

7. Creasey H, Schwartz M, Frederickson H, et al. Quantitative computed tomography in dementia of the Alzheimer type. *Neurology* 1986;36:1563–1568.

8. de Leon MJ, George AE, Reisberg B, et al. Alzheimer's disease: longitudinal CT studies of ventricular change. *AJR Am J Roentgenol* 1989;152:1257–1262.

9. Shellock FG, Morisoli S, Kanal E. MR procedures and biomedical implants, materials, and devices: 1993 update. *Radiology* 1993;189:587–599.

10. Horn R, Ostertun B, Fric M, et al. Atrophy of hippocampus in patients with Alzheimer's disease and other diseases with memory impairment. *Dementia* 1996;7:182–186.

11. Leuchter AF, Spar JE, Walter DO, et al. Electroencephalographic spectra and coherence in the diagnosis of Alzheimer's-type and multi-infarct dementia. *Arch Gen Psychiatry* 1987;44:993–998.

12. Dunkin JJ, Leuchter AF, Newton TF, et al. Reduced EEG coherence in dementia: state or trait marker? *Biol Psychiatry* 1994;35:870–879.

13. Dunkin JJ, Osato S, Leuchter AF. Relationships between EEG coherence and neuropsychological tests in dementia. *Clin Electroencephalogr* 1995;26:47–58.

14. Schuckit MA. An introduction and overview of clinical applications of NeuroSPECT in psychiatry. *J Clin Psychiatry* 1992;53[Suppl]:3–6.

15. Mazziotta JC, Phelps ME. Positron emission tomography studies of the brain. In: Phelps M, Mazziotta J, Schelbert H, eds. *Positron emission tomography and autoradiography: principles and applications for the brain and heart*. New York: Raven, 1986:493–579.

16. Mazziotta JC, Frackowiak RSJ, Phelps ME. The use of positron emission tomography in the clinical assessment of dementia. *Semin Nucl Med* 1992;22:233–246.

17. Silverman DHS, Small GW, Phelps ME. Clinical value of neuroimaging in the diagnosis of dementia: sensitivity and specificity of regional cerebral metabolic and other parameters for early identification of Alzheimer's disease. *Clin Positron Imaging* 1999;2:119–130.

18. Rinne JO, Sahlberg N, Ruottinen H, et al. Striatal uptake of the dopamine reuptake ligand [^{11}C]b-CFT is reduced in Alzheimer's disease assessed by positron emission tomography. *Neurology* 1998;50:152–156.

19. Meltzer CC, Smithy G, DeKosky ST, et al. Serotonin in aging, late-life depression, and Alzheimer's disease: the emerging role of functional imaging. *Neuropsychopharmacology* 1998;18:407–430.

20. Nordberg A, Lundqvist H, Hartvig P, et al. Imaging of nicotinic and muscarinic receptors in Alzhiemer's disease: effect of tacrine treatment. *Dementia Geriatr Cogn Disord* 1997;8:78–84.

21. Weiner MW. NMR spectroscopy for clinical medicine, animal models, and clinical examples. *Ann NY Acad Sci* 1987;508:287–289.

22. Schuff N, Amend D, Ezekiel F, et al. Changes of hippocampal *N*-acetyl aspartate and volume in Alzheimer's disease: a proton MR spectroscopic imaging and MRI study. *Neurology* 1997;49:1513–1521.

23. Huang W, Alexander GE, Daly EM, et al. High brain myo-inositol levels in the predementia phase of Alzheimer's disease in adults with Down's syndrome: a 1H MRS study. *Am J Psychiatry* 1999;156:1879–1886.

24. Belliveau JW, Kennedy DN, McKinstry RC, et al. Functional mapping of the human visual cortex by magnetic resonance imaging. *Science* 1992;254:716–719.

25. Cohen MS, Bookheimer SY. Localization of brain function using magnetic resonance imaging. *Trends Neurosci* 1994;17:268–277.

26. Demb JB, Desmond J, Wagner AD, et al. Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional MRI study of task difficulty and process specificity. *J Neurosci* 1995;15:5870–5878.

27. Gabrieli JDE, Brewer JB, Desmond JE, et al. Separate neural bases of two fundamental memory processes in the human medial temporal lobe. *Science* 1997;276:264–266.

28. Stern CE, Corkin S, González RG, et al. The hippocampal formation participates in novel picture encoding: evidence from functional magnetic resonance imaging. *Proc Natl Acad Sci USA* 1996;93:8660–8665.

29. Schacter DL, Buckner RL, Koutstaal W, et al. Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. *Neuroimage* 1997;6:259–269.

30. Wagner AD, Schacter DL, Rotte M, et al. Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science* 1998;281:1188–1191.

31. Fox PT, Raichle ME. Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proc Natl Acad Sci USA* 1986;83:1140–1144.

32. Ogawa S, Tank DW, Menon R, et al. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proc Natl Acad Sci USA* 1992;89:5951–5955.

33. Backman L, Andersson JLR, Nyberg L, et al. Brain regions associated with episodic retrieval in normal aging and Alzheimer's disease. *Neurology* 1999;52:1861–1870.

34. Jones EK, Simmons A, Williams SCR, et al. Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. *Magn Reson Med* 1999;42:37–41.

35. Buchsbaum MS, Tang CY, Peled S, et al. MRI white matter diffusion anisotropy and PET metabolic rate in schizophrenia. *Neuroreport* 1998;9:425–430.

36. Virta A, Barnett A, Pierpaoli C. Visualizing and characterizing white fiber structure and architecture in the human pyramidal tract using diffusion tensor MRI. *Magn Reson Imaging* 1999;17:1121–1133.

37. Hanyu H, Sakurai H, Iwamoto T, et al. Diffusion-weighted MR imaging of the hippocampus and temporal white matter in Alzheimer's disease. *J Neurol Sci* 1998;156:195–200.

38. Hanyu H, Shindo H, Kakizaki D, et al. Increased water diffusion in cerebral white matter in Alzheimer's disease. *Gerontology* 1997;43:343–351.

39. Small GW, Stern CE, Mandelkern MA, et al. Reliability of drawing regions of interest for positron emission tomographic data. *Psychiatry Res* 1992;45:177–185.

40. Pelizzari CA, Chen GTY, Spelbring DR, et al. Accurate three-dimensional registration of CT, PET and/or MR images of the brain. *J Comput Assist Tomogr* 1989;13:20–26.

41. Woods RP, Mazziotta JC, Cherry SR. MRI-PET registration with automated algorithm. *J Comput Assist Tomogr* 1993;17:536–546.

42. Friston KJ, Ashburner J, Frith CD, et al. Spatial registration and normalisation of images. *Hum Brain Mapping* 1995;2:165–189.

43. Friston KJ, Holmes A, Poline JB, et al. Detecting activations in PET and fMRI: levels of inference and power. *Neuroimage* 1996;4:223–235.

44. Minoshima S, Frey KA, Koeppe RA, et al. A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. *J Nucl Med* 1995;36:1238–1248.

45. Reiman EM, Caselli RJ, Yun LS, et al. Preclinical evidence of Alzheimer's disease in persons homozygous for the epsilon 4 allele for apolipoprotein E [see Comments]. *N Engl J Med* 1996;334:752–758.

46. Rusinek H, de Leon MJ, George AE, et al. Alzheimer disease: measuring loss of cerebral gray matter with MR imaging. *Neuroradiology* 1991;178:109–114.

47. Müller-Gärtner HW, Links JM, Prince JL, et al. Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects. *J Cereb Blood Flow Metab* 1992;12:571–583.

48. Ibáñez V, Pietrini P, Alexander GE, et al. Regional glucose metabolic abnormalities are not the result of atrophy in Alzheimer's disease. *Neurology* 1998;50:1585–1593.

49. McKhann G, Drachman D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939–944.

50. Crook T, Bartus RT, Ferris SH, et al. Age-associated memory impairment: proposed diagnostic criteria and measures of clinical change. Report of a National Institute of Mental Health Work Group. *Dev Neuropsychol* 1986;2:261–276.

51. Larrabee GJ, Crook TH. Estimated prevalence of age-associated memory impairment derived from standardized tests of memory function. *Int Psychogeriatr* 1994;6:95–104.

52. Petersen RC, Smith GE, Waring SC, et al. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol* 1999;56:303–308.

53. Andersen K, Nielsen H, Lolk A, et al. Incidence of very mild to severe dementia and Alzheimer's disease in Denmark: the Odense Study. *Neurology* 1999;52:85–90.

54. Price JL, Morris JC. Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. *Ann Neurol* 1999;45:358–368.

55. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol (Berl)* 1991;82:239–259.

56. Arai T, Ikeda K, Akiyama H, et al. A high incidence of apolipoprotein E ε4 allele in middle-aged non-demented subjects with cerebral amyloid beta protein deposits. *Acta Neuropathol (Berl)* 1999;97:82–84.

57. Snowdon DA, Kemper SJ, Mortimer JA, et al. Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: findings from the nun study. *JAMA* 1996;275:528–532.

58. Beach TG, Honer WG, Hughes LH. Cholinergic fibre loss associated with diffuse plaques in the non-demented elderly: the preclinical stage of Alzheimer's disease? *Acta Neuropathol (Berl)* 1997;93:146–153.

59. Shen J, Barnes CA. Age-related decrease in cholinergic synaptic transmission in three hippocampal subfields. *Neurobiol Aging* 1996;17:439–451.

60. Small GW, Ercoli LM, Huang S-C, et al. PET and genetic risk for Alzheimer disease. *J Nucl Med* 1999;40[Suppl]:70.

61. Small GW, Mazziotta JC, Collins MT, et al. Apolipoprotein E type 4 allele and cerebral glucose metabolism in relatives at risk for familial Alzheimer disease. *JAMA* 1995;273:942–947.

62. de Leon MJ, George AE, Golomb J, et al. Frequency of hippocampal formation atrophy in normal aging and Alzheimer's disease. *Neurobiol Aging* 1997;18:1–11.

63. Boone KB, Miller BL, Lesser IM, et al. Neuropsychological correlates of white-matter lesions in healthy elderly subjects. a threshold effect. *Arch Neurol* 1992;49:549–554.

64. Lopez OL, Becker JT, Rezek D, et al. Neuropsychiatric correlates of cerebral white-matter radiolucencies in probable Alzheimer's disease. *Arch Neurol* 1992;49:828–834.

65. Kesslak JP, Nalcioglu O, Cotman CW. Quantification of magnetic resonance scans for hippocampal and parahippocampal atrophy in Alzheimer's disease. *Neurology* 1991;41:51–54.

66. Pearlson GD, Harris GJ, Powers RE, et al. Quantitative changes in mesial temporal volume, regional cerebral blood flow, and cognition in Alzheimer's disease. *Arch Gen Psychiatry* 1992;49:402–408.

67. Laakso MP, Soininen H, Partanen K, et al. MRI of the hippocampus in Alzheimer's disease: sensitivity, specificity, and analysis of the incorrectly classified subjects. *Neurobiol Aging* 1998;19:23–31.

68. Pasquier F, Lavenu I, Lebert F, et al. The use of SPECT in a multidisciplinary memory clinic. *Dementia Geriatr Cogn Disord* 1997;8:85–91.

69. Pucci E, Belardinelli N, Regnicolo L, et al. Hippocampus and parahippocampal gyrus linear measurements based on magnetic resonance in Alzheimer's disease. *Eur Neurol* 1998;39:16–25.

70. Jack CR, Petersen RC, Xu YC, et al. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology* 1999;52:1397–1403.

71. Wahlund L-O, Julin P, Lindqvist J, Scheltens P. Visual assessment of medial temporal atrophy in demented and healthy control subjects: correlation with volumetry. *Psychiatry Res Neuroimaging* 1999;90:193–199.

72. Killiany RJ, Moss MB, Albert MS, et al. Temporal lobe regions on magnetic resonance imaging identify patients with early Alzheimer's disease. *Arch Neurol* 1993;50:949–954.

73. Laakso MP, Partanen K, Riekkinen P, et al. Hippocampal volumes in Alzheimer's disease, Parkinson's disease, with and without dementia, and in vascular dementia: an MRI study. *Neurology* 1996;46:678–681.

74. Reiman EM, Uecker A, Caselli RJ, et al. Hippocampal volumes in cognitively normal persons at genetic risk for Alzheimer's disease. *Ann Neurol* 1998;44:288–291.

75. Erkinjuntti T, Gao F, Lee DH, et al. Lack of difference in brain hyperintensities between patients with early Alzheimer's disease and control subjects. *Arch Neurol* 1994;51:260–268.

76. Mauri M, Sibilla L, Bono G, et al. The role of morpho-volumetric and memory correlations in the diagnosis of early Alzheimer dementia. *J Neurol* 1998;245:525–530.

77. Bobinski M, de Leon MJ, Convit A, et al. MRI of entorhinal cortex in mild Alzheimer's disease. *Lancet* 1999;353:38–40.

78. Geroldi C, Pihlajamaki M, Laakso MP, et al. APOE-ε4 is associated with less frontal and more medial temporal lobe atrophy in AD. *Neurology* 1999;53:1825–1832.

79. Soininen H, Partanen K, Pitkanen A, et al. Decreased hippocampal volume asymmetry on MRIs in nondemented elderly subjects carrying the apolipoprotein E ε4 allele. *Neurology* 1995;45:1467–1472.

80. Mattman A, Feldman H, Forster B, et al. Regional HmPAO SPECT and CT measurements in the diagnosis of Alzheimer's disease. *Can J Neurol Sci* 1997;24:22–28.

81. Jack CR, Petersen RC, Xu YC, et al. Hippocampal atrophy and apolipoprotein E genotype are independently associated with Alzheimer's disease. *Ann of Neurol* 1998;43:303–310.

82. Motter R, Vigo-Pelfrey C, Kholodenko D, et al. Reduction of beta-amyloid peptide 42 in the cerebrospinal fluid of patients with Alzheimer's disease. *Ann Neurol* 1995;38:643–648.

83. Ashburn TT, Han H, McGuinness BF, et al. Amyloid probes based on Congo red distinguish between fibrils comprising different peptides. *Chem Biol* 1996;3:351–358.

84. Klunk WE, Debnath ML, Pettegrew JW. Chrysamine-G binding to Alzheimer and control brain: autopsy study of a new amyloid probe. *Neurobiol Aging* 1995;16:541]]-548.

85. Lovat LB, O'Brien AA, Armstrong SF, et al. Scintigraphy with ^{123}I-serum amyloid P component in Alzheimer disease. *Alzheimer Dis Assoc Disord* 1998;12:208]]-210.

86. Majocha RE, Reno JM, Friedland RP, et al. Development of a monoclonal antibody specific for beta/A4 amyloid in Alzheimer's disease brain for application to *in vivo* imaging of amyloid angiopathy. *J Nucl Med* 1992;33:2184–2189.

87. Barrio JR, Huang S-C, Cole GM, et al. PET imaging of tangles and plaques in Alzheimer disease. *J Nucl Med* 1999;40[Suppl]:70P–71P.

88. Jacobson A, Petric A, Hogenkamp D, et al. 1,1-Dicyano-2-(6-dimethylamino)naphthalen-2-yl)propene (DDNP): a solvent polarity and viscosity sensitive fluorophore for fluorescence microscopy. *J Am Chem Soc* 1996;118:5572–5579.

89. Wechsler D. *Wechsler memory scale*, rev. manual. San Antonio, TX: Psychological Corp., Harcourt Brace Jovanovich, 1987.

90. Parnetti L, Lowenthal DT, Presciutti O, et al. 1H-MRS, MRI-based hippocampal volumetry, and ^{99m}Tc-HMPAO-SPECT in normal aging, age-associated memory impairment, and probable Alzheimer's disease. *J Am Geriatr Soc* 1996;44:133–138.

91. Kuhl DE, Small GW, Riege WH, et al. Cerebral metabolic patterns before diagnosis of probable Alzheimer's disease. *J Cereb Blood Flow Metab* 1987;7[Suppl 1]:S406.

92. Small GW, La Rue A, Komo S, et al. Predictors of cognitive change in middle-aged and older adults with memory loss. *Am J Psychiatry* 1995;152:1757–1764.

93. Small GW, Ercoli LM, Silverman DHS, et al. Cerebral metabolic and cognitive decline in persons at genetic risk for Alzheimer's disease. *Proc Natl Acad Sci USA* 2000;97:6037–6042.

94. Pietrini P, Dani A, Furey ML, et al. Low glucose metabolism during brain stimulation in older Down's syndrome subjects at risk for Alzheimer's disease prior to dementia. *Am J Psychiatry* 1997;154:1063–1069.

95. Rossor MN, Kennedy AM, Frackowiak RS. Clinical and neuroimaging features of familial Alzheimer's disease. *Ann NY Acad Sci* 1996;777:49–56.

96. Johnson KA, Jones K, Holman BL, et al. Preclinical prediction of Alzheimer's disease using SPECT. *Neurology* 1998;50:1563–1571.

97. Hughes CP, Berg L, Danziger WL, et al. A new clinical scale for the staging of dementia. *Br J Psychiatry* 1982;140:566–572.

98. Mielke R, Heiss W-D. Positron emission tomography for diagnosis of Alzheimer's disease and vascular dementia. *J Neural Transm* 1998;53[Suppl]:237–250.

99. Messa C, Perani D, Lucignani G, et al. High-resolution technetium-99m–HMPAO SPECT in patients with probable Alzheimer's disease: comparison with fluorine-18–FDG PET. *J Nucl Med* 1994;35:210–216.

100. Miller BL, Mena S, Daly J, et al. Temporal-parietal hypoperfusion with single-photon emission computerized tomography in conditions other than Alzheimer's disease. *Dementia* 1990;1:41–45.

101. Bookheimer SY, Strojwas MH, Cohen MS, et al. Brain activation in older people at genetic risk for Alzheimer's disease. *N Engl J Med* 2000;343:450–456.

102. Smith CD, Andersen AH, Kryscio RJ, et al. Altered brain activation in cognitively intact individuals at high risk for Alzhiemer's disease. *Neurology* 1999;53:1391–1396.

103. Reiman EM, Caselli RJ, Chen K, et al. Tracking the decline in cerebral glucose metabolism in persons and transgenic mice at genetic risk for Alzheimer's disease. *Alzheimer Imaging Consortium* 2000.

104. Hoffman JM, Welsh-Bohmer KA, Hanson M, et al. FDG-PET imaging in pathologically verified dementia. *J Nucl Med* 2000.

87

Current and Experimental Therapeutics of Alzheimer Disease

Kenneth L. Davis

Kenneth L. Davis: Department of Psychiatry, Mount Sinai School of Medicine, New York, New York.

An enhanced understanding of the neurobiology and neurochemistry of Alzheimer disease, combined with provocative epidemiologic studies, has led to a plethora of new approaches of treatment. It has been estimated that between 50 and 60 drugs are in or entering clinical trials in Alzheimer disease (1). These approaches range from symptomatic and palliative, to preventive and disease altering. Although no magic bullets have been unveiled, it is clear that since the early 1990s, the therapeutics of Alzheimer disease have inexorably proceeded along a rational course, generating an increasing number of compounds that have entered the marketplace and have shown benefit for patients and caregivers.

ACETYLCHOLINESTERASE INHIBITORS

Part of "87 - Current and Experimental Therapeutics of Alzheimer Disease "

The first drug approved in the United States and Europe with an indication to diminish the intensity of the core symptoms of Alzheimer disease, namely, problems in memory, praxis, and language, was tacrine (Cognex). A few years thereafter, approval through most of Europe and the United States was granted to donepezil (Aricept). In 1999, rivastigmine (Exelon) received approval in Europe and was approved in the United States in 2000. Most recently, galantamine (Reminyl) was approved in Sweden for the treatment of Alzheimer disease. This drug, previously available in Austria under the trade name Nivalin for a host of other indications, is now awaiting approval throughout the rest of Europe and the United States.

Although all these drugs are cholinesterase inhibitors, the mechanism of cholinesterase inhibition and other properties of the compounds make them far less than equivalent. Cholinesterase inhibition can be mediated through numerous different mechanisms, characterized as reversible, irreversible, or pseudoirreversible. Additionally, the relationship among cholinesterase, acetylcholine, and the cholinesterase inhibitor could be either competitive or noncompetitive. The specificity of cholinesterase inhibitors can also vary, with differing affinity for butyrylcholinesterase. Finally, these drugs can also differ in the degree to which they modulate the sensitivity of nicotinic receptors.

The group of cholinesterase inhibitors also differs among classical pharmacokinetic and pharmacodynamic parameters. Degree of protein binding, duration of action, and drug interactions discriminate among the drugs in this class. All these specific differences are delineated in this chapter.

Tacrine

Tacrine (Cognex) is a noncompetitive reversible inhibitor of both butyrylcholinesterase and acetylcholinesterase. In fact, its specificity for butyrylcholinesterase is greater than for acetylcholinesterase. It is of the acridine class. The bioavailability of the drug is variously estimated between 17% and 33%, its peak plasma level occurs relatively rapidly within 1 to 2 hours, and it has a serum half-life of 1½ to 2 hours. The drug has a protein binding of about 75%, and it is metabolized by numerous cytochrome pathways including 1A2 and 2D6. Four times a day dosing is required. The efficacy of tacrine was established in a series of placebo-controlled, double-blind studies (2 ,3 and 4). These studies ultimately led to United States Food and Drug Administration (FDA) approval based on tacrine's ability to improve the core symptoms of Alzheimer disease as reflected in the cognitive subscale of the Alzheimer's Disease Assessment Scale (ADAS-cog) (5) and in the assessment of global change by the clinician (6 ,7 ,8 and 9).

Tacrine has been associated with a high frequency of elevated liver transaminases. At some time in their course of drug administration, approximately 50% of patients receiving tacrine will develop elevation in liver transaminases (10). However, with discontinuation of the drug or even lowering of the dose, most such elevations return to baseline.

Indeed, it is even possible to readminister tacrine to many patients who previously had “transaminitis,” with a subsequent benign course. Given these data, it is not surprising that fatal hepatotoxicity has been extremely rare (11). Extensive experience with tacrine has led to the conclusion that rechallenge with tacrine is possible, unless patients have had jaundice, eosinophilia, or other signs of a hypersensitivity reaction. Nonetheless, frequent monitoring of liver enzymes is a necessary concomitant of administration of tacrine, as is relatively slow drug titration. The most efficacious doses of tacrine are between 120 and 160 mg per day. Patients do not often achieve this dose, and even in those patients in whom a dose of 120 mg is obtained, the minimal time to reach that level is 12 weeks. For all these reasons, tacrine is no longer actively promoted, and it is rarely used.

Donepezil

Donepezil (Aricept) is a piperidine cholinesterase inhibitor that is reversible and has both competitive and noncompetitive features. It is 100% bioavailable, and it reaches peak plasma levels between 3 and 5 hours after administration. The drug has not been shown to have any interaction with food. It is highly protein bound. The drug is metabolized by the cytochrome system, specifically 2D6 and 3A4. Donepezil has a long serum half-life, estimated to be between 70 and 80 hours. The consequence of these characteristics is that donepezil requires only a once-daily dose (12).

Large placebo-controlled, multicenter, randomized trials have established the statistically significant effects of donepezil on ADAS-cog and clinical global measures (13 ,14 ,15 ,16 and 17). The effective doses are 5 and 10 mg administered once per day. Some trials do not show a superiority of 5 mg over 10 mg, although other data would suggest greater superiority for 10 mg (17).

The major adverse events that are associated with donepezil administration are those that can be anticipated from drugs that increase cholinergic activity. These include nausea, vomiting, and diarrhea. To minimize these effects at the higher, more effective, dose of 10 mg, a titration schedule in which patients remain at the 5-mg dose for 6 weeks, before being raised to the 10-mg dose, is recommended.

Other adverse events are less common, but they are also explicable by the cholinomimetic properties of donepezil. The drug has been associated with bradycardia and syncope as well as some sleep disturbance. Increased cholinergic activity is well known to produce bradycardia, and the cholinergic system can have profound effects on sleep architecture including increasing arousal (18 ,19 ,20 ,21 and 22).

Rivastigmine

Rivastigmine (Exelon) is a carbamate that inhibits both acetylcholinesterase and butyrylcholinesterase. Its mechanism of action at cholinesterase is termed pseudoirreversible, meaning that although it binds to the cholinesterase-like irreversible inhibitors do, it is metabolized by cholinesterase, the enzyme it is inhibiting. This circumstance produces a truly “pseudoirreversible” state and accounts for rivastigmine’s half-life of 10 hours, far shorter than would be expected from an irreversible cholinesterase inhibitor. Irreversible inhibitors are active for as long as the time necessary to regenerate cholinesterase, between 2 and 4 weeks. Rivastigmine’s bioavailability is approximately 40%, and its time to peak concentration can be as rapid as half an hour or as long as 2 hours. There is some interaction with food in its absorption. Its binding to plasma proteins is approximately 40%. Its metabolism is totally nonhepatic, and it can be presumed to have minimal drug interactions. Based on its pharmacokinetic and pharmacodynamic characteristics, the drug is given twice a day, with total doses ranging from 6 to 12 mg.

Efficacy of the agent has been established (23). However, only the higher doses of rivastigmine (doses higher than 6 mg) were shown to be efficacious in two pivotal studies, in both ADAS-cog and global measures. High doses also demonstrate efficacy compared with placebo in activities of daily living, as reflected on the progressive deterioration scale. However, for some patients, it is difficult to achieve these high doses. Despite slow dose titration that took up to 12 weeks, approximately 25% of patients receiving more than 6 mg per day of rivastigmine withdrew from the study, and substantially more had some gastrointestinal complaints. The side effects predominantly occurred during dose escalation.

Galantamine

Galantamine is an alkaloid-derived, reversible, competitive acetylcholinesterase inhibitor. It occurs naturally in certain plants. It is relatively selective for acetylcholinesterase, with far less activity at butyrylcholinesterase. The drug is also an agonist at allosteric nicotinic sites, a mechanism of action that it has in common with benzodiazepines that have a similar mechanism of action at the γ-aminobutyric acid$_A$receptor. Activity at this site facilitates release of acetylcholine (24). The drug is less than 10% protein bound, it has a very high bioavailability, and it interacts with food such as to decrease its maximum concentrations (25). Twice-daily dosing is supported by approximately a 9-hour half-life (26). The drug is metabolized in the liver by 2D6 and 3A4 (27).

A series of placebo-controlled, randomized, double-blind studies established the efficacy of galantamine. A dose of 24 mg per day is recommended, although both 16 and 32 mg have been shown to also be efficacious (28 ,29 and 30). A significant difference between drug and placebo has been found on the traditional psychometric and global measures, as well as measures of activities of daily living, ADAS-ADL scale, and behavior, Neuropsychiatric Inventory (NPI).

As with other cholinomimetics, the most common adverse

events are gastrointestinal. These effects are dose related and occur predominantly when the drug's dose is increased. Gastrointestinal side effects can be minimized by dose titration of 8 mg every 4 weeks up to the 16- to 24-mg dose.

Other Cholinesterase Inhibitors

Both metrifonate and extended-release physostigmine have been studied in patients with Alzheimer disease. Indeed, attempts have been made to register these drugs for approval in the United States market, without success. Both drugs have been associated with some degree of efficacy (31 ,32 ,33 and 34). Metrifonate is an organophosphate that is a prodrug for its major metabolite, dichlorvose, which binds irreversibly to acetylcholinesterase. Physostigmine is a reversible inhibitor of both acetylcholinesterase and butyrylcholinesterase.

Development of both these drugs has been stopped, but for different reasons. Metrifonate has been associated with muscle weakness and a possible risk of respiratory muscle dysfunction leading to death. Organophosphate-related delayed neurotoxicity has been well described and has been linked to the binding of a phosphorylated metabolite of organophosphates to neurotoxic esterase (35). A likely related symptom, characterized as a myasthenic-like problem, has also been well described (35), and it seems similar to the problems that had led to the failure of metrifonate to reach the marketplace. Extended-release physostigmine has had a substantial association with nausea and vomiting, with 47% of patients reporting these symptoms during a 12-week trial (31). Unless additional work is done with these compounds to modify this adverse event profile, it is unlikely that either of these drugs will be available for routine use in the clinic.

Course-Altering Properties of Cholinesterase Inhibitors?

That cholinesterase inhibitors are efficacious in the palliative treatment of Alzheimer disease is now beyond question. A far more intriguing issue is whether these drugs alter the course of the disease. Here the data are far more tentative.

The effect of cholinomimetic activity on the processing of amyloid precursor proteins (APP) in various cell culture lines has been studied (36 ,37). Cholinergic stimulation apparently increases the production of nonamyloidogenic APP fragments. In all animals in which potentially amyloidogenic fragments of APP are increased as a consequence of lesioning in various neuronal populations, some, but not all, cholinomimetics normalize that process and diminish the production of amyloidogenic fragments (38). It has also been suggested that the toxicity of beta-amyloid peptide (Aß) itself on neurons is diminished by some cholinesterase inhibitors (39 ,40).

Nicotinic stimulation may be particularly relevant in altering the processes of neurodegeneration. Many epidemiologic studies have demonstrated that the relative risk of Parkinson disease is diminished among smokers compared with nonsmokers (41). However, prevalence studies, not incidence studies, among smokers versus nonsmokers suggest a neuroprotective effect in Alzheimer disease (42 ,43). The absence of incidence data is problematic for the imputation of any epidemiologic data to support the notion that smokers are less likely to be affected by Alzheimer disease than nonsmokers. Still, nicotinic stimulation has been found to protect neurons from ß-amyloid induced neurotoxicity (44 ,45), as well as to enhance the secretion of nonamyloidogenic forms of APP (46).

Ultimately, the question whether cholinomimetic activity, through nicotinic, muscarinic, or other unknown mechanisms, may alter the course of Alzheimer disease rests on clinical data. Such data are scanty and largely indirect. *Post hoc* analyses of patients who participated in the pivotal tacrine studies indicated that patients able to tolerate more than 80 mg per day of the drug had a substantial delay in placement in nursing homes, of the magnitude of approximately 450 days (47). Clearly, there are multiple interpretations of this observed phenomena that need not invoke the effect of tacrine on the progression of Alzheimer disease.

Acetylcholinesterase itself is present in plaques. This enzyme has been shown to enhance the aggregation of ß-amyloid into the more fibrillar form that is deposited in plaques (48 ,49 ,50 and 51). Antibodies to cholinesterase blocks Aß aggregation *in vitro* (52). Whether such effects on aggregation are produced by cholinesterase inhibitors, as occurs with antibodies directed at the cholinesterase molecule *in vitro*, has not been shown. It is possible that the aggregating effects of cholinesterase are facilitated by sites in the enzyme that are totally unaffected by cholinesterase inhibition. Alternatively, cholinesterase inhibition could alter cholinesterase in such a way as to diminish aggregating properties.

Two paradigms that could offer some insight into course-altering properties of cholinesterase inhibitors have been termed *delayed start* and *drug withdrawal*. In the delayed start paradigm, an agent that would alter the course of Alzheimer disease would be expected to have a greater effect in patients who have been started on the drug at time zero than a matched control group started 6 months later. In every study with cholinesterase inhibitors reported to date using a related but flawed delayed start procedure, placebo-treated patients who were given cholinesterase inhibitors 6 months after the group of patients on the drug did not catch up on cognitive measures to the patients who were treated with the drug from the start of the study. However, the interpretability of these data is limited because at the time of switchover from placebo to drug, the studies were no longer blinded. Furthermore, self-selection for switchover, or retention on drug, could occur and further confound these data. In the withdrawal paradigm, patients receiving the drug are randomly discontinued from the drug

on the completion of the trial and compared with the placebo group. Occasionally, this paradigm has demonstrated continued efficacy for some, but not all, cholinesterase inhibitors (52 ,53 and 54). Here too, methodologic problems limit an unequivocal interpretation of these data.

Taken together, no carefully conducted, adequately powered studies address the question of whether cholinesterase inhibitors, at any course in the illness, delay progression. Until such studies are carried out, only tantalizing pieces of the puzzle are open to interpretation. However, as interesting as this question is to clinical neuroscientists, it may be relatively moot to caregivers who struggle with patients with Alzheimer disease. To such people, the nuances of whether plaque and tangle formation may be slowed, and neurons kept alive, are less relevant than the question whether time to a particular milestone of the disease can be delayed by cholinesterase treatment. In fact, such an outcome can occur even if this class of compounds has a solely palliative effect. It can be argued that simply improving the patients' cognitive capacity increases the likelihood that a patient will be maintained at home. Additionally, the effects of some of these compounds, if not all, on such problematic noncognitive behaviors (55) can also lead to a better outcome. This seems increasingly likely given that the cholinesterase inhibitors appear to have their most robust effect in middle-stage disease, or perhaps even later (56 ,57 and 58). This result is completely compatible with postmortem findings of cholinergic parameters that find the cholinergic deficit to be most apparent in middle- or later-stage disease and to be not present in the earliest stages of illness (59).

VITAMIN E AND ANTIOXIDANTS

Part of "87 - Current and Experimental Therapeutics of Alzheimer Disease "

The production of free radical species has been considered a mediating event for many forms of neuronal death or damage. Initiating events as diverse as glutamate-induced neurotoxicity, ischemia, apoptosis, and Aß neurotoxicity all can produce oxidative stress with free radical production (60). Thus, the use of antioxidants and free radical scavengers in the prevention, or delay in the progression, of Alzheimer disease is not without a reasonable rationale. Vitamin E, in part because of its accessibility, has received greatest attention among compounds in this class. Furthermore, *in vitro* cell studies in various cell culture preparations indicate that vitamin E can have a protective effect on ß-amyloid–induced neurotoxicity (61).

A carefully conducted double-blind, placebo-controlled, multicenter investigation of the effect of vitamin E and selegiline provided some support for the efficacy of both these agents in altering the progression of Alzheimer disease (62). In this trial, patients with moderate to severe Alzheimer disease received either 2,000 IU per day of vitamin E, 10 mg per day of selegiline, or the combination of vitamin E and selegiline. An additional treatment arm exposed patients only to placebo. This was a 2-year trial in which the primary outcome measures were nursing home placements, death, or the loss of a well-defined activity of daily living. Cognitive change was also evaluated. All three antioxidant groups showed a statistically significant beneficial effect on all outcome measures except cognition. Surprisingly, a favorable effect on cognition was not found for any agent. Unfortunately, despite randomization, subjects in the treatment arms significantly differed in baseline Mini-Mental State Examination scores. Consequently, the significant results were only obtained when a covariant technique was used to adjust for the difference in baseline cognition across the treatment arms. This circumstance, combined with the negative effect on cognition, raises questions regarding the robustness of these antioxidant treatments. Nonetheless, the inclusion of 2,000 IU per day of vitamin E in the treatment regimen of patients with Alzheimer disease has become relatively commonplace.

Vitamin E ingestion is not without potential toxicity. Thrombophlebitis has been reported in adults in doses far less than the 2,000 IU recommended for patients with Alzheimer disease (63). Coagulopathy can be another vitamin E–associated adverse event (64 ,65 ,66 and 67). Interactions between vitamin E and oral anticoagulants are a real possibility and emphasize the need for monitoring prothrombin times in patients who receive this combination. In contrast to the widespread use of vitamin E, selegiline has not become a routine part of Alzheimer disease therapy because selegiline was not found to be superior to vitamin E, nor was there any benefit of combining vitamin E and selegiline to either drug alone. The adverse event profile for selegiline is far more extensive than vitamin E, and it includes hypotension with subsequent falls, as well as sleep disturbance, psychosis, agitation, and confusion. The potential for a serious interaction between selegiline and antidepressants commonly used to treat comorbid depression in patients with Alzheimer disease further limits the potential utility of selegiline.

ANTIINFLAMMATORY AGENTS

Part of "87 - Current and Experimental Therapeutics of Alzheimer Disease "

Inflammatory processes have well been characterized in the Alzheimer brain. Elevations in cytokine, acute-phase proteins, complement, and activated microglia are all present in Alzheimer disease brain (38 ,68 ,69 and 70). Of potential significance is that the complement cascade can be activated by Aß, ultimately leading to the induction of the membrane attack complex, which can be neurotoxic (71 ,72 and 73). These postmortem findings are given increased meaning by epidemiologic studies that also impute a role for inflammatory mechanisms in Alzheimer disease. The use of nonsteroidal antiinflammatory drugs (NSAIDs) well before the onset of Alzheimer disease has been associated with a decreased incidence of Alzheimer disease in late life. Studies of siblings with differential exposure to NSAIDs reveal a profound delay in the onset of Alzheimer disease in the sibling with exposure to these agents (74 ,75 ,76 and 77).

Particular interest has centered on the inhibition of cyclooxygenase in Alzheimer disease. Although the inflammatory reaction in the Alzheimer brain appears quite broad, a rationale nonetheless exists for inhibition of cyclooxygenase, especially cyclooxygenase 2 (Cox-2). Cox-2 levels are elevated in hippocampal neurons from postmortem examination of patients with Alzheimer disease (78). Additionally, Cox-2 expression is up-regulated in the frontal cortex of the patient with Alzheimer disease. The severity of Alzheimer disease neuropathology correlates with Cox-2 levels (78) and ß-amyloid increase expression of Cox-2 in neuroblastoma lines.

Given these data, it is not surprising that numerous antiinflammatory agents are being, or have been, tested in patients with Alzheimer disease. With the extensiveness of the inflammatory response in the Alzheimer disease brain, a relatively nonspecific antiinflammatory drug such as prednisone seemed a rational approach to treatment. A large, multicenter, double-blind study in which an initial dose of up to 20 mg of prednisone, followed by a maintenance dose of 10 mg for 1 year, was conducted. No evidence of efficacy in delaying the progression of Alzheimer disease was found. Indeed, patients receiving prednisone were more likely to develop behavioral worsening as well as glucocorticoidrelated medical adverse events. Although it is conceivable that a higher dose of prednisone was necessary, the administration of such a dose would seem impossible, based on the medical problems encountered with relatively modest doses of prednisone (79).

Diclofenac, another antiinflammatory agent, was investigated in a 25-week randomized, double-blind, placebo-controlled trial in patients with mild to moderate Alzheimer disease. The patient withdrawal rate from the study was exceedingly high, and it limited the interpretability of the results. Nevertheless, efficacy of this agent was not found. Conversely, indomethacin administered in a 6-month trial was reported to be efficacious, but here, too, the dropout rate was excessive, compromising both the interpretability of the results as well as the ultimate utility of this drug (80).

The most positive results obtained to date from large-scale studies derive from the clinical trials with propentofylline. This drug is an inhibitor of microglia activation. A series of studies demonstrated improvement in global functioning, cognitive measures, and activities of daily living compared with placebo (81 ,82 and 83). However, the effects were exceedingly modest, and attempts to obtain approval for an Alzheimer disease–related indication in the European community have so far been unsuccessful, because the extent of drug effect has not been deemed to be sufficient to warrant approval.

Numerous trials with selective Cox-2 inhibitors are currently ongoing. These results are eagerly awaited. However, to date, despite the relatively compelling rationale for testing antiinflammatory agents in Alzheimer disease, results have not been encouraging. The apparent contradiction between epidemiologic studies showing benefit from prior exposure to NSAIDs and treatment studies with NSAIDs could reflect the period in which NSAIDs were administrated. Conceivably, such drugs will have no effect, or even an adverse effect, once Alzheimer disease has developed, but they may still be effective in delaying onset by drug administration before patients are symptomatic. Hence, a full test of the antiinflammatory approach in Alzheimer disease will require additional studies.

ESTROGEN

Part of "87 - Current and Experimental Therapeutics of Alzheimer Disease "

As with antiinflammatory agents, the basis for estrogen therapy in Alzheimer disease, in part, derives from epidemiologic studies. One such study, the Baltimore longitudinal study of aging, followed 500 women, of whom half were estrogen users, for approximately 16 years. The relative risk of developing Alzheimer disease in the women who were taking estrogen was approximately halved (84). A similar result was obtained in an Italian longitudinal aging study (85). Other epidemiologic surveys have reached similar conclusions (86). The plausibility of these results are enhanced by the finding that estrogen replacement therapy was associated with higher cognitive test scores in healthy elderly women over the age of 65 years, compared with a cohort not receiving such treatment (87 ,88). There is, however, one large 15-year follow-up study of approximately 800 elderly women in which no relationship between estrogen replacement therapy and a host of neuropsychological test scores was found (89).

That estrogen replacement therapy may have a positive effect on the development of Alzheimer disease, or cognition in general, is supported by a series of studies investigating the actions of estrogen on neuronal tissue. For example, ovariectomized rats treated with estrogen show preservation of the integrity of hippocampal neurons and their dendritic arborization (90). Furthermore, activity of choline acetyltransferase is augmented by estrogen treatment (91 ,92). Estrogen may also have antioxidant activity, may facilitate processing of APP toward a nonamyloidogenic pathway, and may promote cell survival (92 ,93). Hence, some role for estrogen in the therapeutics of Alzheimer disease is a reasonable proposition.

Two studies examined the effect of estrogen on both the course and symptoms of Alzheimer disease. Estrogen replacement therapy for 1 year did not slow disease progression among women with mild to moderate Alzheimer disease who had previously undergone a hysterectomy (94). In another randomized, double-blind, placebo-controlled parallel group study, no effects of estrogen on cognitive symptoms was noted (95). Conversely, some benefit of a transdermal estrogen preparation was noted in an 8-week treatment trial in a very small group of women. Furthermore, positive results were found in a few, but not all,

neuropsychological tests (96). Given the effect of estrogen on cholinergic parameters, of note is a retrospective analysis of patients previously exposed to tacrine in the pivotal trials leading to the approval of that drug. Women taking estrogen replacement therapy had a significantly greater response on all outcome measures than those female patients receiving tacrine who were not receiving estrogen replacement therapy. These data raise the possibility that estrogen replacement therapy may augment the cognitive effects of cholinesterase inhibitors (97).

Selective estrogen-receptor modulators (SERMs) have been designed to have agonistic effects on some organ systems and antagonistic effects on others. Should estrogen replacement therapy have beneficial effects in preventing Alzheimer disease, delaying its progress, treating its symptoms, or augmenting other therapies, a SERM with agonist activity in the brain, but without effect on reproductive organs, would have obvious therapeutic potential, including administration to male patients. Many SERMs are currently being tested in numerous conditions. However, as yet no reports of studies on the role of these agents in any aspects of Alzheimer treatment have been published.

GINKGO BILOBA

Part of "87 - Current and Experimental Therapeutics of Alzheimer Disease "

The broad use of vitamin and herbal preparations, facilitated by their general availability without prescription, encouraged a placebo-controlled, double-blind, randomized trial of an extract of *Ginkgo biloba* for dementia (98). This extract, termed *Egb761*, was tested in a 52-week study of mild to severely demented outpatients with various forms of dementia including Alzheimer disease and multiinfarct dementia. One-third of all patients entered into the study did not provide 52-week endpoint data. A small and statistically significant effect was found on the ADAS-cog, but no effect was found on the Clinical Global Index (CGI). Thus, by a prior standard set by the FDA to establish efficacy of an agent in Alzheimer disease (statistically significant drug effect on both a psychometric and a global measure), Egb761 would not have met this standard for receiving an indication for the treatment of Alzheimer disease. Nonetheless, this compound continues to be widely used, even though it has been reported to cause spontaneous bleeding and it may interact with anticoagulant and antiplatelet aggregating agents (99).

APPROACH TO ALTERING AMYLOID DEPOSITION

Part of "87 - Current and Experimental Therapeutics of Alzheimer Disease "

Increasingly, amyloid deposition is seen as one of the earliest components, if not the earliest, of the pathologic process in Alzheimer disease, as well as an initiating event in neuronal death (100 ,101). Furthermore, the elucidation of the cellular consequences of the various mutations associated with Alzheimer disease supports the notion of the centrality of amyloid production in the pathophysiology of Alzheimer disease. Specifically, regardless of whether a mutation occurs in the amyloid precursor protein gene, presenilin 1 or presenilin 2, all mutations increase the concentrations of Aß1-42 in brain, plasma, or cell culture media. A similar outcome is associated with the apolipoprotein E-4 allele compared to E-2 or E-3 (101 ,102 and 103). The well-documented toxicity of Aß, particularly in the aggregated form, adds to the growing consensus that altering Aß production or deposition is a viable approach to the therapeutics of Alzheimer disease.

There are numerous theoretic approaches to altering the Aß concentrations in the brain of patients with Alzheimer disease. The activities of both ß- and γ-secretase are necessary to cleave APP into the Aß fragments that constitute amyloid plaques. Conceivably, inhibiting either γ- or ß-secretase could alter the production of Aß. Alternatively, enhancing the activity of α-secretase could result in the preferential cleavage of APP to nonamyloidogenic forms. Yet another approach focuses on enhancing the breakdown or clearance of Aß in the brain. This approach adopts the view that inflammatory mechanisms in the Alzheimer brain are potentially beneficial and facilitate the removal of Aß from the brain. Finally, the enhanced toxicity of aggregated Aß encourages therapeutics designed to block the aggregation of Aß. All these approaches are in various stages of clinical development.

Numerous groups have cloned and characterized ß-secretase, also termed ß-amyloid cleavage enzymes (BACE) (104 ,105 ,106 ,107 ,108 and 109). The success of this effort encourages combinatorial chemistry and screening efforts designed to identify small lipophilic compounds that could inhibit BACE activity and thereby limit Aß production. The logic of this approach is unquestioned, but the presence of relatively high levels of BACE in the pancreas leads to the question of the role that BACE may play in biological functions whose activity, if inhibited, could cause significant adverse events. Conversely, to produce meaningful changes in the course of Alzheimer disease, or simply to delay the disease onset, safe levels of brain BACE inhibition may readily exist.

Although γ-secretase has not yet been cloned, a γ-secretase inhibitor is currently in clinical trial (104). However, the intimate relationship between presenilins and γ-secretase could have implications for the ultimate safety of this approach. If, in fact, presenilins influence the critically involved Notch pathway (37), a host of potential adverse effects could arise from inhibiting the activity of presenilins. Still, elucidation of the clinical effects of the γ-secretase inhibitors will be eagerly awaited.

Transgenic mice overexpressing Aß have been used as a vehicle to determine whether inoculation with the Aß peptide could produce an immune response that would alter Aß concentrations in a mouse brain (110). Animals inoculated before the deposition of substantial amyloid deposits in the

brain subsequently displayed little amyloid deposition. Even more remarkably, animals in which amyloid deposition had already begun demonstrated an apparent diminution in amyloid plaque load following inoculation. Behavioral data now confirm that Aß peptide immunization reduces cognitive impairment and plaques in animal models of Alzheimer disease (111 ,112 and 113). Vaccination with Aß protects transgenic mice from the learning and age-related memory deficits that normally occur in the mouse model of Alzheimer disease. During testing for potential deleterious effects of the vaccine, all mice performed superbly on the radial-arm water maze test of working memory. Later, at an age when untreated transgenic mice show memory deficits, the Aß-vaccinated transgenic mice showed cognitive performance superior to that of the control transgenic mice and, ultimately, performed as well as nontransgenic mice (111).

Based on these exciting results, Aß inoculations are beginning in humans. A major question that these trials will eventually answer is whether elderly persons can generate an adequately robust immune response to Aß inoculation that will extend to the brain. Additionally, the concern that the adjunctive procedures necessary to generate an immune response to a peptide that already exists in healthy humans would also produce an autoimmune response must be considered. Still, results obtained in transgenic mice are so dramatic that it is essential that Aß inoculations proceed at least preliminarily in humans.

There are numerous theoretic possibilities to altering the aggregation of Aß fibrils into their more toxic aggregated form. Congo red, a dye that readily binds Aß, has been used as a prototypical molecule for the development of analogues that would enter the brain, bind to Aß, and inhibit aggregation (102 ,103). Other approaches have included the development of antibodies directed specifically at Aß or small molecule ligands (114) that also can block aggregation (115 ,116). Acetylcholinesterase has been found to augment Aß aggregation, and antibodies to acetylcholinesterase can, *in vitro*, decrease aggregation (50). Indeed the Alzheimer plaque contains numerous proteins, many of which may facilitate aggregation of Aß. Developing compounds that preferentially bind these plaque-containing molecules could decrease Aß aggregation.

SUMMARY

Part of "87 - Current and Experimental Therapeutics of Alzheimer Disease "

Since the early 1990s, remarkable progress has been made in the current and experimental therapeutics of Alzheimer disease. An area that was recently characterized by therapeutic nihilism can now be regarded with real optimism. It would seem highly likely that the next decade of progress should show the development of compounds that move beyond palliation and could actually either delay onset or substantially alter the course of the illness in such a manner as to bring new hope to the patient with Alzheimer disease.

REFERENCES

1. National Institute on Aging. *Progress report on Alzheimer's disease 1999*. Washington, DC: National Institute on Aging, National Institutes of Health, 1999.

2. Davis KL, Thal LJ, Gamzu ER, et al. A double-blind, placebo-controlled multicenter study of tacrine for Alzheimer's disease: the Tacrine Collaborative Study Group. *N Engl J Med* 1992;29:327:1253–1259.

3. Farlow M, Gracon SI, Hershey LA, et al. A controlled trial of tacrine in Alzheimer's disease: the Tacrine Study Group. *JAMA* 1992;268:2523–2529.

4. Knapp MJ, Knopman DS, Solomon PR, et al. A 30-week randomized controlled trial of high-dose tacrine in patients with Alzheimer's disease: the Tacrine Study Group. *JAMA* 1994;271:985–991.

5. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry* 1984;141:1356–1364.

6. Knopman DS, Knapp MJ, Gracon SI, et al. The Clinician Interview Based Impression (CIBI): a clinician's global change rating scale in Alzheimer's disease. *Neurology* 1994;44:2315–2321.

7. Schneider LS, Olin JT, Doody RS, et al. Validity and reliability of the Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change. *Alzheimer Dis Assoc Disord* 1997;11:S22–S32.

8. Reisberg B, Schneider L, Doody R, et al. Clinical global measure of dementia: position paper from the International Working Group on Harmonization of Dementia Drug Guidelines. *Alzheimer Dis Assoc Disord* 1997;11:8–18.

9. Schneider LS, Olin JT. Clinical global impressions in Alzheimer's clinical trials. *Int Psychogeriatr* 1996;8:277–288.

10. Terrel PS, Bredenkamp VC, Lammers JE. Late-onset alanine aminotransferase increase with tacrine. *Ann Pharmacother* 1996;30:301.

11. Blackard WG Jr, Sood GK, Crowe DR, et al. Tacrine: a cause of fatal hepatotoxicity? *J Clin Gastroenterol* 1998;26:57–59.

12. Ohnishi A, Mihara M, Kamakura H. Comparison of the pharmacokinetics of E2020, a new compound for Alzheimer's disease, in healthy young and elderly subjects. *J Clin Pharmacol* 1993;33:1086–1089.

13. Rogers SL, Farlow MR, Doody RS. A 24-week, double-blind, placebo controlled trial of donepezil in patients with Alzheimer's disease. *Neurology* 1998;50:136–145.

14. Rogers SL, Friedhoff LT. The efficacy and safety of donepezil in patients with Alzheimer's disease: results of a US multicenter, randomized, double-blind, placebo-controlled trial. The Donepezil Study Group. *Dementia* 1996;7:293–303.

15. Rogers SL, Doody RS, Mohs RC, et al. Donepezil improves cognition and global function in Alzheimer disease: a 15-week, double-blind, placebo-controlled study. Donepezil Study Group. *Arch Intern Med* 1998;158:1021–1031.

16. Rogers SL, Friedhoff LT. Long-term efficacy and safety of donepezil in the treatment of Alzheimer's disease: an interim analysis of the results of a US multicenter open label extension study. *Eur Neuropsychopharmacol* 1998;8:67–75.

17. Burns A, Rossor M, Hecker J, et al. The effects of donepezil in Alzheimer's disease: results from a multinational trial. *Dement Geriatr Cogn Disord* 1999;10:237–244.

18. Liu C, Gillette MU. Cholinergic regulation of the suprachiasmatic nucleus circadian rhythm via a muscarinic mechanism at night. *J Neurosci* 1996;16:744–751.

19. O'Hara BF, Edgar DM, Cao VH, et al. Nicotine and nicotinic receptors in the circadian system. *Psychoneuroendocrinology* 1998;23:161–173.

20. Riemann D, Lis S, Fritsch-Montero R, et al. Effect of tetrahydroaminoacridine on sleep in healthy subjects. *Biol Psychiatry* 1996;39:796–802.

21. Holsboer-Trachsler E, Hatzinger M, Stohler R, et al. Effects of the novel acetylcholinesterase inhibitor SDZ ENA 713 on sleep in man. *Neuropsychopharmacology* 1993;8:87–92.

22. Riemann D, Gann H, Dressing H, et al. Influence of the cholinesterase inhibitor galantamine hydrobromide on normal sleep. *Psychiatry Res* 1994;51:253–267.

23. Rosler M, Anand R, Cicin-Sain A. Efficacy and safety of rivastigmine in patients with Alzheimer's disease: international randomized controlled trial. *BMJ* 1999;318:633–640.

24. Maelicke A, Coban T, Storch A, et al. Allosteric modulation of Torpedo nicotinic acetylcholine receptor ion channel activity by noncompetitive agonists. *J Recept Signal Transduct Res* 1997;17:11–28.

25. Mihailova D, Yamboliev I. Pharmacokinetics of galanthamine hydrobromide (Nivalin) following single intravenous and oral administration in rats. *Pharmacology* 1986;32:301–306.

26. Bickel U, Thomsen T, Weber W, et al. Pharmacokinetics of galantamine in humans and corresponding cholinesterase inhibition. *Clin Pharmacol Ther* 1991;50:420–428.

27. Kewitz H, Davis BM, Katz R. Safe and efficient inhibition of acetylcholinesterase in the brain for the treatment of senile dementia of Alzheimer's type: galanthamine versus tacrine. In: *Abstract at First European Congress of Pharmacology*. Milan:1995.

28. Tariot PN, Solomon PR, Morris JC, et al. The Galantamine USA-Study Group: a 5-month randomized placebo-controlled trial of galantamine in AD. *Neurology* 2000;54:2269–2276.

29. Raskind MA, Peskin ER, Wessel T, et al. The Galantamine USA-Study Group: galantamine in AD. A 6-month randomized, placebo controlled trial with a 6-month extension. *Neurology* 2000;54:2261–2268.

30. Wilcock GK, Lilienfeld S, Gaens E. Efficacy and safety of galantamine in patients with mild to moderate Alzheimer's disease: multicentre randomized trial. *BMJ* 2000;321:1445–1449.

31. Van Dyke CH, Newhouse P, Falk WE, et al. Extended-release physostigmine in Alzheimer's disease: a multicenter, double-blind, 12-week study with dose enrichment. *Arch Gen Psychiatry* 2000;57:157–164.

32. Cummings J. Metrifonate: overview of safety and efficacy. *Pharmacotherapy* 1998;18:43–46.

33. Becker RE, Colliver JA, Markwell SJ. Double-blind placebo-control study of metrifonate, an acetylcholinesterase inhibitor of Alzheimer's disease. *Alzheimer Dis Assoc Disord* 1996;10:124–131.

34. Cummings JL, Cyrus PA, Bieber F. Metrifonate treatment of the cognitive deficits of Alzheimer's disease: Metrifonate Study Group. *Neurology* 1998;50:1214–1221.

35. Richardson RJ. Assessment of the neurotoxic potential of chlorpyrifos relative to other organophosphorous compounds: a critical review of the literature. *J Toxicol Environ Health* 1995;44:135–165.

36. Lahiri DK, Farlow MR, Sambamurti K. The secretion of amyloid beta-peptides is inhibited in the tacrine-treated human neuroblastoma cells. *Mol Brain Res* 1998;62:131–140.

37. Buxbaum JD, Oishi M, Chen HI, et al. Cholinergic agonists and interleukin 1 regulate processing and secretion of the Alzheimer beta/A4 amyloid protein precursor. *Proc Natl Acad Sci USA* 1992;89:10075–10078.

38. Haroutunian V, Greig N, Pei XF, et al. Pharmacological modulation of Alzheimer's beta-amyloid precursor protein levels in the CSF of rats with forebrain cholinergic system lesions. *Mol Brain Res* 1997;46:161–168.

39. Svensson AL, Nordberg A. Tacrine and donepezil attenuate the neurotoxic effect of A-beta (25–35) in rat PC12 cells. *Neuroreport* 1998;9:1519–1522.

40. Xiao XQ, Lee NT, Carlier PR, et al. Bis(7)-tacrine, a promising anti-Alzheimer's agent, reduces hydrogen peroxide-induced injury in rat pheochromocytoma cells: comparison with tacrine. *Neurosci Lett* 2000;290:197–200.

41. Lee PN. Smoking and Alzheimer's disease: a review of the epidemiological evidence. *Neuroepidemiology* 1994;13:131–144.

42. Doll R, Peto R, Boreham J, Sutherland I. Smoking and dementia in male British doctors: a prospective study. *BMJ* 2000;320:1097–1102.

43. Launer LJ, Anderson K, Dewey ME, et al. Rates and risk factors for dementia and Alzheimer's disease: results from EURODEM pooled analyses: EURODEM Incidence Research Group and Work Groups. *Neurology* 1999;52:78–84.

44. Kihara T, Shimohama S, Urushitani M, et al. Stimulation of alpha4beta2 nicotinic acetylcholine receptors inhibits beta-amyloid toxicity. *Brain Res* 1998;792:331–334.

45. Kihara T, Shimohama S, Sawada H, et al. Nicotinic receptor stimulation protects neurons against beta-amyloid toxicity. *Ann Neurol* 1997;42:159–163.

46. Kim SH, Kim YK, Jeong SJ, et al. Enhanced release of secreted form of Alzheimer's amyloid precursor protein from PC12 cells by nicotine.*Mol Pharmacol* 1997;52:430–436.

47. Knopman D, Schneider L, Davis K, et al. Long-term tacrine (Cognex) treatment: effects on nursing home placement and mortality, Tacrine Study Group. *Neurology* 1996;47:166–177.

48. Alvarez A, Alarcon R, Opazo C, et al. Stable complexes involving acetylcholinesterase and amyloid-ß peptide change the biochemical properties of the enzyme and increase the neurotoxicity of Alzheimer's fibrils. *J Neurosci* 1998;18:3213–3223.

49. Campos EO, Alvarez A, Inestroza NC. Brain acetylcholinesterase promotes amyloid-ß-peptide aggregation but does not hydrolyze amyloid precursor protein peptides. *Neurochem Res* 1998;23:135–140.

50. Inestrosa NC, Alarcon R. Molecular interactions of acetylcholinesterase with senile plaques. *J Physiol (Lond)* 1998;92:341–344.

51. Alvarez A, Opazo C, Alarcon R, et al. Acetylcholinesterase promotes the aggregation of amyloid-ß-peptide fragments by forming a complex with the growing fibrils. *J Mol Biol* 1997;272:348–361.

52. Reyes AE, Perez DR, Alvarez A, et al. A monoclonal antibody against acetylcholinesterase inhibits the formation of amyloid fibrils induced by the enzyme. *Biochem Biophys Res Commun* 1997;232:652–655.

53. Doody RS. Clinical profile of donepezil in the treatment of Alzheimer's disease. *Gerontology* 1999;49:23–32.

54. Forette F, Anand R, Gharabawi M. A phase II study in patients with Alzheimer's disease to assess the preliminary efficacy and maximum tolerated dose of rivastigmine (Exelon). *Eur J Neurol* 1999;6:423–429.

55. Cummings JL. Cholinesterase inhibitors: a new class of psychotropic compounds. *Am J Psychiatry* 2000;157:4–15.

56. Farlow MR, Brashean A, Huis, et al. The effects of tacrine in patients with mild versus moderate stage Alzheimer's disease. In Iqbal K, Marthmer J, Winblad B, et al., eds. *Research advances in Alzheimer's disease and related disorders*. Chichester, UK: Wiley, 1995:284–292.

57. Imbimbo BP, Lucca U, Lucchelli F, et al. A 25-week, placebo-controlled study of eptastigmine in patients with Alzheimer's disease. *Alzheimer Dis Assoc Disord* 1998;12:313–332.

58. Farlow MR, Evans RM. Pharmacological treatment of cognition in Alzheimer's dementia. *Neurology* 1998;51:S36–S44; discussion S65–S67.

59. Davis KL, Mohs RC, Marin D, et al. Cholinergic markers in elderly patients with early signs of Alzheimer's disease. *JAMA* 1999;281:1401–1406.

60. Pike CJ, Cotman CW. Beta-amyloid neurotoxicity *in vitro:* examination of potential contributions from oxidative pathways [Abstract 423]. In: *Osaka Fifth International Conference on Alzheimer's Disease*. Osaka, Japan:1996.

61. Behl C, Davis J, Cole GM, et al. Vitamin E protects nerve cells from amyloid-ß protein toxicity. *Biochem Biophys Res Commun* 1992;186:944–950.

62. Sano M, Ernesto C, Thomas RG, et al. A controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer's disease. *N Engl J Med* 1997;336:1216–1222.

63. Roberts HJ. Vitamin E and thrombophlebitis. *Lancet* 1978;1:49.

64. Hansten PD, Horn JR. *Drug interactions*. Philadelphia: Lea & Febiger, 1989.

65. Corrigan JJ, Ulfers LL. Effect of vitamin E on prothrombin levels in warfarin-induced vitamin-K deficiency. *Am J Clin Nutr* 1981;34:1701–1705.

66. Corrigan JJ, Marcus FI. Coagulopathy associated with vitamin E ingestion. *JAMA* 1974;230:1300–1301.

67. Heinonen OP, Huttunen JK, Albanes D, et al. Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group: the effects of vitamin E and beta-carotene on the incidence of lung cancer and other cancers in male smokers. *N Engl J Med* 1994;330:1029–1035.

68. Aisen PS, Davis KL. Inflammatory mechanisms in Alzheimer's disease: implications for therapy. *Am J Psychiatry* 1994;151:1105–1113.

69. Campbell IL, Abraham CR, Masliah E, et al. Neurologic disease induced in transgenic mice by cerebral overexpression of interleukin 6. *Proc Natl Acad Sci USA* 1993;90:10061–10065.

70. Bauer J, Strauss S, Schreiter-Gasser U. Interleukin-6 and alpha-2-macroglobulin indicate an acute phase response in Alzheimer's disease cortices. *FEBS Lett* 1991;285:111–114.

71. Eikelenboom P, Hack CE, Rozemuller JM, et al. Complement activation in amyloid plaques in Alzheimer's dementia. *Virchows Arch B Cell Pathol Mol Pathol* 1989;56:259–262.

72. Bergamaschini L, Canziani S, Bottasso B, et al. Alzheimer's beta-amyloid peptides can activate the early components of complement classical pathway in a C1q-independent manner. *Clin Exp Immunol* 1999;115:526–533.

73. Webster S, O'Barr S, Rogers J. Enhanced aggregation and beta structure of amyloid beta peptide after coincubation with C1q. *J Neurosci Res* 1994;39:448–456.

74. McGeer PL, Schulzer M, McGeer EG. Arthritis and anti-inflammatory agents as possible protective factors for Alzheimer's disease: a review of 17 epidemiological studies. *Neurology* 1996;47:425–432.

75. Breitner JC, Welsh KA, Helms MJ, et al. Delayed onset of Alzheimer's disease with nonsteroidal anti-inflammatory and histamine H2 blocking drugs. *Neurobiol Aging* 1995;16:523–530.

76. Aisen PS, Luddy A, Durner M, et al. HLA-DR4 influences glial activity in Alzheimer's disease hippocampus. *J Neurol Sci* 1998;161:66–69.

77. Curran M, Middleton D, Edwardson J, et al. HLA-DR antigens associated with major genetic risk for late-onset Alzheimer's disease. *Neuroreport* 1997;8:1467–1469.

78. Ho L, Pieroni C, Winger D, et al. Regional distribution of cyclooxygenase-2 in the hippocampal formation in Alzheimer's disease. *J Neurosci Res* 1999;57:295–303.

79. Aisen PS, Davis KL, Berg J, et al. A randomized controlled trial of prednisone in Alzheimer's disease. *Neurology* 2000;54:588–593.

80. Rogers J, Kirby LC, Hempelman SR, et al. Clinical trial of indomethacin in Alzheimer's disease. *Neurology* 1993;43:1609–1611.

81. Marcussen J, Rother M, Kittner B, et al. A 12–month, randomized, placebo-controlled trial of propentofylline (HWA 285) in patients with dementia according to DSM III-R: the European Propentofylline Study Group. *Dement Geriatr Cogn Disord* 1997;8:320–328.

82. Kittner B, Rossner M, Rother M. Clinical trials in dementia with propentofylline. *Ann NY Acad Sci* 1997;826:307–316.

83. Rother M, Erkinjuntti T, Rossner M, et al. Propentofylline in the treatment of Alzheimer's disease and vascular dementia: a review of phase III trials. *Dement Geriatr Cogn Disord* 1998;9:36–43.

84. Kawas C, Resnick S, Morrison A, et al. A prospective study of estrogen replacement therapy and the risk of developing Alzheimer's disease: the Baltimore Longitudinal Study of Aging. *Neurology* 1997;48:1517–1521.

85. Baldereschi M, Di Carlo A, Lepore V, et al. Estrogen-replacement therapy and Alzheimer's disease in the Italian Longitudinal Study on Aging. *Neurology* 1998;50:996–1002.

86. Tang MS, Jacobs D, Stern Y, et al. Effect of estrogen during menopause on risk and age at onset of Alzheimer's disease. *Lancet* 1996;348:429–432.

87. Teng EL, Chui HC. The modified Mini-Mental State (3MS) examination. *J Clin Psychiatry* 1987;48:314–318.

88. Steffens DC, Norton MC, Plassman BL, et al. Enhanced cognitive performance with estrogen use in nondemented community-dwelling older women. *J Am Geriatr Soc* 1999;47:1171–1175.

89. Barrett-Connor E, Kritz-Silverstein D. Estrogen replacement therapy and cognitive function in older women. *JAMA* 1993;269:2637–2641.

90. Gibbs RB, Aggarwal P. Estrogen and basal forebrain cholinergic neurons: implications for brain aging and Alzheimer's disease-related cognitive decline. *Horm Behav* 1998;34:98–111.

91. Luine V, Park, Joh T, Reis D, et al. Immunochemical demonstration of increased choline acetyltransferase concentration in rat preoptic area after estradiol administration. *Brain Res* 1980;191:273–277.

92. Luine VN. Estradiol increases choline acetyltransferase in specific basal forebrain nuclei and projection areas of female rats. *Exp Neurol* 1985;89:484–490.

93. Nathan L, Chaudhuri G. Antioxidant and prooxidant actions of estrogens: potential physiological and clinical implications. *Semin Reprod Endocrinol* 1998;16:309–314.

94. Mulnard RA, Cotman CW, Kawas C, et al. Estrogen replacement therapy for treatment of mild to moderate Alzheimer disease: a randomized controlled trial. Alzheimer's Disease Cooperative Study. *JAMA* 2000;283:1007–1015.

95. Henderson VW, Paganini-Hill A, Miller BL, et al. Estrogen for Alzheimer's disease in women: randomized, double-blind, placebo-controlled trial. *Neurology* 2000;54:295–301.

96. Asthana S, Craft S, Baker LD, et al. Cognitive and neuroendocrine response to transdermal estrogen in postmenopausal women with Alzheimer's disease: results of a placebo-controlled, double-blind, pilot study. *Psychoneuroendocrinology* 1999;24:657–677.

97. Schneider LS, Farlow MR, Henderson VW, et al. Effects of estrogen replacement therapy on response to tacrine in patients with Alzheimer's disease. *Neurology* 1996;46:1580–1584.

98. Le Bars PL, Katz MM, Berman N, et al. A placebo-controlled, double-blind, randomized trial of an extract of *Ginkgo biloba* for dementia: North American Egb Study Group. *JAMA* 1997;278:1327–1332.

99. Cupp MJ. Herbal remedies: adverse effects and drug interactions. *Am Fam Physician* 1999;59:1239–1245.

100. Naslund J, Haroutunian V, Mohs R, et al. Correlation between elevated levels of amyloid beta-peptide in the brain and cognitive decline. *JAMA* 2000;283:1571–1577.

101. Selkoe DJ. Translating cell biology into therapeutic advances in Alzheimer's disease. *Nature* 1999;399:A23–A31.

102. Gupta-Bansal R, Brunden KR. Congo red inhibits proteoglycan and serum amyloid P binding to amyloid beta fibrils. *J Neurochem* 1998;70:292–298.

103. Podlinsky MB, Walsh DM, Amarante P, et al. Oligomerization of endogenous and synthetic amyloid beta-protein at nanomolar levels in cell culture and stabilization of monomer by Congo red. *Biochemistry* 1998;37:3602–3611.

104. Vassar R, Bennett BD, Babu-Khan S, et al. ß-Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* 1999;286:735–741.

105. Howlett DR, Simmons DL, Dingwall C, et al. In search of an enzyme: the beta-secretase of Alzheimer's disease is an aspartic proteinase. *Trends Neurosci* 2000;23:565–570.

106. Sauder JM, Arthur JW, Dunbrack RL. Modeling of substrate specificity of the Alzheimer's disease amyloid precursor protein beta-secretase. *J Mol Biol* 2000;300:241–248.

107. Sinha S, Anderson JP, Barbour R, et al. Purification and cloning of amyloid precursor protein beta-secretase from human brain. *Nature* 1999;402:537–540.

108. Hussain I, Powell DJ, Howlett DR, et al. ASP1(BACE2) cleaves the amyloid precursor protein at the beta-secretase site. *Mol Cell Neurosci* 2000;16:609–619.

109. Solans A, Estivill X, de La Luna S. A new aspartyl protease on 21q22.3, BACE2, is highly similar to Alzheimer's amyloid precursor protein beta-secretase. *Cytogenet Cell Genet* 2000;89:177–184.

110. Schenk D, Barbour R, Dunn W, et al. Immunization with amyloid-ß attenuates Alzheimer-disease–like pathology in the PDAPP mouse. *Nature* 1999;400:173–177.

111. Morgan D, Diamond DM, Gottschall PE, et al. Abeta peptide vaccination prevents memory loss in an animal model of Alzheimer's disease. *Nature* 2000;408:982–985.

112. Chen G, Chen KS, Knox J, et al. A learning deficit related to age and beta-amyloid plaques in a mouse model of Alzheimer's disease. *Nature* 2000;408:975–978.

113. Janus C, Pearson J, McLaurin J, et al. Abeta peptide immunization reduces behavioral impairment and plaques in a model of Alzheimer's disease. *Nature* 2000;408:979–982.

114. Kuner P, Bohrmann B, Tjernberg LO, et al. Controlling polymerization of beta-amyloid and prion-derived peptides with synthetic small molecule ligands. *J Biol Chem* 2000;275:1673–1678.

115. Hatters DM, MacPhee CE, Laurence CJ, et al. Human apolipoprotein C-II forms twisted amyloid ribbons and closed loops. *Biochemistry* 2000;39:8276–8283.

116. MacPhee CE, Hatters DM, Sawyer WH, et al. Apolipoprotein C-II39–62 activates lipoprotein lipase by direct lipid-independent binding. *Biochemistry* 2000;39:3433–3440.

88

Alzheimer Disease: Treatment of Noncognitive Behavioral Abnormalities

Murray A. Raskind

Robert F. Barnes

Murray A. Raskind: Veterans Administration Northwest Network Mental Illness Research Education and Clinical Center, Seattle, Washington.

Robert F. Barnes: Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington.

The original patient described by Dr. Alois Alzheimer in 1907 (1) was remarkable for both her progressive cognitive impairment and her prominent noncognitive behavioral abnormalities. Clinical interest in the noncognitive abnormalities in Alzheimer disease (AD) has been substantial because of their high prevalence (2 ,3 and 4) and because noncognitive behavioral problems complicate patient management and often precipitate institutionalization (3 ,5 ,6 ,7 ,8 ,9 ,10 ,11 and 12). The real or apparent resemblance of delusions, hallucinations, depressed mood, agitation, hostility, and other noncognitive behavioral abnormalities of AD to the signs and symptoms expressed in such classic psychiatric disorders as depression, schizophrenia, and mania has prompted the widespread use of psychotropic drugs in the management of AD (13 ,14 ,15 and 16). However, widespread use does not imply established efficacy. In fact, data establishing the efficacy of psychotropic drugs for noncognitive behavioral problems in AD and other dementing disorders remain limited. Although reports of treatment outcome studies incorporating reliable and valid outcome measures, well-defined patient samples, and randomization to an adequate trial of active medication or placebo continue to appear, their number remains small. This chapter reviews informative studies of the psychopharmacologic management of noncognitive behavioral problems in AD. These include depression (or depressive signs and symptoms), psychotic symptoms (delusions and hallucinations), and disruptive agitated behaviors (e.g., physical and verbal aggression, motoric hyperactivity, and uncooperativeness with activities necessary for personal hygiene and safety). Placebo-controlled studies are emphasized, but results of other studies and reports are discussed when they suggest directions for future investigation or are the only studies available.

DEPRESSION IN ALZHEIMER DISEASE

Part of "88 - Alzheimer Disease: Treatment of Noncognitive Behavioral Abnormalities "

Diagnostic Challenges

The diagnosis and treatment of depression complicating the course of AD have received considerable attention. Because depression *per se* can impair cognitive function (17), it is reasonable to hypothesize that effective treatment of depression in the patient with AD may maximize potential cognitive capacity. Furthermore, the consensus is that reduction of depressive signs and symptoms improves quality of life (18). Unfortunately, the apparently straightforward goal of treating depression complicating AD becomes complex when the problems involved in the diagnosis of depression in the context of AD or other dementing illnesses are considered. A fundamental problem is the substantial overlap of signs and symptoms between depression and AD. Common to both disorders are apathy and loss of interest, impaired ability to think and concentrate, psychomotor changes (both retardation and agitation), and sleep disturbance. The ability to diagnose depression in AD is further compromised by the patient's lack of insight and poor recollection of symptoms.

Prevalence

Even if investigators agreed on uniform diagnostic criteria for syndromal depression in AD and used uniform assessment instruments and interviews, discrepant prevalence rates would likely arise from the differential characteristics of the samples of AD patients studied. Prevalence rates of AD with concurrent depression derived from clinical populations are higher than those derived from research registries that select "pure" AD subjects without a history of major

depressive disorder. For example, in an outpatient geriatric clinic, Reifler et al. (19) found that 20% of patients with AD met DSM-III criteria for major depressive episode. In contrast, Burke et al. (20) found no incident case of major depression in an AD research registry population followed longitudinally through the course of illness. The latter study included a longitudinally followed normal control group matched for age, sex, race, and social position. Signs and symptoms of depression occurred in both patients with AD and controls, but criterion-based major depression could not be diagnosed. A similarly low prevalence of major depression in a sample of subjects with AD screened to exclude those with a past history of major psychiatric disorders was reported by Kumar et al. (5). Although depressed mood was more frequent in the AD subjects than in age-matched normal controls, depressed mood in the AD subjects was unaccompanied by classic vegetative signs and symptoms of depression. These investigators, therefore, interpreted depressed mood as reflecting “demoralization” rather than major depressive disorder. Given these problems, it is not surprising that estimates of the prevalence of depression in AD are widely disparate. Perhaps the true prevalence of concurrent depression in AD lies somewhere between these disparate estimates. As early as 1955, Sir Martin Roth (21) addressed the issue of differentiating the common “affective coloring” seen in dementia patients from the relatively uncommon “sustained depressive symptom complex.” He found that the latter syndrome, which can probably be equated with DSM-IV major depressive episode, occurred in only 3% of patients with dementia.

Psychopharmacologic Approaches

It continues to be disappointing that the extensive interest in defining the prevalence of depression complicating AD has generated few interpretable studies evaluating the outcome of antidepressant treatment in such patients. The database consists primarily of anecdotal case reports and non–placebo-controlled outcome studies. These reports suggest that depression complicating AD may respond to tricyclic antidepressants (22 ,23), monoamine oxidase inhibitor (MAOI) antidepressants (24), or selective serotonin reuptake inhibitors (SSRI) (25). In an open trial of nortriptyline, given in doses sufficient to achieve therapeutic plasma concentrations, or electroconvulsive therapy in eight inpatients with AD complicated by depression, Reynolds et al. (26) reported a significant reduction in mean Hamilton Depression Scale (HAM-D) scores (27) from 17 before treatment to 9 after treatment. Although the reduction in depressive signs and symptoms was substantial in AD patients with concurrent depression, it was less robust than in a similarly treated group of elderly nondemented depressed patients. Another open trial of “naturalistic” somatic antidepressant treatment of inpatients with dementia and concurrent depression was reported by Greenwald et al. (28). This study carefully documented the presence of major depressive episode in six patients with AD and four with multiinfarct dementia (MID). Patients were treated for a mean duration of 11 weeks with a variety of conventional somatic antidepressants (doses not reported), electroconvulsive therapy, or both. HAM-D scores significantly and substantially decreased from a mean of 19 on admission to a mean of 5 at discharge. This degree of improvement did not differ significantly from that in an elderly, nondemented, depressed inpatient group treated in a similar naturalistic manner. However, the mean length of stay to achieve comparable improvement in the elderly nondemented, depressed group was substantially shorter than that in the demented, depressed group. Possible differential treatment responses between AD subjects with major depression and MID subjects with major depression were not reported. Both Reynolds et al. (26) and Greenwald et al. (28) interpreted their results as suggesting that major depression complicating dementia is responsive to somatic antidepressant treatments, but both investigators acknowledged that standardized, double-blinded, placebo-controlled studies of antidepressants in dementia patients with major depression are needed.

The SSRI antidepressants are theoretically attractive drugs for the treatment of depression in AD. Decreased numbers of serotoninergic neurons in the dorsal raphe nucleus and decreased concentrations of the serotonin metabolite 5-hydroxyindolacetic acid in forebrain and cerebrospinal fluid are consistent with a serotoninergic deficit in AD (29 ,30). Enhancing serotoninergic neurotransmission by inhibiting serotonin reuptake theoretically might alleviate depression in AD. Furthermore, the adverse effect profile of SSRIs is relatively benign compared with those of the tricyclics and MAOIs. SSRIs are not anticholinergic, nor do they produce orthostatic hypotension. These theoretic advantages likely account for the increasing use of SSRIs in elderly patients (16 ,31 ,32).

Placebo-Controlled Outcome Trials

Despite the widespread use of antidepressants in patients with AD and other dementing disorders (32), it is clear that this use is not grounded in an adequate empiric database. Only three placebo-controlled trials of an antidepressant in AD patients with depression have been published thus far, and only one of these evaluated an antidepressant from the increasingly prescribed class of SSRIs.

Reifler et al. (33) randomly assigned to either imipramine (*n* = 13) or placebo (*n* = 12) subjects who met DSM-III criteria for both primary degenerative dementia of the Alzheimer type and major depressive episode. AD outpatients (mean age, 72) had a mean Mini-Mental State Examination (MMSE) score of 17 [very comparable with the mean MMSE scores of the demented, depressed patients studied openly by Greenwald et al. (28) and Reynolds et al. (26)]

and were suffering from a similarly moderate degree of depression (mean HAM-D score, 19). Imipramine (mean dose, 83 mg/d; mean plasma level of imipramine plus desmethylimipramine, 116 ng/mL) or placebo was prescribed for 8 weeks. Substantial, highly significant, and almost identical improvements occurred in mean HAM-D scores in both the imipramine subjects (19.3 before treatment vs. 11.5 following treatment) and placebo subjects (18.6 before treatment vs. 10.8 following treatment). Imipramine was generally well tolerated in these subjects. However, its central anticholinergic effect likely accounted for subtle decrements in cognitive function in the imipramine group. The improvement in HAM-D scores was similar to that achieved in the open inpatient studies reported by Greenwald et al. (28) and Reynolds et al. (26). This outpatient study demonstrated that depressive signs and symptoms can be reduced in outpatients with AD, but the mechanism of treatment efficacy does not appear to be a specific antidepressant effect of imipramine.

Petracca et al. (34) randomly assigned 24 patients meeting criteria for AD with DSM-III-R depression to treatment with either clomipramine or placebo in a randomized, double-blinded, placebo-controlled crossover study. As in the imipramine trial reported by Reifler et al. (33), both active drug (in this case clomipramine) and placebo were associated with significant improvement in depression ratings. However, in this study, the improvement in depression in the active drug group was modestly greater than that in the placebo group. One subject in the clomipramine group dropped out of the study because of an acute confusional episode likely attributable to the central anticholinergic activity of clomipramine.

Nyth et al. (35) have reported the only placebo-controlled trial of an SSRI in patients with dementia and concomitant depression. In this study, subjects were randomized to 20 mg of citalopram per day or placebo for 6 weeks. Twenty-three patients with dementia (presumably including a substantial number of patients with AD) were “nested” in a larger controlled trial of depressed elderly persons (*n* = 149), the majority of whom were not suffering from dementia. The inclusion criterion for “depression” was an HAM-D score of at least 14. In the 23 subjects with dementia who completed the 6-month trial, modest but significant differences were noted favoring citalopram over placebo in the observer-rated fear/panic, depressed mood, and impaired recent memory items of the dementia rating scale of Gottfries et al. (36). As in the imipramine and clomipramine studies described above, a substantial antidepressant response to placebo occurred in these subjects, but citalopram was modestly superior to placebo and was well tolerated. Although the effects of citalopram on psychometric tests were not reported, observer-rated improvement in memory is supportive of other studies suggesting that SSRIs do not have an adverse effect on cognitive function (37).

These limited data from placebo-controlled studies of antidepressants in outpatients with AD and concurrent depression have several implications. First, the robust responses of the placebo group make it essential to include a placebo group in future antidepressant drug outcome trials in dementia patients if the results are to be interpretable. They also suggest that the “nonpharmacologic” aspects of trial participation contribute to a reduction of depressive signs and symptoms in patients with AD. A recent placebo-controlled trial demonstrating efficacy of behaviorally based psychotherapy for depression in AD outpatients is consistent with this interpretation (38). Second, the absence of SSRI adverse effects on cognitive function or blood pressure regulation is an advantage for this class of antidepressant in older patients (37 ,39). More placebo-controlled trials of SSRIs (and other newer antidepressants, such as venlafaxine and nefazodone) in AD patients with concomitant depression would be informative.

PSYCHOSIS AND DISRUPTIVE AGITATION

Part of "88 - Alzheimer Disease: Treatment of Noncognitive Behavioral Abnormalities "

Prevalence

The prevalence rates of psychotic symptoms (delusions and hallucinations) complicating AD cluster between 20% and 40% in carefully performed cross-sectional studies (2 ,4 ,7 ,40). Because psychotic symptoms can emerge at any time during the course of AD and probably are more prevalent in the later stages of the illness, longitudinal studies reveal even higher prevalence rates (6). Drevets and Rubin (41) longitudinally followed subjects with AD from the early through the later stages of illness and documented the occurrence of psychotic symptoms both cross-sectionally and cumulatively. This study was strengthened by the inclusion of a longitudinally followed age-matched normal control population (psychotic symptoms developed in none of them during the course of the study). Slightly more than 50% of the AD subjects manifested psychotic symptoms at some point during the course of their illness. The subjects were not considered positive for psychotic symptoms if they occurred only rarely or only in the context of a possible delirium. As in other studies of psychotic symptoms in AD (2 ,4 ,7), delusions were usually simple and persecutory, most commonly involving theft. Systematized delusions were uncommon. Hallucinations were most frequently visual, although auditory hallucinations were also common.

Disruptive agitated behaviors such as verbal and physical aggression, motor hyperactivity, uncooperativeness with essential care, persistent irritability, and repetitive vocalizations are highly prevalent in moderately to markedly demented patients with AD (8 ,11). Less severely demented patients with AD who are still able to reside in the community also manifest disruptive behaviors. Ryden (42) surveyed the caregivers of outpatients with AD and found a prevalence of aggressive behavior occurring at least once a week in 31% of subjects and daily in 16% of subjects.

Although psychotic symptoms and “nonpsychotic” disruptive agitated behaviors may not always reflect the same underlying pathophysiologic process or processes, these two classes of noncognitive behavioral problems often are expressed together. Lopez et al. (43) evaluated the presence of belligerence, uncooperativeness, and physical and verbal aggression in psychotic AD patients with delusions and hallucinations (*n* = 17) and nonpsychotic AD patients without delusions and hallucinations (*n* = 17). A greater proportion of psychotic AD patients (11 of 17) manifested these behavioral disturbances than did the nonpsychotic AD patients (1 of 17). A study addressing the relationship between psychotic symptoms and physical aggression in AD patients was reported by Deutsch et al. (40). Delusions (most commonly persecutory) and mistaken identifications (e.g., patients believing that their house is not their home or that strangers are living in the house) frequently preceded and were significantly associated with episodes of physical aggression. However, the presence of delusions could account for episodes of physical aggression in only a minority of cases. Therefore, factors other than apparent psychosis were involved in the precipitation of physically aggressive behavior.

Psychotic Symptoms and Disease Progression

A number of studies have consistently suggested that the presence of psychotic symptoms in AD is associated with more rapid deterioration of cognitive function. Stern et al. (44) were the first to report this phenomenon, and the association between psychotic symptoms and more rapid decline has since been confirmed by other groups. Lopez et al. (43) reported that AD patients with delusions and hallucinations had a more rapid decline in MMSE scores during a 1-year follow-up than did nonpsychotic AD patients, and they appeared to manifest a specific defect in receptive language. Drevets and Rubin (41) reported that the presence of psychotic symptoms predicted an increased rate of cognitive deterioration. Jeste et al. (45) compared the performance over time of delusional and nondelusional AD patients on a neuropsychological test battery. Patients with delusions had a more rapid rate of dementia progression than nondelusional AD patients. It is possible that the presence of psychotic symptoms reflects a more malignant pathobiologic process that adversely affects both behavior and neurodegeneration. The inadvertent inclusion in these studies of patients now more accurately classified as having dementia with Lewy bodies (DLB) also may have contributed to the apparent association between psychotic symptoms and more rapid cognitive deterioration. DLB is characterized by an early incidence and high prevalence of psychotic symptoms (46), and these patients appear to manifest a more rapidly deteriorating course of dementia (47).

Antipsychotic Drug Use for Psychotic and Disruptive Behaviors: Rationale

Psychotropic drugs are widely prescribed to patients with AD in long-term care facilities (16). In fact, repeated documentation of the widespread practice of prescribing antipsychotic and other psychotropic drugs in the long-term care setting, often for extended periods of time (13 ,14 ,15 and 16), has prompted the implementation of federal regulations designed to limit the use of such drugs to short-term treatment regimens with clear indications (48).

The rationale for the use of antipsychotic drugs in AD is partially attributable to the superficial phenomenologic similarities of delusions, hallucinations, and other disruptive behaviors occurring in AD to the symptoms of schizophrenia. Unfortunately, this analogy to psychotic and other behavioral signs and symptoms in schizophrenia is very imperfect. Hallucinations in AD usually are visual, whereas hallucinations in schizophrenia typically are auditory. Delusions in AD often are unelaborated persecutory beliefs, such as delusions of theft. Systematized, complex, and grandiose delusions are uncommon (9). In addition, the memory deficits of AD often appear to play a role in the development of delusional beliefs. For example, patients with AD who forget where an item has been placed and who do not understand their cognitive deficits may assume that it has been stolen. Or, a patient with AD can stubbornly insist in a delusional manner that a long-deceased person who still remains alive in available memory traces is, in fact, alive. Although such beliefs of theft, of a deceased person who continues to exist, or of a spouse who is an impostor may meet formal criteria for delusions, they are phenomenologically dissimilar to the typical delusions of schizophrenia for which the antipsychotic drugs have been demonstrated so effective. These phenomenologic differences may reflect underlying neurobiologic differences, and they may explain why the effect size of antipsychotic drugs in AD is modest (49) and much less robust than in schizophrenia.

Typical Antipsychotic Drugs: Outcome Studies before DSM-III

The interpretation of typical antipsychotic drug outcome studies in patients with dementia that were performed before the introduction of DSM-III is often hampered by the use of unclear diagnostic nomenclature. For example, the term *senile psychosis* connoted dementia with severe cognitive impairment rather than the presence of delusions and hallucinations. In addition, early studies were usually performed in state hospital populations, which included patients with a mixture of degenerative neurologic dementing disorders (at least a subgroup of whom presumably had AD) and patients with chronic schizophrenia who had grown old in the institutional setting. In such early studies targeting psychosis and disruptive agitation, chlorpromazine (50 ,51),

acetophenazine (52), and haloperidol (53) were each modestly more effective than placebo, but adverse effects such as excessive sedation, falls, unsteady gait, and pseudoparkinsonism were more common in the active drug groups.

It is not surprising that early studies of typical antipsychotic drugs in patients with dementia who did not manifest psychotic disruptive behaviors as target symptoms and signs found antipsychotic drugs no more effective than placebo (54 ,55). For example, a comparison of the effects of trifluoperazine and placebo on target symptoms of apathy, withdrawal, and cognitive and behavioral deterioration (loss of ambulation, disorientation, and incontinence) demonstrated no therapeutic effect of the active medication, and the findings were most remarkable for a high prevalence of trifluoperazine-induced sedation and extrapyramidal signs and symptoms (54). One of these studies, however, is particularly instructive concerning the substantial placebo response that can be seen in elderly patients with dementia despite the presence of cognitive impairment. In a comparison of the effects of thiothixene and placebo on cognitive deficits in a group of demented patients (55), the majority of patients in both groups were rated as globally improved (13 of 22 patients receiving thiothixene and 11 of 20 patients receiving placebo).

Typical Antipsychotic Drug Studies: Outcome Studies since DSM-III

Since the introduction of DSM-III, a small number of placebo-controlled studies have evaluated the efficacy of typical antipsychotic drugs in patients with dementia in both outpatient and institutional settings. Because of the use of operationalized diagnostic criteria for AD and MID, one can be more confident that elderly patients with chronic psychiatric disorders beginning in early life, such as schizophrenia, were excluded from these studies. In addition, the studies have clearly targeted psychotic or disruptive agitated behaviors as outcome measures.

Several studies have compared typical antipsychotics with placebo in either state hospital or community nursing home institutional settings. Petrie et al. (56) compared low-dose haloperidol and loxapine with placebo in 64 inpatients of a state psychiatric hospital (mean age, 73 years). The sample included subjects who met diagnostic criteria for either AD or MID. Both antipsychotic medications were more effective than placebo in reducing hallucinations, suspiciousness, hostility, excitement, and uncooperativeness. Global improvement ratings, however, only modestly favored the active medications. Thirty-five percent of haloperidol subjects and 32% of loxapine subjects, in comparison with 9% of placebo subjects, were rated as moderately or markedly improved. Not only were therapeutic responses to active drugs in these elderly demented patients lower than would have been predicted from outcome studies in younger subjects with schizophrenia, but adverse drug effects, including sedation and extrapyramidal signs and symptoms, were frequent. In a placebo-controlled study of antipsychotic drugs in a typical community nursing home sample of elderly demented patients, Barnes et al. (57) randomized 60 behaviorally disturbed patients with dementia (mean age, 83 years) to thioridazine, loxapine, or placebo. All subjects met diagnostic criteria for either AD or MID. Both active antipsychotic drugs were more effective than placebo for reducing excitement and uncooperativeness. Although suspiciousness and hostility tended to improve more with active drugs than with placebo, substantial improvements in these two factors also were documented in subjects receiving placebo. Like Petrie et al. (56), Barnes et al. (57) found that only approximately one-third of patients in the active medication conditions achieved global ratings of either moderate or marked improvement. Finkel et al. (58) compared the typical antipsychotic drug thiothixene with placebo in agitated nursing home patients with dementia (mean age, 85 years). Disruptive problem behaviors included physical aggression (hitting, kicking, and pushing), physical nonaggressive agitation (pacing and repetitive mannerisms), and verbal aggression (screaming and cursing). The presence and nature of psychotic delusions and hallucinations were not specified in this study. Thiothixene was more effective than placebo in this 11-week, parallel-design study, but differences between groups did not become apparent until 6 weeks of treatment had been completed. The positive effects of thiothixene appeared to persist for 3 to 6 weeks after drug discontinuation. Taken together, these three studies suggest that low-dose typical antipsychotic drugs are modestly effective for treating psychotic and nonpsychotic disruptive behaviors in patients with AD residing in long-term care facilities. However, parkinsonian rigidity and sedation occurred in some patients in each of these studies.

In a well-designed and well-executed study of the acute efficacy and adverse effects of a typical antipsychotic in AD outpatients with psychosis and disruptive behaviors, Devanand et al. (59) randomized 71 subjects to 6 weeks of standard-dose haloperidol (2 to 3 mg/d), low-dose haloperidol (0.5 to 0.75 mg/d), or placebo for 6 weeks. When an *a priori* response criterion of at least 25% improvement in one of three quantifications of target symptoms was used, response rates were significantly greater in the standard-dose (55% to 60%) than in either the low-dose (25% to 35%) or placebo (25% to 35%) groups. Although effect sizes were modest, standard-dose haloperidol clearly was effective for both psychosis and agitation. Not surprisingly, moderate to severe extrapyramidal signs developed in a subgroup of standard-dose haloperidol subjects (20%). Low-order but significant correlations were found between haloperidol blood levels and symptomatic improvement, and a stronger correlation between haloperidol blood levels and extrapyramidal signs. Haloperidol was effective for treating psychosis and agitation in AD outpatients in this study, but

the therapeutic window was narrow. A similar narrow therapeutic window in such patients may also exist for several of the newer atypical antipsychotics (see below). In a recently reported multicenter placebo-controlled comparison of haloperidol, trazodone, and behavioral management for disruptive agitation and psychosis in AD outpatients, haloperidol (mean dose, 2 mg/d) was not more effective than placebo (60). The reasons for the discrepant findings between this study and the positive study reported by Devanand et al. (59) in AD outpatients are unclear. One possibility is that the subjects in the study of Devanand et al. may have manifested more severe behavioral symptoms. In an early, placebo-controlled trial of haloperidol (53), it was in the more severely behaviorally disturbed patients that a positive effect of haloperidol was detectable.

Atypical Antipsychotic Drug Studies

A major problem in the use of traditional antipsychotic drugs to manage psychiatric and behavioral problems in AD and other dementing disorders is the frequent emergence of pseudoparkinsonian rigidity, tremor, and bradykinesia. Patients with DLB are particularly susceptible to these adverse effects (61). The atypical antipsychotic drugs such as clozapine, risperidone, olanzapine, and quetiapine offer the theoretic advantage of a reduced or minimal incidence of pseudoparkinsonism. The atypical antipsychotics may also theoretically prove more efficacious than the typical antipsychotics for psychosis and disruptive agitation in AD. The efficacy and tolerability of these drugs in AD patients with psychosis or disruptive agitation recently have been addressed in several large, well-designed, placebo-controlled outcome studies.

Katz et al. (62) reported the results of a large, double-blinded, placebo-controlled study of efficacy and safety of risperidone in institutionalized dementia patients with psychotic and disruptive behaviors (mean age, 83 years). Among the subjects, 73% met criteria for AD, 15% for vascular dementia, and 12% for mixed AD and vascular dementia. At the end of the 12-week study, BEHAVE-AD total scores, in addition to psychosis and aggressiveness subscale scores, were significantly more reduced in patients receiving either 1 or 2 mg of risperidone per day than in those receiving placebo. In contrast, a 0.5-mg low-dose group did not differ from the placebo group at 12 weeks except for a slightly greater reduction in the aggressiveness subscale of the BEHAVE-AD. Extrapyramidal adverse effects did not differ between the 0.5-mg group or the 1-mg group versus placebo. However, significantly more subjects (21%) in the group receiving 2 mg/d manifested extrapyramidal adverse effects than in the placebo group (7%). The efficacy of risperidone in this study was modest and comparable in magnitude with that reported in studies of typical antipsychotics in this type of institutionalized, very elderly sample (57 ,58). The authors concluded that 1 mg of risperidone per day is likely to be the optimal dose for the majority of dementia patients of this age and with cognitive and behavioral symptoms of this degree of severity. Another large, double-blinded, placebo-controlled trial of risperidone, haloperidol, and placebo generally supports the equivalent efficacy of these atypical and typical antipsychotics, but it also suggests an advantage of risperidone in terms of extrapyramidal adverse effects (63). Institutionalized patients with dementia (mean age, 81 years) were randomized to risperidone (mean dose, 1.1 mg/d), haloperidol (mean dose, 1.2 mg/d), or placebo for 12 weeks. Both the risperidone and haloperidol groups had significantly lower aggression cluster scores on the BEHAVE-AD at week 12 in comparison with the placebo group. The severity of extrapyramidal symptoms at endpoint did not significantly differ between the risperidone and placebo groups but was significantly greater in the haloperidol group than in the risperidone group.

The atypical antipsychotic drug olanzapine also has been studied in large, placebo-controlled trials in AD patients with psychosis and other behavioral disturbances. In an 8-week double-blinded, placebo-controlled trial that included 238 outpatients with AD and psychosis (mean age, 79 years), olanzapine (mean dose, 2.4 mg/d) was not significantly more effective than placebo in improving BEHAVE-AD scores, nor were adverse effects such as extrapyramidal signs more common in the olanzapine group (64). In a subsequent large, placebo-controlled study in AD patients residing in nursing facilities and manifesting psychosis and other behavioral disturbances (65), subjects were randomized to placebo or to 5 mg, 10 mg, or 15 mg of olanzapine per day. Clinically significant improvement, defined a reduction from baseline on the Neuropsychiatric Inventory (NPI) total score of 50% or more (66), was demonstrated in 66% of the patients receiving 5 mg/d, 57% receiving 10 mg/d, 43% receiving 15 mg/d, and 36% of the patients receiving placebo. These differences were significant for the groups receiving 5 or 10 mg/d but not for the group receiving 15 mg/d. In addition, the high-dose subjects had a significantly increased incidence of somnolence and abnormal gait. This study suggests that 5 mg per day may be an optimal dose of olanzapine for AD patients with psychosis and other disruptive agitated behaviors.

Taken together, studies of the atypical antipsychotics suggest that the efficacy of these newer agents is comparable with that of the typical antipsychotics and that, somewhat surprisingly, they may have a narrow “therapeutic window.” However, because the incidence of parkinsonism and tardive dyskinesia associated with the atypical antipsychotics is low, it is likely that they will be increasingly used to manage psychosis and disruptive agitation in patients with AD, despite their higher cost.

Maintenance Antipsychotic Drug Therapy in Dementia

When a satisfactory response to an antipsychotic drug has been achieved with an acute treatment regimen, the clinician must next decide how long to maintain the patient on the drug. This question was addressed by Risse et al. (67) in a small but informative antipsychotic drug discontinuation study in behaviorally disturbed elderly patients with dementia who appeared to have benefited from an acute course of antipsychotic medications and had then been maintained on these medications on a long-term basis. Placebo was substituted for maintenance antipsychotic medication in nine men with dementia (mean age, 65 years) who had shown a clear reduction in noncognitive behavioral problems following treatment with antipsychotic medication and who subsequently had been maintained on antipsychotic medication for at least 90 days. At the end of the 6-week placebo-substitution period, disruptive behaviors severe enough to warrant reinstitution of antipsychotic medication had developed in only one patient. Of the remaining eight patients, five actually were less agitated, two were unchanged, and only one was rated as more agitated than when he had been receiving maintenance antipsychotic medication. This small study supports the wisdom of periodic discontinuation of long-term antipsychotic medication to evaluate the need for maintenance. In a larger study performed in 36 community nursing home patients (mean age, 82 years) who met criteria for probable or possible AD, patients were randomly assigned to either continuation of antipsychotic medication or withdrawal from antipsychotic medication and substitution of placebo (68). Of the 22 patients withdrawn from antipsychotic medication, 20 (91%) were able to complete the 4-week, double-blinded withdrawal. In only two cases did the nursing home staff request that the patients be withdrawn from the study because of emergencies involving unacceptable levels of agitation. No significant difference in the incidence of emergent physically aggressive behavior was found between patients withdrawn from antipsychotic medication and those maintained on antipsychotic medication. Half of the patients withdrawn from antipsychotic medication remained off the drugs for an extended period of time after the end of the study, even after the blind had been broken. These two studies demonstrate that an attempt at withdrawal from antipsychotic medication in behaviorally stable patients with dementia is feasible.

Dementia with Lewy Bodies: Implications for Psychopharmacology

It is increasingly clear that a subgroup of patients meeting formal criteria for probable AD (69) are more accurately classified diagnostically as having DLB (46).

The defining neuropathologic feature of DLB is the presence of Lewy bodies in the forebrain. These α-synuclein-containing intracytoplasmic inclusions are the classic histopathologic lesion of Parkinson disease, but substantial numbers of Lewy bodies rarely are expressed outside the substantia nigra in this latter disorder. In DLB, numerous Lewy bodies are found in neocortex, limbic brain areas, and other parts of the forebrain. In most cases of DLB, modest numbers of the amyloid plaques and neurofibrillary tangles characteristic of classic AD are also found. In addition, the presynaptic cholinergic deficit present in AD (70 ,71) is very prominent in DLB (72). The parkinsonian features and cholinergic deficit of DLB have implications for the pharmacologic management of the noncognitive behavioral aspects of this disorder. First, the parkinsonian features of DLB make these patients extremely sensitive to dopaminergic blockade by typical antipsychotics (61). The atypical antipsychotics appear preferable for the management of psychosis in DLB (73). Also, several studies suggest that compensating for the profound cholinergic deficit of DLB with cholinesterase inhibitor therapy improves psychotic and other noncognitive behavioral problems in this disorder (26 ,74 ,75).

Other Pharmacologic Approaches to the Management of Agitated Behaviors in Alzheimer Disease

Despite their somewhat disappointing therapeutic effect size, the consensus is that antipsychotic drugs should be prescribed for clear and troublesome delusions and hallucinations. However, the rationale for prescribing antipsychotic drugs as the drug class of choice for AD patients with disruptive agitation in the absence of clear psychotic symptoms is less compelling. In such patients, attempts to demonstrate efficacy for other types of psychotropic drugs are both reasonable and important. Unfortunately, the database derived from well-designed clinical trials of psychotropic drugs other than the antipsychotics for the management of disruptive behaviors in AD is even less robust than that for the antipsychotic drugs. The following review, therefore, relies heavily on anecdotal reports and non–placebo-controlled studies when data from interpretable placebo-controlled studies are not available.

Benzodiazepines

The use of benzodiazepines in patients with AD and other dementing disorders has been reviewed (76). In a group of “emotionally disturbed” elderly patients (mean age, 81 years), Sanders (77) evaluated the efficacy of oxazepam in comparison with placebo in an 8-week treatment trial. Oxazepam was superior to placebo, particularly for reduction of agitation and anxiety. Interpretation of this study is hampered

by the vagueness of the diagnoses and the likelihood that many of the subjects were not demented. Coccaro et al. (78) compared oxazepam, haloperidol, and the sedating antihistamine diphenhydramine in elderly institutionalized patients. The mean age of these subjects was 75 years, most met criteria for AD, and target signs and symptoms included tension, excitement, aggressiveness, pacing, and increased motor activity. Ratings of target signs and symptoms improved during an 8-week period in all treatment groups. Although statistically significant differences between the groups did not emerge, a trend for greater improvement with diphenhydramine or haloperidol than with oxazepam was noted. The lack of a placebo group in this study complicates interpretation of the modest improvements in objective ratings of disruptive behaviors.

Salzman et al. (79) tapered and then discontinued benzodiazepines in 13 elderly nursing home residents and compared memory function and ratings of depression, anxiety, irritability, and sleep between the subjects who discontinued benzodiazepine and 12 nursing home residents who continued their benzodiazepine regimens. The group in which the drug was discontinued showed greater improvements in memory than did the group that continued to take benzodiazepine, and no differences between the groups were found in measures of depression, anxiety, irritability, or sleep. This study suggests that at least a subgroup of patients maintained for an extended time on short-acting benzodiazepines may benefit from a trial of drug discontinuation. In addition to adverse effects on cognitive function, benzodiazepines have been associated with falls in geriatric psychiatric inpatients (80). Taken together, these studies of benzodiazepines in behaviorally disturbed patients with dementia suggest that the use of benzodiazepines is best limited to short-term treatment of acute anxiety and agitation, and that benzodiazepines are a poor choice for long-term management of disruptive agitation in AD.

Buspirone

Buspirone is a partial 5-hydroxytryptamine subtype 1A ($5\text{-}HT_{1A}$)-receptor agonist with antianxiety activity and a relatively benign adverse effect profile. Two uncontrolled studies of buspirone in dementia patients with agitated behavior have been reported. Sakauye et al. (81) prescribed buspirone to 10 patients with AD complicated by agitated behaviors in an open-label study. A modest but statistically significant overall reduction of agitated behaviors was noted, as was a substantial variability in response, with 4 of the 10 patients demonstrating marked declines in disruptive behaviors. In a similar study, Hermann and Eryavec (82) prescribed buspirone to a group of elderly nursing home residents with heterogeneous types of dementia (AD, MID, and alcoholic dementia). All subjects had demonstrated severe behavioral disturbances, including agitation and aggression, and all had failed to improve with previous trials of other types of psychotropic medications, principally antipsychotic drugs. Six of 16 patients were rated as much or very much improved in terms of agitation and aggressive behavior. Levy (83) used buspirone to treat 20 patients with AD and behavioral disturbances rated as at least moderately troublesome on the BEHAVE-AD in a single-blinded dose-escalation study. After a 2- to 4-week psychotropic drug washout period, subjects were given placebo for 1 week and then progressively increasing weekly doses (15, 30, 45, and 60 mg) of buspirone. A dose–response improvement in anxiety rating occurred. As in the study of Barnes et al. (57) of antipsychotic drugs, placebo had a significant effect on delusions. In these studies of buspirone, adverse effects were unusual. Given the low toxicity of this drug, further placebo-controlled investigations appear warranted.

Serotoninergic Drugs

Selective Serotonin Reuptake Inhibitors

The clear serotoninergic deficit demonstrated in AD (29 ,30) and the relationship between low central nervous system serotoninergic activity and aggressive behaviors in nondemented persons (84) provide the rationale for studies addressing the behavioral efficacy of drugs that enhance central serotoninergic neurotransmission in AD patients with agitated behaviors. In a recently reported multicenter trial, AD subjects selected for the presence of psychosis or disruptive agitation were first treated openly with the cholinesterase inhibitor donepezil and then randomized to the addition of either the SSRI sertraline or placebo (85). Sertraline had a modest positive effect on agitated behaviors (but not psychosis) in comparison with placebo. Two multisite Scandinavian studies have evaluated SSRIs in demented patients with a variety of predominantly nonpsychotic behavioral disturbances. These patients were not reported to have met diagnostic criteria for depression. In demented patients with either AD or vascular dementia, the SSRI citalopram was more effective than placebo for the target symptoms of irritability, fear/panic, depressed mood, and restlessness (86). Improvement was limited to demented patients with AD. No significant effects of citalopram were noted in patients with vascular dementia. Cognitive function was unaffected by either citalopram or placebo, and citalopram was well tolerated by the elderly subjects in this study. In another study of demented patients with AD or vascular dementia (87), the SSRI fluvoxamine tended to be more effective than placebo on the target symptoms of confusion, irritability, anxiety, fear/panic, mood level, and restlessness. The differences between fluvoxamine and placebo, however, failed to reach statistical significance. Although more studies are needed to evaluate the possible roles of the SSRIs for disruptive agitation, results to date do not support their use as first-line agents for this indication.

Other Serotoninergic Drugs

Trazodone is a sedating antidepressant with serotoninergic agonist activity. Simpson and Foster (88) treated four demented

patients who manifested disruptive behaviors with trazodone after antipsychotic drug treatment had proved ineffective. In this anecdotal report, trazodone in doses of 200 to 500 mg daily was associated with decreased agitation and aggressive behavior. Pinner and Rich (89) treated seven demented patients with trazodone for symptomatic aggressive behavior. Again, all subjects had failed to improve with antipsychotic drug therapy. Three of the seven patients demonstrated an apparent marked decrease in aggressive behavior following 4 to 6 weeks of trazodone at doses ranging from 200 to 350 mg/d. In a recently reported multisite study comparing haloperidol, trazodone, behavioral management, and placebo, the results with trazodone (mean dose, 200 mg/d) did not differ from those with placebo (59). In a double-blinded, placebo-controlled crossover study, Lawlor et al. (90) treated 10 patients with AD and behavioral complications (troublesome agitation, depression, psychosis, or anxiety) with trazodone (up to 150 mg/d), buspirone (30 mg/d), or placebo. Trazodone produced a small but significant behavioral improvement in comparison with placebo, whereas buspirone had no apparent effect.

Anticonvulsant Drugs

Because the hyperactive and aggressive behaviors encountered in the manic phase of bipolar disorder at least superficially can resemble agitated behaviors in AD and other dementias, the anticonvulsant drugs effective in the treatment of mania may benefit behaviorally disturbed patients with dementia. Lithium has not been helpful for behavioral symptoms in AD (91). Marin and Greenwald (92) treated two AD patients and one MID patient with carbamazepine in an attempt to reduce combative, agitated behaviors. Within 2 weeks of carbamazepine treatment at doses ranging from 100 to 300 mg/d, behavioral improvement was noted in all subjects. In a larger open study of AD patients who had failed to respond to antipsychotic drugs (93), reduction in hostility, agitation, and uncooperativeness was noted in five of nine patients. In this study, two patients whose agitated behaviors decreased manifested ataxia and confusion, which resolved with reduction of the carbamazepine dose. The mean dose of carbamazepine in this study was 480 mg/d. In contrast to enthusiastic authors of these small reports, Chambers et al. (94) noted no overall benefit from carbamazepine in 19 elderly patients with dementia who were prescribed carbamazepine at doses of 100 to 300 mg/d. Target symptoms in this study were wandering, overactivity, and restlessness. Tariot et al. (95), in a 6-week placebo-controlled, randomized, multisite, parallel-group study, evaluated the efficacy, safety, and tolerability of the anticonvulsant drug carbamazepine in demented patients in long-term care facilities with disruptive agitated behaviors. The modal carbamazepine dose at 6 weeks was 300 mg/d, with a mean serum level of 5.3 µg/mL. Statistical superiority of carbamazepine compared with placebo was attributable primarily to a greater decrease in agitation and aggression in the carbamazepine group.

Sodium valproate is another anticonvulsant drug with demonstrated antimanic activity. Mellow et al. (96) treated four patients with AD who manifested disruptive and agitated behaviors with doses of valproate ranging from 500 mg twice daily to 500 mg three times daily; treatment lasted for 1 to 3 months. Substantial behavioral improvement was noted in two of the four patients, and adverse effects did not appear. A recently reported multisite study of the anticonvulsant valproate sodium has been less encouraging (97). Because of a high frequency of adverse effects, primarily excessive somnolence, the study was terminated prematurely. Although at the end of the study agitation was reduced more in valproate subjects than in placebo subjects, it is possible that this apparent therapeutic effect reflected the high degree of sedation experienced by the subjects in the active treatment group. Methodologically, this study was designed to address “mania-like” symptoms in persons with AD, and the higher doses of valproate typically prescribed for the treatment of mania in a younger population were achieved. Lower doses of valproate may have beneficial therapeutic effects with a more tolerable adverse effect profile.

Cholinergic Enhancement

That drugs that enhance cholinergic neurotransmission in the central nervous system decrease agitation and psychotic symptoms in persons with mild to moderate AD has been an unanticipated finding of large, multisite outcome trials demonstrating modest positive effects of these agents on cognitive function (98 ,99 and 100). The contribution of a presynaptic cholinergic deficit to memory and other cognitive impairments in AD (69 ,70) has been a cornerstone of AD drug development. Interest in a potential contribution of this cholinergic deficit to noncognitive behavioral problems in AD increased after Cummings (101) observed that the agitation and psychotic symptoms characteristic of delirium induced by anticholinergic drug toxicity resemble some noncognitive behavioral symptoms occurring spontaneously in AD (e.g., hallucinations, agitation). Cummings reasoned that enhancing brain cholinergic neurotransmission might improve such noncognitive symptoms. Empiric support for this hypothesis came from a carefully performed single-case study in which the cholinesterase inhibitor physostigmine reduced psychotic symptoms in a patient with AD (102) and from a double-blinded crossover study in which physostigmine and the typical antipsychotic haloperidol equally reduced psychotic and agitated behaviors in 13 patients with advanced AD (103). Further support has come from *post hoc* and secondary outcome analyses of large, multicenter cholinesterase inhibitor outcome trials in AD. In addition to demonstrating modest effects on cognitive function,

the cholinesterase inhibitors tacrine (104), galantamine (105), donepezil (106), metrifonate (107), and (in DLB subjects) rivastigmine (108) significantly improved such noncognitive behaviors as delusions, hallucinations, pacing, and uncooperativeness more than did placebo. However, these large, multicenter cholinesterase inhibitor studies excluded AD patients with substantial noncognitive behavioral problems. It is just such severely disturbed patients who must be studied prospectively to establish the clinical importance of cholinesterase inhibitors in the management of noncognitive behavioral problems in AD.

Perhaps the most impressive data supporting positive effects of cholinergic enhancement on noncognitive symptoms in AD come from a multicenter trial of the selective M1 muscarinic cholinergic agonist xanomeline (109). Although the effects of xanomeline on cognitive function disappointingly did not differ from those of placebo, this investigational cholinergic agonist both decreased psychotic and agitated behaviors present at study entry and reduced emergence of these symptoms during the 6-month duration of the study. Furthermore, a clear dose–response effect was observed over the three doses of xanomeline administered. That xanomeline decreases activity of A10 dopaminergic neurons in preclinical studies suggests a possible interaction between cholinergic enhancement and more traditional dopaminergic antagonist approaches to the reduction of psychotic symptoms.

ß-Adrenergic Antagonists

Despite substantial loss of noradrenergic locus ceruleus neurons in AD, studies measuring the concentrations of norepinephrine and its metabolites in either postmortem brain tissue or cerebrospinal fluid (110 ,111 and 112) suggest that noradrenergic outflow is maintained in AD, apparently through compensatory up-regulation of remaining noradrenergic neurons in the locus ceruleus. In addition, patients with AD manifest an enhanced behavioral agitation response to brain noradrenergic stimulation (113). An uncontrolled study suggested that the centrally active ß-adrenergic antagonist propranolol reduces disruptive agitated behaviors in AD (114). In a recently reported placebo-controlled pilot study of Peskind et al. (115), propranolol was superior to placebo in reducing disruptive agitated behaviors in elderly nursing home residents with AD. These patients' disruptive behaviors had been unresponsive to antipsychotic or other psychotropic medications. Propranolol was well tolerated in this small but very elderly population. These neurobiological findings and pilot clinical data warrant larger-scale placebo-controlled trials of propranolol in AD patients with disruptive agitated behaviors.

CONCLUSION

Part of "88 - Alzheimer Disease: Treatment of Noncognitive Behavioral Abnormalities "

The database supporting guidelines for the pharmacologic management of noncognitive behavioral abnormalities in AD remains limited despite the prevalence of these problems and their impact on patient management. Extrapolating from psychopharmacologic outcome studies in younger, nondemented patients with such diseases as depression and schizophrenia has not been a satisfactory approach to developing effective pharmacologic treatments for noncognitive behavioral disturbances in elderly patients with AD and other dementing disorders. The clear placebo responses noted in several carefully performed pharmacologic trials in AD patients with noncognitive behavioral problems emphasize the necessity for inclusion of a placebo condition if a study is to be interpretable. More must be learned about the neurobiological substrates of psychotic and disruptive agitated behaviors in AD. Such knowledge is essential to the rational development of more effective pharmacotherapeutics for these disturbing and costly problems.

DISCLOSURE

Part of "88 - Alzheimer Disease: Treatment of Noncognitive Behavioral Abnormalities "

Dr. Raskind receives research support and/or consultant honoraria from the following companies: Janssen Pharmaceutica, Novartis, Eli Lilly, Pfizer, and Forest Laboratories.

REFERENCES

1. Alzheimer A. About a peculiar disease of the cerebral cortex [1907 article translated by L. Jarvik and H. Greenson H]. *Alzheimer Dis Assoc Disord* 1987;1:7–8.

2. Cummings JL, Miller B, Hill MA, et al. Neuropsychiatric aspects of multi-infarct dementia and dementia of the Alzheimer type. *Arch Neurol* 1987;44:389–393.

3. Herz LR, Volicer L, Ross V, et al. Pharmacotherapy of agitation in dementia. *Am J Psychiatry* 1992;149:1757–1758.

4. Wragg RE, Jeste DV. Overview of depression and psychosis in Alzheimer's disease. *Am J Psychiatry* 1989;146:577–587.

5. Kumar A, Koss E, Metzler D, et al. Behavioral symptomatology in dementia of the Alzheimer type. *Alzheimer Dis Assoc Disord* 1988;2:363–365.

6. Levy ML, Cummings JL, Fairbanks LA, et al. Longitudinal assessment of symptoms of depression, agitation, and psychosis in 181 patients with Alzheimer's disease. *Am J Psychiatry* 1996;153:1438–1443.

7. Lyketsos CG, Steinberg M, Tschanz JT, et al. Mental and behavioral disturbances in dementia: findings from the Cache County study on memory in aging. *Am J Psychiatry* 2000;157:708–714.

8. Morriss RK, Rovner BW, German PS. Factors contributing to nursing home admission because of disruptive behavior. *Int J Geriatr Psychiatry* 1996;11:243–249.

9. Reisberg B, Borenstein J, Salob SP, et al. Behavioral symptoms in Alzheimer's disease: phenomenology and treatment. *J Clin Psychiatry* 1987;48[Suppl 5]:9–15.

10. Cooper JK, Mungas D, Weiler PG. Relation of cognitive status and abnormal behavior in Alzheimer's disease. *J Am Geriatr Soc* 1990;38:867–870.

11. Souder E, Heithoff K, O'Sullivan PS, et al. Identifying patterns of disruptive behavior in long-term care residents. *J Am Geriatr Soc* 1999;47:830–836.

12. Steele C, Rovner B, Chase GA, et al. Psychiatric symptoms and nursing home placement of patients with Alzheimer's disease. *Am J Psychiatry* 1990;147:1049–1051.

13. Avorn J, Dreyer P, Connelly MA, et al. Use of psychoactive medication and the quality of care in rest homes. *N Engl J Med* 1989;320:227–232.

14. Beers M, Avorn J, Soumerai SB, et al. Psychoactive medication use in intermediate-care facility residents. *JAMA* 1988;260:3016–3020.

15. Buck JA. Psychotropic drug practice in nursing homes. *J Am Geriatr Soc* 1988;36:409–418.

16. Lasser RA, Sunderland T. Newer psychotropic medication use in nursing home residents. *J Am Geriatr Soc* 1998;46:202–207.

17. Cohen RM, Weingartner HW, Smallberg A, et al. Effort and cognition in depression. *Arch Gen Psychiatry* 1982;39:593–597.

18. Pearson JL, Teri L, Reifler BV, et al. Functional status and cognitive impairment in Alzheimer's patients with and without depression. *J Am Geriatr Soc* 1989;37:1117–1121.

19. Reifler BV, Larson E, Teri L, et al. Alzheimer's disease and depression. *J Am Geriatr Soc* 1986;14:855–859.

20. Burke WJ, Rubin EH, Morris JC, et al. Symptoms of "depression" in dementia of the Alzheimer type. *Alzheimer Dis Assoc Dis* 1988;2:356–362.

21. Roth M. The natural history of mental disorders in old age. *J Ment Sci* 1955;101:281–301.

22. Demuth GW, Rand GH. Atypical major depression in a patient with severe primary degenerative dementia. *Am J Psychiatry* 1980;137:1609–1610.

23. Snow SS, Wells CE. Case studies in neuropsychiatry: diagnosis and treatment of coexistent dementia and depression. *J Clin Psychiatry* 1981;42:11:439–441.

24. Jenike MA. Monoamine oxidase inhibitors as treatment for depressed patients with primary degenerative dementia (Alzheimer's disease). *Am J Psychiatry* 1985;162:6:763–764.

25. Volicer L, Rheaume Y, Cyr D. Treatment of depression in advanced Alzheimer's disease using sertraline. *J Geriatr Psychiatry Neurol* 1994;7:227–229.

26. Reynolds CF III, Perel JM, Kupfer DJ, et al. Open-trial response to antidepressant treatment in elderly patients with mixed depression and cognitive impairment. *Psychiatry Res* 1987;21:111–122.

27. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56–62.

28. Greenwald BS, Kramer-Ginsberg E, Marin DB, et al. Dementia with coexistent major depression. *Am J Psychiatry* 1989;146:11:1472–1478.

29. Blenow KAH, Wallin A, Gottfries CG, et al. Significance of decreased lumbar CSF levels of HVA and 5-HIAA in Alzheimer's disease. *Neurobiol Aging* 1991;13:107–113.

30. Hardy J, Adolfsson R, Alafuzoff I, et al. Review: transmitter deficits in Alzheimer's disease. *Neurochem Int* 1985;7:545–563.

31. Kamath M, Finkel SI, Moran MB. A retrospective chart review of antidepressant use, effectiveness, and adverse effects in adults age 70 and older. *Am J Geriatr Psychiatry* 1996;4:167–172.

32. Mamdani MM, Parikh SV, Austin PC, et al. Use of antidepressants among elderly subjects: trends and contributing factors. *Am J Psychiatry* 2000;157:360–367.

33. Reifler BV, Teri L, Raskind M, et al. Double-blind trial of imipramine in Alzheimer's disease patients with and without depression. *Am J Psychiatry* 1989;46:45–49.

34. Petracca G, Teson A, Chermerinski E, et al. A double-blind placebo-controlled study of clomipramine in depressed patients with Alzheimer's disease. *J Neuropsychiatry Clin Neurosci* 1996;8:270–275.

35. Nyth AL, Gottfries CG, Lyby K, et al. A controlled multicenter clinical study of citalopram and placebo in elderly depressed patients with and without concomitant dementia. *Acta Psychiatr Scand* 1992;86:138–145.

36. Gottfries CG, Brane G, Gullberg B, et al. A new rating scale for dementia syndromes. *Arch Gerontol Geriatr* 1982;1:311–330.

37. Bondareff W, Alpert M, Friedhoff AJ, et al. Comparison of sertraline and nortriptyline in the treatment of major depressive disorder in late life. *Am J Psychiatry* 2000;157:729–736.

38. Teri L, Logsdon RB, Uomoto J, et al. Behavioral treatment of depression in dementia patients: a controlled clinical trial. *J Gerontol* 1997;52B:P159–P166.

39. Teri L, Reifler BV, Veith RC, et al. Imipramine in the treatment of depressed Alzheimer's patients: impact on cognition. *J Gerontol* 1991;46:P373–377.

40. Deutsch LH, Bylsma FW, Rovner BW, et al. Psychosis and physical aggression in probable Alzheimer's disease. *Am J Psychiatry* 1991;148:1159–1163.

41. Drevets WC, Rubin EH. Psychotic symptoms and the longitudinal course of senile dementia of the Alzheimer type. *Biol Psychiatry* 1989;25:39–48.

42. Ryden MB. Aggressive behavior in persons with dementia who live in the community. *Alzheimer Dis Assoc Disord* 1988;2:342–355.

43. Lopez OL, Becker JT, Brenner RP, et al. Alzheimer's disease with delusions and hallucinations: neuropsychological and electroencephalographic correlates. *Neurology* 1991;41:906–912.

44. Stern Y, Mayeux R, Sano M, et al. Predictors of disease course in patients with probable Alzheimer's disease. *Neurology* 1987;37:1649–1653.

45. Jeste DV, Wragg RE, Salmon DP, et al. Cognitive deficits of patients with Alzheimer's disease with and without delusions. *Am J Psychiatry* 1992;149:184–189.

46. McKeith IG, Galasko D, Kosaka K, et al. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. *Neurology* 1996;47:1113–1124.

47. Ballard C, Holmes C, McKeith I, et al. Psychiatric morbidity in dementia with Lewy bodies: a prospective clinical and neuropathological comparative study with Alzheimer's disease. *Am J Psychiatry* 1999;156:1039–1045.

48. Health Care Financing Administration. Medicare and Medicaid: requirements for long-term care facilities. Final rule with request for comments. *Federal Register* 1989:5316–5336.

49. Schneider LS, Pollock VE, Lyness SA. A metaanalysis of controlled trials of neuroleptic treatment in dementia. *J Am Geriatr Soc* 1990;38:553–563.

50. Abse DW, Dahlstrom WG. The value of chemotherapy in senile mental disturbances. *JAMA* 1960;174:2036–2042.

51. Seager CP. Chlorpromazine in treatment of elderly psychotic women. *Br Med J* 1955;1:882–885.

52. Hamilton LD, Bennett JL. Acetophenazine for hyperactive geriatric patients. *Geriatrics* 1962;17:596–601.

53. Sugarman AA, Williams BH, Adlerstein AM. Haloperidol in the psychiatric disorders of old age. *Am J Psychiatry* 1964;120:1190–1192.

54. Hamilton LD, Bennett JL. The use of trifluoperazine in geriatric patients with chronic brain syndrome. *J Am Geriatr Soc* 1962;10:140–147.

55. Rada RT, Kellner R. Thiothixene in the treatment of geriatric patients with chronic organic brain syndrome. *J Am Geriatr Soc* 1976;24:105–107.

56. Petrie WM, Ban TA, Berney S, et al. Loxapine in psychogeriatrics: a placebo- and standard-controlled clinical investigation. *J Clin Psychopharmacol* 1982;2:122–126.

57. Barnes R, Veith R, Okimoto J, et al. Efficacy of antipsychotic medications in behaviorally disturbed dementia patients. *Am J Psychiatry* 1982;139:1170–1174.

58. Finkel SI, Lyons JS, Anderson RL, et al. A randomized, placebo-controlled trial of thiothixene in agitated, demented nursing home patients. *Int J Geriatr Psychiatry* 1995;10:129–136.

59. Devanand DP, Marder K, Michaels KS, et al. A randomized, placebo-controlled dose-comparison trial of haloperidol for psychosis and disruptive behaviors in Alzheimer's disease. *Am J Psychiatry* 1998;155:1512–1520.

60. Teri L, Logsdon RG, Peskind E, et al. Treatment of agitation in Alzheimer's disease: a randomized placebo-controlled clinical trial. Presented at the sixth international Stockholm/Springfield symposium on advances in Alzheimer therapy, April 5–8, 2000, Stockholm, Sweden.

61. McKeith I, Fairbairn A, Perry R, et al. Neuroleptic sensitivity in patients with senile dementia of Lewy body type. *BMJ* 1992;305:673–678.

62. Katz IR, Jeste DV, Mintzer JE, et al. Comparison of risperidone and placebo for psychosis and behavioral disturbances associated with dementia: a randomized, double-blind trial. *J Clin Psychiatry* 1999;60:107–115.

63. De Deyn PP, Rabheru K, Rasmussen A, et al. A randomized trial of risperidone, placebo, and haloperidol for behavioral symptoms of dementia. *Neurology* 1999;53:946–955.

64. Satterlee WG, Reams SG, Burns PR, et al. A clinical update on olanzapine treatment in schizophrenia and in elderly Alzheimer's disease patients. *Psychopharmacol Bull* 1995;31:534.

65. Street JS, Clark WS, Mitan S, et al. Olanzapine in the treatment of psychosis and behavioral disturbances associated with Alzheimer's disease. Presented at the annual meeting of the American College of Neuropsychopharmacology, December 14–18, 1998, Las Croabas, Puerto Rico.

66. Cummings JL, Mega M, Gray K, et al. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology* 1994;44:2308–2314.

67. Risse SC, Cubberly L, Lampe TH, et al. Acute effects of neuroleptic withdrawal in elderly dementia patients. *J Geriatr Drug Ther* 1987;2:65–67.

68. Bridges-Parlet S, Knopman D, Steffes S. Withdrawal of neuroleptic medications from institutionalized dementia patients: results of a double-blind, baseline-treatment-controlled pilot study. *J Geriatr Psychiatry Neurol* 1997;10:119–126.

69. McKhann G, Drachman D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: a report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939–944.

70. Davies P. Studies on the neurochemistry of central cholinergic systems in Alzheimer's disease. In: Katzman R, Terry RD, Bick KL, eds. *Alzheimer's disease: senile dementia and related disorders*. New York: Raven Press, 1978:453–459.

71. Whitehouse PJ, Price DL, Struble RG, et al. Alzheimer's disease and senile dementia: loss of neurons in the basal forebrain. *Science* 1982;215:1237–1239.

72. Perry RH, Irving D, Blessed G, et al. Senile dementia of Lewy body type: a clinically and neuropathologically distinct form of Lewy body dementia in the elderly. *J Neurol Sci* 1990;95:119–139.

73. Clark WS, Street J, Sanger T, et al. Reduction of psychotic symptoms in patients with Lewy body-like symptoms treated with olanzapine. *Neurology* 2000;54:A74.

74. Kaufer DI, Catt KE, Lopez OL, et al. Dementia with Lewy bodies: response of delirium-like features to donepezil. *Neurology* 1998;51:1512.

75. Michel BF, Lepine-Haudot M, Delfini M, et al. Clinical efficacy of donepezil on cognitive functions and psychotic manifestations of Lewy body dementia. Presented at the sixth international Stockholm/Springfield symposium on advances in Alzheimer therapy, April 5–8, 2000, Stockholm, Sweden.

76. Stern RG, Duffelmeyer ME, Zemishlani Z, et al. The use of benzodiazepines in the management of behavioral symptoms in dementia patients. *Psychiatr Clin North Am* 1991;14:375–384.

77. Sanders JF. Evaluation of oxazepam and placebo in emotionally disturbed aged patients. *Geriatrics* 1965;20:739–746.

78. Coccaro EF, Zembishlany Z, Thorne A, et al. Pharmacologic treatment of noncognitive behavioral disturbances in elderly demented patients. *Am J Psychiatry* 1990;147:1640–1645.

79. Salzman C, Fisher J, Nobel K, et al. Cognitive improvement following benzodiazepine discontinuation in elderly nursing home residents. *Int J Geriatr Psychiatry* 1992;7:89–93.

80. Aisen PS, DeLuca T, Lawton BA. Falls among geropsychiatry inpatients are associated with PRN medications for agitation. *Int J Geriatr Psychiatry* 1992;7:709–712.

81. Sakauye KM, Camp CJ, Ford PA. Effects of buspirone on agitation associated with dementia. *Am J Geriatr Psychiatry* 1993;1:82–84.

82. Hermann N, Eryavec G. Buspirone in the management of agitation and aggression associated with dementia. *Am J Geriatr Psychiatry* 1993:1:249–253.

83. Levy MA. A trial of buspirone for the control of disruptive behaviors in community-dwelling patients with dementia. *Int J Geriatr Psychiatry* 1994;9:841–848.

84. Stanley B, Molcho A, Stanley M, et al. Association of aggressive behavior with altered serotoninergic function in patients who are not suicidal. *Am J Psychiatry* 2000;157:609–614.

85. McRae T, Griesing T, Whalen E. Managing behavioral symptoms in Alzheimer's disease. Presented at the annual meeting of the American Psychiatry Association, May 13–28, 2000, Chicago.

86. Nyth AL, Gottfries CG. The clinical efficacy of citalopram in treatment of emotional disturbances in dementia disorders. *Br J Psychiatry* 1990_A57:894–901.

87. Olafsson K, Jorgensen S, Jensen HV, et al. Fluvoxamine in the treatment of demented elderly patients: a double-blind, placebo-controlled study. *Acta Psychiatr Scand* 1992;85:453–456.

88. Simpson DM, Foster D. Improvement in organically disturbed behavior with trazodone treatment. *J Clin Psychiatry* 1986;47:191–193.

89. Pinner E, Rich CL. Effects of trazodone on aggressive behavior in seven patients with organic mental disorders. *Am J Psychiatry* 1988;145:1295–1296.

90. Lawlor BA, Radcliffe J, Molchan SE, et al. A pilot placebo-controlled study of trazodone and buspirone in Alzheimer's disease. *Int J Geriatr Psychiatry* 1994;9:55–59.

91. Holton A, George K. The use of lithium in severely demented patients with behavioral disturbance. *Br J Psychiatry* 1985;146:99–100.

92. Marin DB, Greenwald BS. Carbamazepine for aggressive agitation in demented patients. *Am J Psychiatry* 1989;46:805.

93. Gleason RP, Schneider LS. Carbamazepine treatment of agitation in Alzheimer's outpatients refractory to neuroleptics. *J Clin Psychiatry* 1990;51:115–118.

94. Chambers CA, Bain J, Rosbottom R, et al. Carbamazepine in senile dementia and overactivity—a placebo-controlled double blind trial. *IRCS Med Sci* 1982;10:505–506.

95. Tariot PN, Erb R, Podgorski CA, et al. Efficacy and tolerability of carbamazepine for agitation and aggression in dementia. *Am J Psychiatry* 1998;155:54–61.

96. Mellow AM, Solano-Lopez C, Davis S. Sodium valproate in the treatment of behavioral disturbance in dementia. *J Geriatr Psychiatry Neurol* 1993;6:28–32.

97. Tariot PN, Schneider LS, Mintzer JE, et al. Divalproex treatment of mania with dementia. Presented at the annual meeting of the American Psychiatry Association, May 13–28, 2000, Chicago.

98. Corey-Bloom J, Anand R, Veach J (ENA 713 B352 Study Group). A randomized trial evaluating the efficacy and safety of ENA 713 (rivastigmine tartrate), a new acetylcholinesterase inhibitor, in patients with mild to moderately severe Alzheimer's disease. *Int J Geriatr Psychopharmacol* 1998;1:55–65.

99. Davis KL, Thal LJ, Gamzu ER, et al. A double-blind, placebo-controlled multicenter study of tacrine for Alzheimer's disease. *N Engl J Med* 1992;327:1253–1259.

100. Raskind MA, Peskind ER, Wessel T, et al. Galantamine in AD: A 6-month randomized, placebo-controlled trial with a 6-month extension. The Galantamine USA-1 Study Group. *Neurology*; 54:2261–2268.

101. Cummings JL. Cholinesterase inhibitors: a new class of psychotropic compounds. *Am J Psychiatry* 2000;157:4–15.

102. Molchan SE, Vitiello B, Minichiello M, et al. Reciprocal changes in psychosis and mood after physostigmine in a patient with Alzheimer's disease [Letter]. *Arch Gen Psychiatry* 1991;48:1113.

103. Cummings JL, Gorman DG, Shapira J. Physostigmine ameliorates the delusions of Alzheimer's disease. *Biol Psychiatry* 1993;33:536–541.

104. Raskind MA, Sadowsky CH, Sigmund WR, et al. Effect of tacrine on language, praxis, and noncognitive behavioral problems in Alzheimer's disease. *Arch Neurol* 1997;54:836–840.

105. Tariot PN, Solomon PR, Morris JC, et al. A 5-month, randomized, placebo-controlled trial of galantamine in AD. The Galantamine USA-10 Study Group. *Neurology* 54:2269–2276.

106. Gauthier S, Feldman H, Hecker J, et al. Donepezil improves neuropsychiatric symptoms in moderate to severe Alzheimer's disease. Presented at the annual meeting of the American Psychiatry Association, May 13–28, 2000, Chicago.

107. Raskind MA, Cyrus PA, Ruzicka BB, et al. Metrifonate study group: the effects of metrifonate on the cognitive, behavioral, and functional performance of Alzheimer's disease patients. *J Clin Psychiatry* 1999;60:318–325.

108. Del Ser T. Efficacy of cholinesterase inhibitors in Lewy body dementia. Presented at the sixth international Stockholm/Springfield symposium on advances in Alzheimer therapy; April 5–8, 2000, Stockholm, Sweden.

109. Bodick NC, Offen WW, Levey AI, et al. Effects of xanomeline, a selective receptor agonist, on cognitive function and behavioral symptoms in Alzheimer's disease. *Arch Neurol* 1997;4:465–473.

110. Elrod R, Peskind ER, DiGiacomo L, et al. Effects of Alzheimer's disease severity on cerebrospinal fluid norepinephrine. *Am J Psychiatry* 1997;154:25–30.

111. Raskind MA, Peskind ER, Holmes C, et al. Patterns of cerebrospinal fluid catechols support increased central noradrenergic responsiveness in aging and Alzheimer's disease. *Biol Psychiatry* 1999;46:756–765.

112. Zubenko GS. Biological correlates of clinical heterogeneity in primary dementia. *Neuropsychopharmacology* 1992;6:677–694.

113. Peskind ER, Wingerson D, Murray S, et al. Effects of Alzheimer's disease and normal aging on cerebrospinal fluid norepinephrine responses to yohimbine and clonidine. *Arch Gen Psychiatry* 1995;52:774–782.

114. Shankle WR, Neilson KA, Cotman CW. Low dose propranolol reduces aggression and agitation in elderly demented patients. *Alzheimer Dis Assoc Dis* 1995;9:233–237.

115. Peskind ER, Tsuang D, Kiyasu E, et al. A pilot trial of propranolol for disruptive agitation in Alzheimer's disease. Presented at the American College of Psychopharmacology annual meeting, Acapulco, Mexico, December 1999.

89

Cost-Effectiveness of Therapeutics for Alzheimer Disease

Andrea Manca

Linda Davies

Alistair Burns

Andrea Manca: Centre for Health Economics, University of York, York, United Kingdom.

Linda Davies: School of Psychiatry and Behavioural Sciences, University of Manchester, Manchester, United Kingdom.

Alistair Burns: Department of Old Age Psychiatry, Withington Hospital, West Disbury, Manchester, United Kingdom.

Alzheimer disease (AD), the most common form of dementia, is a progressive neuropsychiatric condition, the expression of which is manifested by neuropsychological deficits (aphasia, apraxia, agnosia, and amnesia), neuropsychiatric signs and symptoms (depression, delusions, hallucinations, aggression, and wandering), and problems with self-care (activities of daily living) (1 ,2 and 3). Caring for a person with dementia places a huge strain on both formal (paid, professional) and informal careers (4).

Alzheimer disease is associated with significant and excess morbidity and mortality. Approximately 30% of elderly people with dementia are severely disabled and require intensive or specialized care and support (2). Studies also indicate that 50% of an incident cohort with dementia will be severely disabled within 3 years, and up to 70% within 7 years (5).

The average survival of people with AD has been estimated at 3 to 6 years from diagnosis, and 7 to 9 years from onset of symptoms (6). The length of survival depends on the age at diagnosis, comorbid conditions, setting of care, family situation, and gender (5 ,6). Gray and Fenn (7) estimated that AD accounts for 2.5% to 5% of all life-years lost between the ages of 60 and 95 years.

The demographic trend toward an aging population means that the burden of the condition will increase in the next 25 years. Population estimates suggest that the expected number of people with AD will rise from less than half a million in 1999 to more than 600,000 in 2020 in the United Kingdom (8). Similar increases are predicted in Canada, from 161,000 people in 1991 to 314,000 people in 2011 and 509,000 in 2031 (9), and in the United States, where the number of people with AD was 2.9 to 4.8 million in 1994 and is expected to increase to 9 million by 2040.

Data on the prevalence of dementia and AD are shown in Table 89.1 . The prevalence and incidence of AD increase with age. The prevalence of the disease broadly doubles for every 5 years of age, increasing from less than 1% of the population ages 65 to 69 years to between 10% and 40% of people ages 85 years and over. The age-specific incidence rates of AD are between 51 and 161 cases per 100,000 person-years for ages 65 to 69; they increase to between 1,000 and 2,855 cases per 100,000 person-years for ages 80 to 84 and to between 1,456 and 5,420 cases per 100,000 person-years for ages 85 and over.

TABLE 89.1. AGE-SPECIFIC PREVALENCE OF DEMENTIA AND ALZHEIMER DISEASE (%)

	Eurodem (10)		Framingham (11)		Jorm et al. (12)	
Age	Dementia	AD	Dementia	AD	Dementia	AD
65–69	1.4	0.34	0.9	0.4	1.4	na
70–74	4.1	3.2	2.0	1.1	2.8	na
75–79	5.7		4.3	3.3	5.6	na
80–84	13.0	10.8	8.9	6.9	10.5	na8
85+	21.6–34.7		16.3	12.6	20.8–38.6	na
	Europe (5,6)		**United States (5,6)**		**Canada (6,9)**	
Age	Dementia	AD	Dementia	AD	Dementia	AD
65–69	0.9–1.4	0.3–1	0.8–0.9	0.2–0.8	2.4	1.0
70–74	2.1–4.1	1.1–2.5	1.3–2.0	0.4–1.2		
75–79	4.6–14.6	2.3–8.2	3.6–6.3	2.1–3.7	11.1	4–6.9
80–84	9.6–2.7	4.5–10.6	8.9–12.7	5.1–8.2		
85+	9.6–16.9			8.2–47.2	34.5	10.5–26
85–89	20.4–38.3			16.3–29.7		
90–94	28.3–57.3			40.4–74.3		
95+	42.3–55.8			58.6		

AD, Alzheimer disease.

CURRENT TREATMENT OF ALZHEIMER'S DISEASE

Part of "89 - Cost-Effectiveness of Therapeutics for Alzheimer Disease "

Two theoretically distinct treatment options are available for the treatment of AD. *Symptomatic* treatments are aimed at increasing acetylcholine levels without the expectation that they will affect the underlying course of the disease. *Stabilization* treatments are directed at altering the underlying disorder (characterized by the deposition of amyloid and the presence of neurofibrillary tangles and abnormally phosphorylated tau protein); they do not necessarily produce symptomatic improvement but may delay the progression of the disorder.

Symptomatic Treatment

The most successful agents to provide *symptomatic* improvement are the acetylcholinesterase drugs. AD is associated with a number of neurologic and neurochemical abnormalities, particularly depletion of acetylcholine. Acetylcholinesterase

inhibitors (or anticholinesterases) enhance surviving cholinergic neurotransmission by inhibiting the breakdown of released acetylcholine.

Two first-generation anticholinesterases are physostigmine and tetrahydroaminoacradine (tacrine). Early trials suggested that physostigmine has short-term efficacy in improving memory. However, the results of trials vary substantially. In addition, the drug is associated with a high incidence of side effects (13). Tacrine requires a complex dosing regime and has toxic side effects (3). Systematic reviews suggest that the drug has a modest but not clinically or statistically significant effect on cognition in people with mild to moderate AD (14 ,15). Because administration of the drug has been accompanied by a high rate of adverse events (especially hepatotoxicity), rates of withdrawal from trials have been high.

Second-generation anticholinesterases include donepezil hydrochloride (Aricept, Pfizer) and rivastigmine (Exelon, Novartis).

Clinical trials of rivastigmine show a magnitude of effect similar to that of donepezil in a larger patient database with a longer duration of treatment and less restrictive entry criteria (3 ,16). Overall, patients taking the drug show an improvement in cognition, global clinical state, and carer ratings of activities of daily living.

Drugs that have a *stabilization* effect on the progress of AD are nonsteroidal antiinflammatory agents (NSAIDs), estrogen, and antioxidant agents. There is good evidence of an inflammatory component in AD, and it is well documented that NSAIDs are protective against the development of AD. The evidence of a symptomatic improvement in patients taking these drugs is inconsistent. The beneficial effects of aspirin on cardiovascular and cerebrovascular disease are well documented, but no evidence has been found that it is effective in AD. Estrogen appears to have a significant protective effect against the development of AD and may work through a number of different mechanisms. Small-scale studies have shown minor benefits in terms of improved cognitive function. A growing body of evidence indicates that free radical formation is a mediator of the excessive lipid peroxidation and cell damage seen in AD. Antioxidant vitamins (e.g., vitamins C and E) have been shown to have biological activity in acting as scavengers for free radicals. Sano et al. (17) showed that α-tocopherol in a daily dose of 2000 IU significantly delays the onset of defined milestones in the development of AD. Therefore, of the three stabilization agents currently available, only vitamin E is supported by evidence that it can delay deterioration in the disease. α-Tocopherol has the advantage that it is not toxic (even at high doses), and it is easily available and suitable for all patients. Sano et al. (17) reported no significant side effects of α—tocopherol. Estrogen has significant potential to cause gynecologic cancer and currently is suitable only for women, and antiinflammatory agents can provoke gastric inflammation and bleeding.

ECONOMIC PERSPECTIVE

Part of "89 - Cost-Effectiveness of Therapeutics for Alzheimer Disease "

Given the constraints on health and social care budgets, those responsible for the provision and financing of such

services need to ensure that resources are used efficiently. Economic evidence about the relative costs and outcomes of health and social care helps decision makers determine the best use of scarce health care resources (18 ,19). Two approaches have been used to evaluate the economic consequences of AD. These are (a) cost-of-illness or burden-of-disease studies to assess the costs and consequences of a disease to society and (b) full economic evaluations to compare the costs and consequences of alternative health or social care interventions.

Cost-of-Illness Analyses

From an economic perspective, the aim of cost-of-illness or burden-of-disease studies is to describe and value the costs and consequences of a disease to society. A cost-of-illness study should describe and value the direct costs of health and social care for people with the disease. It should also describe the mortality and morbidity consequences. These should be valued in either monetary terms (as indirect and intangible costs) or by utility-based measures such as quality-adjusted life-years (QALYs).

Cost-of-illness studies can be used to estimate the total burden of disease for a given year. In this approach, known as the *prevalence approach*, the costs and health outcomes of all people with the disease in a given year are included. An alternative approach is to estimate the lifetime costs and consequences for a cohort of people with the disease, from onset to death. This is known as the *incidence approach*.

Cost-of-illness studies do not provide information about the economic benefits of introducing or developing new health or social care programs, and they are of limited use in setting priorities or allocating resources (18 ,19).

Full Economic Evaluation

Economic evaluations compare alternative health or social care interventions and estimates of the relative or incremental costs and benefits of care. In AD, two or more drug therapies to treat the symptoms of the disease or delay progression may be compared for efficacy. The four types of economic evaluation are cost-minimization analysis, cost-effectiveness analysis, cost–utility analysis, and cost–benefit analysis. The analytic framework chosen depends on the perspective of the analysis and the economic questions posed (19).

Cost-Minimization Analysis

In a cost-minimization analysis, the direct costs of two or more health care interventions are compared. This form of analysis does not include a formal economic comparison of the outcomes of health and social care. However, the evidence that patient outcomes do not differ between interventions must be clear and reliable. If such evidence is not available, then the economic evaluation must include a cost-effectiveness, cost–utility, or cost–benefit analysis of patient outcomes.

Cost-Effectiveness Analysis

A cost-effectiveness analysis compares the direct costs of health and social care resources of two or more interventions with patient outcomes, measured in terms of clinical effectiveness. For AD, measures such as years of life with mild or moderate disability or changes in cognitive function are used.

If one intervention, such as a new drug to control symptoms or delay progression, leads to lower direct costs and improved patient outcomes, it is the dominant and preferred option. In other words, it clearly saves resources to provide care and is more beneficial to the patient. More often, a new therapy is associated with improved patient outcomes at additional cost. Incremental cost effectiveness ratios (ICER) provide a measure of the cost of gaining a unit of health improvement, such as cost per life-year gained. The ICER is calculated as follows: (Cost of A - Cost of B)/(Outcome of A - Outcome of B).

Cost-effectiveness analysis is limited by the use of effectiveness measures, which may not capture the total impact of health and social care on quality of life or overall well-being. This is particularly important in AD, in which the impact of the disease and patient care is multidimensional. In this case, an outcome measure is needed that combines several aspects, such as survival and cognitive, physical, and emotional function, into a single index.

Cost–Utility Analysis

Cost–utility analysis is similar to cost-effectiveness analysis, but utility is used as the outcome measure. Cost–utility analysis is used to estimate QALYs. As in cost-effectiveness analysis, incremental cost–utility ratios are calculated to estimate the cost of producing one additional QALY.

Cost–Benefit Analysis

A cost–benefit analysis is based on monetary valuations of the morbidity and mortality consequences of disease or interventions. These allow an estimation of the absolute and relative net social benefit of intervention. This is calculated as the monetary value of the consequences of an intervention minus the direct costs. Any health or social care intervention with a net social benefit greater than zero (i.e., the benefits are greater than the costs) is worth undertaking.

KEY COMPONENTS OF AN ECONOMIC EVALUATION

Part of "89 - Cost-Effectiveness of Therapeutics for Alzheimer Disease "

Perspective of Analysis

Economic studies should consider all costs and outcomes that are a consequence of the illness (cost of illness) or the

health or social care interventions evaluated (economic evaluation). For AD, these may include the costs of hospital care, community-based health care services, social welfare services, and care provided by voluntary agencies or family and friends. People with AD and their families may also receive social welfare or support payments. However, what constitutes a cost from one point of view may not be a cost from another.

For example, the costs of social care services or patient and family out-of-pocket expenses are a cost to society but not to those responsible for provision or funding of hospital based care. In contrast, social welfare payments are a cost to the agency that pays them, but a benefit to the patients and families who receive them. From the point of view of society, social welfare payments are both a cost and a benefit; when added together, they cancel each other out, so they should not be included.

For these reasons, an economic analysis should be clear about the viewpoint or perspective and therefore the range of costs and consequences included. Ideally, a broad perspective that reflects the costs and outcomes to society should be adopted. At a minimum, the perspective of the analysis should include the costs and outcomes to key health and social care providers or funders and to patients and their families.

Time Frame of Analysis

Economic studies should use a time frame that allows full measurement of the relevant costs and benefits. Comparative economic evaluations should monitor resource use, costs, and outcomes for the full period during which the interventions could be expected to have an effect on resource use, survival, and health-related quality of life.

Target Population and Comparators

The population considered in the analysis should be representative of the population to be treated. The interventions compared should be relevant to the health and social care choices faced by decision makers. Unless “do nothing” is a valid management strategy, comparison of a new intervention with placebo is not appropriate for an economic evaluation.

Opportunity Cost

The economic concept of cost is the value of a good or service in terms of its best alternative use, or opportunity cost. Often, the market price or value of the resources used, such as the time of a health care professional, facilities, or medicines, is a reasonable approximation of the opportunity cost or value to society of the services provided.

Measurement and Valuation of Costs

An economic study should describe and quantify the resources used to produce health and social care and support for the patients and their carers. Costs should be estimated from data on the quantity and type of resources used (e.g., number of hospital-based physician visits, number of hospital admissions, number of days per admission) multiplied by the opportunity cost or market price of those resources. If the evaluation compares two or more interventions, care must be taken to ensure that all relevant types of resource use and costs are identified. These include costs of the intervention, follow-up care and support for patients and carers, and management of side effects or adverse events.

These aspects are termed the *direct costs* to produce of health and social care. From a societal perspective, direct costs also include out-of-pocket expenses and the use of resources that do not have a market price, such as the time of family or volunteers. These should be measured and valued because they are potentially important inputs to the production of care. The time costs of volunteers and family members can be valued with average wage rates or the cost of equivalent services with a market price (e.g., private nurses).

Measurement and Valuation of Outcomes

It is crucial that an economic study include the health-related consequences of morbidity and mortality. For AD, these could be the number of years of life lost and the illness-associated reductions in health status and quality of remaining years of life for both patients and informal carers.

These consequences should also be valued to reflect the cost or loss of utility to individuals and society of reductions in the length of life or health. Two approaches have been advocated. The first is to value the consequences in monetary terms as indirect or productivity costs *and* intangible costs. The second is to combine data about length of life and morbidity to provide a single, nonmonetary measure of impact.

Monetary Valuation

Indirect costs represent the value of changes in the amount or type of work done or use of leisure time as a consequence of morbidity or mortality. They are also called *productivity* or *time costs* (18 ,19). With AD, the ability to engage in the normal daily activities of life and leisure is reduced by impaired cognitive function and, in some cases, early death. The physical and mental health of carers may also be affected. Typically, these costs are valued in the same way as the time costs of unpaid carers, by using market values of the time in full health lost, such as average wage rates. However, indirect or productivity costs do not include the costs of patient or carer time used to provide health and social care.

Intangible costs represent the monetary value to individuals and society of health and life *per se*. In practical terms, a determination of intangible costs requires an assessment of the amount of money that individuals would accept as compensation for reductions in health or life expectancy, or the amount they would be prepared to pay for improvements in health or life expectancy.

Nonmonetary Valuation

An alternative approach is to estimate individual and social preferences for life, health, or disability states. This approach combines measures of life-years lost because of early mortality with a value for the morbidity or ill health associated with the remaining years of life. Examples are qualityadjusted life-years (QALYs) and disability-adjusted life-years (DALYs). These are calculated as the number of years of remaining life weighted by the quality or utility of that life. The utility weight is the relative value of society for states less than full health.

Discounting

The costs and consequences of a disease and health and social interventions can occur at different times. For analyses that include a time frame of more than 1 year, it is conventional to discount the costs and outcomes to present values, so that the relative importance of events occurring in the future, rather than the present, is reduced. Discounting is based on the assumption that individuals and society prefer to receive benefits sooner rather than later and to delay costs. There is some debate about whether outcomes and costs should be discounted at the same rate. The rule of thumb is to use a discount rate of 5% for both and repeat the analysis with alternative rates for the costs and outcomes.

HOW SHOULD ECONOMIC DATA BE COLLECTED AND ANALYZED?

Part of "89 - Cost-Effectiveness of Therapeutics for Alzheimer Disease "

To be useful to those concerned with choices in the allocation of health and social care resources, the design of economic evaluations should ensure that the results are timely, relevant, credible, and accurate (20). The economic study can use modeling techniques to synthesize secondary and primary data from several sources, or it can analyze data collected prospectively with a controlled study design. Which of these techniques is used depends on the type of question addressed.

The first type of question assesses the available evidence about the relative costs and outcomes of current and new forms of health and social care. The existing literature and data should be reviewed to determine the following: natural history of the disease; incidence and prevalence of the disease; possible indications and target populations for the new intervention; current treatment patterns; relevant comparators; and the costs and benefits of current treatment or health care. The initial assessment should be based on a synthesis of available data and expert opinion, which requires the development of internally and externally valid and logical models that are consistent and robust. If the quality or completeness of existing data is doubtful, sensitivity analysis should be used to generate minimum and maximum values for key clinical and economic parameters. Best and worst case scenarios should be incorporated to ensure that interactions between key parameters are explored.

If the modeling study indicates that clinical or economic evidence is highly uncertain, the prospective collection of data is required. The objective is to establish whether differences in clinical and economic endpoints are directly attributable to the interventions compared. To this end, well-controlled evaluations with a high level of internal validity are required, such as an integrated economic and clinical controlled trial. Whether randomized, controlled trial methodology or alternative study designs are used depends on the feasibility and relative efficiency of conducting a large pragmatic trial, which is typical of routine practice on a representative sample of patients. If the correlation between resource use and the interventions studied is high, even tightly defined explanatory clinical trials may be appropriate to address the question of efficiency. Alternatively, if the correlation is low and other factors, such as patient characteristics, comorbidities, and organization of health care services, are equally important determinants of service use, then the most pragmatic trial may fail to provide usable economic information.

The costs of the interventions studied should be estimated from activity data, which quantify resources used, and price or unit cost data. All health and social care activity data are potentially important and should be collected, particularly if variability in the intensity of resource use between diseases, patients, or centers is likely to be large.

Costs of Alzheimer's Disease

A number of studies have evaluated the burden of AD in different countries. These have focused primarily on the direct costs of illness and so are partial analyses. The costs have been updated to 1997 figures, with the use of health and social care inflation indices, to provide a common price year for comparison. The costs were then converted to U.S. dollars by means of purchasing power parities (PPPs). The PPP is the rate of currency conversion that ensures that the price level in each country, when expressed in dollars, is the same as that in every other country. The advantage of PPPs over conventional exchange rates is that they reflect the price levels and purchasing power of the currencies converted (21).

Table 89.2 presents data on the direct costs of health and social care for cohorts of people with AD. The variations

in costs within each of the countries and also between them are large. The average cost per person ranges between $14,107 and $50,461 in Canada, $4,747 and $156,794 in the United Kingdom, and $14,216 and $70,939 in the United States. These variations reflect differences in the methods used to collect and analyze the epidemiologic and cost data (31), the range of costs included, timing of the study, and disease severity and setting of care of the sample of people with AD included in the study.

	Mild/Moderate			Severe			All		
Source	Cost	%ICC	%CS	Cost	%ICC	%CS	Cost	%ICC	%CS
United Kingdom									
Gray and Fenn, 1993 (7)									
Community	—	—	89–94%	—	—	54%	4,747	—	—
Long-stay care	—	—	6–11%	—	—	46%	45,405	—	—
Kavanagh et al., 1993[a,b] (22)									
Private household	—	—		32,567	27%	63%	—		
Long-stay care	—	—		60,180	—	37%	—		
Kavanagh and Knapp, 1999[b] (23)									
Private household	26,442	—	77%	28,911	—	67%			—
Long-stay care	65,586	—	23%	66,157	—	33%			—
Souêtre et al., 1999[a] (24)									
Private household	98,322	69%	100%	156,794	68%	100%		—	
Holmes et al., 1998 (25)									
All settings	—			—			62,807	33%	—
Canada									
Østbye and Crosse, 1994 (26)									
Private household[a]	14,107–22,784		48–56%	27,140		59%	—	—	
Long-stay care	—			38,407		—	—	—	
Hux et al., 1998[a] (27)									
Private household	18,120	80%	67%	25,000	79%	14%	—		
Long-stay care	42,657	4%	33%	50,461	8%	86%	—		
All settings	26,780	42%		47,172	13%	—	—		
United States									
Rice et al., 1993[a] (28)									
Community	53,283	81%	—	70,939	69%	—	63,418	73%	—
Long-stay care	50,819	8%	—	64,929	12%	—	64,102	12%	—
Leon et al., 1998[a] (29)									
Community	14,216–23,005	56–60%	47–78%	27,817	50%	27%	19,015	56%	52%
Long-stay care	34,864–37,675	2%	22–53%	40,363	2%	73%	38,424	2%	48%
All settings	18,826–30,780	22–34%	—	36,953	11%	—	28,308	21%	
Ernst et al., 1997 (30)	—					—	37,870–43,604		

CS, cost setting; ICC, informal care cost.
[a]Cost of care with Alzheimer disease/cognitive disability minus the cost of no cognitive disability.
[b]Estimated.

TABLE 89.2. CARE SETTING, INFORMAL CARE, AND COST PER PERSON (U.S. DOLLARS, 1997)

Sources of Data

Data for the cost estimates shown in Table 89.2 and Table 89.3 were derived from a number of sources. One study used

secondary analysis of administrative databases in the United Kingdom (7 ,33). These gave aggregate measures of the use of hospital inpatient care, residential and nursing home admissions, and general practitioner consultations for people with mental disorders in 1984 and 1985. They were supplemented by surveys and expert opinion to generate measures of the use of day care, home-based care, and payments to informal carers. Cost data from other sources were used to estimate total expenditure for people with AD (7 ,33). The primary disadvantages of this approach are that the data for resource use may not be detailed enough to allow a complete measurement of the range of resources used. In addition, coding errors and highly summarized diagnostic categories and comorbidities make it difficult to allocate resource use to specific diseases. These problems may account for the substantially lower community-based costs found in this study.

	Annual Incremental Cost of Care Per Person[a]			
Source	No Cognitive Disability	Mild/Moderate	Severe	All
United Kingdom				
Kavanagh and Knapp, 1999[b] (23)				
Total cost	25,299	35,341	42,886	—
Incremental cost	—	10,042	17,587	—
Souêtre et al., 1999 (24)				
Total cost	4,317	76,315–118,233	156,794	—
Incremental cost	—	71,998–113,916	152,477	—
Canada				
Østbye and Crosse, 1994 (26)				
Incremental cost	—	—	—	18,998
United States				
Ernst and Hay, 1994[a] (32)				
Incremental cost	—	—	—	42,058

[a]Cost of care with Alzheimer disease/cognitive disability minus the cost of no cognitive disability.
[b]Estimated from sample of people with physical or cognitive disabilities.

TABLE 89.3. ESTIMATED INCREMENTAL COSTS OF ALZHEIMER DISEASE (U.S. DOLLARS, 1997)

Four of the costs studies were based on detailed data from large-scale national surveys (22 ,23 ,26 ,27). The U.K. studies both used two surveys of people with disabilities living in private households (*n* = 5,699) and communal establishments (*n* = 3,037) that had been conducted in the middle to late 1980s (22 ,23). The Canadian studies both used the Canadian Study of Health and Ageing, which surveyed a total of 10,263 randomly selected Canadians over the age of 65 (26 ,27). Resource use information was collected by interviewing people with disability and their carers.

The remaining studies in Table 89.2 and Table 89.3 used surveys of selected samples of people. The samples of respondents varied in size from 64 to 679 people. Similar methods were used to identify resource use. Two of the studies collected data prospectively from respondents during 6-month (30) and 2-year periods (28). Most studies included validated measures to determine the presence and severity of cognitive disability.

Estimating resource use from individual patient or carer data is associated with several advantages and disadvantages. First, the use of screening instruments allows a clear identification of people with cognitive disability or AD. However, the use of a variety of instruments may lead to differences in the categorization of people with cognitive problems, so that the comparability of results is reduced. For example, the surveys in the United Kingdom used a broad classification of cognitive disability that resulted in a higher prevalence estimate than that obtained in other epidemiologic research (22).

Secondly, the use of individual patient or carer data allows the health and social care services actually used by the people in the study to be identified. It also allows the use of informal care time and personal expenditures to be measured. However, the use of interviews and questionnaires to determine resource use may be subject to problems with accurate recall or recording the type and quantity of services actually used. This problem is illustrated by the greater range of costs found for people living at home, where a greater range of services may be used, than for people in long-stay care. The highest cost estimates for people living at home were 6 to 7 times greater than the lowest, whereas they were 1.7 to 1.9 times higher for people in long-stay care.

Thirdly, if the study is based on large samples, the impact of random variation in the use of health and social care services is reduced. If a small sample by chance overrepresents people with an extremely high or low use of services, substantial underestimation or overestimation of the costs

can result. For example, the range of costs found in the large U.K. and Canadian studies was relatively small, with the highest estimate between 1.3 and 1.9 times greater than the lowest estimate. However, the range was far greater for the smaller studies. In particular, for people living at home, the highest costs were 5.6 to 6.9 times greater than the lowest costs.

Measurement and Valuation of Informal Care

A key difference between the studies was the use of unpaid care and the method utilized to cost this care. One study did not include the costs of unpaid informal care time (23). When included, informal care costs ranged from 48% to 81% of home-based care for people with mild to moderately severe disease and from 27% to 79% of home-based care for people with severe disease. For people in long-stay care, the proportion of informal care costs ranged from zero to 12%. Excluding the costs of unpaid care does not reduce the variability in the total costs of care. This suggests that in some cases, unpaid care may be a substitute for rather than an addition to formal health and social care services.

Kavanagh et al. (22) used U.K. social security allowances for carers of people with disabilities as a proxy for the costs of informal care. This gave the lowest proportion of informal care costs for people living at home (27%). Three studies used the replacement cost method to estimate the opportunity cost of unpaid carer time (26 ,28 ,29). The time spent by informal carers was estimated and then multiplied by the average wage of professional (paid) caregivers. The informal care costs ranged between $7,900 (50% of costs) (26) and $48,948 (81% of costs) (28).

The other studies used the national average wage (24) or minimum wage (27) to value unpaid carer time. The informal care costs ranged from $14,496 (80% of costs) (27) to $106,620 (69% of costs) (24). These data illustrate the impact of different methods of valuing informal care time on estimates of cost.

Organization and Availability of Care

Rice et al. (28) estimated the expected cost per person with AD. This was based on data for a sample of people living in both community and residential care in northern California. The estimated annual costs per person were higher than the estimates from the other studies reported in Table 89.2 and Table 89.3 . However, the authors stressed that the organization and level of provision of services in California were such that the results should not be generalized to other geographic settings in the United States.

The organization and availability of health and social care vary with time as well as between settings. In many countries, the trend has been toward the provision of mental health and social care services in the community rather than in institutions, with an emphasis on support from family and informal carers (34). This means that earlier studies may overestimate the current costs of institutional care and the potential benefits of reducing the need for such care. At the same time, the costs of community or home-based care and informal carers will be underestimated. This is particularly important if the opportunity costs of informal care are not included or are underestimated in the direct costs of providing health and social care.

Severity of Disease and Setting of Care

The costs of health and social care for people with AD are also affected by two interrelated factors: the severity of disability caused by the disease and the setting of care.

Severity of Disease

The cost data in Table 89.2 indicate a trend toward higher costs of care as the severity of the disease increases. This applies in both community/private home settings and long-stay care settings. Three studies used statistical analysis to compare costs by disease severity. All of them found some degree of statistically significant correlation between some of the costs and disease severity (23 ,27 ,29). Holmes et al (25) used regression analysis to estimate average cost by age and years since diagnosis. In their analysis, total costs were positively related to years since diagnosis. Each additional year since diagnosis was predicted to increase costs by roughly $1,100 per person. However, cost was negatively related to age. Each additional year of age predicted a decrease in costs of about $850. The authors suggested that this finding may have reflected more intensive hospital-based care for younger people with AD.

It has also been suggested that as the cognitive and functional ability of people declines, they can no longer live alone supported by formal health and social care services. They may move to live with family or friends, who provide informal care. In this case, informal care may be a substitute for previous formal care services. If informal care is not adequately costed, then the financial cost decreases, but not the opportunity cost of providing care (35). In contrast, in a secondary analysis of a large-scale disability survey in the United Kingdom, disabled elderly people with more severe cognitive disability received more intensive care and were referred more often to health care services (23).

In an analysis of Canadian data (27) in which a bivariate regression model was used to assess the relationship between severity of disease and costs, each 1-point decrease in Mini-Mental State Examination (MMSE) scores was associated with an average increase in costs of $1,343 (Canadian dollars, 1996 prices). Even when informal care time was valued by using industrial aggregate wage levels rather than minimum wage levels, the relationship between severity and costs remained statistically significant. The relationship between

costs of care and severity of disease is complex. Increases in the costs of care as disease becomes more severe represent in part a greater use of institution-based care as people become more cognitively and functionally disabled by their disease. They may also reflect aging and the effect of comorbidities (23). In addition, informal carers age and may be affected by declining health and less ability to provide care.

Setting of Care

An important determinant of the costs of health and social care is the distribution of people with AD by setting of care. For most of the studies reported in Table 89.2 , the costs of long-stay care are 1.5 to 2.5 times higher than those of home-based care. The exceptions to this are the study by Gray and Fenn (7), in which the costs of long-stay care were 10 times higher than those of community care. The study by Rice et al. (28) indicated roughly equal costs for long-stay and home-based care.

Determinants of the setting of care include the severity of cognitive and functional disability, the presence of other health problems, the ability of informal carers to support the person at home, and the structure of health and social care service provisions. The data in Table 89.2 indicate that the proportion of people cared for in long-stay care settings is between 6% and 53% for people with mild to moderate disease and 33% to 86% for people with severe disease.

Kavanagh and Knapp (23) found cost variations between long-stay care locations in their sample, and they also found the prevalence of severe cognitive disability to be higher in the more expensive settings. In the regression analysis by Holmes et al. (25), the use of institutional care increased with the number of years since diagnosis and the age of the carer. For each additional year of age of the carer, the costs of institutional care were predicted to increase by roughly $264 per year.

Incremental Costs of Alzheimer Disease

Table 89.3 presents estimates of the additional or incremental costs of care associated with AD only, rather than with other illnesses or age. With the exception of those in the study by Souêtre et al. (24), the incremental costs tend to be lower than the full costs reported in Table 89.2 . In the study by Kavanagh and Knapp (23), the costs of people with disabilities, but no cognitive disability, were approximately $25,299. The additional costs of people with cognitive disability in this group can be estimated at $10,042 to $17,587. The studies by Souêtre et al. (24) and Østbye and Crosse (26) estimated the costs of care for people in a similar age group with no disabilities. The study by Ernst and Hay (32) estimated the net costs of care from aggregate data sets and surveys.

Evidence of Cost-Effectiveness

It is clear that the costs of health and social care and informal care for people with AD is high, and evidence suggest that the costs increase with the severity of cognitive disability and need for long-stay care or institutional care. It has been argued that it might be rational to support the introduction of drug treatments to slow down the progression of the disease and delay the onset of institutionalization. This would lead to a saving of costs to offset the acquisition costs of the drugs. However, this proposal has been criticized on the grounds that it would shift the burden of the disease from the public sector budget to private citizens, without a real beneficial effect for society as a whole. A full evaluation from the societal viewpoint of the new drugs used to manage AD is clearly needed (36). A number of economic studies have been published to assess the relative value for money of tacrine, donepezil, and rivastigmine. Table 89.4 gives details of the methods used and comparators of the studies.

TABLE 89.4. ECONOMIC MODELS OF DRUGS FOR ALZHEIMER DISEASE

Study	Country	Drug	Method	Perspective	Costs Included	Discount Rate	Comparators	Time Horizon
Stewart et al., 1998 (37)	United Kingdom	Donepezil	Cost-effectiveness analysis	Societal	Direct costs and costs of informal care	6%	5-mg and 10-mg dose/d vs. placebo	5 y (6-mo cycles)
O'Brien et al., 1999 (38)	Canada	Donepezil	Cost-effectiveness analysis	Societal and government payer	NHC, CCS, CG, drug costs	5%	5-mg dose/d vs. no treatment	5 y (6-mo cycles)
Jönsson et al., 1999 (40)	Sweden	Donepezil	Cost-effectiveness analysis	Not clearly stated	Residential costs, home help, and drug cost	3%	5-mg and 10-mg dose/d vs. no treatment	5 y (6-mo cycles)
Neumann et al., 1999 (39)	United States	Donepezil	Cost-effectiveness analysis	Societal	Direct medical and nonmedical costs; costs of informal care	3%	pooled 5-mg and 10-mg doses/d	1 y (6-mo cycles)
Stein 1997 (41)	United Kingdom	Donopezil	Drug evaluation, NHS report		Acute/geriatric hospital in- and out-patient care, mental hospital care, GP consultations, drug costs	6%	5-mg and 10-mg doses/d and placebo	2, 5, 8, and 10 y
Fenn and Gray, 1999 (33)	United Kingdom	Rivastigmine	Cost-saving analysis	Health and social care system	Long-term care institutions costs, drug costs, costs occurring when living at home	Not stated	1 to 4-mg and 6 to 12-mg doses/d vs. placebo	26 wk, 1 yr, and 2 y
Stein, 1998 (42)	United Kingdom	Rivastigmine	Drug evaluation, NHS report		Acute/geriatric hospital in- and out-patient care, mental hospital care, GP consultations, drug costs	6%	1 to 4-mg and 6 to 12-mg doses/d and placebo	1, 2, and 5 y
Small et al., 1998 (43)	United States	Donepezil	Longitudinal survey	Not clearly stated	Direct medical costs and drug costs	n. a.	drug vs. no drug treatment	6 mo
Henke and Burchmore 1997 (44)	United States	Tacrine	Cost-minimization analysis	Public and private payers	All paid medical and societal services	5%	tacrine 80 mg/d vs. no drug treatment	Lifetime cost for patients newly diagnosed with AD
Lubeck et al., 1994 (45)	United States	Tacrine	Cost-minimization analysis	Not clearly stated	Community and nursing home care, drug costs, treatment-related costs, informal care costs (?)	5%	tacrine 160 mg vs. tacrine (various doses)	Lifetime costs
Wimo et al., 1997 (46)	Sweden	Tacrine	Cost-saving analysis	Public payer	Residential costs, drug costs, laboratory and GP visits costs, costs of diagnostic procedures, cost of informal care	3%	tacrine 160 mg vs. no tacrine treatment (i.e., standard treatment)	Lifetime costs
Wimo et al., 1998 (47)	Sweden	Propentofylline	Cost-saving analysis	Public payer	Direct medical costs	5%	propentofylline vs. usual care without propentofylline	Lifetime costs
Hauber et al., 2000 (48)	United States	Rivastigmine	Cost-saving analysis	Not stated	Not stated	3%	Rivastigmine vs. no treatment	6 mo, 1 yr, and 2 y

AD, Alzheimer disease; CCS, community care services; CG, caregiver's time; GP, general practitioner; n.a., not applicable; NHC, nursing home care; NHS, national health service.

Most of these analyses are modeling exercises that extrapolate the results of randomized, controlled trials to a longer time period and broader population. Only four meet the criteria for full economic evaluations, and these are shown in Table 89.5 , which reports the estimates of the expected net benefits and costs likely to occur from the introduction of these drugs. The four studies analyzed the same drug treatment (donepezil) in four different countries/settings: United Kingdom (37), Canada (38), United States (39), and Sweden (40). Stewart et al. (37) evaluated the cost-effectiveness of donepezil in the United Kingdom for individual patients ages 75 years and over with a diagnosis of either mild or moderate AD. O'Brien et al. (38) considered a hypothetical cohort of people with nonsevere AD (MMSE ≥ 10) in Canada. Finally, Neumann et al. (39) focused on patients in the United States with mild or moderate AD.

TABLE 89.5. FULL ECONOMIC EVALUATIONS OF DRUGS FOR ALZHEIMER DISEASE

Study	Outcome Measure	Year of Costing	Original Currency	Incremental Cost (Health PPP$, 1996)	Health Gain
Stewart et al., 1998 (37)	Expected life-years in condition less than severe (ELY<S)	1997	British pounds	1,333	0.120 ELY<S
O'Brien et al., 1999 (38)	Expected life-years in condition less than severe (ELY<S)	1997	Canadian dollars	−1,292	0.200 ELY<S
Jönsson et al., 1999 (40)	Expected life-years in condition less than severe (ELY<S)	1998	Swedish kronor	−1,962	0.522 ELY<S
Neumann et al., 1999 (39)	Quality-adjusted life-years (QALY)	1997	U.S. dollars	483	0.015 QALYs

PPP, purchasing power parity.

Despite differences in the provision of health care between the United Kingdom, the United States, and Canada, these studies found donepezil to be approximately cost-neutral under several alternative scenarios. Three studies found that the distribution of severity states of patients is the most important variable affecting the cost-effectiveness of drugs. However, it is important to note that these results are preliminary and uncertain and that a number of issues must be considered when the results are interpreted.

Costs

First of all, no prospective measurement of resource use associated with the drug or usual care was made. Costs were estimated from retrospective analysis of available data sets (40) or analysis of published literature (37), sometimes integrated with expert opinion (38 ,39). The range of cost items and the costing methodologies employed in each study were heterogeneous. Some authors included both direct costs and

informal carers' time (37 ,38 and 39), whereas others considered only direct medical costs (40).

Three analyses (37 ,38 ,49) directly or indirectly associated the dynamic of treatment costs with the progression of disease severity, measured with the MMSE. The MMSE score was shown to be strongly correlated with costs of dementia care, but it is unclear to what extent the use of this instrument is robust in modeling studies. It has been suggested (49) that other factors may be strongly correlated with costs, such as indices of activities of daily living and instruments that measure behavioral disturbances.

Outcome Measures

One study used QALYs to measure the benefits derived from introducing the drug (39). In the other studies, benefits were measured in terms of "time spent in condition less than severe." The QALYs were estimated with the Health Utility Index Mark II in a sample of patients and carers. However, this instrument has not been validated in patients with AD, and its ability to detect small improvements in potentially important clinical aspects is doubtful. The QALY data were collected alongside a cross-sectional study, which means that no information was obtained on how the effectiveness of the drugs and utilities varied over the course of the disease. In addition, the sample of patients used to elicit utility values may have been unrepresentative of the population of people with AD (39). There were also potential problems with the use of proxy respondents. However, given the cognitive and behavioral degenerative process associated with AD, the use of alternative respondents may be unavoidable. Additionally, measuring outcomes as "time spent in less than severe state" does not inform health and social care decision makers about the value of quality of life for people with AD and their family and carers.

Effectiveness

Effectiveness data about the new drugs were derived from a limited number of trials that were short in duration and explanatory rather than pragmatic in design. Some of these trials have been criticized elsewhere (50) for having enrolled a carefully selected subgroup of patients with mild-to-moderate AD and excluded those with coexisting illness or concurrent treatment. In real practice, the eligible population may be considerably different. Consequently, only a limited proportion of people may be adequately and safely treated.

Furthermore, the lack of data meant that the duration of the treatment effect of the drug was based on experts' opinions (38 ,39) or was disregarded by assuming that the treatment effect ceased after 6 months (37). The cost-effectiveness of cholinesterase inhibitors depends on the distribution of patients across different severity states (38). In this context, the correct assessment of the duration of the treatment effect of anticholinesterase drugs assumes a central role because it affects the number of people having mild-to-moderate AD at any one time.

Modeling

Some authors have recently challenged the use of Markov models in the evaluation of antidementia drugs (33 ,48). The use of alternative modeling tools, such as statistical models, to extrapolate the results of a short trial to a longer time horizon needs to be explored. Given the considerable uncertainty surrounding the available data, deterministic models in which simplistic sensitivity analysis techniques are used may not be adequate to assess the robustness of the results. The application of stochastic models allows the uncertainty associated with relevant parameters of a model to be incorporated and quantified.

CONCLUSION

Part of "89 - Cost-Effectiveness of Therapeutics for Alzheimer Disease "

As a direct consequence of changes in the age structure of the population, elderly generations are expected to become the largest consumer group of health care resources. Because no cure for AD is yet available, the management of the

disorder focuses on assisting patients in their daily activities and supporting carers. The progressive nature of AD and the aging of the population mean that many people with this condition will require intensive support and long-term residential or nursing home care. A number of factors may trigger the need for long-term institutional care, including the age of family carers, the behavioral problems of patients, and the loss of self-care ability for those who live alone in a private household. Institutionalization has been identified as one of the main cost drivers in the care of people affected by AD (2 ,51).

Some clinical evidence indicates that anticholinesterase drugs may slow the progression of AD or relieve some of the symptoms. If the drugs are effective in controlling symptoms or slowing progression of the illness, they may delay the need for intensive support or institutionalization of patients. The high acquisition cost of the drugs, however, has raised considerable concern about the potential value for money associated with their use, which has prompted a significant number of studies addressing the issue of costs and patient benefits.

To date, a conclusive analysis has not clarified the most appropriate management strategy for the disorder. In the near future, new drugs for the treatment of AD are expected to be licensed, and it would be extremely valuable to be able to compare them in a clear and well-defined framework. In addition, if economic evaluation is to inform health and social care providers and policy makers about the potential impact of new interventions in practice, an estimation of the value for money of these new interventions requires consideration of (a) the perceived and objective risks and benefits of care; (b) the attitudes of people with dementia, carers, and health and social care providers to risk; (c) the utility to these groups of health care interventions; and (d) quantification of the uncertainty surrounding estimates of risk, utility, and costs.

Although the first attempts to analyze currently available antidementia drugs provided limited conclusive results, the contribution of simulation models may help to shed light on several aspects that have not yet been explored. In a context largely characterized by uncertainty surrounding the value of the key variables, modeling techniques can be used to assess the value for money of new management strategies for the treatment of AD and compare them with the alternative policy options. Further primary and secondary research is required to provide robust estimates of the formal and informal care costs associated with the new drugs and the value of health improvements to patients and carers.

REFERENCES

1. Hope T, Keene J, Fairburn CG, et al. Natural history of behavioural changes and psychiatric symptoms in Alzheimer's disease. A longitudinal study. *Br J Psychiatry* 1999;174:39–44.

2. Melzer D, Hopkins S, Pencheon D, et al. In: Stevens A, Raftery J, eds. *Health care needs assessment. The epidemiologically based needs assessment reviews*. Oxford: Radcliffe Medical Press, 1994.

3. Burns A, Russell E, Page S. New drugs for Alzheimer's disease. *Br J Psychiatry* 1999;174:474–479.

4. Donaldson C, Tarrier N, Burns A. The impact of the symptoms of dementia on caregivers. *Br J Psychiatry* 1997;170:62–68.

5. Fratiglioni L. Epidemiology. In: Wimo A, Jonsson B, Karlsson G, et al., eds. *Health economics of dementia*. Chichester: John Wiley and Sons, 1998.

6. Alloul K, Sauriol L, Kennedy W, et al. Alzheimer's disease: a review of the disease, its epidemiology and economic impact. *Arch Gerontol Geriatr* 1998;27:189–221.

7. Gray A M, Fenn P. Alzheimer's disease: the burden of illness in England. *Health Trends* 1993;25:31–37.

8. Manca A, Davies L. An economic analysis of the management of Alzheimer's disease in the UK. *J Br Menopausal Soc* 1999;5:167–173.

9. CSHA Working Group. Canadian Study of Health and Aging: study methods and prevalence of dementia. *Can Med Assoc J* 1994;150:899–913.

10. Hofman A, Rocca WA, Brayne C, et al. Frequency and distribution of Alzheimer's disease in Europe: a collaborative study of 1980–1990 prevalence findings. *Ann Neurol* 1991;30:381–390.

11. Kokmen E, Beard CM, Offord KP, et al. Prevalence of medically diagnosed dementia in a defined United States population: Rochester, Minnesota. *Neurology* 1989;39:773–776.

12. Jorm AF, Korten AE, Henderson AS. The prevalence of dementia: a quantitative integration of the literature. *Acta Psychiatr Scand* 1987;76:465–479.

13. Coelho Filho JM, Birks J, Gromley Evans J. *Physostigmine for Alzheimer's disease* [protocol for a Cochrane Review]. Cochrane Library, Issue 1, 2000. Oxford: Update Software.

14. Qizilbash N, Birks J, Lopez-Arrieta J, et al. *Tacrine for Alzheimer's disease* [Cochrane Review]. Cochrane Library, Issue 1, 2000. Oxford: Update Software.

15. Glennie J. *The efficacy of tacrine and measurement of outcomes in Alzheimer's disease*. Technology overview: pharmaceuticals. CCOHTA Report 5.0, July 1997.

16. Burns A Rossor M, Hacker J et al. The effects of donepezil in Alzheimer's disease—results from a multinational trial. *Dementia Geriatr Cogn Disord* 1999;10:237–244.

17. Sano M, Ernesto C, Thomas RG, et al. A controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer's disease. *N Engl J Med* 1997;336:1216–1222.

18. Gold MR, Siegel JE, Russell LB, et al. *Cost-effectiveness in health and medicine*. New York: Oxford University Press, 1996.

19. Drummond MF, O'Brien B, Stoddart GL, et al. *Methods for the economic evaluation of health care programmes*, second ed. New York: Oxford University Press, 1997.

20. Davies LM Economic evaluation in clinical trials. *Appl Clin Trials* 1998;7:62–70.

21. Organization for Economic Cooperation and Development. OECD Health Data/ECO-Santé-OECD [Data File]. Paris: OECD, 1999.

22. Kavanagh S, Schneider J, Knapp M, et al. Elderly people with cognitive impairment: costing possible changes in the balance of care. *Health Soc Care* 1993;1:69–80.

23. Kavanagh S, Knapp M. Cognitive disability and direct care costs for elderly people. *Br J Psychiatry* 1999;174:539–546.

24. Souêtre EJ, Thwaites RM, Yeardley HL. Economic impact of Alzheimer's disease in the United Kingdom. Cost of care and disease severity for non-institutionalised patients with Alzheimer's disease. *Br J Psychiatry* 1999;174:51–55.

25. Holmes J, Pugner K, Phillips R, et al. Managing Alzheimer's disease: the cost of care per patient. *Br J Health Care Management* 1998;4:332–337.

26. Østbye T, Crosse E. Net economic cost of dementia in Canada. *Can Med Assoc J* 1994;151:1457–1464.

27. Hux MJ, O'Brien BJ, Iskedjian M, et al. Relation between severity of Alzheimer's disease and costs of caring. *Can Med Assoc J* 1998;159:457–465.

28. Rice D P, Fox P J, Max W, et al. The economic burden of Alzheimer's disease. *Health Affairs Millwood* 1993;12:164–176.

29. Leon J, Cheng CK, Neumann PJ. Alzheimer's disease care: costs and potential savings. *Health Affairs Millwood* 1998;17:206–216.

30. Ernst RL, Hay JW, Fenn C et al. Cognitive function and the costs of Alzheimer's disease: an exploratory study. *Arch Neurol* 1997;54:687–693.

31. Ernst RL, Hay JW. Economic research on Alzheimer disease: a review of the literature. *Alzheimer Dis Assoc Disord* 1997;6:135–145.

32. Ernst R L, Hay JW. The U.S. economic and social costs of Alzheimer's disease revised. *Am J Public Health* 1994;84:1261–1264.

33. Fenn P, Gray A. Estimating long-term cost savings from treatment of Alzheimer's disease. A modelling approach. *Pharmacoeconomics* 1999;16:165–174.

34. Winblad B, Hill S, Beermann B, et al. Issues in the economic evaluation of treatment for dementia. Position paper from the International Working Group on Harmonization of Dementia Drugs Guidelines. *Alzheimer Dis Assoc Disord* 1997;11:39–45.

35. Stewart A. *Cost of care for people with dementia aged 75 and over*. PSSRU discussion paper 1997;1303/2.

36. Whitehouse PJ. Pharmacoeconomics of dementia. *Alzheimer Dis Assoc Disord* 1997;11:S22–S33.

37. Stewart A, Phillips R, Dempsey G. Pharmacotherapy for people with Alzheimer's disease: a Markov-cycle evaluation of five years' therapy using donepezil. *Int J Geriatr Psychiatry* 1998;13:445–453.

38. O'Brien B J, Goeree R, Hux M, et al. Economic evaluation of donepezil for the treatment of Alzheimer's disease in Canada. *J Am Geriatr Soc* 1999;47:570–578.

39. Neumann PJ, Hermann RC, Kuntz KM, et al. The cost-effectiveness of donepezil in the treatment of mild or moderate Alzheimer's disease. *Neurology* 1999;52:1138–1145.

40. Jönsson L, Lindgren P, Wimo A, et al. The cost-effectiveness of donepezil therapy in Swedish patients with Alzheimer's disease: a Markov model. *Clin Ther* 1999;21:1320–1340.

41. Stein K. *Donepezil in the treatment of mild to moderate senile dementia of the Alzheimer type (SDAT)*. Bristol: NHS Executive South and West, 1997.

42. Stein K. *Rivastigmine (Exelon™) in the treatment of senile dementia of the Alzheimer type (SDAT)*. Bristol: NHS Executive South and West, 1998.

43. Small GW, Donohue JA, Brooks RL. An economic evaluation of donepezil in the treatment of Alzheimer's disease. *Clin Ther* 1998;20:838–850.

44. Henke C J, Burchmore M J. The economic impact of tacrine in the treatment of Alzheimer's disease. *Clin Ther* 1997;19:330–345.

45. Lubeck DP, Mazonson PD, Bowe T. Potential effects of tacrine on expenditures for Alzheimer's disease. *Med Interface* 1994;October:138.

46. Wimo A, Karlsson G, Nordberg A, et al. Treatment of Alzheimer disease with tacrine: a cost-analysis model. *Alzheimer Dis Assoc Disord* 1997;11:191–200.

47. Wimo A, Witthaus E, Rother M, et al. Economic impact of introducing propentofylline for the treatment of dementia in Sweden. *Clin Ther* 1998;20:552–566.

48. Hauber AB, Gnanasakthy A, Snyder EH, et al. Potential savings in the cost of caring for Alzheimer's disease. Treatment with rivastigmine. *Pharmacoeconomics* 2000;17:351–360.

49. Jönsson L, Lindgren P, Wimo A, et al. Costs of Mini-Mental State Examination-related cognitive impairment. *Pharmacoeconomics* 1999;16:409–416.

50. Birks JS, Melzer D. *Donepezil for mild to moderate Alzheimer's disease* [Cochrane Review]. Cochrane Library, Issue 4, 1999. Oxford: Update Software.

51. Kavanagh S, Schneider J, Knapp M, et al. Elderly people with dementia: costs effectiveness and balance of care. In: Knapp M, ed. *The economic evaluation of mental health care*. Aldershot: Arena, 1995.

90

Neuropsychiatric Manifestations of Hiv-1 Infection and Aids

Dwight L. Evans

Karen I. Mason

Dwight L. Evans: Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania.

Karen I. Mason: Department of Psychiatry and Biobehavioral Sciences, University of California at Los Angeles, Los Angeles, California.

By 1999, the United Nations estimated that more than 33 million people worldwide had been infected by HIV-1 and more than 16 million had died of AIDS. Approximately 90% of these infections occurred in developing countries, particularly in sub-Saharan Africa, which accounts for 65% of HIV cases (1). In the United States and Canada, more than 920,000 adults and children were living with HIV/AIDS by the end of 1999, and more than 400,000 had died. Initially, HIV-1 infection primarily affected men, but today, women comprise more than 16% of all AIDS cases in the United States. Among U.S. women with AIDS, the proportion of African-Americans and other minorities has been rising faster than the proportion of women in other segments of the population. By the end of 1997, African-American women accounted for 56% of AIDS cases among women in the United States. Another change in the demographics of the epidemic is seen in the mode of transmission. In the United States, homosexual transmission has stabilized or declined while infection via heterosexual transmission (5% to 18%) and injection drug use (15% to 23%) has increased significantly.

Neuropsychiatric and neurologic signs and symptoms have been described since the earliest reports on AIDS. The neuropsychiatric manifestations described in these early reports included progressive dementia, depression with pronounced apathy and psychomotor slowing, manic symptoms, and atypical psychosis. Initially, these mental disturbances were attributed to psychological reactions to a systemic illness, the effects of psychosocial stressors associated with the disease, or the consequences of opportunistic infections or neoplasms within the central nervous system (CNS). It is now recognized that the neuropsychiatric manifestations of HIV infection can result from the direct effects of HIV on the brain and CNS or from indirect effects, such as opportunistic infections or tumors associated with immunosuppression, cerebrovascular disease, systemic toxicity, and complications of antiretroviral therapy. As understanding of the broad range of neuropsychiatric manifestations of HIV infection has grown, new classification and diagnostic criteria have replaced the earlier, more inexact terms, such as HIV encephalopathy and AIDS dementia complex, terms that combined diverse cognitive, motor, and affective–behavioral complications. The following sections review neurocognitive functioning in HIV-1 infection and the neuropsychiatric manifestations of such infection and their management.

NEUROPSYCHOLOGICAL MANIFESTATIONS OF HIV-1 INFECTION

Part of "90 - Neuropsychiatric Manifestations of Hiv-1 Infection and Aids "

Neurocognitive Functioning in HIV Infection

HIV-related neurocognitive deficits are manifestations of the direct and indirect effects of HIV on the CNS and can range from subtle changes in attention and psychomotor processing to full-blown dementia (2). Postmortem neuropathologic examination of HIV-seropositive patients has implicated both cortical and subcortical structures, specifically the frontal lobes, subcortical white matter, and basal ganglia (3 ,4 and 5). The caudate nucleus and basal ganglia are primary areas of HIV pathogenesis (6 ,7).

Neuropsychological assessment has played a crucial role in the identification of patterned impairment in HIVinfected persons. It has proved useful in the diagnosis of HIV-related cognitive disturbances and is used widely to quantify changes in cognitive processes associated with treatment. A battery of neuropsychological tests designed to cover the cognitive and behavioral domains affected by

HIV-1 infection has been widely used to assess neurocognitive functioning in HIV-infected persons (8).

Although it is generally agreed that patients in the symptomatic stages of HIV infection and those in whom AIDS has developed exhibit deficits in a variety of cognitive domains (9 ,10 ,11 ,12 and 13), the question of whether significant neurocognitive deficits occur in the asymptomatic stages of infection or in those who are mildly symptomatic remains controversial. Some investigators have found neuropsychological deficits in asymptomatic HIV-seropositive persons (14 ,15 ,16 ,17 ,18 ,19 ,20 and 21 ,22). However, a large number of studies have found subtle or no differences in neuropsychological functioning between asymptomatic HIV-seropositive and HIV-seronegative subjects (23 ,24 ,25 ,26 ,27 ,28 ,29 ,30 ,31 ,32 and 33).

It has been argued that a subgroup of asymptomatic HIV-seropositive patients may show cognitive deficits (12 ,16 ,20 ,34 ,35). In other words, HIV may have a deleterious effect on brain function in certain asymptomatic HIVinfected persons. Evaluating group differences in overall impairment ratings may be the most sensitive and accurate method of assessing neurocognitive impairment in asymptomatic persons (9). In a review of 57 studies that examined neuropsychological functioning in asymptomatic persons, the median percentage of impaired test performances in asymptomatic persons was 35%, and in seronegative persons it was 12% (36). In another review of 36 cross-sectional studies and nine longitudinal studies, about half the reviewed studies indicated no significant difference in neuropsychological test performance between symptomatic and nonsymptomatic participants (37). In the majority of longitudinal studies (66.6%), no cognitive impairment in HIV-positive persons was found at baseline or at subsequent follow-up visits (37). Methodologic differences in study design and analysis probably account for discrepancies between studies.

Asymptomatic persons at risk for neurocognitive impairment have been found to display one of two patterns of deficits: (a) depression, psychomotor slowing, and diminished verbal memory or (b) lowered verbal and nonverbal cognitive functioning in the absence of a mood disturbance (38). Current research indicates that cognitive impairment is uncommon in asymptomatic HIV-seropositive persons (22 ,37), is not associated with deficits in social or occupational functioning (9 ,11), and, when present, is subtle and limited to a few cognitive domains.

It has been estimated that 55% to 86% of persons in the AIDS stages of HIV infection demonstrate significant neurocognitive deficits (9 ,16). Objective impairments in HIV disease include psychomotor slowing, forgetfulness, and difficulties with attention and concentration. In the later stages of HIV infection (symptomatic HIV and AIDS), executive skills and motor speed may also become impaired.

Attention/Concentration

HIV infection has been associated with deficits in attention (9 ,39 ,40 and 41). Specifically, HIV-infected persons show deficits in dual-task or divided-attention paradigms (41 ,42). It has been argued that cognitive slowing may be at the root of the attentional deficits seen in many symptomatic persons (41). Additional work is needed to examine whether other components of attention (e.g., switching, engagement/disengagement, spatial attention) are also disrupted.

Working Memory

One critical cognitive function is “working memory,” the ability to hold information “on line” in the service of performing an impending task (43 ,44). Evidence suggests that persons with HIV infection demonstrate deficits in working memory because of the affinity of HIV-1 for frontal-subcortical circuits (45). Given the anatomic evidence for involvement of frontal and related subcortical structures in executive functioning and HIV infection, it is not surprising that executive processes are affected by HIV infection. In fact, recent studies have found evidence for the *selective* impairment of verbal and spatial working memory processes in HIV-seropositive persons (13 ,45 ,46 ,47 ,48 and 49). Deficits of working memory tend to be observed in the later stages of HIV infection (i.e., after AIDS has been diagnosed) (45).

Learning and Memory

Patients with subcortical disorders (e.g., Parkinson disease, Huntington disease, basal ganglia disease) typically demonstrate deficits of recall in the context of relatively spared recognition memory but show fewer false-positive errors than are typically seen in patients with cortical dysfunction (50 ,51). This pattern, which supposedly reflects a problem with the retrieval of information rather than difficulties with encoding, is also seen in patients with HIV infection (5 ,7 ,12 ,51).

Motor/Psychomotor Speed

Psychomotor slowing appears to be the most common HIV-related neurocognitive deficit and may underlie deficits in higher-order cognitive processes (52). Slowed complex cognitive processing may occur independently of peripheral nerve compromise, basic motor impairment, or psychiatric status (10). Psychomotor and motor impairment in HIV-infected persons has been well documented (5 ,53) and may be evident even in the earliest stages of HIV infection (31 ,39). Reaction time tasks have been particularly helpful in detecting HIV-related cognitive slowing because they allow for a more precise analysis of the effects of HIV on psychomotor processing (42).

Progression of Neurocognitive Deficits

The progression of HIV disease is associated with neurologic deficits (54) and deteriorating neurocognitive performance (9 ,14 ,36 ,55). Neurocognitive impairment is a proximal predictor of AIDS and mortality in HIV-infected persons (56 ,57 and 58). Slowed information processing and memory deficits before the development of AIDS have been associated with mortality independently of Centers for Disease Control and Prevention (CDC) clinical stage, $CD4^+$ T-lymphocyte count, hemoglobin level, antiretroviral treatment, or sociodemographic variables such as age, education, and socioeconomic status (58). Patients with advanced HIV disease who are neurocognitively impaired are at higher risk for death than those at the same stage of infection without cognitive disturbance (56). In addition, prominent psychomotor slowing has been associated with more rapid neurologic progression in persons with HIV-associated dementia (59).

HIV-1-Associated Dementia Complex

In 1986, Navia et al. (60) described a triad of clinical symptoms that were later categorized as AIDS dementia complex (ADC). These symptoms included cognitive impairment, motor dysfunction, and behavioral changes in the context of a diagnosis of AIDS. In this landmark study of 70 patients with AIDS subjected to autopsy, 46 were characterized as demented. The majority of the patients with ADC (63%) had a preexisting diagnosis of AIDS. However, in 37% of the cases, ADC had developed before clinical evidence of AIDS. This finding led to problems with the ADC label because it was evident from the Navia study that a subgroup of patients who did not meet diagnostic criteria for AIDS displayed cognitive symptoms associated with HIV infection. In addition, certain patients displayed some, but not all, of the symptoms necessary to be classified as having ADC.

In 1990, the World Health Organization (WHO) recommended a new diagnostic term, *HIV-1-associated dementia* (HAD), to replace ADC (61). The American Academy of Neurology also adopted this new term for research purposes (62) (Table 90.1). Today, the two terms are often used interchangeably. However, because the progression of HIV infection to AIDS can now be delayed with aggressive antiretroviral therapy, a focus on HIV rather than on AIDS may be a more appropriate way to classify neurocognitive impairment.

I. Sufficient for diagnosis of AIDS
- A. HIV-1-associated dementia complex
  - *Probable* (must have *each* of the following):
    1. Acquired abnormality in at least *two* of the following cognitive abilities (present for at least 1 month): attention/concentration, speed or processing of information, abstraction/reasoning visuospatial skills, memory/learning, and speech/language
       The decline should be verified by reliable history and mental status examination. In all cases, when possible, history should be obtained from an informant, and examination should be supplemented by neuropsychological testing.
    2. At least *one* of the following:
       a. Acquired abnormality in motor function or performance verified by clinical examination (e.g., slowed rapid movements, abnormal gait, limb incoordination, hyperreflexia, hypertonia, or weakness), neuropsychological tests (e.g., fine motor speed, manual dexterity, perceptual motor skills) or both.
       b. Decline in motivation or emotional control or change in social behavior. This may be characterized by any of the following: change in personality with apathy, inertia, irritability, emotional lability, or new onset of impaired judgment characterized by socially inappropriate behavior or disinhibition.
    3. Absence of clouding of consciousness during a period long enough to establish the presence of #1.
    4. Evidence of another etiology, including active central nervous system opportunistic infection or malignancy, psychiatric disorders (e.g., depressive disorder), active alcohol or substance use, or acute or chronic substance withdrawal; must be sought from history, physical and psychiatric examination, and appropriate laboratory and radiologic investigation (e.g., lumbar puncture, neuroimaging). If another potential etiology (e.g., major depression) is present, it is *not* the cause of the above cognitive, motor, or behavioral symptoms and signs.
  - *Possible* (must have *one* of the following):
    1. Other potential etiology present (must have *each* of the following):
       a. As above (see *Probable* #1, 2, and 3).
       b. Other potential etiology is present but the cause of #1 above is uncertain.
    2. Incomplete clinical evaluation (must have *each* of the following):
       a. As above (see *Probable* #1, 2, and 3).
       b. Etiology cannot be determined (appropriate laboratory or radiologic investigations not performed).

[a]From the Working Group of the American Academy of Neurology Task Force, 1991.

TABLE 90.1. CRITERIA FOR A CLINICAL DIAGNOSIS OF HIV-1-ASSOCIATED DEMENTIA[a]

HIV-associated dementia complex is the most common neurologic disorder in HIV infection, affecting 6% to 30% of all infected persons (11 ,63 ,64 and 65). It has an insidious onset, although persons with HAD may experience an acceleration of symptomatology. Before the introduction of highly active antiretroviral therapy (HAART), the median survival time for patients with late-stage HAD was 6 months (60 ,65). HAD affects cognitive, motor, and behavioral functioning, and patients often exhibit apathy, cognitive and motor slowing, and impaired memory, abstract reasoning, and judgment. HAD represents the more severe end of a continuum of HIV-related cognitive deficits; the milder end is represented by the presence of a single cognitive impairment,

such as psychomotor slowing. The deficits observed in this disorder result in impaired social and occupational functioning.
A diagnosis of HAD requires laboratory evidence of HIV-1 infection and the exclusion of other CNS conditions (e.g., toxoplasmosis, CNS lymphoma, cytomegalovirus ventriculitis, cryptococcal meningitis, meningitis associated with syphilis, progressive multifocal leukoencephalopathy). Other causes, such as depression and delirium, which can manifest as cognitive impairment, must also be ruled out.
The American Psychiatric Association has defined dementia due to HIV disease as a dementia that is judged to be the direct pathophysiologic consequence of HIV disease and has outlined its own diagnostic criteria (Table 90.2).

A. The development of multiple cognitive deficits manifested by both
   1. memory impairment (impaired ability to learn new information or to recall previously learned information) and
   2. one (or more) of the following cognitive disturbances:
      a. aphasia (language disturbance)
      b. apraxia (impaired ability to carry out motor activities despite intact motor function)
      c. agnosia (failure to recognize to identify objects despite intact sensory function)
      d. disturbance in executive functioning (i.e., planning, organizing, sequencing, abstracting)
B. The cognitive deficits in criteria A1 and A2 each cause significant impairment in social or occupational functioning and represent a significant decline from a previous level of functioning.
C. Evidence from the history and physical examination or laboratory findings indicate that the disturbance is the direct physiologic consequence of HIV infection affecting the central nervous system.
D. Deficits do not occur exclusively during the course of a delirium.

TABLE 90.2. DIAGNOSTIC CRITERIA FOR DEMENTIA DUE TO HIV DISEASE

HIV-Associated Dementia and HIV Disease Progression

Price and Brew (66) argued that it is not enough simply to characterize HIV-infected persons as demented or not. Diagnosis should also involve staging of the severity of dementia and a description of specific impairments. Therefore, ADC was divided into different stages to differentiate levels of neurocognitive and neuropsychiatric functioning in HIV-infected persons (Table 90.3).

ADC Stage	Characteristics
Stage 0 (normal)	Normal mental and motor function.
Stage 0.5 (equivocal/ subclinical)	Either minimal or equivocal *symptoms* of cognitive or motor dysfunction characteristic of ADC, or mild signs (snout response, slowed extremity movements), but *without impairment of work or capacity to perform ADL;* gait and strength are normal.
Stage 1 (mild)	Unequivocal evidence (symptoms, signs, neuropsychological test performance) of functional intellectual or motor impairment characteristic of ADC, but able to perform *all but the more demanding aspects of work or ADL;* can walk without assistance.
Stage 2 (moderate)	Cannot work or maintain the more demanding aspects of daily life, but able to perform *basic activities of self-care;* ambulatory, but may require a single prop.
Stage 3 (severe)	Major *intellectual incapacity* (cannot follow news or personal events, cannot sustain complex conversation, considerable slowing of all output) or *motor disability* (cannot walk unassisted, requiring walker or personal support, usually also with slowing and clumsiness of arms).
Stage 4 (end stage)	*Nearly vegetative;* intellectual and social comprehension and responses are at a rudimentary level; nearly or absolutely mute; paraparetic or paraplegic with double incontinence.

ADC, AIDS dementia complex; ADL, activities of daily living.
From Price R, Brew B. The AIDS dementia complex. *J Infect Dis* 1988;158:1079–1083, and Sidtis JJ, Gatsonis C, Price RW, et al. Zidovudine treatment of the AIDS dementia complex: results of a placebo-controlled trial. *Ann Neurol* 1993;33:343–349.

TABLE 90.3. AIDS DEMENTIA COMPLEX STAGING SCHEME

The course of HAD is highly variable. HAD begins with subtle deficits in cognitive processes (e.g., difficulties with concentration, forgetfulness, mental slowing). In some patients, HAD progresses rapidly after the diagnosis has been made (within weeks to months), whereas other patients show cognitive stability for months or years or very slow neurologic and neurocognitive progression (5 ,50 ,60). Persons in the early stages of HAD often complain of poor concentration, psychomotor retardation, and forgetfulness (60). They typically present with significantly less impairment in functions of daily living, cognition, and social and occupational functioning than is seen in the later stages of HAD. Hallmark symptoms of early-stage dementia are apathy, lethargy, poor concentration, and social withdrawal. Psychotic symptoms are not typically observed at this stage.

The later stage of HAD corresponds to ADC as originally defined, and survival expectancy at this stage may be 6 months or less. Features includes global cognitive dysfunction, significant functional impairment, and psychotic symptoms. Late-stage symptoms also may include aphasia, confusion, disorientation, delusions, hallucinations, seizures,

and muscular weakness and paralysis (particularly in the lower limbs). Advanced dementia may result in disinhibition, mutism, catatonia, and incontinence (5 ,60). Neuropsychiatric complications of late-stage HAD include depression, mania, and psychosis.

The introduction of HAART has altered the course of HIV infection and AIDS and has diminished the CNS manifestations of HIV, even though protease inhibitors exhibit relatively poor penetration of the blood–brain barrier.

Prevalence of HIV-Associated Dementia

An early prospective study of HIV infection in the United States revealed that HAD develops in approximately 15% of patients with AIDS (65). The WHO cross-cultural study examining the neuropsychiatric consequences of HIV infection represents the best study of the prevalence of HAD based on a diverse clinical sample. This study examined persons at five sites around the world (Bangkok, Thailand; Kinshasa, Zaire; Munich, Germany; Nairobi, Kenya; and Sao Paulo, Brazil). The prevalence rates for dementia in patients with AIDS ranged from 4.4% to 6.9% and did not differ between four of the five recruitment centers (in Bangkok, no cases of HAD were recorded). In addition, no patients met the criteria for HAD while in the asymptomatic stage (11). In 1997, the CDC reported that 5% of adults with an AIDS-defining opportunistic illness had HIV encephalopathy (dementia) (67).

Early HIV-1-Associated Cognitive Motor Impairment

To classify appropriately HIV-1-seropositive persons who do not meet the criteria for dementia, the American Academy of Neurology AIDS Task Force introduced the term *HIV-1-associated minor cognitive/motor disorder* (MCMD) (62) (Table 90.4), a correlate of the WHO diagnosis of HIV-1-associated minor cognitive/motor disorder (61). In comparison with HAD, MCMD involves subtler neurocognitive difficulties that may not necessarily affect a patient's activities of daily living or occupational functioning.

II. Not sufficient for diagnosis of AIDS
- A. HIV-1-associated minor cognitive/motor disorder
  - *Probable* (must have *each* of the following):
    1. Cognitive/motor behavioral abnormalities (must have *each* of the following):
       - a. At least *two* of the following acquired cognitive, motor, or behavioral symptoms (present for at least 1 month) verified by reliable history, when possible, from an informant):
         - (1) Impaired attention or concentration
         - (2) Mental slowing
         - (3) Impaired memory
         - (4) Slowed movements
         - (5) Incoordination
         - (6) Personality change, or irritability or emotional lability
       - b. Acquired cognitive/motor abnormality verified by clinical neurologic examination or neuropsychological testing (e.g., fine motor speed, manual dexterity, perceptual motor skills, attention/concentration, speed of processing of information, abstraction/ reasoning, visuospatial skills, memory/learning, or speech/ language)
    2. Disturbance from cognitive/motor/behavioral abnormalities (see #1) causes mild impairment of work or activities of daily living (objectively verifiable or by report of a key informant).
    3. Does not meet criteria for HIV-1-associated dementia complex or HIV-1-associated myelopathy.
    4. *No* evidence of another etiology, including active central nervous system opportunistic infection or malignancy, or severe systemic illness determined by appropriate history, physical examination, and laboratory and radiologic investigation (e.g., lumbar puncture, neuroimaging). The above features should not be attributable solely to the effects of active alcohol or substance use, acute or chronic substance withdrawal, adjustment disorder, or other psychiatric disorders.
    5. Criteria for minor cognitive motor disorder include at least two of the following acquired cognitive, motor, or behavioral symptoms, generally assessed with neuropsychological evaluation:
       - a. Impaired attention or concentration
       - b. Mental slowing
       - c. Impaired memory
       - d. Slowed movements
       - e. Incoordination
       - f. Personality change, irritability, or emotional lability

[a]From the Working Group of the American Academy of Neurology Task Force, 1991.

TABLE 90.4. CRITERIA FOR A CLINICAL DIAGNOSIS OF HIV-1-ASSOCIATED COGNITIVE/MOTOR COMPLEX3

The American Academy of Neurology AIDS Task Force specified that persons with HIV could be given this diagnosis regardless of their medical status. In other words, patients who are otherwise asymptomatic according to the CDC definitions may meet the criteria for MCMD. MCMD and HAD may represent areas along a common continuum of cognitive impairment, but they may also exist as distinct elements (62). However, the pattern of deficits observed in both disorders is characteristic of patients with dysfunction of frontal-subcortical neuronal circuitry (68), which means that persons with HAD or MCMD demonstrate differential deficits in the retrieval of information relative to the encoding or storage of information (51). HIV seropositivity with MCMD has been associated with a poorer prognosis relative to HIV seropositivity without MCMD (56), which suggests that MCMD may be a precursor of HAD.

Confounding Variables

To diagnose HIV-related cognitive deficits in patients (i.e., to conceptualize cognitive impairment in these patients as

specifically related to the HIV virus), other potential causes of neurocognitive impairment must be ruled out. Depression, anxiety, and substance abuse are conditions seen in HIV disease that may contribute to neurocognitive impairment (69 ,70).

Depression and Apathy

Significant depressive symptomatology has been reported in patients with HIV-1 infection (71 ,72 ,73 and 74). Clinical assessments of persons with HIV infection can be confounded by the overlap of symptoms of HIV infection and somatic symptoms of depression (e.g., lack of motivation, low levels of energy, fatigue, weight loss, poor sleep patterns) (75), so that the severity of depression in HIV-seropositive persons is potentially overestimated (76 ,77). Therefore, cognitive and affective symptoms may be more accurate indicators of an underlying mood disorder in persons with HIV/AIDS because somatic symptoms may reflect advanced stages of disease and be unrelated to depression.

Although symptoms of depression and neuropsychological impairment may occur together in many HIV-infected persons, most studies have demonstrated that neuropsychological abnormalities observed in HIV infection are distinct and cannot be attributed to depression *per se* (18 ,35 ,73 ,77 ,78 ,79 ,80 ,81 ,82 ,83 and 84). It has been shown that depressed patients with HIV-1 infection may exhibit deficits in learning and memory (79 ,81 ,85), but the contributions of depression to the development and degree of these impairments appear to be minimal.

Apathy and reduced motivation are frequently observed in HIV-seropositive patients (45 ,60). Apathy, but not depression, also has been associated with deficits of working memory in HIV-seropositive patients, which suggests that both are manifestations of dysfunction in frontal–subcortical circuitry (45).

Substance Abuse

It is well known that HIV infection can be contracted through the intravenous administration of drugs with shared needles. The persistent use of drugs through other routes (e.g., inhaling crack/cocaine) may also increase the risk for HIV infection; drug abusers may practice prostitution to support their habit or engage in risky sexual behavior while under the influence of drugs or alcohol. The use of drugs such as crack/cocaine has been significantly associated with earlier progression to AIDS in HIV-seropositive persons (86). Injection drug use has also been associated with more rapid progression of HAD (59), and recreational drug use has been found to diminish overall neuropsychological performance and reduce visuomotor processing, executive functioning, motor speed and strength, and sensorimotor perception in persons in the early stages of HIV infection (87).

Many investigators have postulated that drug use may increase the risk of HIV-seropositive persons for cognitive impairment (19); possibly, drug use induces CNS impairment independent of that caused by HIV infection. For example, it is known that chronic cocaine use can result in seizures, cerebrovascular accidents, and movement disorders. Deficits in attention, learning and memory (88 ,89 and 90), word production, visuomotor integration (90), and executive function (91) are specifically affected by cocaine use. These deficits have been related to dysfunction in prefrontal brain regions, the orbitofrontal cortex, and the anterior cingulate gyrus (91) and to cerebral hypoperfusion in the frontal, periventricular, and temporal–parietal areas (90). Although this is a plausible hypothesis, most studies have found that substance abuse does not independently contribute to neuropsychological dysfunction after HIV seropositivity is taken into account (92). The interaction of HIV infection and drug use does not appear to produce additional cognitive deficits (26 ,27 ,80 ,93 ,94 ,95 and 96).

Other Cofactors

Other factors may potentially confound the neuropsychological functioning of HIV-infected persons, including gender, ethnicity, level of education, and medication regimen.

Gender and Ethnicity

Racial and ethnic minorities have been disproportionately affected by HIV/AIDS, and women constitute one of the most rapidly increasing groups at risk for AIDS in the United States (96 ,97). Although neuropsychological assessment has been helpful in elucidating patterns of impairment in persons with HIV/AIDS, most of the study participants have been well-educated, Caucasian, homosexual men. It remains unclear whether neuropsychological instruments are equally valid in assessing HIV-related neurocognitive functioning in populations who differ in terms of ethnicity/culture, language, educational background, and socioeconomic status. Neuropsychological tests and batteries, in addition to being sensitive to the presence of cerebral pathology, are also sensitive to demographic factors, including gender, ethnicity, socioeconomic status, education, and age.

Neuropsychological tests may overdiagnose organic impairment in neurologically intact African-Americans (98); thus, clinicians must be cautious about reporting increased neurocognitive impairment among HIV-infected members of minorities. Ideally, separate norms for specific ethnic subgroups (e.g., African-Americans) should be established. Studies that examined the effect of ethnicity on neuropsychological performance in the context of HIV infection have found that after correction for education, ethnicity accounts for additional variance (99). One possible explanation for these findings is that matching for grade level does not always resolve the effects of disparate educational experiences of test subjects. Also, within the African-American population, the incidence of poverty and homelessness is higher, situations that can have a pervasive influence on the

interactions of persons with their environment and thus their performance on neuropsychological tests. Also, many tests ignore the fact that African-Americans display ideals, values, beliefs, and cultural traditions that contribute to unique psychological processes not often tapped by mainstream neuropsychological instruments (100).

African-Americans who were HIV-seropositive and matched for age, education, gender, and HIV disease stage obtained lower scores than their Caucasian counterparts on neurocognitive measures (101). However, acculturation level accounts for the vast majority of these differences in performance. Acculturation level refers to the degree to which a person's beliefs, values, and behaviors are consistent with European-American culture, the dominant cultural basis for most standardized neuropsychological measures. Therefore, it is important to examine acculturation level to improve the accuracy with which HIV-related neurocognitive deficits are diagnosed in ethnic minorities. Manly et al. (101) suggest that measures that assess attention, psychomotor speed, and retention may be of greater utility in assessing HIV-associated cognitive deficits across cultural groups.

African-Americans show a pattern of performance deficits in neuropsychological functioning that are similar to those reported in majority populations. For example, studies have found deficits in verbal and nonverbal memory in symptomatic men (77) and women (102). As in ethnic minority populations, the incidence of HIV-1 infection in women has continued to rise. However, few studies have explored the neuropsychological sequelae of HIV infection in women. It has been reported that HIV-infected women report more psychological symptoms than HIV-infected men (103 ,104) and that the early signs of HIV infection in women are often overlooked and underrecognized in comparison with the symptoms of men (105). HIV disease progression does not appear to differ according to sex (86). Gender does not appear to be related to rates of $CD4^+$ cell decline, first $CD4^+$ count below 200/mm^3, clinical progression to clinical AIDS, or mortality (86).

However, neurocognitive functioning does appear to differ between HIV-seropositive men and women. Whereas HIV-infected men and women show similar vulnerability to cognitive dysfunction, AIDS in women is associated with a poorer neuropsychological test performance (52). Also, the degree of impairment in symptomatic HIV-seropositive women has been found to be greater than that in either HIV-seropositive homosexual men or HIV-seropositive intravenous drug users (106). These differences do not appear to be secondary to differences in age, education, or mood, which suggests that neuropsychological impairment may become apparent earlier in HIV-seropositive women than in HIV-seropositive men.

Few studies have examined neurocognitive functioning solely in HIV-seropositive women. The studies that have been conducted seem to demonstrate a similar pattern of deficits in both women and men. In a longitudinal study, HIV-seropositive women demonstrated slower reaction times and motor speed and less improvement in verbal memory than did seronegative women at 6-month follow-up (107). Cross-sectional analyses have found no differences between asymptomatic women and seronegative controls (30 ,102). However, with progression to AIDS, deficits in attention and memory became evident (102). In a preliminary investigation, Costa et al. (108) found no differences between asymptomatic HIV-seropositive and HIV-seronegative women in quantitative EEG activity during various conditions, nor between the two groups on any of the neurocognitive or neuropsychiatric measures administered.

Psychosocial Variables

Fewer years of education, lower estimated premorbid intelligence, lower occupational attainment, and lower socioeconomic status may put patients at particular risk for HIV-related neurocognitive impairment (38 ,109 ,110). Satz (109) found the rate of impairment in asymptomatic HIV-seropositive participants to be comparable with that of seronegative controls among participants with more than 12 years of education. The rate of impairment in asymptomatic HIV-seropositive participants with 12 or fewer years of education was more than twice that of seronegative controls (38% vs. 17%). Stern et al. (110) found that HIV-seropositive participants with a low cognitive reserve (based on measures of education level, occupational attainment, and vocabulary) performed worse on neuropsychological measures than did HIV-seronegative and HIV-seropositive participants with a high cognitive reserve. According to the cognitive reserve theory, the threshold for neuropsychological symptoms of persons with a greater cognitive or brain reserve may be higher after acquired brain injury (109). These studies suggest that a low cognitive reserve increases one's risk for cognitive dysfunction and underscore the need to consider demographic variables in any evaluation of neuropsychological function in HIV infection.

Clinical Significance of Early Cognitive/Motor Disturbance and Its Relationship to Future Disease

The identification of HIV-related neurocognitive and neuropsychiatric disturbances has potential medical and vocational implications for patients. In particular, cognitive disturbances can affect the ability to adhere to antiretroviral medication regimens and can affect occupational functioning.

Adherence to Medication Regimen

The administration of protease inhibitors in combination with nucleoside analogues has successfully suppressed HIV replication in many persons with HIV-1 infection. However, the long-term effectiveness of medication regimens that include protease inhibitors depends on strict adherence

to the prescribed drug regimen. Poor adherence can lead to changes in plasma levels of HIV RNA and the development of treatment-resistant viral mutations (111). In other words, resistance to a particular drug and cross-resistance to drugs within a particular class can develop in persons who comply poorly with their medication regimen. Understanding the various factors that contribute to medication adherence is critical to optimizing the treatment of persons with HIV/AIDS.

Factors affecting adherence may include relationships with health care providers, complexity of antiretroviral regimens, depression, substance abuse, cultural beliefs, and neurocognitive functioning. Factors affecting medication adherence may vary across HIV disease stage. For example, it is well established that persons in the later stages of HIV infection exhibit neuropsychological deficits associated with frontal–subcortical brain dysfunction. Therefore, in more advanced HIV infection, poor compliance may be related to neurocognitive compromise. Poor adherence to antiviral medication has been associated with poor performance on measures of divided attention, learning and memory, executive function, and psychomotor ability (42 ,112 ,113). Hinkin et al. (42) examined the relationship between neurocognitive and neuropsychiatric symptoms, substance abuse, medication complexity and side effects, and a variety of psychosocial factors in predicting adherence to antiretroviral medication. Medication adherence was assessed by computerized monitoring (medication event monitor system, or MEMS), in which a computer chip embedded in the cap of a pill bottle records the date and time when the bottle is opened. Preliminary data from this group revealed that medication adherence is associated with executive function and prospective memory. Apathy, but not depression, was also found to predict poor adherence.

Employment

It has been proposed that neurocognitive impairment is clinically significant when it affects everyday functioning (114). For some HIV-infected persons, cognitive difficulties result in occupational problems even in the early stages of infection, when cognitive impairment is mild (114 ,115 and 116). When employed persons rated their job performance relative to that before their HIV diagnosis, self-perceived decreases in work abilities were five times more prevalent in those with neurocognitive impairment (114). Impaired persons also performed worse on standardized work samples and displayed more discrepancies between expected levels of functioning (based on work history) and observed levels (based on work samples), which suggested an acquired decrease in vocational abilities (117). Furthermore, HIV-seropositive persons who demonstrated neurocognitive impairment were two to three times more likely to be unemployed than were those without cognitive impairment, even after control for medical status (114 ,115). A poorer performance on tasks of learning and executive functioning seems to be a good predictor of loss of employment status (115).

Pharmacologic Treatment of HIV Infection

Highly active antiretroviral therapy has changed the epidemiology of HIV disease progression. In 1996, the annual AIDS incidence decreased for the first time in the United States. In 1997, this pattern continued as the number of new AIDS diagnoses decreased (97). However, AIDS prevalence increased from 1996 to 1997, probably because of longer survival times after diagnosis. This decline in the incidence of AIDS and AIDS deaths and the observed delay in progression to AIDS are in part a consequence of HAART. In a HAART regimen, three or more antiretroviral drugs, such as a nucleoside analogue reverse transcriptase inhibitor, a protease inhibitor, and a non-nucleoside reverse transcriptase inhibitor, are usually combined.

Before the advent of HAART, monotherapy with zidovudine (AZT) was reported to improve neurocognitive functioning, slow progression to dementia (118 ,119 and 120), decrease neuropathologic features of AIDS (4), and prevent mild cognitive impairment associated with HIV (121).

In a patient population largely comprising persons on monotherapy (81%), treated patients showed superior information processing on a reaction time task (a sensitive measure of HIV-related cognitive slowing) in comparison with untreated participants (48). High doses of AZT are reportedly more effective in improving neurocognitive functioning (120), and long-term use of AZT has been associated with improved cognitive performance in subjects with early symptomatic HIV infection and AIDS (122).

The introduction of protease inhibitors has resulted in new approaches to the treatment of HIV infection. Patients on an HAART regimen perform better than do those treated with less intensive antiretroviral therapy (e.g., regimens that do not include a protease inhibitor) (116). However, persons on combined antiretroviral therapy, regardless of whether a protease inhibitor is included, have shown improved psychomotor speed in comparison with antiretroviral-naïve patients and patients treated with monotherapy (123). Improvement in neurocognitive functioning has been associated with a reduction in viral load (116). HAART has resulted in undetectable levels of HIV in the blood. Even though the CNS penetration of protease inhibitors is poor, multiple drug regimens lower serum viral load, slow disease progression, and in some cases improve HIV-associated cognitive motor complex, reverse HIV encephalopathy, and improve cognitive impairment (124 ,125). Future research will focus on the mechanisms of HAART-associated improvement in neuropsychological functioning and the possibility that the CNS may act as a reservoir for HIV.

PSYCHIATRIC MANIFESTATIONS OF HIV-1 INFECTION

Part of "90 - Neuropsychiatric Manifestations of Hiv-1 Infection and Aids "

Psychiatric Symptoms in HIV-1 Infection

Major depression is the most common psychiatric disease among HIV-1-infected persons (74) Earlier controlled studies showed that the prevalence of major depression and other mood disorders is higher in asymptomatic HIV-1-infected homosexual men than in the general population (126 ,127) but is similar to the prevalence in HIV-1-seronegative homosexual men (71 ,128 ,129). In several early studies, from 4% to 9% of *both* HIV-1-infected and uninfected homosexual men reported a major depression in the month before study evaluation, and in the study of Perkins and colleagues (73), a major depression developed in 6% of both HIV-1-infected and uninfected subjects during a 6-month follow-up period. Evidence also indicates that similar proportions (from zero to 5%) of HIV-1-infected and uninfected persons meet DSM-III-R criteria for current anxiety disorders (73). Thus, after more than 15 years of research, the available data suggest that the prevalence of major depression is high in asymptomatic HIV-1-infected gay men in comparison with the prevalence in men of similar age in the population at large, but no higher than that in seronegative gay men of similar age and somewhat lower than that in patients with serious medical illnesses, such as cancer and heart disease (130 ,131 ,132 and 133). These findings underscore the issue that mood disorders should not be considered a "normal" phenomenon in HIV-1-infected persons. Rather, they should be assessed carefully and treated appropriately.

Diagnosing major depression in HIV-1-infected patients can be complicated because several symptoms of major depression (i.e., fatigue, sleep disturbance, and weight loss) are also common symptoms of progression of HIV-1 disease (134 ,135). However, although complaints of fatigue and insomnia in asymptomatic HIV-1-infected homosexual men are significantly associated with depressed mood and other symptoms of major depression, they are not associated with low $CD4^+$ counts or decreased neuropsychological functioning (136 ,137). Thus, complaints of fatigue and insomnia in otherwise asymptomatic HIV-1-infected patients are highly suggestive of an underlying mood disorder, and patients with such complaints should be assessed routinely for major depression.

Factors that influence the development of major depression in HIV-1-infected persons have not been well studied. However, several studies suggest that the prevalence of a past history of major depression is relatively high in HIV-infected persons and may be a risk factor for the development of major depression (74). The first of these was the study of Perkins and colleagues (73), who found a relationship between major depression in asymptomatic HIV-1-infected homosexual men and a prior history of major depression. Coping with the threat of AIDS also may be related to the overall level of depressed and dysphoric mood. Leserman and colleagues (138) reported that a depressed and anxious mood was less frequent in asymptomatic HIV-1-infected men using active coping strategies to deal with the threat of AIDS (e.g., fighting spirit, reframing stress to maximize personal growth, planning a course of action, seeking social support) than in those using passive coping strategies (denial or feeling helpless). Like the studies of persons with other potentially life-threatening diseases, early studies of HIV-1-seropositive persons found that they usually are able to adjust successfully to their infection and that most are able to maintain hope over time. More recently, the availability of HAART has led to a still greater sense of hope. Therefore, coping strategies in HIV-infected persons may influence the development of depression or anxiety.

Studies of psychiatric and psychosocial factors in HIV-infected women are emerging as the demographics of the HIV epidemic change. Early studies are difficult to interpret because study methodology and populations differed considerably (74). A high rate of major depression was found in women using intravenous drugs, but this rate did not differ from that of men using intravenous drugs (139 ,140); high rates of major depression were found in both seropositive and seronegative men and women using intravenous drugs. However, a gender difference was found; the prevalence of depressive and anxiety symptoms, but not syndromes, was higher in women than in men. This finding held for both seropositive and seronegative subjects. In a related study of Boland et al. (141), HIV status was not related to depressive symptoms at baseline in a large, multicenter, prospective sample of U.S. women. Both seronegative and seropositive women had a high prevalence of depressive symptoms on the Center for Epidemiological Studies Depression Scale (CES-D); however, diagnoses of depressive disorder were not obtained. In the study of Morrison et al. (142), who examined the prevalence of depressive and anxiety disorders in a large cohort of rural U.S. women, the prevalence of major depressive disorder and anxiety symptoms was significantly higher in seropositive women without active substance abuse than in seronegative controls (142).

As reviewed earlier, the effects of HIV-1 on the CNS may result in a variety of neurocognitive and related psychiatric symptoms in the later stages of illness. In addition, persons infected with HIV-1 may be at further risk for the development of psychiatric symptoms because of diseases secondary to AIDS that also have CNS effects and because of medications used to treat HIV-1 infection. Although psychiatric symptoms in HIV-1-infected persons in the later stages of illness may represent new-onset psychiatric disorders, it is more likely that these symptoms reflect the direct CNS effects of HIV-1, HIV-1-related CNS disturbances, and CNS effects of medications used in the treatment of AIDS. Thus, although Leserman and colleagues (138) found an increase in depressive symptoms approximately 1.5 years before the

onset of AIDS, in the study of Rabkin et al. (140), the onset of syndromal depression and anxiety did not increase despite worsening HIV infection during a 4-year period.

Evidence from earlier stages of the epidemic suggests that HIV-1 may cause organic mood disturbance. In a 17-month retrospective chart review of patients with AIDS, Lyketsos and colleagues examined associated historical and clinical features in an attempt to separate organic and functional symptoms (145). They used a family history of mood disorder as a "marker" for functional mood disorders. They further assumed that coexisting dementia and a low $CD4^+$ count are "markers" of HIV-1-related mood disorders. None of the patients with a personal or family history of mood disorder had coexistent dementia, and all but one of the patients without a personal or family history of mood disorder had coexistent dementia. In addition, among the 8% of patients who experienced manic episodes, the $CD4^+$ count was significantly higher in those without a personal history of mood disorder. Although these findings suggest that mania may be a consequence of the direct or indirect effects of HIV-1 on the brain, controlled studies have yet to find this relationship (74). Vitamin B_{12} deficiency may also place HIV-1-infected patients at risk for organic mood disturbance. Between 20% and 30% of patients with AIDS and 7% of asymptomatic HIV-1-infected patients have been reported to have a vitamin B_{12} deficiency. Furthermore, vitamin B_{12} deficiency has previously been shown to be associated with depression and can occur in the absence of hematologic or neurologic signs (146). Although the relationship between vitamin B_{12} level and depressive symptomatology in HIV-1-infected patients is not clear (147), it is prudent that the medical evaluation of depressive symptoms in HIV-1-infected patients include an assessment of serum B_{12} levels.

Although relatively uncommon, psychosis may result from HIV-1 involvement of the CNS. Earlier case studies of symptomatic HIV-1-infected persons have reported psychotic symptoms, including delusions, bizarre behavior, and hallucinations. Mood disturbances, including euphoria, irritability, and labile or flat affect, have often accompanied psychotic symptoms in these patients. Similarly, anxiety and agitation were reported in almost half of the reported cases. In addition, some evidence indicates that psychosis may be a symptom of the terminal stages of AIDS; half of the patients described had a progressively worsening course, with dementia or death occurring within a few months after the onset of psychotic symptoms. Psychosis may be more frequently found in patients with significant AIDS-related neurocognitive impairment than in patients in earlier stages of the disease. In one retrospective chart review of 46 patients identified with HIV-1-associated dementia, Navia and Price (148) found that psychotic symptoms had developed in 15%. Relatedly, data from the San Diego HIV Neurobehavioral Research Center (149) suggest that HIV-1-infected patients with psychosis trend toward greater neurocognitive impairment than do nonpsychotic HIVinfected controls. Thus, new-onset psychosis may be secondary to HIV-related encephalopathy. Accordingly, a complete organic workup should be considered for HIV-1-infected patients with significant disturbance of mood or psychosis.

Treatment of Psychiatric Disorders in HIV-1 Infection

Available evidence suggests that mood symptoms and syndromes in the asymptomatic stage of HIV-1 infection are not secondary to the effects of HIV-1 on the brain and should be evaluated and treated as in the general population. Although confirmatory data are lacking, this probably also holds true in the symptomatic stages of the disease.

Controlled Studies with Antidepressant Drugs

Only a small proportion of the published studies of the treatment of mood disorders in patients with HIV-1 infection have been double-blinded, randomized, placebo-controlled studies. In the study of Rabkin et al. (150), imipramine was effective in 97 HIV-infected patients. At 6 weeks, they found a response rate of 74% in the imipramine group versus 26% in the placebo group. No changes in $CD4^+$ helper/inducer cell counts were found in the imipramine-treated subjects. However, adverse anticholinergic side effects led to discontinuation of imipramine within 6 months in more than one-third of the responders. Elliott and co-workers (151) blindly and randomly assigned 75 HIV-seropositive patients to treatment with imipramine, paroxetine, or placebo. Of the 75 enrolled subjects, 75% completed 6 weeks of treatment; only 45% completed the full 12-week trial. The two antidepressants were found to be equally efficacious at 6, 8, and 12 weeks, and both were significantly more efficacious than placebo. Side effects of the tricyclic antidepressants markedly influenced attrition. The dropout rate in the imipramine group was 48%, 20% in the paroxetine group, and 24% in the placebo group. Zisook and colleagues (152) reported that fluoxetine was more effective than group therapy in a randomized, double-blinded, placebo-controlled study of 47 HIV-seropositive men with major depression. Rabkin and colleagues (153) recently completed a double-blinded, placebo-controlled, 8-week trial with fluoxetine in 120 HIV-seropositive subjects. Among the subjects who completed the 8-week trial, 74% of the fluoxetine group responded to treatment, in comparison with 47% of the placebo group. When intention-to-treat analysis was used, the differences between the treatment groups were less remarkable (57% of the fluoxetine-treated subjects responded compared with 41% of the placebo group). Drug treatment did not alter levels of $CD4^+$ cell counts. Thus, the available data suggest that the selective serotonin re-uptake inhibitors (SSRIs) are effective and well

tolerated in the treatment of HIV-associated major depression.

Open Trials with Antidepressant Drugs

In a study related to the one described above, Rabkin and colleagues (154) enrolled HIV-infected subjects with depression who had failed imipramine treatment (i.e., they had relapsed, could not tolerate the side effects, or did not respond to the drug) in a 12-week open-label trial of fluoxetine. Although depression at baseline as measured on the Hamilton Depression Scale (HAM-D) was more severe in this sample (12.5) than in the imipramine study sample (15.8), 83% of subjects treated with fluoxetine (at doses of 15 to 60 mg/d) responded and exhibited significant reductions in their HAM-D scores. Fluoxetine treatment did not alter $CD4^+$ counts. Fluoxetine was tolerated better than imipramine. In an open-label trial of 28 depressed HIV-infected subjects, Rabkin et al. (155) found a 70% response rate among subjects who had 8 weeks of treatment with sertraline. Side effects resulted in a loss of 18% of the total sample. Sertraline did not alter the counts of either $CD4^+$ cells or natural killer cells.

Ferrando and colleagues (156) conducted a 6-week open-label, parallel-group trial of the SSRIs paroxetine, fluoxetine, and sertraline in 33 symptomatic HIV-infected patients with depression. Seventy-three percent of subjects completed the trial, and of these, 83% responded to their assigned treatment. Most of the subjects who dropped out did so because of complaints of agitation, anxiety, and insomnia during weeks 1 through 3. Both depression and somatic symptoms perceived to be related to HIV infection improved with SSRI treatment. Differences in efficacy between the three SSRIs could not be ascertained reliably because of the study design and small sample size. More recently, these authors performed a small open trial comparing fluoxetine (n = 21) and sertraline (n = 9) in HIV-infected women (157). Sixty percent of the women completed the trial, and of these, 78% were responders (e.g., HAM-D score decreased 50% or more). Grassi et al. (158) performed a 6-week open trial of the efficacy of paroxetine in 10 HIV-seropositive patients with major depression. Significant improvement in HAM-D scores was noted between weeks 2 through 6 of the study. Recently, a 73% response rate was demonstrated with nefazodone in a small open trial of 15 outpatients. Few adverse side effects were noted (159). As discussed in the next section, clinicians must be aware of the potential for drug interactions between antidepressants that potentially inhibit the CPY450 3A4 isoenzyme system and protease inhibitors and for adverse effects of herbal therapy. Although the outcomes of the open-label studies have generally been consistent with those of the available double-blinded, randomized, placebo-controlled trials, these findings must nonetheless be interpreted cautiously.

Effects of Psychostimulants and Novel Agents

Fernandez et al. (160) compared desipramine with methylphenidate in a treatment trial of 15 subjects. With either agent, subjects showed a response rate of approximately 50%; however, subjects treated with desipramine experienced more adverse events, including dry mouth, anxiety, and insomnia. In the open trial of Wagner et al. (161) of 24 patients with AIDS, debilitation, low levels of energy, and a DSM-III-R diagnosis of depressive disorder, the response rate to dextroamphetamine treatment was 75%. Improvements in mood and energy coincided, and analysis revealed significant reductions in HAM-D scores after as little as 2 weeks of treatment. Although systematic follow-up evaluations were not conducted, anecdotal evidence suggested that the treatment effect (improved mood and energy) was maintained for up to 2 years in some subjects. In a small treatment series of Rabkin et al. (154), all seven patients treated with a combination of fluoxetine and dextroamphetamine (5 to 25 mg/d) responded during the 12-week course of treatment, another preliminary observation that warrants further study. Placebo-controlled trials are necessary to confirm these promising observations.

HIV-associated reductions in testosterone levels have been found to correlate with changes in mood, appetite, and energy and with sexual dysfunction. In a double-blinded, placebo-controlled trial (6-week trial followed by 12-week open-label maintenance), testosterone injections were effective in improving both mood and libido, energy, and body muscle mass in 70 HIV-seropositive men with hypogonadal symptoms who completed the trial (162). The authors also found that exercise may augment improvement in psychological and nutritional status in HIV-seropositive patients receiving testosterone therapy (163). In an 8-week open-label pilot study of 45 HIV-positive subjects, the adrenal steroid dihydroepiandrosterone (DHEA) appeared promising for improving mood in addition to anabolic and androgenic parameters (153).

Treatment Considerations for Mood Disorder and Other Psychiatric Conditions

Antidepressant therapy is effective and can improve the quality of life of HIV-infected persons. SSRIs and newer agents may be particularly well suited for use in depressed HIV-seropositive patients because these agents do not produce the significant side effects (e.g., anticholinergic, α-adrenergic, histaminergic, and cardiac effects) caused by the tricyclic agents and other older classes of antidepressants (164). Further study is required.

Pain is frequently undertreated in patients with HIV infection (165), although it is well established that antidepressants are effective agents for the treatment of chronic pain, particularly antidepressants with noradrenergic properties (132). However, patients with AIDS experience adverse effects

more frequently with tricyclic antidepressants than do patients with AIDS-related complex and asymptomatic HIV-seropositive patients (166). Better-tolerated antidepressants with effects on serotoninergic and noradrenergic neurotransmitter systems include venlafaxine, mirtazepine, and paroxetine; these may prove useful and are awaiting controlled studies. Finally, placebo-controlled clinical trials have not demonstrated that stimulants are effective in treating primary depression (167), but they may be useful adjunctive agents in depressed patients with HIV infection, as noted above.

Treatment of Psychotic Symptoms

The treatment of psychotic disorders in HIV-infected patients has been less well studied than the treatment of mood disorders. Several reports have noted that HIV-seropositive patients may be more sensitive to the extrapyramidal side effects associated with dopamine-receptor antagonists (149 ,168). This is thought to be related to the subcortical motor slowing associated with HIV infection. In a case series of 21 patients with psychotic symptoms (12 had mania with psychotic features), risperidone was found to be efficacious and caused fewer side effects than did conventional antipsychotic drugs (169), although some anecdotal data suggest that the sensitivity of AIDS patients to both older and newer antipsychotic agents may be increased (170). Controlled studies are needed in this area.

In general, the same strategies used to treat the general population with psychotropic drugs are appropriate in HIV-infected patients with psychotic symptoms or mood disorders. Pharmacologic knowledge can be used to therapeutic advantage and to avoid potential untoward effects. Factors such as drug interactions related to psychotropic drug metabolism and protein binding, half-life, and effects on appetite require careful consideration, particularly in more debilitated HIV-infected patients. Potential interactions between psychotropic drugs and the antiretroviral agents used in HIV therapy with multiple drugs warrant consideration by the clinician because a potential for drug interactions exists. Psychotropic drugs, non-nucleoside reverse transcriptase inhibitors, and protease inhibitors may serve as substrates for various cytochrome P-450 enzymes in the liver. Each of these classes of compounds may possess enzyme-inducing or enzyme-inhibiting properties, and drugs such as the protease inhibitor ritonavir may simultaneously modify a number of isoenzymes (171).

The nontraditional herbal psychotropic agents used to treat psychiatric syndromes must also be monitored closely in HIV-positive patients. An open-label study recently revealed that the efficacy of the protease inhibitor indinavir (which is a metabolized by the 3A4 isoenzyme system) is markedly reduced by the concomitant administration of St. John’s wort (a potent inducer of the 3A4 system) (172). The reduction in indinavir levels was estimated to be sufficient to cause drug resistance and treatment failure. Finally, given that psychotropic drugs such as antidepressants can improve the quality of life of HIV-positive persons, further research is needed to determine whether effective treatment can improve medical outcomes in selected subsets of HIV-infected patients.

Psychoneuroimmunology of HIV-1 Infection

The multifactorial nature of HIV infection has led researchers to examine the influence of stress, depression, and other psychosocial factors on the course of this disease. A growing literature points to the potentially harmful effects of stress and depression on cellular immunity (173 ,174 ,175 and 176), and to the potentially negative impact of these psychosocial factors on the course of several types of cancer (177 ,178 ,179 and 180). Among patients with breast cancer, severe life stress has been associated with a greater probability of relapse (179), and psychosocial interventions to improve coping skills have resulted in increased numbers and function of natural killer cells and longer survival in patients with breast cancer or melanoma (177 ,178 ,181). These studies await replication.

Early findings on the significance of depression in predicting a decline in immune status and disease progression were inconsistent. Investigators found no relationship between psychosocial and psychiatric factors such as depressive disorders, distress, and stressors, on the one hand, and measures of HIV-1 disease progression, including $CD4^+$ and $CD8^+$ cell counts, on the other. However, a relationship was noted between the number and severity of HIV-related symptoms and levels of depressive disorders, distress, and stressors (182 ,183 and 184). In a 1996 metaanalysis, Zorilla et al. (185) found that depressive symptoms were longitudinally related to self-reported symptoms of HIV infection, but not to changes in $CD4^+$ T-lymphocyte counts or other commonly accepted markers of HIV disease progression. The mixed findings of this metaanalysis may have resulted from the inclusion of studies in which cross-sectional designs and relatively brief follow-up periods were used.

Prospective studies conducted for longer time intervals have found that depression may significantly predict HIV disease progression. In the San Francisco Men’s Health Study, a 9-year longitudinal study of 395 seropositive gay men, researchers found that subjects classified as depressed at study entry on the CES-D (186) progressed more rapidly to AIDS (187). The median time to first AIDS diagnosis was 6.2 years for subjects who were depressed at study entry and 7.6 years for nondepressed subjects. This finding held after control for baseline demographic variables, $CD4^+$ T-lymphocyte count, HIV-related medical symptoms, and health habits. At 5 years, this cohort showed no significant association of baseline depression with AIDS diagnosis; however baseline depression was associated with a decline in $CD4^+$ T lymphocytes (188). After 7 years of follow-up,

subjects with elevated depressive symptoms at every visit had a 1.7 times greater risk of mortality than did those who had never had an elevated depression score (189).

Initial analysis of 1,809 HIV-seropositive gay men in the Multicenter AIDS Cohort (MAC) study found no relationship between baseline depression, measured by the CES-D, and progression of HIV infection during 8 years of follow-up (190). Disease progression was defined as time to AIDS, death, or decline in $CD4^+$ T lymphocytes. In a subsequent report on years 2 through 6, a robust increase of 30% to 104% above baseline levels (depending on CES-D depression cut point) was noted in self-reported depressive symptoms beginning 1.5 years before a clinical diagnosis of AIDS and continuing beyond the diagnosis of AIDS (191). The authors interpreted these findings as an indication that depression may increase toward the later stages of HIV infection and thus be a manifestation of the HIV disease process. However, a subsequent survival analysis of these data, in which the level of depressive symptoms during the 6 months before AIDS diagnosis was used, showed no relationship between depression and time to death (144). A limitation of both these prospective cohort studies is the method of ascertainment of depression. The CES-D is not a clinical diagnostic tool; its sensitivity for DSM-III major depression is 80% to 90%, and its specificity is 70% to 80% (192).

The Coping in Health and Illness Project (CHIP) is a prospective study of initially asymptomatic HIV-infected gay men who are followed every 6 months; extensive clinical interviews are used to define depression and stressful life events. An analysis of this cohort at study entry showed a significant effect of stress on parameters of cellular immunity (143); furthermore, depressive symptoms, especially in the presence of severe stress, were related to declines in several lymphocyte subsets (e.g., CD16 and CD56 natural killer cells and CD8 cytotoxic suppressor cells) during a 2-year period (193). At 5.5 years, an increased risk for AIDS was associated with a higher cumulative level of depressive symptoms, measured by a modified Hamilton Depression Rating Scale (HDRS) excluding somatic symptoms that could be related to HIV disease. For every increase of one severe depressive symptom (3-point increment on the HDRS), the risk for AIDS doubled (194). This result, however, lost statistical significance after control for stressful life events and social support, which overlapped with depressive symptoms. The small number of subjects with elevated scores may partially account for this outcome.

The CHIP investigators are also directly investigating the effect of stressful life events on clinical outcome. Evans et al. (195) found that for each severe stressful event per 6-month study interval, the risk for early HIV disease progression doubled in men studied for up to 3.5 years. At 5.5 years (194) and 7.5 years (196) of follow-up of the CHIP cohort, Leserman et al. reported that more cumulative stressful events were associated with faster progression to AIDS. At both time points in follow-up, every increase in cumulative average stress equivalent to one severe stressor or two moderate stressors doubled the risk for AIDS. At 7.5 years, 47% of those above the median for stressful life events had progressed to AIDS, versus 27% of those below the median. Higher levels of serum cortisol were also associated with faster progression to AIDS, but variations in cortisol did not account for the stress findings (196).

Other studies also lend support to the hypothesis that stressful events may hasten the progression of HIV infection. In the study of Kemeny and Dean (197), the stress of bereavement before study entry was associated with a more rapid decline in $CD4^+$ count during 3 to 4 years of follow-up in 85 gay men. Bereavement did not predict progression to AIDS or mortality rate. Ironson et al. (198) found that in men with whose distress was greater at the time of HIV serostatus notification, the likelihood for the development of HIV-related clinical symptoms at 2-year follow-up was greater. In a recent study of 67 asymptomatic HIV-infected African-American women, trauma (e.g., death of child, assault, rape), particularly among those with posttraumatic stress disorder, was associated with a greater decrease in the $CD4^+/CD8^+$ ratio during 1 year of follow-up (199).

It is noteworthy that studies in which relatively short follow-up periods and questionnaire methods are used to assess life stress have generally not shown an association of stress with reduction in $CD4^+$ T-lymphocyte counts. Studies that examine actual stressors (e.g., bereavement) or interview-based contextual ratings of cumulative stressful events are more likely to show such results than studies based on questionnaire assessments of stress.

Other psychosocial variables (social support and coping) have been linked to HIV disease progression. Recently, Leserman and colleagues reported faster progression to AIDS during 5.5 and 7.5 years of follow-up in men with less cumulative average social support satisfaction and more cumulative average denial (194 ,196). These findings echo those of some earlier research showing potentially harmful effects of denial and potentially beneficial effects of social support (200 ,201 and 202). In the study of Antoni and colleagues (203), HIV-infected gay men scoring above, rather than below, the median on passive coping strategies (e.g., denial, behavioral and mental disengagement) had lower $CD4^+/CD8^+$ ratios and lymphocyte proliferative responses to phytohemagglutinin at 3 weeks and 1 year after serostatus notification. An increase in denial from before to after serostatus notification was also associated with a greater probability of development of symptoms and AIDS during a 2-year follow-up study of gay men (198). In the study of Solano and colleagues (201) of 100 male and female HIV-infected subjects, those who became symptomatic after 1 year had shown more denial and less “fighting spirit” at baseline.

The findings of other studies regarding the effects of social support have been less consistent. Larger social networks

and greater emotional support predicted longer survival during 5 years in men who were symptomatic or had AIDS; however, larger social networks were associated with faster progression to AIDS in those who were asymptomatic at entry (202). Loneliness was associated with a more rapid decline in $CD4^+$ levels but was unrelated to AIDS or mortality during 3 years of follow-up in 205 symptomatic HIV-infected men (204). Other prospective studies have reported no significant associations of social support with HIV disease progression or decline in $CD4^+$ T lymphocytes (205).

In summary, the evidence is substantial that psychosocial factors such as depression and stressful life events may adversely affect disease progression in persons infected with HIV. It must be noted that most of the cited studies of psychosocial moderators of HIV infection have been conducted in men, primarily before the advent of protease inhibitors. Therefore, we need additional studies of women and patients currently on HAART.

CONCLUSION

Part of "90 - Neuropsychiatric Manifestations of Hiv-1 Infection and Aids "

Considerable preclinical and clinical research has been conducted in an effort to describe the neuropsychiatric manifestations of HIV-1 disease and increase our understanding of its underlying neuropathologic mechanisms. The virus enters the CNS early in the course of disease and causes both direct and indirect CNS effects. Subtle abnormalities can be detected on pathologic, neuroimaging, and neuropsychological studies before the onset of AIDS-defining illnesses, although the clinical significance of these findings continues to be unclear. In symptomatic AIDS, neuropsychiatric and neurologic complications are prevalent, and these can often be among the first manifestations of AIDS.

Since the earliest years of the HIV epidemic, most persons infected with HIV-1 have coped well. Major depression continues to be the most prevalent common psychiatric diagnosis in HIV-1-seropositive men; the prevalence is high compared with epidemiologic estimates in the general population, but it is similar to that in seronegative gay men and no higher than that in patients with other serious medical illnesses. The interrelationships between the CNS, endocrine system, and immune system are being actively investigated. Moreover, recent studies suggest that stress and depression may adversely affect immune function and accelerate the progression of HIV-1 disease. However, these studies require confirmation by comprehensive, longitudinal investigations in which similar methodologies are used. Future study is also necessary to increase our understanding of the neuropsychiatric manifestations of HIV-1 infection in women and its special effects on neurologic development in infants and children.

Recent controlled trials of psychopharmacologic treatment have yielded positive results for the alleviation of depression, and preliminary evidence also indicates a reduction in neurocognitive impairment. Future neuropsychopharmacologic approaches will likely focus on both direct and indirect effects of HIV-1 in the brain in an effort to develop novel interventions that may alter the course of disease and symptomatic treatments to improve clinical outcome and quality of life. The long-term impact of HAART on HIV-related CNS disease and associated neuropsychiatric manifestations will also be extensively studied.

ACKNOWLEDGMENTS

Part of "90 - Neuropsychiatric Manifestations of Hiv-1 Infection and Aids "

The authors thank Carol Roberts, B.S.N., for editorial assistance in the preparation of this manuscript.

DISCLOSURE

Part of "90 - Neuropsychiatric Manifestations of Hiv-1 Infection and Aids "

Dr. Evans has received research support from SmithKline Beecham and serves as a consultant to a number of pharmaceutical companies, including Abbott Laboratories, Eli Lilly, Janssen Pharmaceutica, Organon, Pfizer, SmithKline Beecham, TAP Pharmaceuticals, Wyeth-Ayerst Laboratories, and Forest Laboratories.

REFERENCES

1. World Health Organization. *UNAIDS*. Geneva: World Health Organization, 1999.

2. Atkinson JH, Grant I. Natural history of neuropsychiatric manifestations of HIV disease. *Psychiatr Clin North Am* 1994;17:17–33.

3. Aylward EH, Henderer JD, McArthur JC, et al. Reduced basal ganglia volume in HIV-1-associated dementia: results from quantitative neuroimaging. *Neurology* 1993;43:2099–2104.

4. Gray F, Belec L, Keohane C, et al. Zidovudine therapy and HIV encephalitis: a 10-year neuropathological survey. *AIDS* 1994;8:489–493.

5. Navia BA, Cho E-S, Petito CK, et al. The AIDS dementia complex: II. Neuropathology. *Ann Neurol* 1986;19:525–535.

6. Aylward EH, Brettschneider PD, McArthur JC, et al. Magnetic resonance imaging measurement of gray matter volume reductions in HIV dementia. *Am J Psychiatry* 1995;152:987–994.

7. van Gorp WG, Mandelkern M, Gee M, et al. Cerebral metabolic dysfunction in AIDS: findings in an AIDS sample with and without dementia. *J Neuropsychiatry Clin Neurosci* 1992;4:280–287.

8. Butters N, Grant I, Haxby J, et al. Assessment of AIDS-related cognitive changes: recommendations of the NIMH workshop on neuropsychological assessment approaches. *J Clin Exp Neuropsychol* 1990;12:963–978.

9. Heaton RK, Grant I, Butters N, et al., and the HIV Neurobehavioral Research Center Group. The HNRC 500—neuropsychology of HIV infection at different disease stages. *J Int Neuropsychol Soc* 1995;1:231–251.

10. Llorente AM, Miller EN, D'Elia LF, et al. Slowed information processing in HIV-1 disease. *J Clin Exp Neuropsychol* 1998;20:60–72.

11. Maj M, Satz P, Janssen R, et al. WHO neuropsychiatric AIDS study, cross-sectional phase II: neuropsychological and neurological findings. *Arch Gen Psychiatry* 1994;51:51–61.

12. Peavy G, Jacobs D, Salmon DP, et al. Verbal memory performance of patients with human immunodeficiency virus infection: evidence of subcortical dysfunction. The HIV Neurobehavioral Research Center Group. *J Clin Exp Neuropsychol* 1994;16:508–523.

13. Stout JC, Salmon DP, Butters N, et al. Decline in working memory associated with HIV infection. The HIV Neurobehavioral Research Center Group. *Psychol Med* 1995;25:1221–1232.

14. Bornstein RA, Nasrallah HA, Para MF, et al. Duration of illness and neuropsychological performance in asymptomatic HIV infection. *J Neuropsychiatry Clin Neurosci* 1994;6:160–164.

15. Bornstein RA, Nasrallah HA, Para MF et al. Neuropsychological performance in asymptomatic HIV infection. *J Neuropsychiatry Clin Neurosci* 1992;4:386–394.

16. Grant I, Atkinson JH, Hesselink JR et al. Evidence for early central nervous system involvement in the acquired immunodeficiency syndrome (AIDS) and other human immunodeficiency virus (HIV) infections. Studies with neuropsychological testing and magnetic resonance imaging. *Ann Intern Med* 1987;107:828–836.

17. Marder K, Stern Y, Malouf R, et al. Neurologic and neuropsychological manifestations of human immunodeficiency virus infection in intravenous drug users without acquired immunodeficiency syndrome. *Arch Neurol* 1992;49:1169–1175.

18. Marsh NV, McCall DW. Early neuropsychological change in HIV infection. *Neuropsychology* 1994;8:44–48.

19. Wellman MC. Neuropsychological impairment among intravenous drug users in AIDS stages of HIV infection. *Int J Neurosci* 1992;64:183–194.

20. Wilkie F, Eisdorfer C, Morgan R, et al. Cognition in early HIV infection. *Arch Neurol* 1990;47:433–440.

21. Wilkie FL, Morgan R, Fletcher MA, et al. Cognition and immune function in HIV-1 infection. *AIDS* 1992;6:977–981.

22. Stern RA, Perkins DO, Evans DL. Neuropsychiatric manifestations of HIV-1 infection and AIDS. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:1545–1558.

23. Miller EN, Selnes OA, McArthur JC, et al. Neuropsychological performance in HIV-1-infected homosexual men: The Multicenter AIDS Cohort Study (MACS). *Neurology* 1990;40:197–203.

24. McArthur JC, Cohen BA, Selnes OA, et al. Low prevalence of neurological and neuropsychological abnormalities in otherwise healthy HIV-1 infected individuals: results from the Multicenter AIDS Cohort Study. *Ann Neurol* 1989;26:601–611.

25. Royal III, W, Updike M, Selnes OA, et al. HIV-1 infection and nervous system abnormalities among a cohort of intravenous drug users. *Neurology* 1991;41:1905–1910.

26. Selnes OA, Miller EN, McArthur J, et al. HIV-1 infection: no evidence of cognitive decline during the asymptomatic stages. *Neurology* 1990;40:204–208.

27. Selnes OA, McArthur JC, Royal W III, et al. HIV-1 infection and intravenous drug use: longitudinal neuropsychological evaluation of asymptomatic subjects. *Neurology* 1992;42:1924–1930.

28. Selnes OA, Galai N, McArthur JC, et al. HIV infection and cognition in intravenous drug users: long-term follow-up. *Neurology* 1997;48:223–230.

29. Sinforiani E, Mauri M, Bono G, et al. Cognitive abnormalities and disease progression in a selected population of asymptomatic HIV-positive subjects. *AIDS* 1991;5:1117–1120.

30. Stern RA, Arruda JE, Somerville JA, et al. Neurobehavioral functioning in asymptomatic HIV-1 infected women. *J Int Neuropsychol Soc* 1998;4:172–178.

31. Stern RA, Singer N, Silva SG, et al. Neurobehavioral functioning in a nonconfounded group of asymptomatic HIV-seropositive homosexual men. *Am J Psychiatry* 1992;149:1099–1102.

32. Swanson B, Bieliauskas LA, Kessler HA, et al. Infrequent neuropsychological impairment in asymptomatic persons infected with the human immunodeficiency virus. *Clin Neuropsychol* 1991;5:183–189.

33. van Gorp WG, Satz P, Hinkin C, et al. Metacognition in HIV-1 seropositive asymptomatic individuals: self-ratings versus objective neuropsychological performance. *J Clin Exp Neuropsychol* 1991;13:812–819.

34. Grant I, Heaton RK. Human immunodeficiency virus-type 1 (HIV-1) and the brain. *J Consult Clin Psychol* 1990;58:22–30.

35. Mapou RL, Law WA, Martin A, et al. Neuropsychological performance, mood, and complaints of cognitive and motor difficulties in individuals with the human immunodeficiency virus. *J Neuropsychiatry Clin Neurosci* 1993;5:86–93.

36. White DA, Heaton RK, Monsch AU, and the HIV Neurobehavioral Research Center Group. Neuropsychological studies of asymptomatic human immunodeficiency virus-type 1 infected individuals. *J Int Neuropsychol Soc* 1995;1:304–315.

37. Newman SP, Lunn S, Harrison MJG. Do asymptomatic HIV-seropositive individuals show cognitive deficit? *AIDS* 1995;9:1211–1220.

38. van Gorp WG, Hinkin C, Satz P, et al. Subtypes of HIV-related neuropsychological functioning: a cluster analysis approach. *Neuropsychology* 1993;7:62–72.

39. Law WA, Martin A, Mapou RL, et al. Working memory in individuals with HIV infection. *J Clin Exp Neuropsychol* 1994;16:173–182.

40. Martin EM, Robertson LC, Edelstein HE, et al. Performance of patients with early HIV-1 infection on the Stroop Task. *J Clin Exp Neuropsychol* 1992;14:857–868.

41. Sorenson DJ, Martin EM, Robertson LC. Visual attention in HIV-1 infection. *Neuropsychology* 1994;8:424–432.

42. Hinkin CH, Castellon SA, Hardy DJ, et al. Neuropsychological predictors of medication adherence in HIV/AIDS: a preliminary report. *J Int Neuropsychol Soc* 2000;6:135.

43. Baddeley A. *Human memory: theory and practice*. London: Erlbaum, 1990.

44. Rabbitt P. Introduction: methodologies and models in the study of executive function. In: Rabbitt P, ed. *Methodology of frontal and executive function*. Brighton: Psychology Press, 1997:1–38.

45. Castellon SA, Hinkin CH, Wood S, et al. Apathy, depression, and cognitive performance in HIV-1 infection. *J Neuropsychiatry Clin Neurosci* 1998;10:320–329.

46. Bartok JA., Martin EM, Pitrak DL, et al. Working memory deficits in HIV-seropositive drug users. *J Int Neuropsychol Soc* 1997;3:451–456.

47. Martin EM, Pitrak DL, Pursell KJ, et al. Delayed recognition memory span in HIV-1 infection. *J Int Neuropsychol Soc* 1995;1:575–580.

48. Martin EM, Cook JA, Cohen M, et al. Working memory function in HIV-seropositive women: the women's interagency HIV study. *J Int Neuropsychol Soc* 1999;5:155.

49. Sahakian BJ, Elliott R, Low N, et al. Neuropsychological deficits in tests of executive function in asymptomatic and symptomatic HIV-1 seropositive men. *Psychol Med* 1995;25:1233–1246.

50. Butters N, Tarlow S, Cermak LS, et al. A comparison of the information-processing deficits of patients with Huntington's chorea and Korsakoff's syndrome. *Cortex* 1976;12:134–144.

51. Delis DC, Peavy G, Heaton R, et al. Do patients with HIV-associated minor cognitive/motor disorder exhibit a "subcortical" memory profile? Evidence using the California Verbal Learning Test. *Assessment* 1995;2:151–165.

52. Becker JT, Dew MA, Lopez OL. Neuropsychological test performance among HIV-infected women seeking primary medical care. *Proceedings of the National Conference on Women and HIV* 1997;[P1.19]:160(abst).

53. Tross S, Price RW, Navia B, et al. Neuropsychological characterization of the AIDS dementia complex: a preliminary report. *AIDS* 1998;2:81–88.

54. Marder K, Liu X, Stern Y, et al. Neurologic signs and symptoms in a cohort of homosexual men for 4.5 years. *Neurology* 1995;45:261–267.

55. Miller EN, Satz P, Visscher B. Computerized and conventional neuropsychological assessment of HIV-1 infected homosexual men. *Neurology* 1991;41:1608–1616.

56. Ellis RJ, Deutsch R, Heaton RK, et al., and the San Diego HIV Neurobehavioral Research Center Group. Neurocognitive impairment is an independent risk factor for death in HIV infection. *Arch Neurol* 1997;54:416–424.

57. Mayeux R, Stern Y, Tang MX, et al. Mortality risks in gay men with human immunodeficiency virus infection and cognitive impairment. *Neurology* 1993;43:176–182.

58. Wilkie FL, Goodkin K, Eisdorfer C, et al. Mild cognitive impairment and risk of mortality in HIV-1 infection. *J Neuropsychiatry Clin Neurosci* 1998;10:125–132.

59. Bouwman FH, Skolasky RL, Hes D, et al. Variable progression of HIV-associated dementia. *Neurology* 1998;50:1814–1820.

60. Navia BA, Jordan BD, Price RW. The AIDS dementia complex: I. Clinical features. *Ann Neurol* 1986;19:517–524.

61. World Health Organization. World Health Organization consultation on the neuropsychiatric aspects of HIV-1 infection. *AIDS* 1990;4:935–936.

62. American Academy of Neurology. Nomenclature and research case definitions for neurologic manifestations of human immunodeficiency virus-type 1 (HIV-1) infection. *Neurology* 1991;41:778–785.

63. Day JJ, Grant I, Atkinson JH, et al. Incidence of dementia in a 2-year follow-up on AIDS and ARC patients on an initial phase II AZT placebo-controlled study: San Diego cohort. *J Neuropsychiatry* 1992;4:15–20.

64. Janssen RS, Nwanyanwu OC, Selik RM, et al. Epidemiology of human immunodeficiency virus encephalopathy in the United States. *Neurology* 1992;42:1472–1476.

65. McArthur JC, Hoover DR, Bacellar H, et al. Dementia in AIDS patients: incidence and risk factors. *Neurology* 1993;26:601–611.

66. Price R, Brew B. The AIDS dementia complex. *J Infect Dis* 1988;158:1079–1083.

67. Centers for Disease Control and Prevention. *HIV/AIDS Surveillance Report* 1997;9:18.

68. Cummings JL. Frontal-subcortical circuits and human behavior. *Arch Neurol* 1993;50:873–880.

69. Fernandez F, Levy JK, Ruiz P. The use of methylphenidine in HIV patients: a clinical perspective. In: Grant I, Martin A. eds. *Neuropsychology of HIV Infection*. New York: Oxford University Press, 1994:295–309.

70. McKusick L. The impact of AIDS on practitioner and client. *Am Psychol* 1988;43:935–940.

71. Atkinson J, Grant I, Kennedy C, et al. Prevalence of psychiatric disorders among men infected with human immunodeficiency virus: a controlled study. *Arch Gen Psychiatry* 1988;45:859–964.

72. Perry SW. Organic mental disorders caused by HIV: update on early diagnosis and treatment. *Am J Psychiatry* 1990;147:696–710.

73. Perkins DO, Stern RA, Golden R, et al. Mood disorders in HIV infection: prevalence and risk factors in a non-epicenter of the AIDS epidemic.*Am J Psychiatry* 1994;151:233–236.

74. Rabkin JG. Prevalence of psychiatric disorders in HIV illness. *Int Rev Psychiatry* 1996;8:157–166.

75. Kalichman SC, Sikkema KJ, Sandmlai A. Assessing persons with human immunodeficiency virus (HIV) infection using the Beck Depression Inventory: disease processes and other potential confounds. *J Pers Assess* 1995;64:86–100.

76. Drebing CE, van Gorp WG, Hinkin C, et al. Confounding factors in the measurement of depression in HIV. *J Pers Assess* 1994;62:68–83.

77. Richardson MA, Satz P, Myers HF, et al. Effects of depressed mood versus clinical depression on neuropsychological test performance among African-American men impacted by HIV/AIDS. *J Clin Exp Neuropsychol* 1999;21:769–783.

78. Bornstein RA, Pace P, Rosenberger P, et al. Depression and neuropsychological performance in asymptomatic HIV infection. *Am J Psychiatry* 1993;150:922–927.

79. Claypoole KH, Elliott AJ, Uldall KK, et al. Cognitive functions and complaints in HIV-1 individuals treated for depression. *Appl Neuropsychol* 1998;5:74–84.

80. Goodwin GM, Pretsell DO, Chiswick A, et al. The Edinburgh cohort of HIV-positive injecting drug users at 10 years after infection: a case–control study of the evolution of dementia. *AIDS* 1996;10:431–440.

81. Goggin KJ, Zisook S, Heaton RK, et al., and the HIV Neurobehavioral Research Center Group. Neuropsychological performance of HIV-1 infected men with major depression. *J Int Neuropsychol Soc* 1997;3:457–464.

82. Grant I, Olshen RA, Atkinson JH, et al. Depressed mood does not explain neuropsychological deficits in HIV-infected persons. *Neuropsychology* 1993;7:53–61.

83. Harker JO, Satz P, Jones FD, et al. Measurement of depression and neuropsychological impairment in HIV-1 infection. *Neuropsychology* 1995;9:110–117.

84. Moore LH, van Gorp WG, Hinkin CH, et al. Subjective complaints versus actual cognitive deficits in predominantly symptomatic HIV-1 seropositive individuals. *J Neuropsychiatry Clin Neurosci* 1997;9:37–44.

85. Kalechstein AD, Hinkin CH, van Gorp WG, et al. Depression predicts procedural but not episodic memory in HIV-1 infection. *J Clin Exp Neuropsychol* 1998;20:529–535.

86. Webber MP, Schoenbaum EE, Gourevitch MN, et al. A prospective study of HIV disease progression in female and male drug users. *AIDS* 1999;13:257–262.

87. Claypoole KH, Townes BD, Collier AC, et al. Cognitive risk factors and neuropsychological performance in HIV infection. *Int J Neurosci* 1993;70:12–27.

88. Ardila A, Rosselli M, Strumwasser S. Neuropsychological deficits in chronic cocaine abusers. *Int J Neurosci* 1991;57:73–79.

89. O'Malley S, Adamse M, Heaton RK, et al. Neuropsychological impairments in chronic cocaine abusers. *Am J Drug Alcohol Abuse* 1992;18:131–144.

90. Strickland TL, Mena I, Villanueva-Meyer J, et al. Cerebral perfusion and neuropsychological consequences of chronic cocaine abuse. *J Neuropsychiatry Clin Neurosci* 1993;5:1–9.

91. Bolla KI, Cadet JL, London ED. The neuropsychiatry of chronic cocaine abuse. *J Neuropsychiatry Clin Neurosci* 1998;10:280–289.

92. Janssen RS, Saykin AJ, Cannon L, et al. Neurological and neuropsychological manifestations of HIV-1 in association with AIDS-related complex but not asymptomatic HIV infection. *Ann Neurol* 1988;26:592–600.

93. Bornstein RA, Fama R, Rosenberger P, et al. Drug and alcohol use and neuropsychological performance in asymptomatic HIV infection. *J Neuropsychiatry Clin Neurosci* 1993;5:254–259.

94. Bono G, Mauri M, Sinforiani E, et al. Longitudinal neuropsychological evaluation of HIV-infected intravenous drug users. *Addiction* 1996;91:263–268.

95. Selnes OA, Galai N, Bacellar H, et al. Cognitive performance after progression to AIDS: a longitudinal report from the Multicenter AIDS Cohort Study. *Neurology* 1995;45:267–275.

96. Centers for Disease Control and Prevention. Update: AIDS among women—United States. *MMWR Morb Mortal Wkly Rep* 1995;44:81–84.

97. Centers for Disease Control and Prevention. *HIV/AIDS Surveillance Report* 1999;11:5–37.

98. Ford-Booker P, Campbell A, Combs S, et al. The predictive accuracy of neuropsychological tests in a normal population of African-Americans. *J Clin Exp Neuropsychol* 1993;15:64.

99. Bing EG, Miller EN, Munoz A, et al. Sociodemographic factors influencing neuropsychological test results. *Proceedings of the seventh international conference on AIDS*, Florence, June 1991:79(abst).

100. Campbell A, Rorie K, Dennis G, et al. Neuropsychological assessment of African-Americans: conceptual and methodological considerations. In: Jones RL, ed. *Handbooks of tests and measurements for black populations* (vol 2). Hampton: Cobb and Henry, 1996.

101. Manly JJ, Miller SW, Heaton RK, et al., and the HIV Neurobehavioral Research Center Group. The effect of African-American acculturation on neuropsychological test performance in normal and HIV-positive individuals. *J Int Neuropsychol Soc* 1998;4:291–302.

102. Mason KI, Campbell A, Hawkins P, et al. Neuropsychological functioning in HIV seropositive African-American women with a history of drug use. *J Natl Med Assoc* 1998;90:665–674.

103. Bromberg J, Grijalva K, Skurnick J, et al. Psychologic differences between HIV+ women and HIV+ men in discordant couples: a report from the heterosexual HIV transmission study (HATS). *Proceedings of the seventh international conference on AIDS*, Florence, June 1991.

104. Franke GH, Jager H, Thomann B, et al. Assessment and evaluation of psychological distress in HIV-infected women. *Psychol Health* 1992;6:297–312.

105. Kelly PJ, Holman S. The new face of AIDS. *Am J Nursing* 1993;93:26–32.

106. Schroeder MM, Handelsman L, Torres L, et al. Early and late cognitive event-related potentials mark stages of infection in the drug-user risk group. *Biol Psychiatry* 1994;35:54–69.

107. Durvasula R, Miller EN, Wyatt GE. Longitudinal assessment of neuropsychological performance in HIV+ women (baseline and 6 months). *Proceedings of the international conference on AIDS*, 1999:470.

108. Costa L, Arruda JE, Stern RA, et al. Asymptomatic HIV-infected women: Preliminary study of quantitative EEG activity and performance on a continuous performance test. *Percept Mot Skills* 1997;85:1395–1408.

109. Satz P. Brain reserve capacity on symptom onset after brain injury: a formulation and review of evidence for threshold theory. *Neuropsychology* 1993;7:272–295.

110. Stern RA, Silva SG, Chaisson N, et al. Influence of cognitive reserve on neuropsychological functioning in asymptomatic human immunodeficiency virus-1 infection. *Arch Neurol* 1996;53:148–153.

111. Vanhove GF, Schapiro JM, Winters MA, et al. Patient compliance and drug failure in protease inhibitor monotherapy. *JAMA* 1996;276:1955–1956.

112. Albert SM, Weber C, Todak G, et al. An observed performance test of medication management ability in HIV: relation to neuropsychological status and adherence outcomes. Neuroscience of HIV infection. *J Neurovirol* 1998;4:338.

113. Durvasula RS, Golin C, Satz P, et al. Neuropsychological performance and antiretroviral therapy (ART) adherence in HIV-positive women. Presented at the 27th annual meeting of the International Neuropsychological Society, Boston, Massachusetts. *J Int Neuropsychol Soc* 1999;5:97(abst).

114. Heaton RK, Velin RA, McCutchan JA, et al., and the HIV Neurobehavioral Research Center Group. Neuropsychological impairment in human immunodeficiency virus-infection: implications for employment. *Psychosom Med* 1994;56:8–17.

115. van Gorp WG, Baerwald JP, Ferrando SJ, et al. The relationship between employment and neuropsychological impairment in HIV infection. *J Int Neuropsychol Soc* 1999;5:534–539.

116. Ferrando S, van Gorp W, McElhiney M, et al. Highly active antiretroviral treatment in HIV infection: benefits for neuropsychological function. *AIDS* 1998;12[Suppl]:F65–F70.

117. Heaton RK, Marcotte TD, White DA, et al. Nature and vocational significance of neuropsychological impairment associated with HIV infection. *Clin Neuropsychol* 1996;10:1–14.

118. Bell JE, Donaldson YK, Lowrie S, et al. Influence of risk group and zidovudine therapy on the development of HIV encephalitis and cognitive impairment in AIDS patients. *AIDS* 1996;10:493–499.

119. Schmitt FA, Bigley JW, McKinnis R, et al., and the AZT Collaborative Working Group. Neuropsychological outcome of zidovudine (AZT) treatment of patients with AIDS and AIDS-related complex. *N Engl J Med* 1988;319:1573–1578.

120. Sidtis JJ, Gatsonis C, Price RW, et al., and the AIDS Clinical Trials Group. Zidovudine treatment of the AIDS dementia complex: results of a placebo-controlled trial. *Ann Neurol* 1993;33:343–349.

121. Egan V, Deary IJ, Brettle RP. The Edinburgh cohort of HIV-positive drug users: the effects of depressed mood and drug use upon neuropsychological function. *Br J Health Psychol* 1996;1(Part 3):231–244.

122. Baldeweg T, Catalan J, Lovett E, et al. Long-term zidovudine reduces neurocognitive deficits in HIV-1 infection. *AIDS* 1995;9:589–596.

123. Sacktor NC, Lyles RH, Skolasky RL, et al., for the Multicenter AIDS Cohort Study (MACS). Combination antiretroviral therapy improves psychomotor speed performance in HIV-seropositive homosexual men. *Neurology* 1999;52:1640–1647.

124. Chang L, Ernst T, Leonido-Yee M, et al. Highly active antiretroviral therapy reverses brain metabolite abnormalities in mild HIV dementia. *Neurology* 1999;53:782–789.

125. Skolnick AA. Protease inhibitors may reverse AIDS dementia. *J Natl Med Assoc* 1998;279:419.

126. Reiger DA, Boyd JH, Rae DS, et al. One-month prevalence of mental disorders in the United States. *Arch Gen Psychiatry* 1988;45:977–986.

127. Robins LN, Helzer JE, Weissman MM, et al. Lifetime prevalence of specific psychiatric disorders in three sites. *Arch Gen Psychiatry* 1984;41:949–958.

128. Williams JBW, Rabkin JG, Remien RH, et al. Multidisciplinary baseline assessment of homosexual men with and without human immunodeficiency virus infection: standardized clinical assessment of current and lifetime psychopathology. *Arch Gen Psychiatry* 1991;48:124–130.

129. Evans DL, Perkins DO. The clinical psychology of AIDS. *Curr Opin Psychiatry* 1990;3:96–102.

130. Evans DL, McCartney CF, Nemeroff CB, et al. Depression in women treated for gynecological cancer: clinical and neuroendocrine assessment. *Am J Psychiatry* 1986;143:447–461.

131. Musselman DL, Evans DL, Nemeroff CB. The relationship of depression to cardiovascular disease: epidemiology, biology and treatment. *Arch Gen Psychiatry* 1998;55:580–592.

132. Evans DL, Staab J, Ward H, et al. Depression in the medically ill: management considerations. *Depression Anxiety* 1997;4:199–208.

133. Evans DL, Staab JP, Petitto JM, et al. Depression in the medical setting. *J Clin Psychiatry* 1999:60:40–55.

134. Moeller AA, Oechsner M, Backmund HC, et al. Self-reported sleep quality in HIV infection: correlation to the stage of infection and zidovudine therapy. *AIDS* 1991;4:1000–1003.

135. Norman SE, Chediak AD, Freeman C, et al. Sleep disturbances in men with asymptomatic human immunodeficiency (HIV) infection. *Sleep* 1992;15:150–155.

136. Baum S, Perkins DO, Gray-Silva S, et al. Asymptomatic HIV infection and insomnia. *Proceedings of the American Psychiatric Association*, New Research Program and Abstracts 1993(abst).

137. Perkins DO, Leserman J, Gray-Silva S, et al. Fatigue and HIV infection: relationship to depressive symptoms and indicators of HIV disease. *Proceedings of the American Psychiatric Association*, New Research Program and Abstracts, 1993(abst).

138. Leserman J, Perkins DO, Evans DL. Coping with the threat of AIDS: the role of social support. *Am J Psychiatry* 1992;149:1514–1520.

139. Deleted in press.

140. Rabkin JG, Ferrando SJ, Jacobsberg LB, et al. Prevalence of axis I disorders in an AIDS cohort: a cross-sectional controlled study. *Compr Psychiatry* 1997;38:146–154.

141. Boland RJ, Moore J, Solomon L, et al. Depression and HIV in women: prevalence and associations in a national cohort. Presented at the 42nd annual meeting of the Academy of Psychosomatic Medicine, Palm Springs, CA, 1999.

142. Morrison MF, Petitto J, TenHave T, et al. Prevalence of depressive and anxiety disorders in women with HIV infection. American College of Neuropsychopharmacology 1998.

143. Evans DL, Leserman J, Perkins DO, et al. Stress-associated reductions of cytotoxic T lymphocytes and natural killer cells in asymptomatic HIV infection. *Am J Psychiatry* 1995;152:543–550

144. Lyketsos CG, Hoover DR, Guccione M. Depression and survival among HIV-infected persons. *JAMA* 1996;275:35–36.

145. Lyketsos CG, Hanson AL, Fishman M, et al. Manic syndrome early and late in the course of HIV. *Am J Psychiatry* 1993;150:326–327.

146. Evans DL, Edelsohn G, Golden R. Organic psychoses without anemia or spinal cord symptoms in vitamin B_{12}-deficient patients. *Am J Psychiatry* 1983;140:218–221.

147. Robertson KR, Stern RA, Hall CD, et al. Vitamin B_{12} deficiency and nervous system disease in HIV infection. *Arch Neurol* 1993;50:807–811.

148. Navia BA, Price RW. The acquired immunodeficiency syndrome dementia complex as presenting or sole manifestation of human immunodeficiency virus infection. *Arch Neurol* 1987;44:65–69.

149. Sewell DD, Jeste DV, Atkinson JH, et al. HIV-associated psychosis: a study of 20 cases. San Diego HIV Neurobehavioral Research Center Group. *Am J Psychiatry* 1994;151:237–242.

150. Rabkin JG, Rabkin R, Harrison W, et al. Effect of imipramine on mood and enumerative measures of status in depressed patients with HIV illness. *Am J Psychiatry* 1994;151:516–523.

151. Elliott AJ, Uldall KK, Bergam K, et al. Randomized, placebo-controlled trial of paroxetine versus imipramine in depressed HIV-positive outpatients. *Am J Psychiatry* 1998;155:367–372.

152. Zisook S, Peterkin J, Goggin KJ, et al. Treatment of major depression in HIV-seropositive men. HIV Neurobehavioral Research Center Group. *J Clin Psychiatry* 1998;59:217–224.

153. Rabkin JG, Wagner G, Rabkin R. Fluoxetine treatment for depression in patients with HIV and AIDS: a randomized, placebo-controlled trial. *Am J Psychiatry* 1999;156:101–107.

154. Rabkin JG, Rabkin R, Wagner G. Effects of fluoxetine on mood and immune status in depressed patients with HIV illness. *J Clin Psychiatry* 1994;55:92–97.

155. Rabkin JG, Wagner G, Rabkin R. Effects of sertraline on mood and immune status in patients with depression and HIV illness: an open trial. *J Clin Psychiatry* 1994;55:433–439.

156. Ferrando SJ, Goldman JD, Charness WE. Selective serotonin reuptake inhibitor treatment of depression in symptomatic HIV infection and AIDS: improvement in affective and somatic symptoms. *Gen Hosp Psychiatry* 1997;19:89–97.

157. Ferrando SJ, Rabkin JG, de Moore GM, et al. Antidepressant treatment of depression in HIV-seropositive women. *J Clin Psychiatry* 1999;60:741–746.

158. Grassi B, Gambini O, et al. Efficacy of paroxetine for treatment of depression in the context of HIV infection. *Pharmacotherapy* 1997;30:70–71.

159. Elliott AJ, Karina KK, Bergam K, et al. Antidepressant efficacy in HIV-seropositive outpatients with major depressive disorder: an open trial of nefazodone. *J Clin Psychiatry* 1999;60:226–231.

160. Fernandez F, Levy JK, Samley HR, et al. Effects of methylphenidate in HIV-related depression: a comparative trial with desipramine. *Int J Psychiatry Med* 1995;25:53–67.

161. Wagner GJ, Rabkin JG, Rabkin R. Dextroamphetamine as a treatment for depression and low energy in AIDS patients: a pilot study. *J Psychosom Res* 1997;42:407–411.

162. Rabkin JG, Ferrando SJ, Wagner G, et al. A double-blind placebo-controlled trial of testosterone for HIV-positive men with hypogonadal symptoms. *Arch Gen Psychiatry* 2000;57:141–147.

163. Wagner G, Rabkin J, Rabkin R. Exercise as a mediator of psychological and nutritional effects of testosterone therapy in HIV+ men. *Med Sci Sports Exercise* 1998;30:811–817.

164. Wagner GJ, Rabkin JG, Rabkin R. A comparative analysis of standard an alternative antidepressants in the treatment of human immunodeficiency virus patients. *Compr Psychiatry* 1996;37:402–408.

165. Breitbart W, Rosenfeld BD, Passick SD, et al. The undertreatment of pain in ambulatory AIDS patients. *Pain* 1996;65:243–249.

166. Fernandez F, Levy JK, Mansell PW. Response to antidepressant therapy in depressed persons with advanced HIV infection. Presented at the International Conference on AIDS, Montreal, June 4–9, 1989(abst).

167. Satel SL, Nelson JC. Stimulants in the treatment of depression: a critical overview. *J Clin Psychiatry* 1989;50:241–249.

168. Hriso E, Kuhn T, Masdeu JC, et al. Extrapyramidal symptoms due to dopamine blocking agents in patients with AIDS encephalopathy. *Am J Psychiatry* 1991;148:1558–1561.

169. Singh AN, Golledge H, Catalan J. Treatment of HIV-related psychotic disorders with risperidone: a series of 21 cases. *J Psychosom Res* 1997;42:489–493.

170. Meyer JM, Marsh J, Simpson G. Differential sensitivities to risperidone and olanzapine in a human immunodeficiency virus patient. *Biol Psychiatry* 1998;44:791–794.

171. Tseng AL, Foisy MM. Significant interactions with new antiretrovirals and psychotropics. *Ann Pharmacother* 1999;33:461–473.

172. Priscitelli SC, Burstein AH, Chaitt D, et al. Indinavir concentrations and St. John's wort. *Lancet* 2000;355:547–548.

173. Evans DL, Leserman J, Golden RN, et al. Immune correlates of stress and depression. *PsychopharmacoI Bull* 1989;25:319–324.

174. Stein M, Miller AH, Trestman RL. Depression, the immune system, and health and illness. *Arch Gen Psychiatry* 1991;48:171–177.

175. Herbert TB, Cohen S. Stress and immunity in humans: a meta-analytic review. *Psychosom Med* 1993;55:364–379.

176. Herbert TB, Cohen S. Depression and immunity: a meta-analytic review. *Psychol Bull* 1993;113:472–486.

177. Fawzy FI, Fawzy NW, Hyun CS, et al. Malignant melanoma. Effects of an early structured psychiatric intervention, coping, and affective state on recurrence and survival 6 years later. *Arch Gen Psychiatry* 1993;50:681–689.

178. Spiegel D, Kraemer HC, Bloom JR, et al. Effect of psychosocial treatment on survival of patients with metastatic breast cancer. *Lancet* 1989;2:888–891.

179. Ramirez AJ, Craig TJK, Watson JP, et al. Stress and relapse of breast cancer. *Br Med J* 1989;298:291–293.

180. Kiecolt-Glaser JK, Glaser R. Psychoneuroimmunology and cancer: fact or fiction? *Eur J Cancer* 1999;35:1603–1607.

181. Fawzy FI, Kemeny ME, Fawzy NW, et al. A structured psychiatric intervention for cancer patients. II. Changes over time in immunological measures. *Arch Gen Psychiatry* 1990;47:729–735.

182. Rabkin JG, Williams JBW, Reimien RH, et al. Depression, distress, lymphocyte subsets, and human immunodeficiency virus symptoms on two occasions in HIV-positive homosexual men. *Arch Gen Psychiatry* 1991;48:111–119.

183. Perry S, Fishman B, Jacobsberg L, et al. Relationship over 1 year between lymphocyte subsets and psychosocial variables among adults with infection by human immunodeficiency virus. *Arch Gen Psychiatry* 1992;49:396–401.

184. Kessler RC, Foster C, Joseph J, et al. Stressful life events and symptom onset in HIV infection. *Am J Psychiatry* 1991;148:733–738.

185. Zorilla EP, McKay JR, Luborsky L, et al. Relation of stressors and depressive symptoms to clinical progression of viral illness. *Am J Psychiatry* 1996;153:626–635.

186. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Measures* 1977;1:386–401

187. Page-Schaefer K, Delorenze GN, Satariano W, et al. Comorbidity and survival in HIV-infected men in the San Francisco Men's Health Survey. *Ann Epidemiol* 1996;6:420–430.

188. Burack JH, Barrett DC, Stall RD, et al. Depressive symptoms and CD4 lymphocyte decline among HIV-infected men. *JAMA* 1993;270:2568–2573.

189. Mayne TJ, Vittinghoff E, Chesney MA, et al. Depressive affect and survival among gay and bisexual men infected with HIV. *Arch Intern Med* 1996;156:2233–2238.

190. Lyketsos CG, Hoover DR, Guccione M, et al. Depressive symptoms as predictors of medical outcomes in HIV infection. *JAMA* 1993;270:2563–2567.

191. Lyketsos CG, Hoover DR, Guccione M, et al. Changes in depressive symptoms as AIDS develops. *Am J Psychiatry* 1996;153:1430–1437.

192. Zimmerman M, Coryell W. The validity of a self-report questionnaire for diagnosing major depressive disorder. *Arch Gen Psychiatry* 1988;45:738–740.

193. Leserman J, Petitto JM, Perkins DO, et al. Severe stress, depressive symptoms, and changes in lymphocyte subsets in human immunodeficiency virus-infected men. *Arch Gen Psychiatry* 1997;54:279–285.

194. Leserman J, Jackson ED, Petitto JM, et al. Progression to AIDS: the effects of stress, depressive symptoms, and social support. *Psychosom Med* 1999;61:397–406.

195. Evans DL, Leserman J, Perkins DO, et al. Severe life stress as a predictor of early disease progression in HIV infection. *Am J Psychiatry* 1997;154:630–634.

196. Leserman J, Petitto JM, Golden RN, et al. Impact of stressful life events, depression, social support, coping and cortisol on progression to AIDS. *Am J Psychiatry* 2000;157:1221–1228.

197. Kemeny ME, Dean L. Effects of AIDS-related bereavement on HIV progression among New York City gay men. *AIDS Education Prevention* 1995;7:36–47.

198. Ironson G, Friedman A, Klimas N, et al. Distress, denial, and low adherence to behavioral interventions predict faster disease progression in gay men infected with human immunodeficiency virus. *Int J Behav Med* 1994;1:90

91

Dementia with Lewy Bodies

Ian G. McKeith

David Burn

John O'Brien

Robert Perry

Elaine Perry

Ian G. McKeith and John O'Brien: Institute for the Health of the Elderly, Wolfson Research Centre, Newcastle General Hospital, Newcastle upon Tyne, United Kingdom.

David Burn: Department of Neurology, Regional Neurosciences Centre, Newcastle upon Tyne, United Kingdom.

Robert Perry: Department of Neuropathology, Newcastle General Hospital, Newcastle upon Tyne, United Kingdom.

Elaine Perry: MRC Neurochemical Pathology Unit, Newcastle General Hospital, Newcastle upon Tyne, United Kingdom.

The concept of dementia with Lewy bodies (DLB) has been slowly gaining momentum since 1961, when Okazaki et al. (1) published case reports about two male patients, ages 69 and 70 years, who presented with dementia and died shortly thereafter with severe extrapyramidal rigidity. Autopsy showed Lewy body pathology in the cerebral cortex. During the next 20 years, a total of 34 such cases was reported, all by Japanese workers, who adopted the term *diffuse Lewy body disease* to describe the typical distribution of Lewy bodies in subcortical and cortical regions (2).

During the following decade, cortical Lewy body pathology was found in up to 20% of all cases of elderly demented patients reaching autopsy (3 ,4 and 5), and in 1990, Hansen et al. (6) reported that 36% of patients given a clinical diagnosis of Alzheimer disease (AD) had Lewy bodies at autopsy—the Lewy body variant of AD (6). The significance of these reports was unclear—whether this was a new type of dementing disorder that had not previously existed, or whether the Lewy bodies had previously been overlooked. Reexamination of original material from a cohort of autopsy material collected during the 1960s in Newcastle upon Tyne revealed that 17% of the cases had cortical Lewy bodies, a prevalence similar to that found today. It appears, then, that DLB is not a new disorder but is one that has only recently been recognized, even though it is the second most common form of degenerative dementia in old age, only AD being more common.

Lewy bodies are spherical, intracytoplasmic, eosinophilic, neuronal inclusions; they have a dense hyaline core and a halo of radiating filaments composed of abnormally truncated and phosphorylated intermediate neurofilament proteins that include ubiquitin and associated enzymes. They were first described by the German neuropathologist Friederich Lewy when he was working in Alzheimer's laboratory in Munich between 1910 and 1912. Subcortical Lewy bodies are easily seen with conventional hematoxylin and eosin staining. The presence of Lewy bodies in pigmented brainstem nuclei (the substantia nigra in particular), coupled with neuronal loss and gliosis, comprise the characteristic pathologic findings in the prototypal Lewy body disease, which is Parkinson disease (PD).

Cortical Lewy bodies lack the characteristic core and halo appearance of their brainstem counterparts and were therefore difficult to detect until the late 1980s, when the development of anti-ubiquitin immunocytochemical staining methods allowed their true prevalence to be appreciated (4). More recently, α-synuclein antibodies have been shown to label purified Lewy bodies, and the α-synuclein antibodies PER1 and PER2 strongly stain Lewy bodies and Lewy neurites (7 ,8) (Fig. 91.1). Axon pathology in DLB involves not only α-synuclein but also ß- and γ-synucleins (9). α-Synuclein antibodies reveal presynaptic axon pathology in various regions of the hippocampus, and γ-synuclein antibodies detect axonal spheroid-like inclusions in the dentate molecular layer.

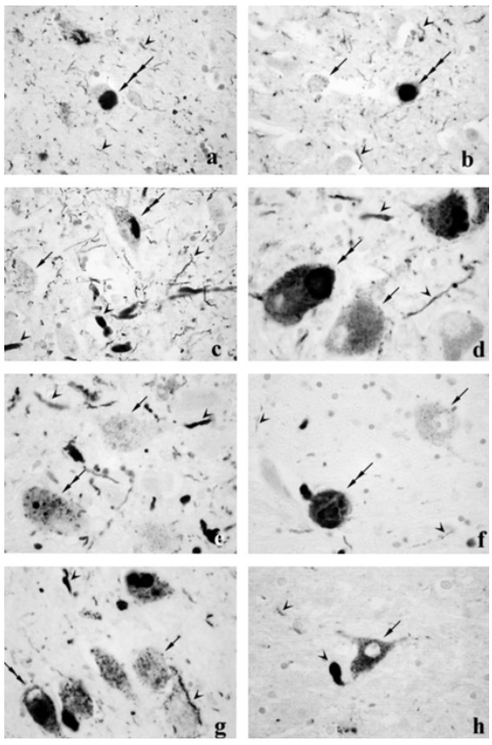


FIGURE 91.1. Examples of α-synuclein pathology from cases with Parkinson disease or dementia with Lewy bodies. Immunohistochemistry was performed with α-synuclein monoclonal antibodies from Novocastra Laboratories on formalin-fixed, paraffin-embedded sections pretreated with formic acid, Vectastain Elite ABC peroxidase kit, diaminobenzidine tetrahydrochloride (DAB), and hematoxylin counterstain. *Single arrow*, granular neuronal inclusions. *Double arrow*, combined granular and solid irregular neuronal inclusions. *Triple arrow*, Lewy body-like neuronal inclusions. *Arrowhead*, thin and thick neurites. Original magnification of A through H was ×63. **A:** Transentorhinal cortex. **B:** Anterior cingulate cortex. **C:** Hippocampus, CA2 segment. **D:** Substantia nigra in lower midbrain. **E:** Dorsal raphe in lower midbrain. **F:** Pedunculopontine nucleus in lower midbrain. **G:** Nucleus basalis of Meynert. **H:** Thalamus, central lateral nucleus. See color version of figure.

Current opinions on the classification of Lewy body disorders is that a spectrum of disease exists, with the clinical presentation varying according to the site of Lewy body formation and neuronal loss (10) (Table 91.1). Although “pure” presentations are seen in clinical practice, heterogenous combinations of parkinsonism, dementia, and signs of autonomic failure are most frequent. Recommendations have recently been made as to which brain regions should be examined for Lewy bodies, and a simple, semiquantitative scoring system has been devised in which a score of 1 is assigned if any Lewy bodies are seen in a given area and a score of 2 if more than five are seen per field (11). These

scores are then added to generate three possible pathologic categories—brainstem-predominant, limbic (or transitional), and neocortical DLB. It has not yet been established to what extent these three patterns of pathologic distribution correlate with different clinical profiles. Extensive neocortical pathology is not necessary for the development of dementia or other psychiatric symptoms, all of which may occur in the presence of limbic disease alone.

Region Primarily Affected	Clinical Syndrome	Classification
Substantia nigra	Extrapyramidal movement disorder	Parkinson disease
Limbic cerebral cortex	Cognitive decline and neuropsychiatric symptoms	Dementia with Lewy bodies
Sympathetic neurons in spinal cord	Autonomic failure	Primary autonomic failure
Dorsal vagal nuclei	Dysphagia	Lewy body dysphagia
Pedunculopontine nucleus[a]	Sleep disturbance	REM sleep behavior disorder

REM, rapid eye movement.
[a]Precise clinicopathologic correlate for this is yet to be established, although involvement of the pedunculopontine nucleus is highly probable from current data.
Adapted from Lowe JS, Mayer RJ, Landon M, Pathological significance of Lewy bodies in dementia; In: Perry R, McKeith I, Perry E, eds. *Dementia with Lewy bodies*. New York: Cambridge University Press, 1996:195–203.

TABLE 91.1. PRIMARY LEWY BODY DISORDERS

A heated debate has surrounded the interpretation of the Alzheimer-type changes that are also seen in most patients with DLB. Numerous senile plaques are found in most, although these are morphologically indistinguishable from those of pure AD (12). Two reports have appeared of a relative sparsity of ß-amyloid proteins 1 through 40 in DLB in comparison with AD (13 ,14). However, the plaques are seldom tau-immunoreactive, and indeed in 80% to 90% of cases of DLB, no evidence of significant neocortical tau pathology, paired helical filaments, or neurofibrillary tangles can be found (15). Whether or not DLB is considered to be a variant of AD depends on the pathologic definition of AD used (16). Thus, 77% of cases with Lewy body pathology and dementia had “plaque-only” AD, a concept derived from definitions of AD that depend heavily on plaque density. By contrast, 80% to 90% of DLB cases failed to fulfill definitions of AD that require numbers of neocortical neurofibrillary tangles above a certain threshold (17). The new NIA/Reagan Foundation criteria for AD appear to be responsible for a significant shift in this direction, with a proposed requirement for frequent neurofibrillary tangles equivalent to Braak stages 5 and 6 (18). DLB and pure AD are, according to such criteria, pathologically distinct in the majority of cases. Gomez-Isla et al. (19) recently concluded that the lack of a relationship between the extent of Alzheimer-type changes and Lewy body formation in DLB suggests that DLB is a distinct disease rather than a variant of AD.

In summary, at least three anchor points appear to be recognizable along a spectrum of neurodegenerative disorders. PD is a disorder of predominantly subcortical Lewy body neurofilament inclusions, which are the most visible markers of an extensive neuritic degeneration involving α-synuclein and ubiquitin. A more extensive distribution of Lewy bodies typifies DLB, in which significant ß-amyloidosis and senile plaque formation that fall short of what is seen in AD are also usually present. AD, in contrast, is characterized by a combination of ß-amyloidosis and neocortical neurofibrillary tangles—the latter representing dysregulation of microtubule assembly proteins—tau-related cytoskeletal abnormalities that are not found in most cases of DLB.

Alzheimer disease and DLB do share the features of ß-amyloidosis, senile plaque formation, and severe depletion of acetylcholine, which is even greater in DLB than in AD. Interestingly, they both also share an increased frequency of apolipoprotein e4 allele (20), which is not seen in nondemented cases of PD. Additional vascular changes are seen in up to 30% of cases of AD and DLB.

CLINICAL FEATURES

Part of "91 - Dementia with Lewy Bodies "

Most cases of DLB coming to autopsy are men (21 ,22 ,23 ,24 ,25 ,26 ,27 ,28 and 29). Although the mean age at onset and survival are similar to those in AD, survival times in DLB are sometimes skewed by rapidly progressive illness. The reduction in survival is probably partly attributable to neuroleptic sensitivity reactions (23). More recently, no differences in age at onset or survival were reported by Heyman et al. (30) and Walker et al. (31) between 24 and 32 patients with DLB and 74 and 43 patients with AD, respectively, although cognitive decline appeared to be faster in DLB (32).

Dementia is usually, but not always, the presenting feature of DLB; a minority of patients present with psychosis in the absence of dementia, some with mood disorders or psychosis, and others with orthostatic hypotension and falls. Fluctuation occurs in half to three-fourths of patients, but the range reported is wide, probably because this is such a difficult symptom to define. Fluctuation, irrespective of definition, is not commonly seen in AD. Visual hallucinations are present in one-third to one-half of DLB patients, although in some series the prevalence is 80%. Auditory hallucinations may occur in 20% of DLB subjects but seldom in AD. Depressive symptoms are common in both disorders, but a 38% prevalence in DLB is significantly greater than that in AD and similar to the rates reported in PD. Shimomura et al. (33) reported disproportionately more severe visuoperceptual, visuoconstructive, and visuospatial dysfunction and disproportionately milder impairment of memory. Ballard et al. (34) reported that although recent memory function is better preserved, visuospatial praxis is more impaired, a finding that potentially provides a psychological tool for differentiating DLB from AD.

A history of bursts of vigorous movements of the arms and legs with vocalization during sleep and associated with dream recall is highly suggestive of rapid-eye-movement (REM) sleep behavior disorder. Although REM sleep disorder may occur in, or indeed precede, a range of neurodegenerative disorders, including PD and multiple-system atrophy, in the context of degenerative dementia it suggests DLB (35).

The reported frequency of extrapyramidal signs in DLB varies greatly. Furthermore, the presence of extrapyramidal signs in DLB and their value in discriminating DLB from AD is unresolved. A number of issues must be considered in this context. First, the “background” population prevalence of parkinsonism is very common in the age range in which both DLB and AD occur. In one recent community-based study of 467 residents, parkinsonism (defined as the presence of at least two of the following: bradykinesia, gait disturbance, rigidity and tremor) was found to affect nearly

15% of people 65 to 74 years of age, 30% of those 75 to 84, and 52% of those 85 and older (36). Second, a wide range of frequencies (5% to 90%) of extrapyramidal signs has been reported in patients with AD (37). Although this may be related in part to differences in disease severity and study duration, it also likely reflects imprecision in the clinical definition of so-called extrapyramidal signs. Thus, predominantly cortically determined signs, such as ideomotor apraxia, paratonic rigidity (Gegenhalten), and frontal gait disorder, may be mistaken for bradykinesia, parkinsonian rigidity, and parkinsonian gait, respectively. Such motor disturbances produce “pseudoparkinsonism,” which is fundamentally different from the true parkinsonism determined by basal ganglia pathology (38). Finally, the reported rates for parkinsonism in DLB undoubtedly partly reflect case ascertainment biases. Patients collected through neurologic departments, which primarily receive referrals for movement disorders, are more likely to exhibit extrapyramidal signs than are DLB cases identified through memory clinics and psychogeriatric services.

Overall, probably fewer than half of DLB cases have extrapyramidal signs at presentation, and a fourth continue to have no evidence of them throughout their illness. Clinicians must therefore be prepared to diagnose DLB in the absence of parkinsonism—if they do not, their case detection rates will be unacceptably low.

When extrapyramidal signs do occur in DLB, a number of studies have contrasted them with the signs in PD in an attempt to characterize parkinsonian syndrome and identify potential diagnostic markers for DLB (39 ,40). In comparison with PD, less resting tremor and myoclonus, greater disease symmetry (especially at presentation), and a poor response to levodopa have all been reported for DLB, albeit inconsistently. It should be emphasized that any differences reported reflect group differences. The positive predictive value of any particular sign, or combination of signs, in differentiating DLB from PD in an individual patient has not been established. In common with PD, rigidity and bradykinesia, hypophonic speech, masked facies, stooped posture, and festinant gait have all been reported for DLB.

Finally, recurrent falls affect up to a third of DLB cases, a proportion significantly greater than in AD, as does neuroleptic sensitivity, detected in 61% of all DLB patients who receive neuroleptics but in only 15% of AD patients. Two studies have examined interrater reliability and found agreement rates and κ values to be acceptable for some symptoms of DLB, such as delusions, hallucinations, parkinsonism, and falls, but unacceptably low for others, particularly fluctuation (41 ,42).

The recent consensus criteria for the clinical diagnosis of DLB are shown in Table 91.2 (11). Emphasis is placed on the particular characteristics of the dementia syndrome—attentional deficits and prominent frontal–subcortical and visuospatial dysfunction. Fluctuation is no longer essential for the diagnosis, although it is frequently present. It seems probable that the fluctuating attentional deficit is linked to dysregulation of central cholinergic mechanisms controlling the level of consciousness (see section below on neurochemical clinical–pathologic relationships). The important hallucinatory symptoms are specified as visual, recurrent, and detailed, usually occurring most days of the week; they are typically colorful, three-dimensional images of animals and children. Insight into the unreal nature of these hallucinations is usually absent while they occur but is gained after the event. Ballard et al. (43) reported that more than 90% of patients with DLB experience such hallucinations and that in comparison with those of AD, the hallucinations are more persistent and the images are more likely to be accompanied by vocalization. Spontaneous parkinsonism not attributable to medication is a key symptom in most patients with DLB. If two of these three symptoms (fluctuations, visual hallucinations, and parkinsonism)

are present, a diagnosis of probable DLB is made; if only one is present, a diagnosis of possible DLB is allowed.

TABLE 91.2. CONSENSUS CRITERIA FOR THE CLINICAL DIAGNOSIS OF PROBABLE AND POSSIBLE DEMENTIA WITH LEWY BODIES

Consensus criteria for the clinical diagnosis of *probable* and *possible* dementia with Lewy bodies (DLB)

1. The central feature required for a diagnosis of DLB is progressive cognitive decline of sufficient magnitude to interfere with normal social or occupational function. Prominent or presistent memory impairment may not necessarily occur in the early stages but is usually evident with progression. Deficits on tests of attention and of frontal–subcortical skills and visuospatial ability may be especially prominent.
2. Two of the following core features are essential for a diagnosis of *probable* DLB; one is essential for *possible* DLB.
   a. Fluctuating cognition with pronounced variations in attention and alterness
   b. Recurrent visual hallucinations that are typically well formed and detailed
   c. Spontaneous motor features of parkinsonism
3. Features supportive of the diagnosis are the following:
   a. Repeated falls
   b. Syncope
   c. Transient loss of consciousness
   d. Neuroleptic sensitivity
   e. Systematized delusions
   f. Hallucinations in other modalities
   (Depression and REM sleep behavior disorder have been suggested as additional supportive features.)
4. A diagnosis of DLB is less likely in the presence of
   a. Stroke disease, evident as focal neurologic signs or on brain imaging
   b. Evidence on physical examination and investigation of any physical illness, or other brain disorder, sufficient to account for the clinical picture

DLB, dementia with Lewy bodies; REM, rapid eye movement.
Adapted from McKeith IG, Galasko D, Kosaka K et al., Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. *Neurology* 1996;47:1113–1124.

The sensitivity and specificity of the consensus clinical criteria against autopsy findings have been examined in several studies (Table 91.3). All find the diagnostic specificity to be relatively high, comparable with that of existing clinical criteria for AD and PD. This high specificity suggests that the DLB clinical criteria are appropriate for confirmation of the diagnosis (few false-positives). Sensitivity of case detection is reported as more variable and generally lower. The tendency to clinical underdiagnosis was noted during the second international workshop on DLB (44). However, two studies prospectively applying consensus criteria (as opposed to retrospective inspection of previous case records) did detect more than 80% of autopsy-confirmed DLB cases (45 ,46). A prospective validation study in Newcastle reported on a sample of 50 hospital-referred demented cases followed to autopsy (46). The sensitivity and specificity for a clinical diagnosis of probable DLB were 0.83 and 0.95, respectively.

TABLE 91.3. AUTOPSY VALIDATION OF CONSENSUS CRITERIA FOR DEMENTIA WITH LEWY BODIES

	Sensitivity	Specificity
Mega et al., 1996 (41)	0.75	0.79
Litvan et al., 1998 (116)	0.18	0.99
Papka et al., 1998 (117)	0.43	—
McShane et al., 1998[a] (45)	0.70	0.89
Holmes et al., 1999 (118)	0.22	1.00
Luis et al., 1999 (119)	0.57	0.90
Lopez et al., 1999 (120)	0.00	1.00
Verghese et al., 1999 (121)	0.61	0.84
McKeith et al., 2000[a] (122)	0.83	0.95

[a]Clinical diagnoses made prospectively, not by chart review.

DIFFERENTIAL DIAGNOSIS

Part of "91 - Dementia with Lewy Bodies "

Four main categories of disorders should be considered in the differential diagnosis of DLB. These are the following:

1. Other dementia syndromes. Sixty-five percent of autopsy-confirmed DLB cases meet the NINCDS/ADRDA clinical criteria for probable or possible AD (47), which is the most frequent clinical misdiagnosis applied to patients with DLB presenting with a primary dementia syndrome. For this reason, DLB should routinely be excluded when a diagnosis of AD is made. Up to a third of DLB cases are additionally misclassified as vascular dementia on the Hachinski ischemic index by virtue of the fluctuating nature and course of the illness. However, pyramidal and focal neurologic signs are usually absent. The development of myoclonus in patients with a rapidly progressive form of DLB may lead the clinician to suspect sporadic Creutzfeldt–Jakob disease (11).
2. Other causes of delirium. In patients with intermittent delirium, appropriate examination and laboratory tests should be performed during the acute phase to maximize the chances of detecting infective, metabolic, inflammatory, or other etiologic factors. Pharmacologic causes are particularly common in elderly patients. Although the presence of any of these features makes a diagnosis of DLB less likely, comorbidity is not unusual in elderly patients, and the diagnosis should not be excluded simply on this basis.
3. Other neurologic syndromes. In patients with a prior diagnosis of PD, the onset of visual hallucinations and fluctuating cognitive impairment may be attributed to side effects of antiparkinsonian medications, and this possibility must be tested by dose reduction or withdrawal. Other neurodegenerative akinetic-rigid syndromes associated with a poor response to levodopa, cognitive impairment, and postural instability include progressive supranuclear palsy, corticobasal degeneration, and progressive subcortical gliosis. Normal-pressure hydrocephalus must be considered in a patient with so-called lower-body parkinsonism, cognitive impairment, and urinary incontinence. Syncopal episodes in DLB are often incorrectly attributed to transient ischemic attacks despite an absence of focal neurologic signs. Recurrent disturbances in consciousness accompanied by complex visual hallucinations may suggest complex partial seizures (temporal lobe epilepsy), and vivid dreaming with violent movements during sleep may meet the criteria for REM sleep behavior disorder. Both these conditions have been reported as uncommon presenting symptoms of autopsy-confirmed DLB. Patients with REM sleep behavior disorder differ clinically from those without, performing worse on attentional tasks (48).
4. Other psychiatric disorders. If parkinsonian features or cognitive decline develops spontaneously in a patient, or if a patient shows excessive sensitivity to neuroleptic medication in the course of late-onset delusional disorder, depressive psychosis, or mania, the diagnosis of DLB should be considered.

NEUROTRANSMITTER ABNORMALITIES

Part of "91 - Dementia with Lewy Bodies "

Neurochemical activities have been widely investigated in AD and PD, including in some instances PD with dementia, but fewer reports are available on DLB. These are summarized in Table 91.4 as they relate to neurotransmitter systems.

TABLE 91.4. NEUROTRANSMITTER ACTIVITIES IN DLB, AD, ANDPD[a]

	DLB	AD	PD
I. CHOLINERGIC SYSTEM			
ChAT			
Cerebral cortex	↓↓	↓	↓↓[b]
Hippocampus	↓	↓↓	↓
Striatum	↓	↓	→
Thalamus	↓	→/↓	→
AChE			
Cortex	↓	↓	↓
BUChE			
Cortex		↓	
VAChT			
Cortex		↓	↓[b]
Muscarinic receptors			
M1			
Cortex	↑	→	↑[b]
Striatum	↓	↑	→
M2			
Cortex		↓	
Nicotinic receptors			
α7/αBT binding			
Cortex	→	→	
Thalamus	↓	↓	
α4/high-affinity			
agonist site			
Cortex	↓	↓	↓
Striatum	↓	→/↓	↓↓
Thalamus	→/↓	→/↓	→
II. MONOAMINERGIC SYSTEMS			
DOPAMINERGIC			
Presynaptic			
Dopamine			
Striatum	↓	→	↓↓
Cortex	↓	→	↓
Dopamine transporter[c]	↓	→	↓↓
Receptors			
D1 receptor[c]	→		→
D2 receptor[c]	↑	→	→/↓
D3 receptor[c]	→/↓		→
SEROTONINERGIC			
Presynaptic			
Serotonin			
Striatum	↓		↓
Cortex	↓	↓	↓
Serotonin transporter			
Cortex	↓	↓	↓
Receptors			
5-HT_{2A} receptor			
Cortex	→/↓	↓	↓
NORADRENERGIC			
Noradrenaline			
Striatum	↓		↓↓
Cortex		↓	
MAO-B		↑	

5-HT, 5-hydroxytryptamine; AChE, acetylcholinesterase; AD, alzheimer disease; BUChE, butyrycholinesterase; ChAT, choline acetyltransferase; DLB, dementia with Lewy bodies; MAO-B, monoamineoxidase B; PD, Parkinson disease; VAChT, vesicular acetylcholine transporter.
[a]Summary of neurochemical findings, reviewed Perry et al. (123); see also the following recent references: 51,54,56,59,124,125.
[b]Denotes more extensive in PD + dementia.
[c]Striatal activities.

Reductions in presynaptic cholinergic activities, particularly in the cerebral neocortex, are more marked in DLB than in AD and are similar to those in PD with dementia (49). As in PD, the cortical cholinergic deficit appears to reflect neuronal loss in the basal nucleus of Meynert (50).

The cortical cholinergic pathology is independent of the extent of Alzheimer pathology, being equally great in DLB cases with and without this type of pathology (51). Cholinergic deficits in DLB extend beyond the cortex to the striatum and certain nuclei of the thalamus (52). Also, in contrast to AD and similar to PD with dementia, DLB is associated with elevation of the muscarinic receptor subtype M1 (53), a finding that has recently been confirmed by immunoabsorption studies (54). However, muscarinic M1 receptors are not uncoupled to the same extent as in AD (52; Perry et al., *in preparation*). Changes in nicotinic receptors in the cortex include a loss of the high-affinity agonist binding site (likely to reflect the α_4 subunit), but no change in the α_7 subunit or α-bungarotoxin binding (54a). In contrast, little change in nicotine binding occurs in the thalamus, but highly significant reductions in α-bungarotoxin binding are seen in the reticular nucleus (55). Similar nicotinic receptor abnormalities occur in AD and (as far as has been investigated) in PD, although the loss of nicotine binding in the striatum is greater in PD, in keeping with the more extensive reduction in basal ganglia dopaminergic projections (56). Although the loss of high-affinity nicotinic receptor binding in AD has been related to synapse loss, measured by synaptophysin levels (57), synaptophysin loss occurs in DLB only when the pathology includes the Alzheimer type (58).

The involvement of the dopaminergic system is the other consistent neurochemical feature of DLB (Table 91.4); however, as might be expected from the variations in extrapyramidal features (absent/mild to severe, as in classic PD), the extent of dopamine or dopamine transporter loss in the striatum varies widely. Earlier reports that dopamine loss was in some cases severe despite the absence of neurologic symptoms (49), a finding that was interpreted to indicate compensatory striatal pathology, need to be replicated in prospectively assessed cases because symptoms may have been overlooked in psychogeriatric clinics; furthermore, neuroleptic medication reduces striatal dopamine. Although in PD striatal dopamine deficits are more marked in caudal regions, particularly putamen, in DLB the loss of dopamine transporter is similar at different rostral caudal levels (59). Whereas in PD dopamine D2 receptors are up-regulated, at least in earlier stages of the disease, receptors are not increased in DLB and in particular are not up-regulated as a result of neuroleptic medication (60). In addition to striatal dopamine deficits, dopamine losses in cortical areas also occur (Table 91.4).

The significance of the serotoninergic, noradrenergic, and neuropeptide (e.g., somatostatin and corticotropinreleasing factor) abnormalities that occur in DLB, as in AD and PD, has not generally been evaluated in pathologic or clinical terms. The clinical significance of some of the cholinergic abnormalities that have so far been examined in prospectively assessed cases is discussed in the section on clinical–pathologic relationships .

GENETICS

Familial cases of DLB have been reported, although the majority of cases appear to be sporadic. Following the discovery of two separate missense mutations in the α-synuclein gene on chromosome 4 in a small number of families with pathologically confirmed early-onset PD, mutation screening was undertaken in this gene in both familial and sporadic cases of DLB (61). These studies failed to reveal any nucleotide changes within the exons screened.

Although, like cases of PD, most cases of DLB appear to be sporadic, this does not exclude a potentially significant genetic influence in the etiology of either condition. Such an influence may be via so-called susceptibility genes. Because DLB shares pathologic overlap with PD and AD, susceptibility genes of interest for both conditions have been considered in candidate gene approaches. For PD and DLB, allelic frequencies of the cytochrome P-450 gene *CYP2D6* (debrisoquine-4-hydroxylase) have been examined. The results of these studies have been conflicting. An increased frequency of the *CYP2D6*B* allele has been reported in DLB (62), whereas another study found no association (63). Others found no difference between the frequency of this allele in AD and DLB despite an increased frequency in PD (64). The current balance of evidence suggests that the *CYP2D6*B* allele is not a major genetic determinant of DLB.

The ε4 type of apolipoprotein E is significantly elevated in both DLB and AD, with a concomitant reduction in the Eε3 type of apolipoprotein. Interestingly, although the ε4 allele was associated with an increased risk for the development of DLB, it did not appear to affect the burden of pathology, measured by senile plaque and neurofibrillary tangle density in the neocortex (65). In AD, the ε4 allele is associated with neuronal loss in the substantia nigra (66). In PD, the ε4 allele increases the risk for drug-induced hallucinations (67).

Most recently, a significant difference has been reported for the allelic distribution of a pentanucleotide repeat within the promoter region of the nitric oxide synthase gene *(NOS2A)* in a comparison of autopsy-proven DLB cases with controls (68). Nitric oxide functions normally in the brain as a physiologic neuronal mediator, but excessive production of nitric oxide (e.g., after ischemic brain injury) can cause cell death through the generation of potent oxidants. Furthermore, with use of a murine MPTP model of parkinsonism, it has been shown that inhibitors of nitric oxide synthase may provide protective benefit in the treatment of PD (69).

Associations between polymorphisms within the α_2-macroglobulin, α_1-antichymotrypsin, and presenilin 1 genes and DLB have not been demonstrated (after accounting for apolipoprotein E ε4 allele frequency) (70 ,71 and 72).

Accumulating evidence suggests that patients with DLB are more responsive to cholinergic therapy than those with AD at a similar stage of dementia. The polymorphism causing the K allele in the gene for butyrylcholinesterase has been reported to be associated with AD, although this finding has not been replicated by others. In DLB, an increased number of butyrylcholinesterase K homozygotes have been found. It has been suggested that this genotype may partly explain the enhanced responsiveness to cholinesterase inhibitors in DLB (73). Although abnormalities in butyrylcholinesterase are evident in AD, including elevated enzymatic activity associated with both amyloid plaques and neurofibrillary tangles, the enzyme has not been examined in DLB.

CLINICAL–PATHOLOGIC RELATIONSHIPS

Part of "91 - Dementia with Lewy Bodies "

Cognitive and Neuropsychiatric

In DLB, a consistent gradient of LB density has been noted, as follows: substantia nigra > entorhinal cortex > cingulate gyrus > insula > frontal cortex > hippocampus > occipital cortex. Paralimbic and neocortical LB densities are highly correlated with each other but not with nigral pathology, which suggests that DLB should not be considered merely a severe form of PD (74). One study of pathologic burden versus clinical severity examined correlations between two simple measures of cognitive ability and a range of lesion counts and neurochemical measures in the midfrontal cortex of DLB cases (75). Severity of dementia was significantly correlated with LB density, plaque density, and severity of cholinergic deficit, but not with neurofibrillary tangle density or synaptophysin levels. In contrast, in AD cases, tangle density and synaptophysin levels were most highly correlated with clinical severity. This suggests that the dementias of DLB and AD may have different pathologic but similar neurochemical substrates. Other studies have failed to find robust correlations between LB density and clinical features (74 ,76).

Neurologic

Cell loss in the ventrolateral tier of the substantia nigra is the dominant pathologic correlate for parkinsonism in both DLB and PD. This is associated with gliosis and Lewy body formation. Dopaminergic neurons in this area of the substantia nigra project predominantly to the putamen, which, in turn, is an integral component of the so-called basal ganglia “motor loop” (77). The net effect of loss of the modulating effects of dopamine within the putamen is increased neuronal activity in the globus pallidus (internal segment). Because output from the globus pallidus (internal segment) is inhibitory to the ventrolateral thalamic nucleus, this leads to excessive inhibition of thalamic activity and thus reduced feedback to the motor cortex. The perturbation in this loop, resulting from dopamine deficiency, is believed to be the basis of the neural substrate for bradykinesia.

Although systematic studies have not yet been performed,

some evidence suggests that responsiveness to levodopa in DLB may be less than that in PD. Because the presynaptic lesion is similar in the two disorders, the answer may therefore lie postsynaptically. Evidence to support this notion comes from postmortem neurochemical studies comparing dopaminergic activities in DLB with those in PD and AD (59). In these studies, dopamine D2 receptor binding was reduced in the caudal putamen and was significantly lower than in PD at all levels.

Although the increased falls reported in DLB may be multifactorial, it is likely that more widespread involvement of brainstem nondopaminergic nuclei is a contributing factor. Degeneration of the predominantly cholinergic pedunculopontine nucleus is a likely explanation because neuronal loss in this structure has been associated with postural instability (78). In addition, degeneration of the pedunculopontine nucleus has been implicated as the pathophysiologic basis for REM sleep behavioral disorder, which is also reported in DLB (79).

Neurochemical

As yet, only a few clinical–neurochemical relationships have been identified in DLB. In earlier reports of the loss of cholinergic activity from the cortex, correlations were identified, as in AD, with the severity of cognitive impairments (17 ,75).

In regard to noncognitive or neuropsychiatric symptoms, patients with visual hallucinations have significantly lower levels of choline acetyltransferase than do nonhallucinators (80); recently, they have also been found to have lower levels of nicotinic α-bungarotoxin receptor binding in visual association cortex (Ballard et al., *in preparation*). Muscarinic M1 binding in temporal cortex is increased in patients experiencing persistent delusions (81). Delusional misidentification has also been associated with lower levels of α-bungarotoxin binding in this region (Ballard et al., submitted). Disturbances in consciousness are associated with a tendency for choline acetyltransferase to be lower in the thalamic reticular nucleus (53) and with a relative preservation of the high-affinity nicotinic receptor in the cortex (Ballard et al., submitted). Although reductions in this receptor correlate with attentional deficits, it appears that the ability to return periodically to normal levels of consciousness (fluctuations) depends on a degree of integrity of the nicotinic receptor. It has been suggested that greater EEG slowing is related to the greater cholinergic deficit in DLB than in AD (82). A hypothesis relating the function of cerebral acetylcholine in the integrative processes that generate conscious awareness has recently been proposed (83).

In regard to noncholinergic transmitter abnormalities, sensitivity to neuroleptic medication has been related to a lack of dopamine D2 receptor up-regulation, and depression to relatively preserved serotonin transporter levels (Ballard et al., submitted).

CLINICAL INVESTIGATIONS

Part of "91 - Dementia with Lewy Bodies "

As with any patient presenting with cognitive impairment, obtaining a full history and performing a mental and physical examination are essential steps toward making a firm clinical diagnosis. As with suspected cases of AD, the level and extent of laboratory investigations vary according to the clinical picture, associated comorbidity, and physical examination findings. However, because of the particular associations of DLB with fluctuations in attention and cognition and visual hallucinations, both very commonly associated with a variety of other organic disorders, the investigation of a suspected case of DLB requires a very careful laboratory evaluation. This usually includes routine hematology and biochemistry, determinations of erythrocyte sedimentation rate and creatine phosphate, thyroid function tests, measurements of B_{12} and folate levels, syphilis serology, and urinalysis. A chest roentgenogram may also be considered routine in view of the high incidence of lung carcinomas in the elderly, especially smokers. As in the diagnosis of AD, neuroimaging investigations are often helpful, both in excluding other intracranial disorders (including cerebrovascular disease) that may be responsible for the cognitive impairment and in providing supportive features for the diagnosis.

The EEG findings may be abnormal in up to 90% of DLB patients; loss of alpha rhythm and transient slow-wave activity in the temporal lobe areas are the most common changes (82). Patients with AD are less likely to have transient slow waves, and slowing of the dominant rhythm is less marked. However, the positive predictive value of the EEG in suspected cases of DLB has not been assessed in a prospective clinicopathologic study. Increasingly, some form of structural imaging is becoming essential to apply diagnostic criteria rigorously, such as the NINCDS/ADRDA criteria for AD, the NINCS/ADRDA criteria for vascular dementia, and the consensus criteria for DLB.

Structural Imaging Changes

Few studies have investigated computed tomographic (CT) or magnetic resonance imaging (MRI) changes in DLB. In a longitudinal study of AD subjects who came to postmortem examination, Förstl et al. (24) reported more pronounced frontal lobe atrophy on CT in eight subjects with LB pathology in a comparison with pure AD cases. However, using MRI, Harvey et al. (84) found no differences in frontal lobe volumes between AD and DLB subjects, a finding replicated in a different and larger cohort by Barber et al. (85). Although further studies are awaited, frontal lobe atrophy does not seem to be a particular feature of DLB. Similarly, DLB does not seem to differ from AD in terms of degree of ventricular enlargement or presence of white matter changes on MRI (86).

The strong association between AD and atrophy of the

medial temporal lobe, whether assessed by a linear measurement of medial temporal lobe width on CT (87) or visual or volumetric ratings of hippocampal atrophy on MRI (88 ,89), led to an investigation of whether similar changes are associated with DLB. Jobst et al. (87) found medial temporal lobe atrophy of similar magnitude to that in AD in two of their four cases of DLB. However, with the use of MRI, both case reports and controlled studies have shown DLB to be associated with relative preservation of temporal lobe structures in comparison with AD (84 ,85 ,90 ,91). Volumetric analysis of subregions within the temporal lobe indicates that the differences lie in medial temporal lobe structures (i.e., hippocampus and parahippocampal gyrus) rather that in the lateral temporal lobe, which does not show any differences between AD and DLB. Volumetric analysis, although essential for research studies and investigating clinical correlates, is currently too time-consuming to be adopted into routine clinical practice. Using visual ratings, which can be performed quickly (1 minute per scan) and simply, Barber et al. (85) found that 38% of DLB subjects but no AD subjects had a normal rating of temporal lobe atrophy, which suggests that at least in some cases relative preservation of the hippocampus and medial temporal lobe may support a diagnosis of DLB. Sample medial temporal lobe images are shown in Fig. 91.2 . The reason for this variability in temporal lobe atrophy in DLB is unknown, although based on a very limited autopsy examination of four cases, Harvey et al. (84) suggested that temporal lobe atrophy on MRI may be a marker of concurrent AD pathology in DLB. However, although cross-sectional imaging may be helpful in some cases, it clearly is not diagnostic. It is yet to be determined whether accurate longitudinal assessment of regional volume change on MRI increases the accuracy of diagnosis, as may be the case for AD (92).

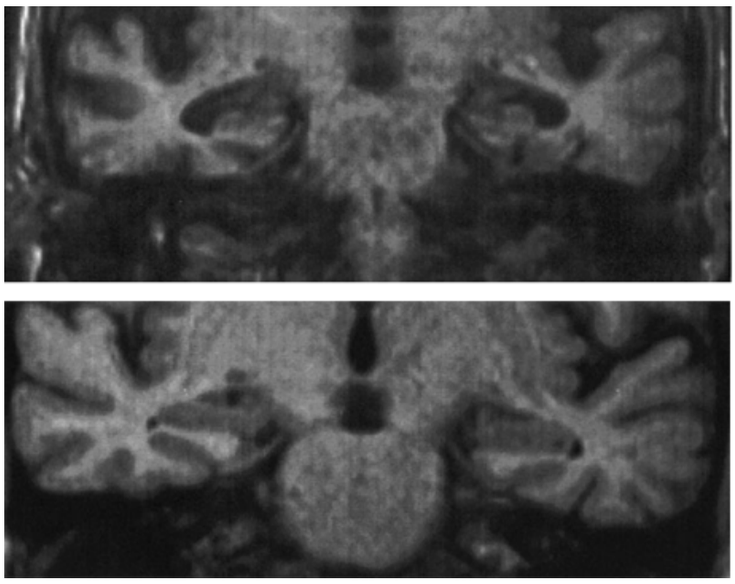


FIGURE 91.2. Coronal magnetic resonance imaging slices of patients with Alzheimer disease (AD) and dementia with Lewy bodies (DLB). Note severe atrophy of hippocampus and medial temporal lobe structures bilaterally in subject with AD. In contrast, the appearance of the medial temporal lobe in the subject with DLB is normal for age.

In summary, the limited evidence available suggests that structural imaging in DLB reveals generalized atrophic changes similar to those of AD in most cases, although approximately 40% of DLB subjects show preservation of medial temporal lobe structures.

Functional Imaging Changes

Single-photon emission tomography (SPET) with the use of blood flow markers such as Tc-HMPAO (hexamethylpropylene amine oxime) has been extensively investigated in dementia. In AD, the classic appearance is one of posterior bilateral symmetric temporoparietal hypoperfusion (87 ,93), which contrasts with the frontal hypoperfusion characteristically seen in frontal lobe dementia (94). Vascular dementia is associated with a mottled, uneven, patchy appearance, reflecting the variable anatomic localization of vascular disease (95). In PD, the blood flow in basal ganglia is decreased, and when PD is associated with dementia, bilateral parietal changes similar to those seen in AD are reported (96 ,97). In the few SPET studies of DLB, patterns of blood flow changes similar to those of AD have been found, although Donnemiller et al. (98) found a subtle difference in perfusion patterns, with a greater degree of occipital hypoperfusion in DLB than in AD, and Defebvre et al. (99) found decreased frontal perfusion in DLB. Perfusion of the medial temporal lobe may be less impaired in DLB than in AD (100), consistent with the structural imaging findings described above of preservation of the same structures in DLB. Higuchi et al. (101) have suggested that glucose hypometabolism in occipital cortex is a diagnostic aid in distinguishing DLB from AD, and Imamura et al. (102) found that hypometabolism in this region in conjunction with relatively mild hypometabolism in temporal and parietal cortex is associated with visual hallucinations in DLB.

The more powerful, although still research-based, use of SPET involves the use of specific ligands for different neurochemical systems. Ligands have been developed for presynaptic and postsynaptic dopaminergic and cholinergic systems. Donnemiller et al. (98), using positron emission tomography (PET), found significant differences between DLB and AD in ß-carbomethoxy-iodophenyl-tropane (CIT) binding (a ligand for the dopamine transporter), a difference that would be predicted from the known neurochemical differences between AD and DLB. However, one disadvantage of CIT is that imaging has to be delayed for 24 hours after injection. A single case report has suggested that a ligand with faster imaging kinetics, FP (fluoropropyl)-CIT, can distinguish DLB from AD (103). A reduced density of dopamine D2 receptors in basal ganglia, demonstrated with [^{123}I]iodobenzamide, has been reported in DLB (104). With use of a marker of the vesicular acetylcholine

transporter, significant differences between AD subjects and controls and between PD subjects with and without dementia have been found (105). In summary, current evidence suggests that SPET studies of blood flow show similar appearances in DLB and AD, although SPET may still be useful in distinguishing DLB, like AD, from frontal lobe dementia or vascular dementia. New chemical imaging techniques, although not yet clinically available, show great promise in differentiating DLB from other disorders and are an exciting area of current research.

TREATMENT

Part of "91 - Dementia with Lewy Bodies "

Neuropsychiatric Symptoms

Some of the key clinical features of DLB are similar to those induced in normal subjects by anticholinergic, specifically antimuscarinic, agents. Clouding of consciousness, confusion, and visual hallucinations are recognized effects of anticholinergic drug toxicity, and cumulative effects of subcortical and cortical cholinergic dysfunction probably play a major role in the spontaneous generation of similar fluctuating symptoms in DLB. As discussed in the earlier section on neurochemical clinical-pathologic relationships , reductions in choline acetyltransferase are correlated with cognitive impairment (75), and hallucinations may be related to hypocholinergic and (relatively) hypermonoaminergic neocortical neurotransmitter function (80).

Several reports have indicated that patients who respond well to cholinesterase inhibitor treatments are more likely to have DLB than AD at autopsy (106 ,107). This finding is consistent with the neurochemical profile of DLB and the fact that postsynaptic cortical muscarinic receptors are functionally intact. Case reports suggest that cholinesterase inhibitors can reduce psychotic symptoms in DLB (108), and a recently completed placebo-controlled study of rivastigmine in patients meeting consensus criteria for probable DLB found significant improvements in both neuropsychiatric features and cognition (109).

It is possible that cholinergic drugs may emerge as the antipsychotic treatment of choice in dementing diseases of the elderly, such as DLB. Not only are typical neuroleptics inappropriate (23); according to more recent reports, so are atypical drugs such as olanzapine (110 ,111).

Neurologic Symptoms

Because of the combination of parkinsonism with neuropsychiatric symptoms, the management of extrapyramidal signs in DLB is rather like walking a tightrope. The best outcome will invariably be a compromise between a relatively mobile but psychotic patient and a nonpsychotic but immobile patient.

Many currently available antiparkinsonian drug treatments, including monoamine oxidase B inhibitors, anticholinergic agents, and dopamine agonists, have an unacceptably high risk of precipitating or exacerbating hallucinations and confusion. In addition, most of these agents can cause or worsen orthostatic hypotension and lead to an increased incidence of falls. Although it has never been formally assessed, avoidance of these drugs in DLB patients would seem prudent. Indeed, for a DLB patient presenting with parkinsonism, in whom dementia and neuropsychiatric symptoms then develop, the first therapeutic decision should be to review the antiparkinsonian medication and undertake a gradual withdrawal if necessary.

At a practical level, the antiparkinsonian drug with the best risk-to-benefit ratio currently available for the treatment of extrapyramidal signs in DLB is levodopa. However, although the efficacy of levodopa for the treatment of PD is beyond question, how effective is this drug in the management of DLB? The evidence regarding the responsiveness of DLB patients to levodopa is conflicting. Some reports suggest that up to 100% of patients with DLB may improve, but the numbers of patients have usually been small in these series, and the degree of functional change and duration of the response have not been specified (39 ,40). Furthermore, the effects of levodopa therapy on neuropsychiatric symptoms are poorly documented in this group. Given the postsynaptic changes in dopamine D2 receptors noted in postmortem studies, it might be postulated that levodopa responsiveness is diminished in DLB. Clearly, a number of issues regarding the efficacy of dopaminergic treatment in DLB have not been resolved, and further trials are in progress to address these issues.

The most important point in the management of patients with DLB is to exercise caution in (or preferably avoid) prescribing neuroleptic medications, which are the mainstay of antipsychotic treatment in other groups of patients. Severe neuroleptic sensitivity reactions can precipitate irreversible parkinsonism, further impair the level of consciousness, and induce autonomic disturbances reminiscent of neuroleptic malignant syndrome (22 ,23). They occur in 40% to 50% of DLB patients treated with neuroleptics and are associated with a twofold to threefold increase in mortality. Acute blockade of D2 receptors is thought to mediate these effects, and despite some promising initial reports, atypical and novel antipsychotics such as risperidone and olanzapine seem to be just as likely to cause neuroleptic sensitivity reactions as the older drugs.

Until safe and effective medications become available, the mainstay of clinical management is undoubtedly to educate patients and carers about the nature of their symptoms and suggest strategies to cope with them. The clinician must ascertain which symptoms are most troublesome for the patient and explain the risks and benefits associated with changes in medication (112).

SUMMARY AND CONCLUSION

Part of "91 - Dementia with Lewy Bodies "

Dementia with Lewy bodies appears to be distinct both clinically and neuropathologically from AD and is the second most common form of degenerative dementia, accounting for up to 20% of cases in the elderly. It is characterized by fluctuating cognitive impairment, spontaneous parkinsonism, and recurrent visual hallucinations. Consensus clinical and neuropathologic criteria were published in 1996, and the clinical criteria have been shown to be highly specific, although they may still lack sensitivity. As in PD, the defining neuropathologic feature is the presence of Lewy bodies and neurites (positive for α-synuclein and ubiquitin) in a range of subcortical nuclei and in cortical regions, particularly cingulate and entorhinal. Alzheimer-type pathology is variable, ranging from none in the neocortex to, most commonly, extensive ß-amyloidosis and, rarely, additional neurofibrillary tangle formation. The main neurotransmitter abnormalities include the following: extensive reduction in presynaptic cholinergic activities in the cortex, related to psychotic features such as hallucinations; elevations of muscarinic receptors; abnormalities in nicotinic receptors, related to disturbances in consciousness; and both presynaptic and postsynaptic dopamine abnormalities, related to extrapyramidal dysfunction, including sensitivity to neuroleptic medication. The recognition of DLB is clinically important in view of the high incidence (60%) of adverse and life-threatening reaction to antipsychotic medications, the difference in prognosis, and the differential treatment response to cholinergic therapy. Neuroimaging changes have not been described in DLB to any extent, but some show promise as potential markers to differentiate DLB from AD. These include relative preservation of temporal lobe structures on MRI and loss of presynaptic and postsynaptic dopaminergic markers on SPET. Patients with DLB respond positively to cholinesterase inhibitors with reductions in neuropsychiatric symptoms such as hallucinations, delusions, and agitation.

This chapter has summarized the major advances that have taken place within the last decade in understanding the mechanisms underlying DLB, in diagnosing the disease, and in treating some of the clinical features. In relation to neuropsychopharmacology, the disease provides a unique opportunity to understand mechanisms underlying symptoms such as hallucinations and disturbances in consciousness because such symptoms can occur in the absence of significant Alzheimer pathology and be correlated with quantifiable neurotransmitter abnormalities.

In terms of understanding the core pathologic mechanisms, however, as in AD and PD, the objective of disease prevention still appears to be a long way off. Transgenic mice expressing wild-type α-synuclein have synuclein-positive inclusions in cortex and substantia nigra and decreased dopamine levels in basal ganglia (113). It is interesting that this model, relevant to PD and DLB, expresses the phenotype in relation to transmitter abnormality (at least in regard to dopamine), whereas the equivalent model considered relevant to AD—overexpression of mutated amyloid precursor protein (APP)—does not. However, one report of numerous and widespread α-synuclein-negative Lewy bodies in an asymptomatic patient raises the question of just how pathogenic abnormalities of this synaptic protein really are (114).

With improved methods of diagnosing DLB now available, epidemiologic studies are needed. It would be interesting to determine whether tobacco use is associated with a decreased risk for the development of DLB, as it is for PD. The potential for neuroprotection based on stimulation of nicotinic receptors is increasingly being recognized (115).

Better genotype–phenotype correlation is also needed for DLB. For example, do functional polymorphisms within the dopamine D2-receptor gene influence levodopa responsiveness? In addition, does butyrylcholinesterase K homozygote status predict therapeutic response to cholinesterase inhibitors? Improved knowledge in this area could conceivably rationalize the use of drug treatments for DLB.

The involvement of brainstem nuclei in DLB other than the substantia nigra needs to be explored further. The quality of future clinicopathologic correlations will be enhanced by the prospective acquisition of clinical data in longitudinal studies with the use of standardized and validated instruments.

The role of levodopa in the treatment of the extrapyramidal syndrome associated with DLB needs to be defined better. For example, levodopa-induced “on/off” responses and dyskinesias have not been reported in DLB, as they have in PD. This may be because parkinsonism is generally less severe and may take longer to develop than DLB or because distinct striatal pathology (e.g., loss of cholinergic activity) is present. Novel therapies aimed at relieving parkinsonism that do not exacerbate neuropsychiatric features are needed. Drugs acting via nondopaminergic neurotransmitter systems may be applicable in this area (e.g., adenosine A_{2A} antagonists).

ACKNOWLEDGMENTS

Part of "91 - Dementia with Lewy Bodies "

The secretarial assistance of Maureen Middlemist and Lorraine Hood is gratefully acknowledged.

REFERENCES

1. Okazaki H, Lipkin LE, Aronson SM. Diffuse intracytoplasmic inclusions (Lewy type) associated with progressive dementia and quadrapesis in flexion. *J Neuropathol Exp Neurol* 1961;20:237–244.

2. Kosaka K, Yoshimura M, Ikeda K, et al. Diffuse type of Lewy body disease: progressive dementia with abundant cortical Lewy bodies and senile changes of varying degree—a new disease? *Clin Neuropathol* 1984;3:185–192

3. Perry RH, Irving D, Blessed G, et al. Clinically and neuropathologically distinct form of dementia in the elderly. *Lancet* 1989;1;166.

4. Lennox G, Lowe J, Landon M, et al. Diffuse Lewy body disease: correlative neuropathology using anti-ubiquitin immunocytochemistry. *J Neurol Neurosurg Psychiatry* 1989;52:1236–1247.

5. Dickson DW, Davies P, Mayeux R, et al. Diffuse Lewy body disease. Neuropathological and biochemical studies of six patients. *Acta Neuropathol* 1986;75:8–15.

6. Hansen L, Salmon D, Galasko D, et al. The Lewy body variant of Alzheimer's disease: a clinical and pathologic entity. *Neurology* 1990;40:1–8.

7. Spillantini MG, Schmidt ML, Lee VMY, et al. α-Synuclein in Lewy bodies. *Nature* 1997;388:839–848.

8. Trojanowski JQ. Dementia with Lewy bodies: histopathological aspects of differential diagnosis. *Neurobiol Aging* 1998; 19:S4.

9. Galvin JE, Uryu K, Lee VM, et al. Axon pathology in Parkinson's disease and Lewy body dementia hippocampus contains alpha-, beta-, and gamma-synuclein. *Proc Natl Acad Sci USA* 1999;96:13450–13455.

10. Lowe JS, Mayer RJ, Landon M. Pathological significance of Lewy bodies in dementia. In: Perry R, McKeith I, Perry E, eds. *Dementia with Lewy bodies*. New York: Cambridge University Press, 1996:195–203.

11. McKeith IG, Galasko D, Kosaka K, et al. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. *Neurology* 1996;47:1113–1124.

12. McKenzie JE, Edwards RJ, Gentleman SM, et al. A quantitative comparison of plaque types in Alzheimer's disease and senile dementia of the Lewy body type. *Acta Neuropathol* 1996;91:526–529.

13. Mann DM, Brown SM, Owen F, et al. Amyloid beta protein (A beta) deposition in dementia with Lewy bodies: predominance of A beta 42(43) and paucity of A beta 40 compared with sporadic Alzheimer's disease. *Neuropathol Appl Neurobiol* 1998;24:187–194.

14. Lippa CF, Ozawa K, Mann DM, et al. Deposition of beta-amyloid subtypes 40 and 42 differentiates dementia with Lewy bodies from Alzheimer disease. *Arch Neurol* 1999;56:1111–1118.

15. Harrington CR, Perry RH, Perry EK, et al. Senile dementia of Lewy body type and Alzheimer type are biochemically distinct in terms of paired helical filaments and hyperphosphorylated tau protein. *Dementia* 1994;5:215–228.

16. Hansen LA, Masliah E, Galasko D, et al. Plaque-only Alzheimer disease is usually the Lewy body variant, and vice versa. *J Neuropathol Exp Neurol* 1993;52:648–654.

17. Perry RH, Irving D, Blessed G, et al. Senile dementia of Lewy body type. A clinically and neuropathologically distinct form of Lewy body dementia in the elderly. *J Neurol Sci* 1990;95:119–139.

18. Ball M, Braak H, Coleman P, et al. Consensus recommendations for the post-mortem diagnosis of Alzheimer's disease. *Neurobiol Aging* 1997;18:S1–S2.

19. Gomez-Isla T, Growdon WB, McNamara M, et al. Clinicopathologic correlates in temporal cortex in dementia with Lewy bodies. *Neurology* 1999;53:2003–2009.

20. Benjamin R, Leake A, Edwardson JA, et al. Apolipoprotein E genes in Lewy body and Parkinson's disease. *Lancet* 1994;343:1565.

21. Byrne EJ, Lennox G, Lowe J, et al. Diffuse Lewy body disease: clinical features in 15 cases. *J Neurol Neurosurg Psychiatry* 1989;52:709–717.

22. McKeith IG, Perry RH, Fairbairn AF, et al. Operational criteria for senile dementia of Lewy body type (SDLT). *Psychol Med* 1992;22:911–922.

23. McKeith I, Fairbairn A, Perry R, et al. Neuroleptic sensitivity in patients with senile dementia of Lewy body type. *Br Med J* 1992;305:673–678.

24. Förstl H, Burns A, Luthert P, et al. The Lewy-body variant of Alzheimer's disease. Clinical and pathological findings. *Br J Psychiatry* 1993;162:385–392.

25. Galasko D, Katzman R, Salmon DP, et al. Clinical and neuropathological findings in Lewy body dementias. *Brain Cogn* 1996;31:166–175.

26. Weiner MF, Risser RC, Cullum CM, et al. Alzheimer's disease and its Lewy body variant: a clinical analysis of postmortem verified cases. *Am J Psychiatry* 1996;153:1269–1273.

27. Klatka LA, Louis ED, Schiffer RB. Psychiatric features in diffuse Lewy body disease: findings in 28 pathologically diagnosed cases. *Neurology* 1996;46:A180.

28. Hely MA, Reid WGJ, Halliday GM, et al. Diffuse Lewy body disease: clinical features in nine cases without coexistent Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 1996;60:531–538.

29. Ala TA, Yang K-H, Sung JH, et al. Hallucinations and signs of parkinsonism help distinguish patients with dementia and cortical Lewy bodies from patients with Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 1997;62:16–21.

30. Heyman A, Fillenbaum GG, Gearing M, et al. Comparison of Lewy body variant of Alzheimer's disease with pure Alzheimer's disease: consortium to establish a registry for Alzheimer's disease, part XIX. *Neurology* 1999;52:1839–1844.

31. Walker Z, Allen RL, Shergill S, et al. Three years survival in patients with a clinical diagnosis of dementia with Lewy bodies. *Int J Geriatr Psychiatry* 2000;15:267–273.

32. Olichney JM, Galasko D, Salmon DP, et al. Cognitive decline is faster in Lewy body variant than in Alzheimer's disease. *Neurology* 1998;51:351–357.

33. Shimomura T, Mori E, Yamashita H, et al. Cognitive loss in dementia with Lewy bodies and Alzheimer's disease. *Arch Neurol* 1998;55:1547–1552.

34. Ballard CG, Ayre G, O'Brien J, et al. Simple standardised neuropsychological assessments aid in the differential diagnosis of dementia with Lewy bodies from Alzheimer's disease and vascular dementia. *Dementia Geriatr Cogn Disord* 1999;10:104–108.

35. Boeve BF, Silber MH, Ferman TJ, et al. REM sleep behavior disorder and degenerative dementia: an association likely reflecting Lewy body disease. *Neurology* 1998;51:363–370.

36. Bennett DA, Beckett LA, Murray AM, et al. Prevalence of parkinsonian signs and associated mortality in a community population of older people. *N Engl J Med* 1996;334:71–76.

37. Mitchell SL. Extrapyramidal features in Alzheimer's disease. *Age Ageing* 1999;28:401–409.

38. Kurlan R, Richard IH, Papka M, et al. Movement disorders in Alzheimer's disease: more rigidity of definitions is needed. *Mov Disord* 2000;15:24–29.

39. Louis ED, Klatka LA, Liu Y, et al. Comparison of extrapyramidal features in 31 pathologically confirmed cases of diffuse Lewy body disease and 34 pathologically confirmed cases of Parkinson's disease. *Neurology* 1997;48:376–380.

40. Gnanalingham KK, Byrne EJ, Thornton A, et al. Motor and cognitive function in Lewy body dementia: comparison with Alzheimer's and Parkinson's diseases. *J Neurol Neurosurg Psychiatry* 1997;62:243–252.

41. Mega MS, Masterman DL, Benson F, et al. Dementia with Lewy bodies: reliability and validity of clinical and pathologic criteria. *Neurology* 1996;47:1403–1409.

42. McKeith IG, Fairbairn AF, Bothwell RA, et al. An evaluation of the predictive validity and inter-rater reliability of clinical diagnostic criteria for senile dementia of Lewy body type. *Neurology* 1994;44:872–877.

43. Ballard C, McKeith I, Harrison R, et al. A detailed phenomenological comparison of complex visual hallucinations in dementia with Lewy bodies and Alzheimer's disease. *Int Psychogeriatr* 1997;9:381–388.

44. McKeith IG, Perry EK, Perry RH. Report of the second dementia with Lewy body international workshop. *Neurology* 2000;53:902–905.

45. McShane RH, Esiri MM, Joachim C, et al. Prospective evaluation of diagnostic criteria for dementia with Lewy bodies. *Neurobiol Aging* 1998;19:S204.

46. McKeith IG, Ballard CG, Perry RH, et al. Prospective validation of consensus criteria for the diagnosis of dementia with Lewy bodies. *Neurology* 2000;54:1050–1058.

47. McKeith IG, Fairbairn AF, Perry RH, et al. The clinical diagnosis and misdiagnosis of senile dementia of Lewy body type (SDLT). *Br J Psychiatry* 1994;165:324–332.

48. Ferman TJ, Boeve BF, Smith GE, et al. REM sleep behavior disorder and dementia: cognitive differences when compared with AD. *Neurology* 1999;52:951–957.

49. Perry EK, Marshall E, Thompson P, et al. Monoaminergic activities in Lewy body dementia: relation to hallucinosis and extrapyramidal features. *J Neural Transm* 1993;6:167–177.

50. Lippa CF, Smith TW, Perry E. Dementia with Lewy bodies: choline acetyltransferase parallels nucleus basalis pathology. *J Neural Transm* 1999;106:525–535.

51. Tiraboschi P, Hansen LA, Alford M, et al. Cholinergic dysfunction in diseases with Lewy bodies. *Neurology* 2000;54:407–411.

52. Perry EK, Court J, Goodchild R, et al. Clinical neurochemistry: developments in dementia research based on brain bank material. *J Neural Transm* 1998;105:915–933.

53. Perry EK, Smith CJ, Court JA, et al. Cholinergic, nicotinic and muscarinic receptors in dementia of Alzheimer, Parkinson and Lewy body types. *J Neural Transm* 1990;2:149–158.

54. Shiozaki K, Iseki E, Uchiyama H, et al. Alterations of muscarinic acetylcholine receptor subtypes in diffuse Lewy body disease: relation to Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 1999;67:209–213.

54a. Martin-Ruiz C, Court J, Lee M, et al. Nicotinic receptors in dementia of Alzheimer, Lewy body and vascular types. *Acta Neurol Scand* 2000;102:34–51.

55. Court J, Spurden D, Lloyd S, et al. Neuronal nicotinic receptors in dementia with Lewy bodies and schizophrenia: alpha-bungarotoxin and nicotine binding in the thalamus. *J Neurochem* 1999;73:1590–1597.

56. Court J, Piggott M, Lloyd S, et al. Nicotine binding in human striatum: elevation in schizophrenia and reductions in dementia with Lewy bodies, Parkinson's disease and Alzheimer's disease, and in relation to neuroleptic medication. *Neuroscience* 2000;98:79–87.

57. Sabbagh MN, Corey-Bloom J, Alford M, et al. Correlation of nicotinic receptor binding with clinical and neurochemical markers in Alzheimer's disease and dementia with Lewy bodies. *Neurobiol Aging* 1998;19:S207.

58. Hansen LA, Daniel SE, Wilcock GK, et al. Frontal cortical synaptophysin in Lewy body diseases: relation to Alzheimer's disease and dementia. *J Neurol Neurosurg Psychiatry* 1998;64:653–656.

59. Piggott MA, Marshall EF, Thomas N, et al. Striatal dopaminergic markers in dementia with Lewy bodies, Alzheimer's disease and Parkinson's disease: rostrocaudal distribution. *Brain* 1999;122:1449–1468.

60. Piggott MA, Perry EK, Marshall EF, et al. Nigrostriatal dopaminergic activities in dementia with Lewy bodies in relation to neuroleptic sensitivity: comparison with Parkinson's disease. *Biol Psychiatry* 1998;44:764–774.

61. El-Agnaf OM, Curran MD, Wallace A, et al. Mutation screening in exons 3 and 4 of alpha-synuclein in sporadic Parkinson's and sporadic and familial dementia with Lewy bodies. *Neuroreport* 1998;9:3925–3927.

62. Saitoh T, Xia Y, Chen X, et al. The *CYP2D6B* mutant allele is overrepresented in the Lewy body variant of Alzheimer's disease. *Ann Neurol* 1995;37:110–112.

63. Bordet R, Broly F, Destee A, et al. Genetic polymorphism of cytochrome P-450 2D6 in idiopathic Parkinson's disease and diffuse Lewy body disease. *Clin Neuropharmacol* 1994;17:484–488.

64. Atkinson A, Singleton AB, Steward A, et al. *CYP2D6* is associated with Parkinson's disease but not with dementia with Lewy bodies or Alzheimer's disease. *Pharmacogenetics* 1999;9:31–35.

65. Benjamin R, Leake A, Ince PG, et al. Effects of apolipoprotein E genotype on cortical neuropathology in senile dementia of the Lewy body and Alzheimer's disease. *Neurodegeneration* 1995;4:443–448.

66. Camicioli R, Kaye J, Payami H, et al. Apolipoprotein E epsilon4 is associated with neuronal loss in the substantia nigra in Alzheimer's disease. *Dementia Geriatr Cogn Disord* 1999;10:437–441.

67. de la Fuente-Fernandez R, Nunez MA, Lopez E. The apolipoprotein E epsilon 4 allele increases the risk of drug-induced hallucinations in Parkinson's disease. *Clin Neuropharmacol* 1999;22:226–230.

68. Xu W, Liu L, Emson P, et al. The CCTTT polymorphism in the *NOS2A* gene is associated with dementia with Lewy bodies. *Neuroreport* 2000;11:297–299.

69. Liberatore GT, Jackson-Lewis V, Vukosavic S, et al. Inducible nitric oxide synthase stimulates dopaminergic neurodegeneration in the MPTP model in Parkinson's disease. *Nat Med* 1999;5:1403–1409.

70. Singleton AB, Lamb H, Leake A, et al. Failure to find an association between an intronic polymorphism in the presenilin 1 gene and late onset Alzheimer's disease and dementia with Lewy bodies. *Neuroreport* 1997;8:3637–3639.

71. Lamb H, Christie J, Singleton AB, et al. Apolipoprotein E and alpha-1 anti-chymotrypsin polymorphism genotyping in Alzheimer's disease and in dementia with Lewy bodies: distinctions between diseases. *Neurology* 1998;50:388–391.

72. Singleton AB, Gibson AM, McKeith IG, et al. α_2-Macroglobulin polymorphisms in Alzheimer's disease and dementia with Lewy bodies. *Neuroreport* 1999;10:1507–1510.

73. Singleton AB, Gibson AM, Smith G, et al. Butyrylcholinesterase K: an association with dementia with Lewy bodies. *Lancet* 1998;351:1818.

74. Gomez-Tortosa E, Newell K, Irizarry MC, et al. Clinical and quantitative pathological correlates of dementia with Lewy bodies. *Neurology* 1999;53:1284–1291.

75. Samuel W, Alford M, Hofstetter C, et al. Dementia with Lewy body versus pure Alzheimer's disease: differences in cognition neuropathology, cholinergic dysfunction and synapse density. *J Neuropathol Exp Neurol* 1997;56:499–508.

76. Perry RH, Jaros E, Irving D, et al. What is the neuropathological basis of dementia associated with Lewy bodies? In: Perry R, McKeith IG, Perry E, eds. *Dementia with Lewy bodies*. New York: Cambridge University Press, 1996:212–223.

77. Alexander GE, Crutcher MD, DeLong MR. Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Prog Brain Res* 1990;85:119–146.

78. Garcia-Rill E. The basal ganglia and the locomotor regions. *Brain Res* 1986;396:47–63.

79. Schenck CH, Bundlie SR, Mahowald MW. Delayed emergence of a parkinsonian disorder in 38% of 29 older men initially diagnosed with idiopathic rapid eye movement sleep behavior disorder. *Neurology* 1996;46:388–393.

80. Perry EK, Marshall E, Kerwin J, et al. Evidence of a monoaminergic–cholinergic imbalance related to visual hallucinations in Lewy body dementia. *J Neurochem* 1990;55:1454–1456.

81. Ballard C, Piggott M, Johnson M, et al. Delusions associated with elevated muscarinic M1-receptor binding in dementia with Lewy bodies. *Ann Neurol* 2000;48:868–876.

82. Briel RC, McKeith IG, Barker WA, et al. EEG findings in dementia with Lewy bodies and Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 1999;66:401–403.

83. Perry E, Walker M, Grace J, et al. Acetylcholine in mind: a neurotransmitter correlate of consciousness? *Trends Neurosci* 1999;22:273–280.

84. Harvey GT, Hughes J, McKeith IG, et al. Magnetic resonance imaging differences between dementia with Lewy bodies and Alzheimer's disease: a pilot study. *Psychol Med* 1999;29:181–187.

85. Barber R, Gholkar A, Scheltens P, et al. Medial temporal lobe atrophy on MRI in dementia with Lewy bodies. *Neurology* 1999;52:1153–1158.

86. Barber R, Scheltens P, Gholkar A, et al. White matter lesions on magnetic resonance imaging in dementia with Lewy bodies, Alzheimer's disease, vascular dementia, and normal aging. *J Neurol Neurosurg Psychiatry* 1999;67:66–72.

87. Jobst KA, Barnetson LPD, Shepstone BJ. Accurate prediction of confirmed Alzheimer's disease and the differential diagnosis of dementia: the use of NINCDS-ADRDA and DSM IIIR criteria, SPET, X-ray CT and ApoE4 in medial temporal lobe dementias. *Int Psychogeriatr* 1998;10:271–302.

88. Jack CR Jr, Petersen RC, O'Brien PC, et al. MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease. *Neurology* 1992;42:183–188.

89. O'Brien JT, Desmond P, Ames D, et al. Temporal lobe magnetic resonance imaging can differentiate Alzheimer's disease from normal ageing, depression, vascular dementia and other causes of cognitive impairment. *Psychol Med* 1997;27:1267–1275.

90. Robles A, Rodriguez RM, Aldrey JM, et al. Diagnostico clinico de la demencia asociada a cuerpos de Lewy corticales. *Rev Neurol (Barc)* 1995;23:62–66.

91. Hashimoto M, Kitagaki H, Imamura T, et al. Medial temporal and whole-brain atrophy in dementia with Lewy bodies: a volumetric MRI study. *Neurology* 1998;51:357–362.

92. Fox MC, Freeborough PA, Rossor MN. Visualisation and quantification of rates of atrophy in Alzheimer's disease. *Lancet* 1996;348:94–97.

93. O'Brien JT, Eagger S, Syed GS, et al. A study of regional cerebral blood flow and cognitive performance in Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 1992;55:1182–1187.

94. Neary D, Snowden JS, Shields RA, et al. Single photon emission tomography using ^{99m}Tc-HM-PAO in the investigation of dementia. *J Neurol Neurosurg Psychiatry* 1987;50:101–109.

95. Read SL, Miller BL, Mena I, et al. SPECT in dementia: clinical and pathological correlation. *J Am Geriatr Soc* 1995;43:1243–1247.

96. Pizzolato G, Dam M, Borsato N, et al. [^{99m}Tc]-HM-PAO-SPECT in Parkinson's disease. *J Cereb Blood Flow Metab* 1988;8:S101–S108.

97. Habert MO, Spampinato U, Mas JL, et al. A comparative technetium 99m hexamethylpropylene amine oxime SPET study in different types of dementia. *Eur J Nucl Med* 1991;18:3–11.

98. Donnemiller E, Heilmann J, Wenning GK, et al. Brain perfusion scintigraphy with ^{99m}Tc-HMPAO or ^{99m}Tc-ECD and ^{123}I-beta-CIT single-photon emission tomography in dementia of the Alzheimer-type and diffuse Lewy body disease. *Eur J Nucl Med* 1997;24:320–325.

99. Defebvre LJ, Leduc V, Duhamel A, et al. Technetium HMPAO SPECT study in dementia with Lewy bodies, Alzheimer's disease and idiopathic Parkinson's disease. *J Nucl Med* 1999;40:956–962.

100. Ishii K, Imamura T, Sasaki M, et al. Regional cerebral glucose metabolism in dementia with Lewy bodies and Alzheimer's disease. *Neurology* 1998;51:125–130.

101. Higuchi M, Tashiro M, Arai H, et al. Glucose hypometabolism and neuropathological correlates in brains of dementia with Lewy bodies. *Exp Neurol* 2000;162:247–256.

102. Imamura T, Ishii K, Hirono N, et al. Visual hallucinations and regional cerebral metabolism in dementia with Lewy bodies (DLB). *Neuroreport* 1999;10:1903–1907.

103. Walker Z, Costa DC, Ince P, et al. *In vivo* demonstration of dopaminergic degeneration in dementia with Lewy bodies. *Lancet* 1999;354:646–647.

104. Walker Z, Costa DC, Janssen AG, et al. Dementia with Lewy bodies: a study of post-synaptic dopaminergic receptors with iodine-123 iodobenzamide single-photon emission tomography. *Eur J Nucl Med* 1997;24:609–614.

105. Kuhl DE, Minoshima S, Fessler JA, et al. *In vivo* mapping of cholinergic terminals in normal aging, Alzheimer's disease, and Parkinson's disease. *Ann Neurol* 1996;40:399–410.

106. Levy R, Eagger S, Griffiths M, et al. Lewy bodies and response to tacrine in Alzheimer's disease. *Lancet* 1994;343:176.

107. Wilcock GK, Scott MI. Tacrine for senile dementia of Alzheimer's or Lewy body type. *Lancet* 1994;344:544.

108. Shea C, Macknight C, Rockwood R. Donepezil for treatment of dementia with Lewy bodies: a case series of nine patients. *Int Psychogeriatr* 1998;10:229–239.

109. McKeith IG, Del Ser T, Spano PF, et al. Efficacy of rivastigmine in dementia with Lewy bodies: results of a randomised placebo-controlled international study. *Lancet* 2000;356:2031–2036.

110. Burke WJ, Pfeiffer RF, McComb RD. Neuroleptic sensitivity to clozapine in dementia with Lewy bodies. *J Neuropsychiatry Clin Neurosci* 1998;10:227–229.

111. Walker Z, Grace J, Overshot R, et al. Olanzapine in dementia with Lewy bodies: a clinical study. *Int J Geriatr Psychiatry* 1999;14:459–466.

112. Harrison RH, McKeith IG. Senile dementia of Lewy body type—a review of clinical and pathological features: implications for treatment. *Int J Geriatr Psychiatry* 1995;10:919–926.

113. Masliah E, Rockenstein E, Veinbergs I, et al. Dopaminergic loss and inclusion body formation in alpha-synuclein mice: implications for neurodegenerative disorders. *Science* 2000;287:1265–1269.

114. van Duinen SG, Lammers GJ, Maat-Schieman ML, et al. Numerous and widespread alpha-synuclein-negative Lewy bodies in an asymptomatic patient. *Acta Neuropathol (Berl)* 1999;97:533–539.

115. Clementi F, Court J, Perry E. Involvement of neuronal nicotinic receptors in disease. In: Clementi F, ed. *Handbook of experimental pharmacology*. 144:751–812.

116. Litvan I, McIntyre A, Goetz CG, et al. Accuracy of the clinical diagnoses of Lewy body disease, Parkinson's disease and dementia with Lewy bodies. *Arch Neurol* 1998;55:969–978.

117. Papka M, Schiffer R, Rubio A. Diagnosing Lewy body disease; accuracy of clinical criteria in detecting Lewy body pathology. *Neurobiol Aging* 1998;19:S203.

118. Holmes C, Cairns N, Lantos P, et al. Validity of current clinical criteria for Alzheimer's disease, vascular dementia and dementia with Lewy bodies. *Br J Psychiatry* 1999;174:45–51.

119. Luis CA, Barker WW, Gajaraj K, et al. Sensitivity and specificity of three clinical criteria for dementia with Lewy bodies in an autopsy-verified sample. *Int J Geriatr Psychiatry* 1999;14:526–533.

120. Lopez OL, Litvan I, Catt KE, et al. Accuracy of four clinical diagnostic criteria for the diagnosis of neurodegenerative dementias. *Neurology* 1999;53:1292–1299.

121. Verghese J, Crystal HA, Dickson DW, et al. Validity of clinical criteria for the diagnosis of dementia with Lewy bodies. *Neurology* 1999;53:1974–1982.

122. McKeith IG, Ballard CG, Perry RH, et al. Predictive accuracy of clinical diagnostic criteria for dementia with Lewy bodies—a prospective neuropathological validation study. *Neurology* 1998;50:181.

123. Perry E, Goodchild R, Piggott M. Neurotransmitter systems in dementia. In: Hodges J, ed. *Early onset dementia (in press)*.

124. Förstl H. The Lewy body variant of Alzheimer's disease: clinical, pathological and conceptual issues. *Eur Arch Psychiatry Clin Neurosci* 1999;249:64–67.

125. Ince PG, Perry EK, Morris CM. Dementia with Lewy bodies. A distinct non-Alzheimer dementia syndrome? *Brain Pathol* 1998;8:299–324.

92

Molecular Pathophysiology of Stroke

Steven H. Graham

Robert W. Hickey

Steven H. Graham: Department of Neurology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

Robert W. Hickey: Department of Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

Cerebral ischemia occurs when the amount of oxygen and other nutrients supplied by blood flow is insufficient to meet the metabolic demands of brain tissue. In ischemic stroke, the blood supply to the brain is disrupted by cerebrovascular disease. For decades, extensive research and clinical approaches to combat stroke have focused on the vascular aspects of cerebral ischemia. Therapeutic advances, including carotid endarterectomy, thrombolytic therapy, anticoagulation for cardiogenic stroke, antiplatelet agents, and the treatment of risk factors such as hypertension and hyperlipemia, have had significant effects on the morbidity and mortality of stroke.

The final event in cerebral ischemia is the death of neurons, resulting in irreversible loss of neurologic function. The advent of animal and tissue culture models of ischemia has led to many new insights into the mechanisms by which ischemic neurons die. If ischemia is complete and prolonged, neuronal death is inevitable. However, it has become increasingly clear that many secondary biochemical changes that exacerbate injury occur in response to the initial insult. In models of cerebral ischemia in rodents, as much as 50% or more of ischemic brain may be spared from infarction by preventing these secondary biochemical events. Understanding of the mechanisms by which neuronal cell death takes place has resulted in a number of therapeutic strategies that aim to prevent secondary biochemical changes and thus decrease the damage that results from cerebral ischemia. These basic mechanisms may also have relevance to other neurodegenerative diseases associated with excessive neuronal death.

This chapter summarizes many of the mechanisms that have been demonstrated to exacerbate the neuronal death caused by hypoxia and hypoglycemia. Ischemic neuronal death may involve the activation of enzymes and receptors that are constitutively expressed in brain. These existing receptors and enzymes do not require energy or the synthesis of new protein to exacerbate *necrotic* cell death. New evidence suggests that ischemic injury may also be exacerbated by the inducible proteins that mediate *programmed* cell death. These mechanisms are appealing targets for therapeutic intervention because they may occur hours or days after the initiation of ischemia.

CELL DEATH: NECROSIS VERSUS PROGRAMMED CELL DEATH

Part of "92 - Molecular Pathophysiology of Stroke "

It has been observed that an orderly expression of new gene products is required to produce programmed cell death during development of the roundworm *Caenorhabditis elegans* (1). This observation has led to intense interest in the hypothesis that the expression of similar death-promoter genes could be important in the pathogenesis of human disease (2). Support for this hypothesis is derived from the existence of oncogenes, death-regulating genes that are either deleted or overexpressed in cancer. Genetic mechanisms that control cell death are clearly relevant to mitotic cells in development, cancer, and the maintenance and turnover of regenerating adult tissues. Neurons may also die by these mechanisms. A classic example is the withdrawal of nerve growth factor from dorsal root sympathetic neurons that results in delayed death, which requires the transcription of new messenger RNA (mRNA) and the synthesis of new proteins (3). Before their death, these neurons undergo morphologic changes associated with *apoptosis* (4), a term originally used to describe the morphologic characteristics associated with programmed cell death.

Programmed cell death is a mechanism by which the organism can remove unnecessary or redundant cells during development or in mature tissues where cell turnover is required. Programmed cell death has several key characteristics: (a) The death process is active, and the expression of new proteins is often involved. (b) Cellular energy stores

are normal until the final stages of cellular death; therefore, energy failure is a late, secondary event in programmed cell death. (c) The activation of endonucleases results in numerous double-stranded DNA breaks at the boundaries between histosomes. The result is DNA fragments in multiples of 400 base pair size that produce characteristic “laddering” on DNA gels. (d) Morphologic changes characteristic of apoptosis, including cytoplasmic and nuclear budding (“apoptotic bodies”), are present. (e) In contrast to necrosis, programmed cell death results in neuronal death with little or no accompanying inflammation. Thus, “collateral damage” to neighboring cells is avoided.

In contrast to programmed cell death, necrotic cell death is characterized by energy failure, which results in inhibition of protein synthesis. Therefore, new gene products may not be expressed. Histologic characteristics of necrotic cell death are cytoplasmic and nuclear swelling, loss of integrity of cell organelles, rupture of the cell membrane, and dissolution of all cell structures. *In vivo*, necrotic cell death is often accompanied by intense inflammation with recruitment of inflammatory cells. This inflammatory response can injure adjacent normal cells. The characteristics of programmed cell death and necrosis are summarized in Table 92.1 .

	Necrosis	Programmed Cell Death
Process	Passive	Active
Energy failure	Primary	Secondary
Protein translation	Blocked	Exacerbates cell death
Morphology	Coagulative necrosis	Apoptosis
DNA fragmentation	None or random, resulting in either no migration or a smear on DNA gels	Occurs at histosome boundaries, resulting in multiples of 400 base fragments producing laddering on DNA gels
Inflammation	Prominent	Little or none

TABLE 92.1. CHARACTERISTICS OF NECROSIS AND PROGRAMMED CELL DEATH

NATURE OF NEURONAL DEATH IN CEREBRAL ISCHEMIA

Part of "92 - Molecular Pathophysiology of Stroke "

In ischemia, a mismatch between energy supply and demand may result in energy failure. Without adequate energy, protein synthesis cannot occur, and the genes that execute programmed cell death may not be expressed. The predominant histologic feature of stroke is infarction. Infarction is synonymous with necrosis (i.e., cytoplasmic swelling, dissolution of organelles and plasma membranes, and inflammation are present). In the permanent middle cerebral artery occlusion model in the rat, loss of glucose utilization is rapid and complete within a few hours (5), with little time or energy available for the synthesis of new gene products. Thus, one would surmise that necrosis is the primary mode of cell death in this model.

Under other circumstances, however, cerebral ischemia may produce neuronal death with many of the characteristics of programmed cell death. In models of transient ischemia, for example, energy failure is transient, and neuronal death develops more slowly than in permanent focal ischemia, with many features of apoptotic cell death. Under these circumstances, cleavage of genomic DNA into fragments of various sizes on DNA gels, characteristic of programmed cell death, occurs (6 ,7). However, the most convincing evidence that the production of new gene products may be important in the pathogenesis of neuronal death after transient ischemia is that protein synthesis inhibitors block delayed death of neurons (8 ,9 and 10). Thus, depending on the duration and severity of ischemia, stroke may produce cell death with features of necrosis or apoptosis.

MECHANISMS OF NECROTIC CELL DEATH

Part of "92 - Molecular Pathophysiology of Stroke "

The primary pathologic mechanism in stroke is the depletion of energy stores; however, considerable evidence indicates that excitatory amino acids (EAAs) exacerbate ischemic injury. EAAs such as glutamate are released by approximately 40% of all synapses in the central nervous system (11). Under physiologic conditions, EAAs participate in many neurologic functions, including memory, movement, sensation, cognition, and synaptic plasticity (12 ,13). However, EAAs can also have a pathologic effect. EAA-mediated toxicity was first demonstrated by Olney and co-workers (14) by peripheral administration of an EAA agonist that selectively killed neurons in the arcuate nucleus of the hypothalamus. These neurons contain high concentrations of glutamate receptors. Choi (15) demonstrated that micromolar extracellular glutamate and other EAAs produce rapid increases in intraneuronal cytosolic Ca^{2+} concentrations.

This increase in intracellular calcium concentration is rapidly lethal to primary neuronal cultures. The importance of calcium entry and excitotoxicity is supported by data demonstrating a direct correlation between extracellular calcium stimulation and neuronal death following exposure to glutamate (16).

The increase in intraneuronal Ca^{2+} in response to extracellular EAAs *in vitro* is mediated by the opening of a receptor-gated ion channel, the *N*-methyl-D-aspartate (NMDA) channel (17). The NMDA channel, named after its highest-affinity ligand, primarily gates calcium entry into the neuron. Treatment with antagonists that compete with glutamate and other EAAs for the receptor (competitive NMDA antagonists) or antagonists that bind to the ion channel itself (noncompetitive antagonists) can block calcium entry into neurons and prevent cell death induced by glutamate (18 ,19). Glycine is a co-agonist that is required in addition to glutamate to open the NMDA Ca^{2+} channel (20). Antagonists that bind to the glycine site on the NMDA receptor also block excitotoxicity *in vitro* (21). In addition to rescuing cells from EAA toxicity by blockade of the EAA receptors, it is possible to rescue neurons in culture by removal of extracellular calcium and sodium from the culture media following glutamate exposure (18). Conversely, inhibition of the sodium–calcium exchanger that normally facilitates extrusion of calcium results in an increase in neuronal death (18 ,22). Similarly, dantrolene, which attenuates decompartmentalization of intracellular stores of calcium, can reduce glutamate neurotoxicity in cortical neurons (23). Finally, neurons containing high concentrations of calcium-binding proteins, such as calbindin or parvalbumin, are relatively resistant to excitotoxic injury (24 ,25). These data provide compelling evidence that EAA-induced increases in intracellular Ca^{2+} are toxic to neurons in culture.

Compelling evidence is also available to indicate that excitotoxicity mediated by the NMDA receptor contributes to injury from cerebral ischemia *in vivo*. A rapid and large increase in the concentration of extracellular amino acids can be monitored by microdialysis after cerebral ischemia (26). Although NMDA antagonists are not effective in global ischemia models in which temperature is carefully controlled (27), a large number of studies have found that they decrease infarction volume in both permanent and temporary middle cerebral artery occlusion models in rodents (28). Blocking the translation of a gene that encodes a subunit of the NMDA receptor with intraventricular injection of antisense oligonucleotides also decreases infarction volume after middle cerebral artery occlusion in the rat (29). These data and many other studies support the hypothesis that excitotoxicity contributes to ischemic injury *in vivo*.

Several calcium-dependent or calcium-induced enzymes mediate the toxic effects of increased intracellular calcium (Fig. 92.1). These include nitric oxide synthase, cyclooxygenase, phospholipase A_2, and calpain 1. Calpain 1 is a calcium-activated protease that has been specifically linked to glutamate receptors in the rat hippocampus (30). Calpain 1 participates in the conversion of xanthine dehydrogenase to xanthine oxidase, which metabolizes xanthine to its reactive oxygen species, superoxide (31). Similarly, phospholipase A_2 is activated by calcium and facilitates the release of arachidonic acid from injured cell membranes (32). Arachidonic

acid is then metabolized by the enzyme cyclooxygenase into a prostaglandin, PGH_2. The cyclooxygenase enzyme may produce a superoxide ion as a by-product of arachidonic acid metabolism (33). In addition, intracellular calcium can activate calcium-dependent isoforms of nitric oxide synthase to produce nitric oxide (34). The nitric oxide then combines with the superoxide produced as the by-product of cyclooxygenase, xanthine oxidase, or other sources to form the highly reactive species peroxynitrite, which exacerbates tissue damage (35). Therefore, EAA-mediated elevation of intracellular calcium concentrations activates both cyclooxygenase and nitric oxide synthase, which then synergistically contribute to ischemic brain injury (36).

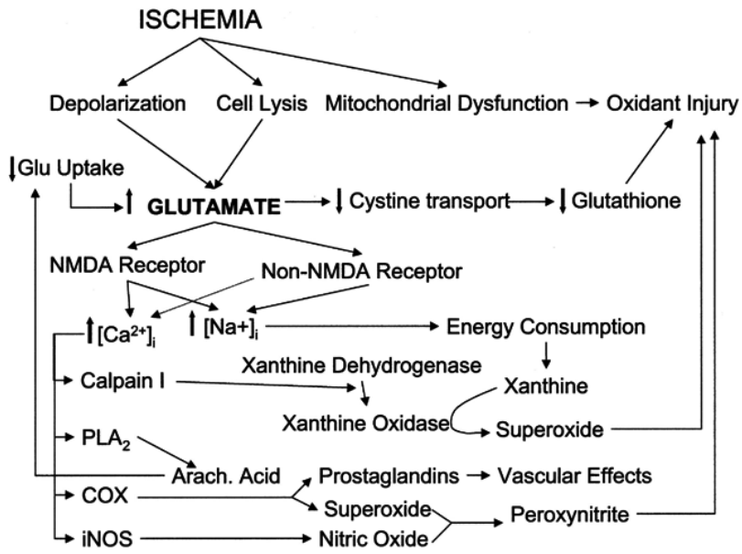


FIGURE 92.1. Schematic diagram illustrating the mechanisms by which ischemia and excitotoxicity injure neurons.

Extracellular EAAs may activate other receptors besides the NMDA channel. EAA receptors can be categorized as ionotropic or metabotropic receptors. Ionotropic receptors are coupled directly to membrane ion channels, whereas metabotropic receptors are coupled to G proteins and modulate intracellular second messengers such as inositol triphosphate, calcium, and cyclic nucleotides. More than 20 genes have been identified that encode subunits of these receptors. The subunits combine in a variety of confirmations to yield receptors with specific pharmacologic and electrophysiologic characteristics (37). Ionotropic receptors can be categorized based on their sensitivity to the selective agonists NMDA, AMPA (α-amino-3-hydroxy-5-methyl-4-isoxasole-proprionic acid), and kainate. The ionotropic receptors depolarize membranes by facilitating an influx of positively charged ions. The NMDA receptor facilitates an influx of both sodium and calcium, whereas the non-NMDA receptors (AMPA and kainate receptors) primarily facilitate an influx of sodium. However, some of the kainate and AMPA receptors are comprised of subunits that allow calcium permeability (38). This may be relevant to ischemic injury because in neurons after cerebral ischemia, glutamate receptor 2 (GR_2), a subunit necessary for non-NMDA receptors to exclude Ca^{2+}, is relatively depleted (39). Accordingly, these non-NMDA subunits may become calcium-permeable after ischemia. The metabotropic receptors may also increase intracellular calcium by mobilizing calcium from stores in the endoplasmic reticulum. Studies with antagonists of the metabotropic receptor show that, depending on their subunit specificity, some, but not all, drugs of this class are neuroprotective in models of focal ischemia (40 ,41).

In addition to the direct downstream effects of enzymes that are activated by elevation of intracellular calcium, a number of complex interactions and positive feedback loops augment the contribution of EAAs to ischemic brain injury. For example, free arachidonic acid can potentiate NMDA-evoked currents in neurons (42) and inhibit reuptake of glutamate by astrocytes (43). In addition, platelet-activating factor, a phospholipase A_2 metabolite, can stimulate the release of glutamate (44). Acidotic conditions favor the release of free iron, which can then participate in the metabolism of peroxide into the hydroxyl radical (Fenton reaction) (45). In addition, glutamate can interfere with the function of the cystine transporter. Inhibition of the cystine transporter results in decreased intracellular concentrations of glutathione and diminished intracellular endogenous antioxidant stores (46).

In vivo, excitotoxicity may be ameliorated by additional strategies besides inhibition of the NMDA receptor (Fig. 92.2). Glutamate release into synaptic cleft, where it interacts with EAA receptors, is primarily mediated by the release of glutamate from the synaptic pool. Thus, a large component of excessive neuronal excitation may be the result of synaptic release of EAAs. Neuronal depolarization of presynaptic neurons in turn depends on activation of non-NMDA receptor-gated channels and other depolarizing neurotransmitter receptors. The excitatory action of depolarizing neurotransmitter receptors is countered by hyperpolarizing receptor-gated ion channels, such as the GABA (γ-aminobutyric

acid) receptor. Propagation of the action potential induced by depolarization of the neuronal cell body requires voltage-dependent sodium channels. Finally, the release of glutamate itself depends on P- and Q-type voltage-dependent calcium channels. Glutamate release into the synaptic cleft can bind to the NMDA receptor and open the calcium channel. As a result, calcium enters the cell driven by its concentration gradient. However, intraneuronal calcium may increase by other mechanisms. Postsynaptic voltage-dependent calcium channels may allow calcium entry into the neuron when cells are depolarized, and glutamate released into the extracellular cleft may activate non-NMDA receptor-gated channels and depolarize the neuron. Also, Na^+ may enter the cell via the NMDA receptor-gated channel and depolarize the neuron. Thus, excitotoxicity may be ameliorated at a number of sites *in vivo*.

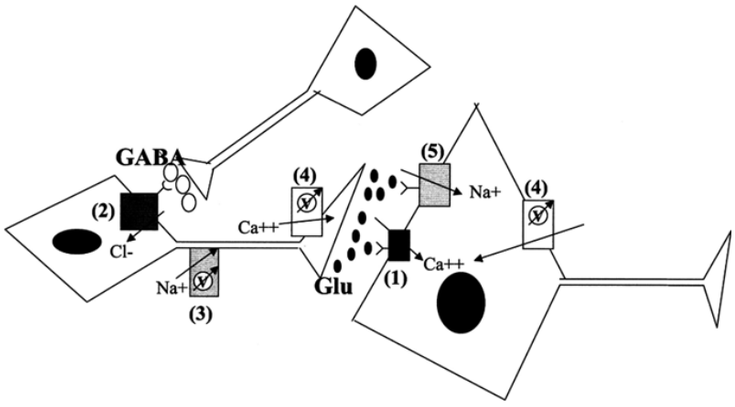


FIGURE 92.2. A simplified neuronal circuit diagram illustrating the ion channels that determine the synaptic release of glutamate and intraneuronal Ca^{2+} concentrations in response to ischemia. 1, *N*-methyl-D-aspartate (NMDA) receptor-gated ion channel; 2, α-aminobutyric acid (GABA) receptor-gated Cl^- channel; 3, voltage-dependent Na^+ channel; 4, voltage-dependent Ca^{2+} channel; 5, non-NMDA receptor-gated ion channel. See color version of figure.

Many drugs that can inhibit excitotoxicity at each of these steps have been developed. GABA agonists such as clomethazole have been shown to be neuroprotective *in vivo* and are currently undergoing clinical trials (47 ,48). In rodent models of stroke, BW1003,619 and phosphenytoin prevent prolonged opening of the voltage-dependent sodium channel, ameliorate increases in extracellular glutamate, and decrease infarction volume (49 ,50 and 51). Drugs that prevent prolonged opening of P- and Q-type calcium channel antagonists are also neuroprotective in animal models of stroke (52). In contrast to their very limited effects in primary neuronal tissue culture models, non-NMDA antagonists are very effective in both global and focal ischemia models in rodents. Indeed, such agents have a longer time window of efficacy than do NMDA antagonists when administered after the onset of ischemia (53 ,54). Likewise, voltage-dependent calcium channel antagonists are not effective *in vitro*; however, the voltage-dependent calcium channel antagonist nimodipine is effective in reducing infarction volume in temporary focal ischemia in rats (55).

Blockade of excitotoxicity via all these pharmacologic strategies has proved effective in temporary focal ischemia models in rodents, the model that most closely resembles human stroke. Unfortunately, results with these agents in human trials have to date been very disappointing, for several possible reasons. First, drugs that affect neurotransmission in the brain have many undesirable side effects, which in turn have led to reductions to doses that may have been ineffective. Side effects include effects on respiration and cardiac rhythm. In addition, agents that directly antagonize the NMDA receptor may injure a circumspect population of neurons in the cingulate and retrosplenial cortex in rodents (56), and may induce hallucinations and psychosis in humans (57). Another obvious reason for the lack of efficacy in these drugs in clinical trials is the time interval between the onset of ischemia and the administration of drug. When given before the onset of ischemia, these treatments can spare 50% or more of ischemic rat brain tissue from eventual infarction. When given after the onset of ischemia, they are progressively less effective; however, such agents are effective up to 2 hours after the onset of middle cerebral artery occlusion in the rat. In the clinical trials, most patients were enrolled 6 to 12 hours after the onset of ischemia, long after the time that these drugs were effectively administered in animal studies.

Whatever the reason for the failure of these anti-excitotoxic drugs in human trials, it has become clear that it may be more practical to select treatment approaches that target mechanisms that are active at longer intervals after ischemia. Accordingly, efforts to understand the delayed mechanisms of neuronal injury have been increased, in particular the role of programmed cell death in ischemic neurons.

MECHANISMS OF PROGRAMMED CELL DEATH

Part of "92 - Molecular Pathophysiology of Stroke "

Many of the key molecular events in programmed cell death have now been determined (Fig. 92.3). Just as calcium entry into the neuron is a key step in excitotoxicity, the release of cytochrome *c* from the mitochondria is a key event in initiating apoptosis in many cell types. Cytosolic cytochrome

c complexes with APAF-1 and procaspase 9 (58). As a result, procaspase 9 is cleaved into its active form, caspase 9. Caspase 9 then cleaves and activates other caspases, including caspase 3.

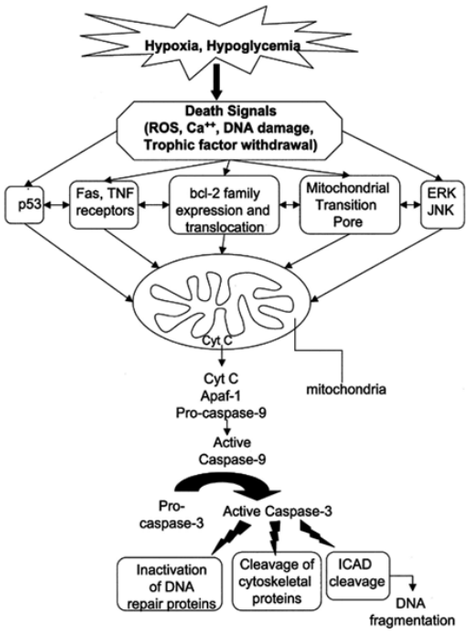


FIGURE 92.3. A schematic diagram illustrating the molecular mechanisms that control programmed cell death.

Caspases are a family of proteases that play a key role in executing programmed cell death. They were first identified by their homology with *CED3*, the key gene that irreversibly commits neurons in *C. elegans* to programmed cell death. A dozen mammalian caspases have been identified that have a variety of roles in executing programmed cell death and other cellular functions. Among the caspases, caspase 3 has the closest homology with *CED3* and appears to play a key role as the final committed step in programmed cell death. Caspase 3 executes programmed cell death via cleavage of many other proteins. These proteolytic targets of caspase 3 include cytoskeletal protein(s), DNA repair proteins such as PARP, and other proteins (59). Caspase 3 also cleaves ICAD, an inhibitor of CAD, an endonuclease that cleaves DNA between histosomes. The result is cleavage of DNA between histosomes, a hallmark of programmed cell death (60).

The egress of cytochrome ^{c} from the mitochondria into the cytosol is controlled by several mechanisms. Genes of the bcl-2 family play an important role in controlling cytochrome *c* egress. Anti-apoptotic bcl-2 family members, such as bcl-2 itself and bcl-x-long, inhibit the egress of cytochrome ^{c} (61 ,62). Pro-apoptotic members of the bcl-2 family, such as bcl-x-short and bax, can form dimers with themselves or with anti-apoptotic bcl-2 family members. The balance between the pro-apoptotic and anti-apoptotic bcl-2 family proteins in the mitochondrial membrane determines whether permeability of the membrane will increase to allow egress of cytochrome *c* into the cytoplasm. Under some circumstances, cytochrome *c* exits the mitochondria via the mitochondrial permeability transition pore. This pore can open in response to prolonged depolarization, produced by such stimuli as an increase in intracellular calcium (63). Furthermore, pro-apoptotic bcl-2 family members such as bax may also interact with this pore (64). However, bcl-2 family members themselves may form pores in membranes (65), and some evidence indicates that bax induces egress of cytochrome *c* from the mitochondria independently of the mitochondrial permeability transition pore (66). Initiation of the mitochondrial apoptosis is also controlled by expression and translocation of other numerous bcl-2 family members. For example, translocation of bax from the cytosol to the mitochondria initiates programmed cell death (67). Bad is phosphorylated before being translocated to the mitochondria (68). More than 20 additional proteins are found in the bcl-2 family, including many that are also involved in mitochondrial homeostasis. Thus, a key event in apoptosis, egress of cytochrome *c* from the mitochondria, is controlled by bcl-2 family proteins.

The molecular mechanisms by which programmed cell death is initiated are numerous and complex. Programmed cell death may be activated via cell surface receptors, including the Fas receptor and tumor necrosis factor-α (TNF-α) (69 ,70). Activation of these receptors triggers activation of caspase 8, which in turn cleaves the bcl-2 family protein bid (71). The cleaved bid then translocates from the cytoplasm to the mitochondria, where it initiates cytochrome *c* egress (72). Other mechanisms by which the initiation of programmed cell death is controlled include the ERK (externally regulated kinase) and JNK protein kinase cascades (73). Finally, DNA base oxidation and other DNA damage may initiate programmed cell death via expression of the p53 transcription factor. These and other mechanisms may be involved in the initiation of programmed cell death in ischemic neurons.

PROGRAMMED CELL DEATH AFTER CEREBRAL ISCHEMIA

Part of "92 - Molecular Pathophysiology of Stroke "

Evidence indicates that many of the mechanisms that initiate programmed cell death are activated in ischemic neurons under certain conditions. The mRNA of the Fas ligand is induced by forebrain ischemia (74). Expression of the Fas ligand and associated proteins and infarction volumes was smaller in LPR mice that expressed a dysfunctional Fas ligand than in wild-type controls (75). The Fas receptor is also up-regulated after cerebral ischemia in rat brain (76). TNF-α mRNA transcription is induced as an early response after cerebral ischemia (77). Expression of the TNF receptor is also increased after cerebral ischemia (78). TNF-binding protein, a protein that binds and inhibits TNF-α, reduced infarction volume after middle cerebral artery occlusion in rats (79). However, ischemic injury was exacerbated in TNF-α-receptor null mice, which suggests that TNF signaling pathways may instead have beneficial effects in ischemic injury under some circumstances (80). Caspase 8, which is activated by both the Fas and TNF receptors, is expressed and activated after cerebral ischemia (81). Changes in expression and activity of both the ERK and JNK kinase pathways occur following cerebral ischemia. The M-terminal kinases of c-Jun are activated after ischemia and phosphorylate c-Jun (82). The increased expression of ERK after focal ischemia and inhibitor of NEK-1, another kinase in the ERK pathway, protect the brain against focal cerebral ischemia (83 ,84). Single-stranded DNA damage is an early event in cerebral ischemia reperfusion injury and may trigger expression of p53 (85 ,86). Expression of p53 is induced after cerebral ischemia (87).

A number of studies in cerebral ischemia support a role for bcl-2 family genes in controlling ischemic neuronal death. In rodent models of ischemia, anti-apoptotic members of the bcl-2 family, including bcl-2 and bcl-x long, are expressed in neurons that are ischemic yet survive. Expression of pro-apoptotic members of the family, such as bax, is increased in neurons that are ischemic and die, such as

CA1 hippocampal neurons in models of global ischemia (88 ,89). Overexpressing anti-apoptotic members of the bcl-2 family protects neurons against ischemia. Transgenic mice that overexpress bcl-2 in neurons have a smaller infarction volume after temporary focal ischemia than do wild-type controls (90). Similarly, overexpression of bcl-2 by means of herpes simplex viral vectors protects neurons against ischemia *in vivo* (91 ,92). These studies show that overexpression of bcl-2 protein before ischemia is neuroprotective. To address whether bcl-2 that is endogenously expressed after ischemia has a protective role, antisense oligonucleotides were used to prevent translation of bcl-2 induced after ischemia. Rats treated with bcl-2 antisense oligonucleotides had a larger infarction volume than did rats treated with sense oligonucleotides or vehicle after temporary focal ischemia. These results suggest that expression of endogenous bcl-2 increases survival of ischemic neurons (93).

Abundant *in vivo* evidence also suggests that caspase activity exacerbates ischemic injury. Transgenic mice that express a dominant negative mutation of caspase 1 had smaller infarctions than did their wild-type litter mates (94). Furthermore, intraventricular infusion of peptide inhibitors of caspases decreased infarction volume in rats subjected to temporary middle cerebral artery occlusion. These peptide inhibitors of caspases also blocked damage in response to injection of excitotoxins (95). Caspase 3 mRNA and protein expression is induced in CA1 neurons after global ischemia. Caspase 3 is activated, and treatment with a specific peptide inhibitor of caspase 3 ameliorated neuronal death in the global ischemia model (96). These and other studies support a role for caspases in ischemic neuronal injury.

CONCLUSION

Part of "92 - Molecular Pathophysiology of Stroke "

If ischemia is of long duration or severe, death is rapid and necrotic. However, if ischemia is transient or incomplete, the genes that execute programmed cell death may be activated. Which process occurs depends on the duration and severity of the ischemic insult (97 ,98). In rodent models of ischemia, the duration and severity of ischemia are controlled, so one can produce ischemic neuronal death with either characteristic. Indeed, a spectrum of biochemical and morphologic changes occurs with characteristics of both death processes. In the human disease, the duration and severity of ischemia depend on the exact location of the arterial occlusion and whether or not reperfusion occurs. Reperfusion can occur spontaneously or as the result of thrombolytic therapy. Just as in the animal models, arterial occlusions produce regions where blood flow is nearly completely absent, and also surrounding zones with incomplete ischemia. Therefore, transient and incomplete ischemia occurs in the human disease and so provides the conditions necessary for programmed cell death.

Our knowledge of the mechanisms by which ischemic neurons die has increased considerably. It is now clear that the toxic effects of EAAs exacerbate injury resulting from ischemia. Antagonizing excitotoxicity via a variety of approaches can ameliorate injury in animal models of ischemia; however, these treatments appear to be too toxic and are effective for too short an interval after the onset of ischemia to be practical treatments in humans. When ischemia is transient or less severe, programmed cell death is activated. These events occur hours or days after the onset of ischemia and thus may be more practical targets for treatment. Further work is needed to determine the most effective and practical therapeutic strategies to prevent neuronal death after ischemia.

REFERENCES

1. Hengartner MO, Horvitz HR. *C. elegans* cell survival gene *ced*-9 encodes a functional homolog of the mammalian proto-oncogene bcl-2. *Cell* 1994;76:665–676.

2. Thompson CB. Apoptosis in the pathogenesis and treatment of disease. *Science* 1995;267:1256–1262.

3. Martin DP, Schmidt RE, DiStefano PS, et al. Inhibitors of protein synthesis and RNA synthesis prevent neuronal death caused by nerve growth factor deprivation. *J Cell Biol* 1988;106:829–844.

4. Saunders J Jr. Death in embryonic systems. *Science* 1966;154:604–612.

5. Shiraishi K, Sharp FR, Simon RP. Sequential metabolic changes in rat brain following middle cerebral artery occlusion: a 2-deoxyglucose study. *J Cereb Blood Flow Metab* 1989;9:765–773.

6. MacManus JP, Buchan AM, Hill IE, et al. Global ischemia can cause DNA fragmentation indicative of apoptosis in rat brain. *Neurosci Lett* 1993;164:89–92.

7. Li Y, Sharov VG, Jiang N, et al. Ultrastructural and light microscopic evidence of apoptosis after middle cerebral artery occlusion in the rat. *Am J Pathol* 1995;146:1045–1051.

8. Goto K, Ishige A, Sekiguchi K, et al. Effects of cycloheximide on delayed neuronal death in rat hippocampus. *Brain Res* 1990;534:299–302.

9. Tortosa A, Rivera R, Ferrer I. Dose-related effects of cycloheximide on delayed neuronal death in the gerbil hippocampus after bilateral transitory forebrain ischemia. *J Neurol Sci* 1994;121:10–17.

10. Linnik MD, Zobrist RH, Hatfield MD. Evidence supporting a role for programmed cell death in focal cerebral ischemia in rats. *Stroke* 1993;24:2002–2008.

11. Fonnum F. Glutamate: a neurotransmitter in mammalian brain. *J Neurochem* 1984;42:1–11.

12. Gasic GP, Hollmann M. Molecular neurobiology of glutamate receptors. *Annu Rev Physiol* 1992;54:507–536.

13. Lipton SA, Kater SB. Neurotransmitter regulation of neuronal outgrowth, plasticity and survival. *Trends Neurosci* 1989;12:265–270.

14. Olney JW, Ho OL, Rhee V. Cytotoxic effects of acidic and sulphur-containing amino acids on the infant mouse central nervous system. *Exp Brain Res* 1971;14:61–76.

15. Choi DW. Ionic dependence of glutamate neurotoxicity. *J Neurosci* 1987;7:369–379.

16. Ogura A, Miyamoto M, Kudo Y. Neuronal death *in vitro*: parallelism between survivability of hippocampal neurones and sustained elevation of cytosolic Ca^{2+} after exposure to glutamate receptor agonist. *Exp Brain Res* 1988;73:447–458.

17. Choi DW, Koh JY, Peters S. Pharmacology of glutamate neurotoxicity in cortical cell culture: attenuation by NMDA antagonists. *J Neurosci* 1988;8:185–196.

18. Hartley DM, Choi DW. Delayed rescue of *N*-methyl-D-aspartate receptor-mediated neuronal injury in cortical culture. *J Pharmacol Exp Ther* 1989;250:752–758.

19. Choi DW. Methods for antagonizing glutamate neurotoxicity. *Cerebrovasc Brain Metab Rev* 1990;2:105–147.

20. Thomson AM. Glycine modulation of the NMDA receptor/channel complex. *Trends Neurosci* 1989;12:349–353.

21. Smith SE, Meldrum BS. The glycine-site NMDA receptor antagonist, *r*-(+)-*cis*-beta-methyl-3-amino-1-hydroxypyrrolid-2-one, l-687,414 is anticonvulsant in baboons. *Eur J Pharmacol* 1992;211:109–111.

22. Andreeva N, Khodorov B, Stelmashook E, et al. Inhibition of Na^+/Ca^{2+} exchange enhances delayed neuronal death elicited by glutamate in cerebellar granule cell cultures. *Brain Res* 1991;548:322–325.

23. Frandsen A, Schousboe A. Dantrolene prevents glutamate cytotoxicity and Ca^{2+} release from intracellular stores in cultured cerebral cortical neurons. *J Neurochem* 1991;56:1075–1078.

24. Sloviter RS. Calcium-binding protein (calbindin—D28 kd) and parvalbumin immunocytochemistry: localization in the rat hippocampus with specific reference to the selective vulnerability of hippocampal neurons to seizure activity. *J Comp Neurol* 1989;280:183–196.

25. Weiss JH, Koh J, Baimbridge KG, et al. Cortical neurons containing somatostatin- or parvalbumin-like immunoreactivity are atypically vulnerable to excitotoxic injury *in vitro. Neurology* 1990;40:1288–1292.

26. Benveniste H, Drejer J, Schousboe A, et al. Elevation of the extracellular concentrations of glutamate and aspartate in rat hippocampus during transient cerebral ischemia monitored by intracerebral microdialysis. *J Neurochem* 1984;43:1369–1374.

27. Buchan A, Li H, Pulsinelli WA. The *N*-methyl-D-aspartate antagonist, mk-801, fails to protect against neuronal damage caused by transient, severe forebrain ischemia in adult rats. *J Neurosci* 1991;11:1049–1056.

28. McCulloch J. Excitatory amino acid antagonists and their potential for the treatment of ischaemic brain damage in man. *Br J Clin Pharmacol* 1992;34:106–114.

29. Wahlestedt C, Golanov E, Yamamoto S, et al. Antisense oligodeoxynucleotides to NMDA-R_1 receptor channel protect cortical neurons from excitotoxicity and reduce focal ischaemic infarctions. *Nature* 1993;363:260–263.

30. Siman R, Noszek JC, Kegerise C. Calpain 1 activation is specifically related to excitatory amino acid induction of hippocampal damage. *J Neurosci* 1989;9:1579–1590.

31. Stark K, Seubert P, Lynch G, et al. Proteolytic conversion of xanthine dehydrogenase to xanthine oxidase: evidence against a role for calcium-activated protease (calpain). *Biochem Biophys Res Commun* 1989;165:858–864.

32. Chan PH, Fishman RA, Schmidley JW, et al. Release of polyunsaturated fatty acids from phospholipids and alteration of brain membrane integrity by oxygen-derived free radicals. *J Neurosci Res* 1984;12:595–605.

33. Kukreja RC, Kontos HA, Hess ML, et al. PGH synthase and lipoxygenase generate superoxide in the presence of NADH or NADPH. *Circ Res* 1986;59:612–619.

34. Vane JR, Mitchell JA, Appleton I, et al. Inducible isoforms of cyclooxygenase and nitric oxide synthase in inflammation. *Proc Natl Acad Sci USA* 1994;91:2046–2050.

35. Beckman JS, Koppenol WH. Nitric oxide, superoxide, and peroxynitrite: the good, the bad, and the ugly. *Am J Physiol* 1996;271:C1424–C1437.

36. Nogawa S, Forster C, Zhang F, et al. Interaction between inducible nitric oxide synthase and cyclooxygenase 2 after cerebral ischemia. *Proc Natl Acad Sci USA* 1998;95:10966–10971.

37. Nakanishi S. Molecular diversity of glutamate receptors and implications for brain function. *Science* 1992;258:597–603.

38. Hollmann M, Hartley M, Heinemann S. Ca^{2+} permeability of Ka-AMPA-gated glutamate receptor channels depends on subunit composition. *Science* 1991;252:851–853.

39. Pellegrini-Giampietro DE, Zukin RS, Bennett MV, et al. Switch in glutamate receptor subunit gene expression in CA1 subfield of hippocampus following global ischemia in rats. *Proc Natl Acad Sci USA* 1992;89:10499–10503.

40. Bond A, Ragumoorthy N, Monn JA, et al. Ly379268, a potent and selective group II metabotropic glutamate receptor agonist, is neuroprotective in gerbil global, but not focal, cerebral ischaemia. *Neurosci Lett* 1999;273:191–194.

41. Lam AG, Soriano MA, Monn JA, et al. Effects of the selective metabotropic glutamate agonist Ly354740 in a rat model of permanent ischaemia. *Neurosci Lett* 1998;254:121–123.

42. Miller B, Sarantis M, Traynelis SF, et al. Potentiation of NMDA receptor currents by arachidonic acid. *Nature* 1992;355:722–725.

43. Volterra A, Trotti D, Cassutti P, et al. High sensitivity of glutamate uptake to extracellular free arachidonic acid levels in rat cortical synaptosomes and astrocytes. *J Neurochem* 1992;59:600–606.

44. Clark GD, Happel LT, Zorumski CF, et al. Enhancement of hippocampal excitatory synaptic transmission by platelet-activating factor. *Neuron* 1992;9:1211–1216.

45. Ying W, Han SK, Miller JW, et al. Acidosis potentiates oxidative neuronal death by multiple mechanisms. *J Neurochem* 1999;73:1549–1556.

46. Murphy TH, Miyamoto M, Sastre A, et al. Glutamate toxicity in a neuronal cell line involves inhibition of cystine transport leading to oxidative stress. *Neuron* 1989;2:1547–1558.

47. Snape MF, Baldwin HA, Cross AJ, et al. The effects of chlormethiazole and nimodipine on cortical infarct area after focal cerebral ischemia in the rat. *Neuroscience* 1993;53:837–844.

48. Wahlgren NG, Ranasinha KW, Rosolacci T, et al. Clomethiazole acute stroke study (CLASS): results of a randomized, controlled trial of clomethiazole versus placebo in 1,360 acute stroke patients. *Stroke* 1999;30:21–28.

49. Graham SH, Chen J, Sharp FR, et al. Limiting ischemic injury by inhibition of excitatory amino acid release. *J Cereb Blood Flow Metab* 1993;13:88–97.

50. Smith SE, Lekieffre D, Sowinski P, et al. Cerebroprotective effect of BW619C89 after focal or global cerebral ischaemia in the rat. *Neuroreport* 1993;4:1339–1342.

51. Boxer PA, Cordon JJ, Mann ME, et al. Comparison of phenytoin with noncompetitive *N*-methyl-D-aspartate antagonists in a model of focal brain ischemia in rat. *Stroke* 1990;21:III47–III51.

52. Goldin SM, Subbarao K, Margolin LD, et al. Neuroprotective use-dependent blockers of Na^+ and Ca^{2+} channels controlling presynaptic release of glutamate. *Ann NY Acad Sci* 1995;765:210–229.

53. Sheardown MJ, Suzdak PD, Nordholm L. AMPA, but not NMDA, receptor antagonism is neuroprotective in gerbil global ischaemia, even when delayed 24 hours. *Eur J Pharmacol* 1993;236:347–353.

54. Buchan AM, Lesiuk H, Barnes KA, et al. AMPA antagonists: do they hold more promise for clinical stroke trials than NMDA antagonists? *Stroke* 1993;24:I148–I152.

55. Kawaguchi K, Henshall DC, Simon RP. Parallel dose-response studies of the voltage-dependent Na^+ channel antagonist BW619C89 and the voltage-dependent Ca^{2+} channel antagonist nimodipine in rat transient focal cerebral ischaemia. *Eur J Pharmacol* 1999;364:99–105.

56. Olney JW, Labruyere J, Wang G, et al. NMDA antagonist neurotoxicity: mechanism and prevention. *Science* 1991;254:1515–1518.

57. Woods JH, Koek W, France CP. Behavioral effects of NMDA antagonists.

58. Li P, Nijhawan D, Budihardjo I, et al. Cytochrome *c* and DATP-dependent formation of APAF-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell* 1997;91:479–489.

59. Cohen GM. Caspases: the executioners of apoptosis. *Biochem J* 1997;326:1–16.

60. Enari M, Sakahira H, Yokoyama H, et al. A caspase-activated DNAase that degrades DNA during apoptosis and its inhibitor ICAD. *Nature* 1998;391:43–50.

61. Yang J, Liu X, Bhalla K, et al. Prevention of apoptosis by bcl-2: release of cytochrome *c* from mitochondria blocked. *Science* 1997;275:1129–1132.

62. Kluck RM, Bossy-Wetzel E, Green DR, et al. The release of cytochrome *c* from mitochondria: a primary site for bcl-2 regulation of apoptosis. *Science* 1997;275:1132–1136.

63. Hirsch T, Marzo I, Kroemer G. Role of the mitochondrial permeability transition pore in apoptosis. *Biosci Rep* 1997;17:67–76.

64. Narita M, Shimizu S, Ito T, et al. Bax interacts with the permeability transition pore to induce permeability transition and cytochrome *c* release in isolated mitochondria. *Proc Natl Acad Sci USA* 1998;95:14681–14686.

65. Antonsson B, Conti F, Ciavatta A, et al. Inhibition of bax channel-forming activity by bcl-2. *Science* 1997;277:370–372.

66. Eskes R, Antonsson B, Osen-Sand A, et al. Bax-induced cytochrome *c* release from mitochondria is independent of the permeability transition pore but highly dependent on Mg^{2+} ions. *J Cell Biol* 1998;143:217-224.

67. Wolter KG, Hsu YT, Smith CL, et al. Movement of bax from the cytosol to mitochondria during apoptosis. *J Cell Biol* 1997;139:1281–1292.

68. D'Agata V, Magro G, Travali S, et al. Cloning and expression of the programmed cell death regulator bad in the rat brain. *Neurosci Lett* 1998;243:137–140.

69. Nagata S. Apoptosis regulated by a death factor and its receptor: fas ligand and fas. *Philos Trans R Soc Lond B Biol Sci* 1994;345:281–287.

70. Beutler B, Bazzoni F. TNF, apoptosis and autoimmunity: a common thread? *Blood Cells Mol Dis* 1998;24:216–230.

71. Yin XM, Wang K, Gross A, et al. Bid-deficient mice are resistant to fas-induced hepatocellular apoptosis. *Nature* 1999;400:886–891.

72. Gross A, Yin XM, Wang K, et al. Caspase cleaved bid targets mitochondria and is required for cytochrome *c* release, while bcl-xl prevents this release but not tumor necrosis factor-r1/fas death. *J Biol Chem* 1999;274:1156–1163.

73. Seger R, Krebs EG. The MAPK signaling cascade. *FASEB J* 1995;9:726–735.

74. Matsuyama T, Hata R, Yamamoto Y, et al. Localization of fas antigen MRNA induced in postischemic murine forebrain by *in situ* hybridization. *Brain Res Mol Brain Res* 1995;34:166–172.

75. Martin-Villalba A, Herr I, Jeremias I, et al. Cd95 ligand (fas-l/apo-1l) and tumor necrosis factor-related apoptosis-inducing ligand mediate ischemia-induced apoptosis in neurons. *J Neurosci* 1999;19:3809–3817.

76. Felderhoff-Mueser U, Taylor DL, Greenwood K, et al. Fas/cd95/apo-1 can function as a death receptor for neuronal cells *in vitro* and *in vivo* and is upregulated following cerebral hypoxic-ischemic injury to the developing rat brain. *Brain Pathol* 2000;10:17–29.

77. Wang X, Yue TL, Barone FC, et al. Concomitant cortical expression of TNF-alpha and IL-1 beta MNRAs follows early-response gene expression in transient focal ischemia. *Mol Chem Neuropathol* 1994;23:103–114.

78. Botchkina GI, Meistrell ME 3rd, Botchkina IL, et al. Expression of TNF and TNF receptors (p55 and p75) in the rat brain after focal cerebral ischemia. *Mol Med* 1997;3:765–781.

79. Nawashiro H, Martin D, Hallenbeck JM. Neuroprotective effects of TNF binding protein in focal cerebral ischemia. *Brain Res* 1997;778:265–271.

80. Bruce AJ, Boling W, Kindy MS, et al. Altered neuronal and microglial responses to excitotoxic and ischemic brain injury in mice lacking TNF receptors. *Nat Med* 1996;2:788–794.

81. Velier JJ, Ellison JA, Kikly KK, et al. Caspase-8 and caspase-3 are expressed by different populations of cortical neurons undergoing delayed cell death after focal stroke in the rat. *J Neurosci* 1999;19:5932–5941.

82. Herdegen T, Claret FX, Kallunki T, et al. Lasting N-terminal phosphorylation of c-jun and activation of c-jun N-terminal kinases after neuronal injury. *J Neurosci* 1998;18:5124–5135.

83. Alessandrini A, Namura S, Moskowitz MA, et al. Mek1 protein kinase inhibition protects against damage resulting from focal cerebral ischemia. *Proc Natl Acad Sci USA* 1999;96:12866–12869.

84. Kitagawa H, Warita H, Sasaki C, et al. Immunoreactive akt, pi3-k and erk protein kinase expression in ischemic rat brain. *Neurosci Lett* 1999;274:45–48.

85. Chen J, Jin K, Chen M, et al. Early detection of DNA strand breaks in the brain after transient focal ischemia: implications for the role of DNA damage in apoptosis and neuronal cell death. *J Neurochem* 1997;69:232–245.

86. Canman CE, Chen CY, Lee MH, et al. DNA damage responses: p53 induction, cell cycle perturbations, and apoptosis. *Cold Spring Harb Symp Quant Biol* 1994;59:277–286.

87. Chopp M, Li Y, Zhang ZG, et al. P53 expression in brain after middle cerebral artery occlusion in the rat. *Biochem Biophys Res Commun* 1992;182:1201–1207.

88. Chen J, Zhu RL, Nakayama M, et al. Expression of the apoptosis-effector gene, *bax*, is up-regulated in vulnerable hippocampal CA1 neurons following global ischemia. *J Neurochem* 1996;67:64–71.

89. Krajewski S, Mai JK, Krajewska M, et al. Upregulation of Bax protein levels in neurons following cerebral ischemia. *J Neurosci* 1995;15:6364–6376.

90. Martinou JC, Dubois-Dauphin M, Staple JK, et al. Overexpression of bcl-2 in transgenic mice protects neurons from naturally occurring cell death and experimental ischemia. *Neuron* 1994;13:1017–1030.

91. Linnik MD, Zahos P, Geschwind MD, et al. Expression of bcl-2 from a defective herpes simplex virus-1 vector limits neuronal death in focal cerebral ischemia. *Stroke* 1995;26:1670–1674.

92. Lawrence MS, McLaughlin JR, Sun GH, et al. Herpes simplex viral vectors expressing bcl-2 are neuroprotective when delivered after a stroke. *J Cereb Blood Flow Metab* 1997;17:740–744.

93. Chen J, Simon RP, Nagayama T, et al. Suppression of endogenous bcl-2 expression by antisense treatment exacerbates ischemic neuronal death. *J Cereb Blood Flow Metab* 2000;20:1033–1039.

94. Hara H, Fink K, Endres M, et al. Attenuation of transient focal cerebral ischemic injury in transgenic mice expressing a mutant ice inhibitory protein. *J Cereb Blood Flow Metab* 1997;17:370–375.

95. Hara H, Friedlander RM, Gagliardini V, et al. Inhibition of interleukin 1beta converting enzyme family proteases reduces ischemic and excitotoxic neuronal damage. *Proc Natl Acad Sci USA* 1997;94:2007–2012.

96. Chen J, Nagayama T, Jin K, et al. Induction of caspase-3-like protease may mediate delayed neuronal death in the hippocampus after transient cerebral ischemia. *J Neurosci* 1998;18:4914–4928.

97. Ankarcrona M, Dypbukt JM, Bonfoco E, et al. Glutamate-induced neuronal death: a succession of necrosis or apoptosis depending on mitochondrial function. *Neuron* 1995;15:961–973.

98. Bonfoco E, Krainc D, Ankarcrona M, et al. Apoptosis and necrosis: two distinct events induced, respectively, by mild and intense insults with *N*-methyl-D-aspartate or nitric oxide/superoxide in cortical cell cultures. *Proc Natl Acad Sci USA* 1995;92:7162–7166.

93

Current and Experimental Treatment of Stroke

Daniel L. Small

Paul Morley

Alastair M. Buchan

Daniel L. Small and Paul Morley: Institute for Biological Sciences, National Research Council of Canada, Ottawa, Ontario, Canada.

Alastair M. Buchan: Department of Clinical Neuroscience, Foothills Hospital, Calgary, Alberta, Canada.

Stroke is the third leading cause of death in the United States; approximately 730,000 Americans have a new or recurrent stroke each year. That's one every minute and it costs the health care system $30 billion annually. The incidence of stroke is expected to rise dramatically as the population ages because stroke risk increases with age. The risk of stroke doubles for each decade after age 55. Stroke is a major factor in the late-life dementia that affects more than 40% of Americans over the age of 80. One in four men will have had a disabling stroke by the age of 80 and one in five women by the age of 85.

Currently Alteplase (recombinant tissue plasminogen activator [rtPA]) is the only treatment for acute stroke approved in the United States. Alteplase is a thrombolytic agent that restores cerebral blood flow by removing the vascular occlusion. Alteplase is only an appropriate therapy for a small proportion of stroke patients (˜2%) because it must be given early to achieve efficacy and functional recovery following delayed reperfusion. Less than 15% of patients, however, are admitted to the hospital within the 3-hour safety window (1). Several other new treatments are being tested in the clinic and even more are in preclinical development. Antiplatelet therapy and thrombolytics are aimed at improving cerebral blood flow but there are other therapeutic strategies such as neuroprotectants, antiinflammatory agents, free radical scavengers, and neurotrophic agents. In this chapter we survey the current status of clinical trials for stroke and review these therapeutic strategies.

PATHOPHYSIOLOGY AND PRECLINICAL MODELS OF STROKE

Part of "93 - Current and Experimental Treatment of Stroke "

Stroke in the clinic is represented predominantly by ischemic stroke (80%), in which there is a loss of cerebral blood flow owing to vascular occlusion. The remaining 20% of strokes result from cerebral hemorrhage. Experimentally, there are two main types of *in vivo* stroke models, global and focal ischemia. In global models, the entire forebrain is made ischemic for a brief period (5 to 20 minutes) and then reperfused. After a significant delay (>3 days), selective neuronal death evolves in layers III, V, and VI of the cortex, the CA1 region of the hippocampus, and the striatum (2 ,3 and 4). These models are representative of cardiac arrest and coronary artery bypass surgery rather than clinical stroke. In focal models, the middle cerebral artery is occluded for a more prolonged period (60 to 120 minutes in temporary models and permanently in permanent occlusion models). Focal models are representative of clinical stroke and produce histologic damage similar to ischemic stroke in humans. The focal model produces a pannecrotic core surrounded by a narrow penumbral border. The penumbra is at risk of becoming necrotic but is potentially salvageable given the appropriate therapeutic intervention (3 ,4). The evolution of delayed neuronal death in global and focal ischemia occurs by a cascade of events that unfold temporally and spatially (5). Blocking biochemical processes in the excitotoxic cascade is the rationale behind the various therapeutic strategies for treating stroke and will be the framework for this chapter as we describe therapeutics that have been tested, or are currently being tested, in the clinic.

THERAPEUTIC STRATEGIES FOR TREATING STROKE

Part of "93 - Current and Experimental Treatment of Stroke "

Neuronal damage and death do not occur immediately after ischemia, which suggests temporal thresholds for reperfusion as with tPA and possibly therapeutic thresholds for neuroprotection. Thrombolytic therapy attempts to reestablish blood flow in ischemic regions with the aim of preventing or minimizing cell damage (6 ,7). Various agents have been used in clinical trials to restore cerebral blood flow

(Table 93.1). All of the agents designed to restore blood flow are inappropriate for treating hemorrhagic stroke because of the danger of exacerbation of cerebral bleeding (8). In fact, most of the drugs that restore blood flow are associated with an increased risk of hemorrhage particularly when administered late (>6 hours). The drugs used to restore blood flow can generally be grouped into antithrombotics, antiplatelet agents, fibrinogen depleting agents, or thrombolytics (Table 93.1).

Drugs to improve blood flow
- Antithrombotic
  - Heparin
  - Nadroparin (low molecular weight heparin)
  - Tinzaparin (low molecular weight heparin)
  - Danaparoid (low molecular weight heparinoid, Org 10172)
- Anti-platelet
  - Aspirin
  - Abciximab
- Fibrinogen depleting
  - Ancrod
- Improve capillary flow
  - Pentoxifylline
- Thrombolytics
  - Pro-urokinase
  - Tissue plasminogen activator
  - Streptokinase
  - Urokinase

Drugs to protect brain tissue (neuroprotective agents)
- Calcium channel blockers
  - Nimodipine
  - Flunarizine
- Free radical scavengers—antioxidants
  - Ebselen
  - Tirilazad
  - NPY-059
- GABA agonists
  - Clomethiazole
- Glutamate antagonists
  - AMPA antagonists
    - GYKI 52466
    - NBQX
    - YM90K
    - YM872
    - ZK-200775 (MPQX)
  - Kainate antagonist
    - SYM 2081
  - NMDA antagonists
    - Competitive NMDA antagonists
      - CGS 19755 (Selfotel)
    - NMDA channel blockers
      - Aptiganel (Cerestat)
      - Dextrorphan
      - Dextromethorphan
      - Magnesium
        - Memantine
        - MK-801
        - NPS 1506
        - Remacemide
        - AR-R15896AR
        - HU-211
      - Glycine site antagonists
        - ACEA 1021
        - GV150526
      - Polyamine site antagonists
        - Eliprodil
        - Ifenprodil
- Growth factors
  - Fibroblast Growth factor (bFGF)
- Leukocyte adhesion inhibitor
  - Anti-ICAM antibody (Enlimomab)
  - Hu23F2G
- Nitric oxide inhibitor
  - Lubeluzole
- Opioid antagonists
  - Naloxone
  - Nalmefene
- Phosphatidylcholine precursor
  - Citicoline (CDP-choline)
- Serotonin agonists
  - Bay × 3072
- Sodium channel blockers
  - Fosphenytoin
  - Lubeluzole
  - 619C89
- Potassium channel opener
  - BMS-204352

[a]From: *http://www.neuro.wustl.edu/stroke/stroke-drug-categories.htm#ACUTE STROKE THERAPY.*

TABLE 93.1. CLINICAL TRIALS FOR ACUTE STROKE TREATMENT[a]

Antithrombotics such as heparin and warfarin have been tested in numerous trials. There is no evidence to date to support the use of warfarin in the treatment of acute stroke (9). Heparin, followed by coumadin, is an extremely important modality to prevent secondary strokes, particularly cardioembolic stroke, in certain circumstances arterial occlusion following dissection, and nearly always following venous infarction. In a large-scale randomized trial comparing heparin with the antiplatelet aspirin, heparin was associated with only three fewer deaths per 1,000 at 14 days and no difference at 6 months (10). Although there were fewer recurrent strokes within 14 days, more of them were hemorrhagic. Moreover, 12,500 IU heparin was associated with significantly more transfused or fatal extracranial bleeds, hemorrhagic strokes, and deaths than 5,000 IU (10). Low molecular weight heparinoids have been developed that have greater bioavailability and less effect on platelet function than heparin, thus reducing complications such as hemorrhage and thrombocytopenia (11). Results from the trials of two of the low molecular weight heparinoids have been reported. ORG 10172 (Danaparoid) was tested in the Trial of ORG 10172 in Acute Stroke (TOAST) (12). Despite an apparent positive response at 7 days, there was no significant improvement in favorable outcome at 3 months (12). In contrast, Nadroparin (Fraxiparine) was effective in improving outcomes at 6 months when treatment was initiated within 48 hours of symptom onset for a period of 10 days (13). Unfortunately, this study has not been replicated. The results of a 1,500-patient phase III trial (Tinzaparin in Acute Ischemic Stroke Trial [TAIST]) for a third low molecular weight heparinoid, tinzaparin (Innohep), are completed and are currently being analyzed (14).

Antiplatelet agents such as aspirin have also been tested in numerous trials (Table 93.2). Two large-scale phase III trials (Chinese Acute Stroke Trial [CAST] and International Stroke Trial [IST]), in which there were 20,000 patients in each, found that there was a small but worthwhile improvement in the main outcome measures (absolute risk reduction of <1% [need to treat 111 to benefit one]) at 6 months and that aspirin should be started as soon as possible after the onset of symptoms (10 ,15). In the IST trial, there were fewer deaths and recurrent ischemic strokes with no significant excess of hemorrhagic strokes among the aspirin-allocated patients (10). Recently, the Abciximab in Ischemic Stroke Investigators (16) reported that Abciximab (ReoPro), a potent parenterally administered platelet glycoprotein IIb/IIIa

antagonist, when administered up to 24 hours after stroke onset, resulted in an improved functional outcome. At 3 months, there was a trend toward a higher rate of minimal residual disability (Barthel Index ≥95 or modified Rankin scale ≤1) among Abciximab patients compared with those who received placebo. Antiplatelets may prove a viable therapy for both prevention and intervention.

Multicentre Acute Stroke Trial-Italy (MAST-I)
Stroke Prevention in Atrial Fibrillation II (SPAF II)
International Stroke Trial (IST)
Acute Ischemic Stroke Trial—Oral Aspirin versus Intravenous Heparin on Stroke
Progression (AIST-ASH)
Chinese Acute Stroke Trial (CAST)
Carotid Artery Stenosis with Asymptomatic Narrowing: Operation Versus
Aspirin (CASANOVA)

[a]From: *http://www.neuro.wustl.edu/stroke/therapy/Therapy_1Page5.html*

TABLE 93.2. CLINICAL TRIALS INVOLVING ASPIRIN AS A THERAPEUTIC FOR STROKE[a]

Ancrod (Viprinex/Arvin) is a fibrinogenolytic agent that comes from the venom of the Malaysian pit viper (17). Ancrod has recently completed a North American phase III trial (Stroke Treatment with Ancrod Trial [STAT]). Treatment was a 5-day paradigm initiated within 3 hours of onset (18). There was an increase in symptomatic intracranial hemorrhage (5.2% versus 2% for placebo), and there was no difference in mortality at 3 months (18). There is a European trial (ESTAT) for Ancrod that has yet to be completed but BASF, who was jointly conducting the phase III trials, has decided not to continue the trials after an independent group looked at the interim results and failed to see any evidence of efficacy (19).

Thrombolysis has been performed with a variety of compounds administered by various routes. From numerous trials, one thing is clear of all agents in this category: They must be given early or serious hemorrhagic complications are likely (7). Urokinase (uPA) is an endogenous proteolytic enzyme, secreted as a proenzyme, which converts circulating plasminogen to plasmin, producing a systemic lytic state. Intraarterial administration of pro-urokinase (the single chain precursor of urokinase) was evaluated in the Prolyse in Acute Cerebral Thromboembolism Trial (PROACT) with high- or low-dose heparin (20) and in the PROACT II trial with low-dose heparin (21). Although prourokinase and intravenous high-dose heparin resulted in an increased rate of hemorrhage, intraarterial pro-urokinase with low-dose heparin significantly improved the proportion of good outcomes from 25% to 40% and hemorrhage was seen in 10% of patients, consistent with other thrombolytic trials.

Streptokinase is derived from B-hemolytic streptococci and it converts plasminogen to plasmin following the formation of a complex with plasminogen. Metaanalysis of the three major trials of streptokinase, Multicentre Acute Stroke Trial–Europe (MAST-E), the Australian Streptokinase (ASK) trial and the Multicentre Acute Stroke Trial–Italy (MAST-I), showed an increased risk of hemorrhage and death (22 ,23). The trials were all terminated prematurely.

Intravenous administration of tissue plasminogen activator (tPA) within 3 hours of symptom onset has been approved for the treatment of acute ischemic stroke in the United States for 4 years and it remains the only approved treatment for acute ischemic stroke. There have been four large trials of Alteplase, the European Cooperative Acute Stroke Studies (ECASS I and II), the National Institutes of Neurologic Disorders and Stroke (NINDS) trial, and the Alteplase Thrombolysis for Acute Noninterventional Therapy in Ischaemic Stroke (ATLANTIS). The metaanalysis of these trials shows a significant reduction in death and disability (160/1,000 or a number needed to treat of 6 versus 9/1,000 with a number needed to treat of 111 for aspirin). More trials are needed to gain a better understanding of the timing, safety, and efficacy of Alteplase in different patient populations. However, a recent study has demonstrated the efficacy of tPA for acute stroke in a community setting (24). It remains unclear whether intravenous (IV) or intraarterial (IA) tPA delivery is superior in acute ischemic stroke treatment (7 ,25 ,26 ,27 ,28 ,29 and 30). IV therapy is faster and more convenient, whereas IA therapy allows for mechanical disruption and higher drug concentration at the clot site (7). A protocol was recently tested in which severe stroke patients with little or no CT changes were given IA tPA following IV tPA if there was no clinical improvement and a persistent occlusion was identified on transcranial Doppler (30). Combined therapy was performed on nine patients and no intracerebral hemorrhages or significant systemic bleeding complications occurred. Marked clinical improvement was noted in four patients, suggesting that combined IA and IV tPA therapy is feasible and safe and that future studies should consider combining rather than comparing these two delivery strategies (30). There is an NIH-approved trial (Interventional Management Study, IMS trial) comparing IV versus combined IV and IA tPA therapy that will be underway by the time this is published.

NEUROPROTECTIVE AGENTS

Part of "93 - Current and Experimental Treatment of Stroke "

Once the excitotoxic cascade has begun, the first therapeutic strategy, and the one that has raised the most hope in the last decade, utilizes neuroprotective agents targeting mediators in the excitotoxic cascade (Fig. 93.1). The cascade can be separated spatially and temporally. The cascade begins with an increase in presynaptic calcium influx, followed by glutamate release, postsynaptic glutamate receptor activation resulting in sodium and calcium influx, which results in postsynaptic depolarization and further calcium influx.

Voltage-gated sodium and potassium channels are targets that affect depolarization, whereas calcium channels mediate calcium influx and affect depolarization. Most of the neuroprotective agents tested in the clinic have targeted either voltage-gated calcium channels or glutamate receptors, particularly the NMDA receptor subtype. In addition, GABA receptor agonists attenuate excitotoxicity (31) and free radical scavengers are neuroprotectants aimed at the later stages of the excitotoxic cascade (32). After the excitotoxic cascade has progressed, an inflammatory response occurs in which there is infiltration of leukocytes and monocytes (33). Microglia and astrocytic glial cells are activated and macrophages begin responding to chemoattractants. Still other therapeutic strategies have targeted leukocyte adhesion (34 ,35) and nitric oxide production (36 ,37). Once much of the damage has occurred and neuroprotection is no longer a viable strategy, neural regeneration and trophism becomes an option. This approach has been mounted with infusion of growth factors but trials to date have been unsuccessful (38). We now look at trials of compounds targeted to these various processes in the excitotoxic cascade.

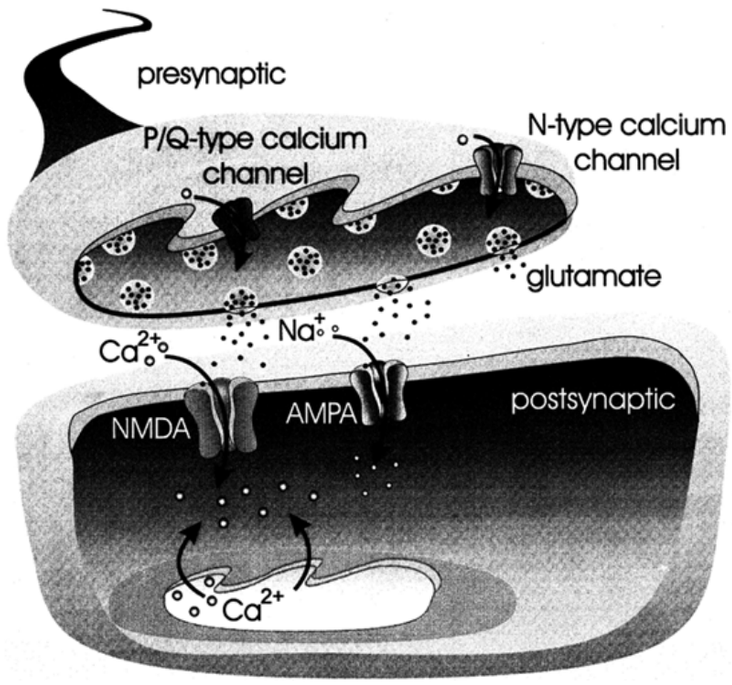


FIGURE 93.1. Thereapy utilizing neuroprotective agents targeting mediators in the excitotoxic cascade. See color version of figure.

CALCIUM CHANNEL ANTAGONISTS

Part of "93 - Current and Experimental Treatment of Stroke "

Clinically, L-type calcium channel antagonists have been used extensively for the treatment of cardiovascular disorders such as hypertension; however, very few have been tested in the arena of stroke therapy. Although the metaanalysis of oral nimodipine (Nimotop) trials demonstrated no significant effects on neurologic or functional outcome, the drug was administered within a very large therapeutic window (i.e., 24 to 48 hours) following symptom onset (39). More recently, a more aggressive trial (Very Early Nimodipine Use in Stroke, VENUS) designed to treat with nimodipine within 6 hours of symptom onset was terminated before completion because the results indicated that there was no beneficial role of early administration of oral nimodipine at 3 months (40). Although nimodipine failed to demonstrate efficacy in treating acute stroke, it was approved for the treatment of subarachnoid hemorrhage more than a decade ago. The results of a phase III trial of flunarizine (Sibelium) for acute stroke suggested that it also did not improve neurologic or functional outcome at 3 months when administered early (<6 hours) (41).

The N-type calcium channel antagonist, SNX-111 (Ziconotide), preferentially blocks presynaptic calcium channels and inhibits neurotransmitter release. In both focal and global animal models of cerebral ischemia SNX-111 is highly neuroprotective, even when administered after a delay of 24 hours following reperfusion (42). SNX-111 has also been tested in clinical trials for acute stroke. Although the stroke trials of SNX-111 were discontinued because of severe hypotension that exacerbated the ischemic damage (43), SNX-111 has progressed to phase III trials for head trauma (44 ,45) and has just been approved by the FDA for the treatment of pain (46). Spider toxin antagonists of the P/Q-type neuronal calcium channels are neuroprotective *in vitro* but their *in vivo* toxicity in animals, primarily respiratory depression causing death, has limited their clinical development. However, efforts to generate small peptide analogues of these spider toxins, which exhibit efficacy *in vitro*, are ongoing.

GLUTAMATE RECEPTOR ANTAGONISTS

Part of "93 - Current and Experimental Treatment of Stroke "

Numerous clinical trials have been carried out for NMDA receptor antagonists based on preclinical testing in animal models of cerebral ischemia (47). All the phase III trials to date have failed. Optimism for the use of NMDA receptor antagonists in the treatment of acute ischemia has waned and has even prompted some pharmaceutical companies to abandon efforts to develop therapeutics for acute stroke. The experience with NMDA receptor antagonists in the clinic has been that most NMDA receptor antagonists result in psychosis as a common adverse effect (48). NMDA receptor antagonists with greater specificity for various binding sites on the receptor, or selectivity for a given receptor subunit, are being developed which demonstrate greater safety and fewer adverse effects (49 ,50 ,51 and 52).

NMDA receptors are heteromeric pentamers composed of at least one NR1 subunit and one or more of the four different NR2 subunits, NR2A, NR2B, NR2C, or NR2D. There are a number of sites on the NMDA receptor at

which antagonists can bind (Fig. 93.2). Competitive antagonists bind to the same site as NMDA or glutamate. Glycine and polyamines each bind as activators of the NMDA receptor and there are antagonists of these two sites, respectively. There are also noncompetitive antagonists that bind to the inside of the pore of the channel and sterically inhibit the influx of ions (53 ,54 and 55). These compounds are called use-dependent. Competitive NMDA receptor antagonists block channel activity best when glutamate levels are low. These antagonists would be expected to be less efficacious during ischemia compared to a normally functioning brain because glutamate levels rise during ischemia (56); therefore, the therapeutic index of these agents would be expected to be low.

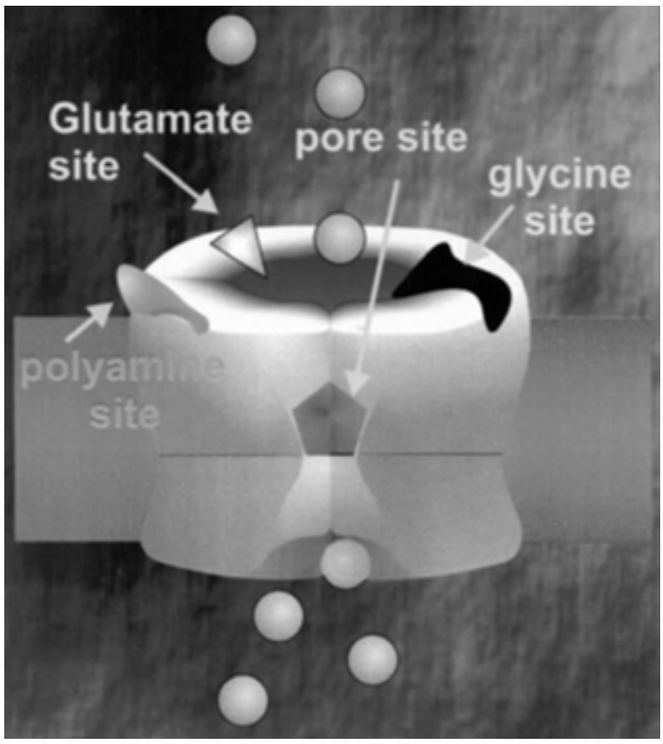


FIGURE 93.2. Sites on the NMDA receptor at which antagonists can bind. See color version of figure.

Not surprisingly, two phase III trials (Acute Stroke Trials Involving Selfotel Treatment [ASSIST] and a head injury trial) of the competitive NMDA receptor antagonist CGS 19755 (Selfotel) were suspended given that they were not effective and caused neurotoxic side effects (57 ,58 and 59). The principal side effects were agitation, hallucinations, and psychotomimetic effects similar to those seen with phencyclidine (PCP) (60 ,61). Furthermore, among the patients with severe stroke, there was a trend toward a significant difference in the number of deaths between those treated with Selfotel (33%) and those in the placebo group (22%) (58).

Two years later, phase III trials for the noncompetitive NMDA receptor antagonist Aptiganel (Cerestat/CNS-1102) were discontinued (62). In a subsequent analysis of the phase III trial, some potential therapeutic benefit was identified in a subset of the stroke population. Cambridge Neuroscience plans to further investigate these beneficial effects in partnership with Boehringer Ingelheim (63). The most common side effects with this high-affinity channel blocker were motor retardation, elevated blood pressure, perceptual disturbances, and hallucinations, similar to those seen with Selfotel (60). The rest of the noncompetitive NMDA receptor antagonists exhibit much lower affinity for the ion channel pore than Cerestat. Although memantine has been shown to be neuroprotective in both *in vitro* and *in vivo* models (64) and memantine is progressing for the treatment of dementia (65), it is not currently being developed for the treatment of acute stroke to our knowledge. Remacemide and its active desglycinyl metabolite are well-tolerated at relatively high doses in humans, have demonstrated significant neuroprotective efficacy in animal models of cerebral ischemia, and were doing well in phase II trials for acute stroke (66); however, there is some question as to whether optimal neuroprotective doses would be achieved within the early hours of treatment in humans (51). Thus, the low-affinity compound ARR-15896 was developed as a backup to remacemide and is currently in early Phase II trials (66). Hu-211 (dexanabinol, Cypros) is a nonpsychogenic cannabinoid with both low affinity use dependent block of NMDA receptors, as well as inhibition of tumor necrosis factor-α and antioxidant properties. It exhibits widespread neuroprotective actions in several animal models of stroke and head trauma (67) and has just recently completed a small phase II trial for head trauma clearing the way for a large multinational clinical trial of several hundred severe head trauma patients to commence later this year (68). Magnesium, also a low-affinity noncompetitive NMDA receptor antagonist, is currently in phase III clinical trials for the treatment of acute stroke (Intravenous Magnesium Efficacy in Stroke [IMAGES]). Patients are randomized to receive placebo or intravenous magnesium (16 mmol over 15 minutes followed by 65 mmol over 24 hours) within 12 hours of symptom onset (69). Recruitment of the 2,700 patients is ongoing and should be completed by 2002. Magnesium infusions are well tolerated in humans and have been demonstrated to be neuroprotective in animal models of cerebral ischemia (49 ,69 ,70). In a smaller trial, magnesium-treated patients improved neurologically and the need for institutional care 6 months after the stroke was reduced (49 ,71).

NPS-1506 is another open channel blocker being developed for the treatment of acute stroke. NPS-1506 acts at a novel site within the pore and is an orally active small molecule developed by drawing on the knowledge of spider toxins with pore-blocking abilities (72). NPS-1506 is currently in phase Ib and to date is devoid of the side effects normally attributed to noncompetitive antagonists. In animal models of cerebral ischemia, significant protection is achieved even when NPS-1506 is administered up to 2 hours after onset of stroke. At neuroprotective doses there were no PCP-like behavioral effects, no learning or memory impairment, no neuronal vacuolization, and no significant sedation or cardiovascular side effects (72 ,73).

Polyamine site NMDA receptor antagonists like eliprodil (SL 82-0715) and ifenprodil have also not done well in

clinical trials (54). Their failure has been attributed to the cardiovascular side effects they possess by virtue of their affinity for sodium and calcium channels in addition to NMDA receptors (74 ,75). Eliprodil had electrocardiographic effects that limited dosing such that efficacy was not obtained and phase III trials were abandoned in 1997 (74). Another feature of the polyamine site antagonists is that they are relatively specific for NR2B receptor subunits (76). The affinity of eliprodil for NR2B subunits is more than 100-fold greater than that for NR2A or NR2C receptor subunits. A number of other NR2B subunit selective antagonists are currently under development for the treatment of stroke (Table 93.3). Ro 25-6981 and CP-101,606 are both high-affinity NR2B-selective antagonists with very slow kinetics (76 ,77 and 78). CP-101,606 is neuroprotective in animal models of cerebral ischemia but the slow kinetics may limit the rate at which neuroprotective doses may be achieved in human stroke. The fact that there were no psychotomimetic effects of CP-101,606 in phase II trials for moderate head injury and hemorrhagic stroke and that the drug was well tolerated suggest that CP-101,606 might be well tolerated in occlusive stroke (79). Compounds such as Co 101244/PD 174494 are being developed with high-affinity antagonists such as CP-101,606 but more rapid kinetics like isoxsuprine in anticipation of the slower kinetics of CP-101, 606 presenting problems in efficacy in treating stroke (80). Something else which may increase the chances of success of Co-101244/PD174494 is that it also had much less affinity for α1-adrenergic receptors and displayed a reduction in inhibition of potassium channels, something that most other high-affinity compounds possess (80).

Subunit Compound	NR2A	NR2B	NR2C	Cortical Neurons (Mostly NR2B)	Rank Order Kinetics
Isoxsuprine (Sigma I-0880)	—	0.7	—	—	Fast
Nylidrin[b]	27	0.19	33	0.2	Medium
Ifenprodil (Synthelabo)[b]	84	0.31	130	0.12	Medium
	40[a]	0.25[a]			
Eliprodil (Synthelabo)	>100	1	>100	0.6	Fast
	50[a]	0.7[a]			
Co-101244/PD174494		0.02	—	—	Medium
Haloperidol	>300	3.1	>300	1.5	?
Ro 25-6981 (Roche)[b]	60[a]	0.009[a]	—	—	Slow
CP 101,606 (Pfizer)[b]	100[a]	0.1	—	—	Slow
		0.06[a]			
Glutamate	4[a]	1.3[a]	—	—	—

[a]All numbers are IC50s in μM and are from Whittemore et al., 1996 unless indicated by[a], in which case they are from Trube et al., 1996. The more potent, the slower the kinetics but also the more protective. From Whittemore ER, Ilyin VT, Woodward RM, Subtype-selective antagonism of NIMDA receptors by Nylidrin. *Soc Neurosci Abat* 1996;22:506–507 and Trube G, Elrhard P, Huber G. The selectivity of RO 25-6981 for NMDA receptor subtypes expressed in Xenopus oocytes. *Soc Neurosci Abat* 1996;22:691–693.
[b]Glutamate dependence: lower affinity with lower glutamate, than with higher glutamate concentrations. All IC50s are done with high glutamate.

TABLE 93.3. NR2B SUB UNIT-SPECIFIC NMDA RECEPTOR ANTAGONISTS

Glycine site antagonists exhibit better safety profiles than NMDA receptor antagonists that bind to other sites (48). Most of the clinical attention on glycine site antagonists has been focused on two therapeutic candidates, ACEA 1021 (Licostinel) and GV150526. ACEA 1021 (5-nitro-6,7-dichloro-2,3-quinoxalinedione) is a member of a family of compounds called kappagems. ACEA-1021 has demonstrated neuroprotective efficacy in animal models of focal and global ischemia (81) and it exhibits minimal adverse CNS or cardiovascular side effects (50 ,82). The compound, originated by ACEA, was being developed by CoCensys for Novartis, but Novartis stopped participating in its development after crystals of ACEA-1021 were found in the urine of some participants in a phase I study (83). CoCensys retains the rights to the drug and is continuing to evaluate it in phase II trials. One approach to dealing with the problems of ACEA-1021 excretion has been to treat in combination with probenecid, which nonselectively inhibits secretion of anionic compounds (84). The combination of ACEA-1021 and probenecid resulted in significantly larger infarct reductions in animal models of cerebral ischemia suggesting that the limiting factor in ACEA-1021 efficacy is related to the steady-state levels that can be elevated by combination therapy with probenecid (84).

GV150526 is significantly neuroprotective in animal models of cerebral ischemia and like ACEA-1021, has shown good tolerability with minimal CNS side effects in the Glycine Antagonist in Neuroprotection (GAIN) phase I and II trials (The GAIN Investigators) (52). Minor abnormalities in liver function were noted with higher maintenance doses, but these changes were asymptomatic and resolved

within 10 days (52). The results of the dose-escalation phase II clinical trial for GV150526 (85) were reported last year and were followed with the GAIN Americas and the GAIN International phase III trials. Approximately 1,600 patients were recruited into randomized double-blinded placebo controlled trials. Two substudies in each trial were planned, to measure lesion volume by MRI-DWI (Magnetic Resonance Imaging-Diffusion Weighted Imaging) and to measure health-related quality of life outcomes. The results of the GAIN International were presented at the 25th International American Heart Association meeting and were less than anticipated and completely neutral. Unlike most trials with NMDA receptor antagonists, the problem with GV150526 was not safety but rather efficacy (85). The results from the GAIN American study are not yet available.

In contrast to NMDA receptor antagonists, there is little experience using AMPA and KA receptor antagonists to treat acute stroke. Fewer binding sites have been characterized on the AMPA receptor compared to the NMDA receptor. Competitive quinoxalinedione antagonists bind to the AMPA binding site and there are also noncompetitive GYKI [GYKI 52466,1-(4-aminophenyl)-4-methyl-7,8-methylenedioxy-5H-2,3-benzodiazepine HCl] site-binding antagonists, benzodiazepine binding-site antagonists, and pore-blocking antagonists (86). The pore-blocking antagonists are mostly from spider toxins and are not appropriate for clinical development. There is one noncompetitive AMPA/KA receptor antagonist, EGIS-9637, which demonstrates neuroprotective efficacy in both focal and global animal ischemia models, but the clinical development has not begun and little is known of the tolerability (87). One of the more common CNS effects of AMPA receptor antagonists is sedation. All of the AMPA receptor antagonists in clinical trials for the treatment of acute stroke are competitive antagonists based on the quinoxalinediones such as NBQX. NBQX and other quinoxalinediones are neuroprotective in animal models of ischemia, even when administered up to 12 hours following reperfusion (88). The main problem with quinoxalinediones has been nephrotoxicity owing to the poor solubility of these compounds (89). One AMPA receptor antagonist, ZK-200775, progressed to phase IIa but these trials were discontinued in 1998 owing to excessive sedation (90). Yamanouchi Pharmaceuticals have developed a series of competitive AMPA receptor antagonists that are neuroprotective in animal models and exhibit improved solubility (91 ,92 and 93). In phase I trials YM90K was well tolerated and in spite of a high rate of urinary excretion there were only mild changes in kidney function markers with a single intravenous dose of 36 mg and repeated doses of 24 mg over 3 hours (94). These phase I trials were followed by phase II trials using the AMPA receptor antagonist, YM872, which has a solubility 800-fold greater than NBQX or YM90K at pH 7 (95). Phase IIa trials are ongoing. The only adverse effect reported in the YM872 phase I study was euphoria in some patients at the higher dose levels (96). The challenge with YM872 will be overcoming the short half-life of the drug to maintain the elevated plasma levels necessary to effect neuroprotection.

GABA RECEPTOR AGONISTS

Part of "93 - Current and Experimental Treatment of Stroke "

Clomethiazole (Zendra), a GABA receptor agonist, has just completed large-scale phase III clinical trials for the treatment of acute stroke, Clomethiazole Acute Stroke Study (CLASS and CLASS-I, H and T) (97). Clomethiazole was well tolerated and appeared safe. Sedation was the most common adverse event that led to treatment withdrawal in almost 16% of treated, compared to only 4% in the placebo group. In the main efficacy analysis 56.1% of patients taking clomethiazole and 54.8% of placebo patients reached relative functional independence (97). The difference was not significantly different but in a subgroup analysis of 545 patients who had total anterior circulation syndrome (TACS), the percentage of those reaching relative functional independence was 40.8% on clomethiazole and 29.8% in the placebo group (98). This subgroup efficacy has prompted the testing of clomethiazole in large ischemic cerebral infarctions (CLASS-I), ischemic infarctions in those who receive tPA (CLASS-T), and intracerebral hemorrhage (CLASS-H). Previous failed clinical trials for stroke may have shown efficacy had there been subgroup targeting like that being carried out in the CLASS-I study. However, it is difficult to categorize subgroups of strokes while operating under the extreme time constraints stroke treatment requires. A recent protocol for grading strokes has been developed (ASPECTS, which addresses the need for rapid diagnosis and assessment of infarct (99).

OTHER THERAPEUTIC TARGETS

Part of "93 - Current and Experimental Treatment of Stroke "

Voltage-Gated Channels

Potassium channel openers may rescue injured penumbral neurons, thus reducing the size of the infarct and improving neurologic outcome. The novel potassium channel opener, BMS-204352, is being assessed in the Potassium-channel Opening Stroke Trials (POST) in which medication is administered within 6 hours of onset of symptoms. The primary endpoint is change in lesion volume at 12 weeks determined by diffusion/perfusion MRI.

Opioid Receptor Antagonists

The competitive κ-selective opioid receptor antagonist Cervene (Nalmefene) is neuroprotective in animal models of

ischemia and clinically safe (100). In a phase II trial in patients with acute ischemic stroke treated with 6, 20, or 60 mg Cervene there was no significant functional improvement compared to placebo after 3 months (100). A secondary analysis showed an improved 3-month outcome in patients less than 70 years of age.

Free Radicals, Nitric Oxide, and Inflammation

During reperfusion, oxygen-free radicals contribute to damage neurons when antioxidant defense mechanisms are not optimal. Superoxide, hydrogen peroxide, and hydroxyl radicals are responsible for damaging membranes and degrading DNA. In experimental models, agents such as vitamin E, glutathione, superoxide dismutase, iron chelators, lazaroids, NPY-059, and catalase have been tested as free radical scavengers. One of the most tested agents, Triliazad mesylate (Freedox), an inhibitor of lipid peroxidation, is neuroprotective in stroke models (37). In a phase II study of over 400 patients treated with 6 mg/kg per day tirilazad for 3 days within 6 hours of onset, no significant outcome relative to placebo control was observed (101). The Randomized Trial of High Dose Tirilazad in Acute Stroke (RANTTAS II) phase III trials of a higher dose was stopped because of a lack of efficacy (102). Similarly, the antioxidant Ebselen, which has glutathione peroxidase-like activity, showed no efficacy (103). NPY-059 is currently in early phase II trials by Centaur.

Lubeluzole (Prosynap) is a benzothiazole that blocks glutamate-induced nitric oxide-mediated neurotoxicity in rats. In a phase II trial patients presenting within 6 hours were randomized to a placebo group or to receive iv Lubeluzole at 7.5 or 15 mg over 1 hour as a loading dose, followed by a continuous infusion of 10 or 20 mg per day for 5 days, respectively. This trial was terminated early owing to high mortality in the high-dose group (36). Clinical development was subsequently stopped after a large phase III trial (Lub-Int-13) failed to show efficacy.

Citicoline provides cytidine and choline as precursors in the synthesis of the integral membrane component phosphatidylcholine. During ischemia phosphatidylcholine is broken down into free fatty acids, which in turn generate free radicals. Citicoline has been studied as a neuroprotectant as well as a promoter of neural plasticity and repair after stroke through its actions to promote membrane synthesis, to decrease the levels of free fatty acids and to promote the regrowth of axons and nerve terminals. In a phase II study, oral citicoline (CerAxon; 500, 1,000, and 2,000 mg per day) was administered within 24 hours for a 6-week period (104). A dramatic improvement in functional independence outcome (61% versus 39% in the placebo group) was seen in the 500 mg and 2,000 mg (but not the 1,000 mg) dose groups. A second study of 500 mg per day citicoline versus placebo failed to demonstrate significant differences in outcome or mortality. The results from the 899 patient phase III trial comparing 2,000 mg citicoline to placebo (105) failed to meet the primary endpoint, an improvement in neurologic function. However, citicoline was shown to be safe and there was a higher percentage of citicoline-treated patients with reduced infarct volumes compared to the placebo group.

Treatment with a monoclonal anti-ICAM-1 antibody inhibits leukocyte adhesion by blocking leukocyte attachment and migration through cerebral endothelial cells and reduces neurologic deficits in rodent ischemia models (34). The Enlimomab Acute Stroke Trial (EAST) completed in 1996 showed that the anti-ICAM-1 antibody treated patients had a worse outcome than placebo-treated patients (35). LeukArrest (Hu23F2G) is another human antibody that inhibits neutrophil movement from the blood into the brain by specifically targeting CD11/CD18 molecules on the surface of neutrophils. A recently completed phase II trial showed LeukArrest to be safe and suggested an improved neurologic outcome. The Hu23F2G Phase III Stroke Trial (HALT) with 310 patients was begun in early 1999.

CONCLUSION AND FUTURE DIRECTIONS: THE ARSENAL OF A UTOPIAN STROKE CARE CENTER

Part of "93 - Current and Experimental Treatment of Stroke "

Given the worldwide efforts focusing on cloning the human genome and the parallel efforts to identify disease specific genes and proteins through genomics and proteomics, we have no way of knowing how many stroke specific genes and proteins will be identified. Estimates of the total number of disease genes are between 10,000 and 30,000. Because there are currently only 450 clinical targets of therapeutic intervention worldwide, there are at least 9,000 new genes for which we have no idea of function and no pharmacologic tools with which to study them. One thing is certain; we will be faced with an explosion of information in 2001, the expected completion date of the human genome project.

In terms of the strategies described in the preceding, efforts to increase blood flow may include angiogenesis using a recombinant adenovirus expressing vascular endothelial growth factor (VEGF) or perhaps other gene therapy delivered vascularly. Still more effort is being focused on NMDA receptors, but in developing safer compounds such as ARL 15896 and GV150526 with a better therapeutic ratio than agents that previously failed in trials. Trials are well underway for some AMPA receptor antagonists like YM872 and the results of these will reveal whether this strategy is a viable one worthy of further attention.

Growth factor therapies are targeted to improve neuronal survival, repair, and plasticity following acute ischemic stroke. A phase III trial of fibroblast growth factor (Biblast, bFGF, Trafermin) was discontinued because of no efficacy

relative to the placebo group. The glial derived growth factor, GFG2, is being developed by Cambridge NeuroScience (CNSI) for treating neurodegenerative disorders and may perhaps treat stroke (106). Another avenue of therapy is to promote the growth of new vessels through angiogenesis, but this too has yet to be realized in trials for stroke. Neurogenesis is an area that is generating attention in animal models but the growth of new neurons occurs very late and experimentally only in the dentate; therefore, the relevance to the clinical situation is somewhat obscure (107). A viable alternative to promoting the growth of the brain's own cells is to treat with stem cells that have the potential of generating new neurons. Recently this was successfully attempted *in vivo* (108).

Although the recent Heart Outcomes Prevention Evaluation (HOPE) trial, which looked at ramipril versus vitamin E, failed to demonstrate effects of vitamin E on any of the cardiovascular outcome measures, including myocardial infarction, unstable angina, congestive heart failure, stroke, and death from cardiovascular causes (109), two new trials, Vitamin Intervention for Stroke Prevention or VISP (110) and Vitamins to Prevent Stroke or VITATOPS (111) have recently begun to determine whether daily B_{12}, B_6, and folate, which reduce homocysteine levels, are better than the best medical practice and surgical management alone to reduce the recurrence of vascular events in stroke patients.

The Women's Estrogen for Stroke Trial (WEST) will assess whether postmenopausal estrogen therapy (1 mg estradiol per day) in women with recent transient ischemic attacks or nondisabling stroke can reduce the risk of death or recurrent stroke (112).

Inasmuch as we have laid out spatially distinct strategies, there is an overlapping temporal association of these strategies. This temporal aspect can be exploited in the form of co-therapies. Because ischemic death occurs via a cascade involving several processes, it is likely that targeting one single process will prove inadequate and instead, targeting several of the processes with a cocktail of therapeutic agents given in a specific temporal sequence will prove efficacious. The development of safe therapeutics for the intervention of stroke should provide a brighter future for stroke survivors with the increase in stroke awareness and the number of centers with dedicated stroke care units.

ACKNOWLEDGMENTS

Part of "93 - Current and Experimental Treatment of Stroke "

The authors wish to thank Eugene Palmer for his helpful suggestions and critical reading of the manuscript.

REFERENCES

1. Jorgensen HS, Nakayama H, Kammersgaard LP, et al. Predicted impact of intravenous thrombolysis on prognosis of general population of stroke patients: simulation model. *Br Med J* 1999;319:288–289.

2. Pulsinelli WA, Brierley JB, Plum F. Temporal profile of neuronal damage in a model of transient forebrain ischemia. *Ann Neurol* 1982;5:491–498.

3. Small DL, Buchan AM. Animal models. *Br Med Bull* 2000;56:307–317.

4. Tamura A, Kawai K, Takagi K. Animal models used in cerebral ischemia and stroke research. In: Ter Horst GJ, Korf J, eds. *Clinical pharmacology of cerebral ischemia*. Totowa, NJ: Humana Press, 1997:265–294.

5. Small DL, Morley P, Buchan AM. Biology of ischemic cerebral cell death. *Progr Cardiovasc Dis* 1999a;42:185–207.

6. Lyden PD. Thrombolysis for acute stroke. *Progr Cardiovasc Dis* 1999;42:175–183.

7. Muir KW, Roberts M. Thrombolytic therapy for stroke. A review with particular reference to elderly patients. *Drugs Aging* 2000;16:41–54.

8. Zivin JA. Thrombolytic stroke therapy. Past, present and future. *Neurology* 1999;53:14–19.

9. Rha J-H. Acute ischemic stroke trial: oral aspirin vs intravenous heparin on stroke progression (AIST-ASH). *Stroke* 2000;31:259.

10. International Stroke Trial Collaborative Group. The International Stroke Trial (IST): a randomised trial of aspirin, subcutaneous heparin, both, or neither among 19435 patients with acute ischaemic stroke. *Lancet* 1997;349:1569–1581.

11. Meuleman DG. Orgaran (Org 10172): its pharmacological profile in experimental models. *Haemostasis* 1992;22:58–65.

12. Publications Committee for the Trial of ORG 10172 in Acute Stroke Treatment (TOAST) Investigators. Low molecular weight heparinoid, ORG 10172 (danaparoid), and outcome after acute ischemic stroke: a randomized controlled trial. *JAMA* 1998;279:1265–1272.

13. Kay R, Wong KS, Yu XL, et al. Low-molecular-weight heparin for the treatment of acute ischemic stroke. *N Engl J Med* 1995;333:1588–1593.

14. Bath PM. Tinzaparin in Acute Ischemic Stroke Trial (TAIST) *Stroke* 2000;31:255.

15. CAST (Chinese Acute Stroke Trial) Collaborative Group CAST: randomised placebo-controlled trial of early aspirin use in 20,000 patients with acute ischaemic stroke. *Lancet* 1997;349:1641–1649.

16. Abciximab in Ischemic Stroke Investigators. Abciximab in acute ischemic stroke: a randomized, double-blind, placebo-controlled, dose-escalation study. *Stroke* 2000;31:601–609.

17. Pollak VE, Glas-Greenewalt P, Olinger CP, et al. Ancrod causes rapid thrombolysis in patients with acute stroke. *Am J Med Sci* 1990;299:319–325.

18. Levy DE. Stroke Treatment with Ancrod Trial (STAT). *Stroke* 1999;30:234.

19. Scrip database, BASF halts Phase III stroke trial. *World Pharmaceutical News*, PJB Publications Ltd, April 3, 2000.

20. Del Zoppo GJ, Higashida RT, Furlan AJ, et al. The Prolyse in Acute Cerebral Thromboembolism Trials (PROACT): results of 6 mg dose tier. *Stroke* 1996;27:164.

21. Del Zoppo G, Higashida RT, Furlan AJ, et al. A phase II randomized trial of recombinant pro-urokinase by direct arterial delivery in acute middle cerebral artery stroke. *Stroke* 1998;29:4–11.

22. Hommel M, Boissel JP, Cornu C, et al. Termination of trial of streptokinase in severe acute ischaemic stroke. *Lancet* 1995;345:57.

23. Wardlaw JM, Warlow CP, Counsell CE. Systematic review of evidence of thrombolytic therapy for acute ischaemic stroke. *Lancet* 1997;350:607–614.

24. Buchan AM, Barber PA, Newcommon N, et al. Effectiveness of t-PA in acute ischemic stroke. Outcome relates to appropriateness. *Neurology* 2000;54:679–684.

25. Mohr JP. Thrombolytic therapy for ischemic stroke. From clinical trials to clinical practice. *JAMA* 2000;283:1189–1191.

26. Katzan IL, Furlan AJ, Lloyd LE, et al. Use of tissue-type plasminogen activator for acute ischemic stroke. The Cleveland area experience. *JAMA* 2000;283:1151–1158.

27. Albers GW, Bates VE, Clark WM, et al. Intravenous tissue-type plasminogen activator for treatment of acute stroke. The standard treatment with alteplase to reverse stroke (STARS) study. *JAMA* 2000;283:1145–1150.

28. Naff NJ, Turim S, Williams M, et al. Intraventricular thrombolysis trial. *Stroke* 2000;31:257.

29. Adams HP, Brott TG, Furlan AJ, et al. Guidelines for thrombolytic therapy for acute stroke: a supplement to the guidelines for the management of patients with acute ischemic stroke. *Circulation* 1996;94:1167–1174.

30. Demchuk AM, Felberg RA, Christou I, et al. Combined intravenous and intraarterial thrombolytic therapy for severe stroke. *Stroke* 2000;31:283.

31. Lyden PD. GABA and neuroprotection. *Intl Rev Neurobiol* 1997;40:233–258.

32. Hall ED. Free radicals in stroke In: Miller LP, ed. *Stroke therapy: basic, preclinical and clinical directions*. New York: Wiley-Liss, 1999:245–270.

33. Garcia JH, Pantoni L. Cellular inflammation in experimental ischemic stroke. In: Miller LP, ed. *Stroke therapy: basic, preclinical and clinical directions*. New York: Wiley-Liss, 1999:271–298.

34. Zhang RL, Chopp M, Jiang N, et al. Anti-intercellular adhesion molecule-1 antibody reduces ischemic cell damage after transient but not permanent middle cerebral artery occlusion in the Wistar rat. *Stroke* 1995;26:1438–1443.

35. Enlimomab Acute Stroke Trial: Final Results. *Neurology* 1997;4:A270.

36. Diener HC, Hacke W, Hennerici M, et al. Lubeluzole in acute ischemic stroke. A double-blind placebo-controlled phase II trial. *Stroke* 1996;27:76–81.

37. Xue D, Sivka A, Buchan AM. Tirilizad reduces cortical infarction after temporary but not permanent focal cerebral ischemia in rats. *Stroke* 1992;23:894–899.

38. Cramer SC, Finklestein SP. Reparative approaches: growth factors and other pharmacological treatments. In: Miller LP, ed. *Stroke therapy: basic, preclinical and clinical directions*. New York: Wiley-Liss, 1999:321–336.

39. Mohr JP, Orgogozo JM, Harrison MJG, et al. Metaanalysis of oral nimodipine trials in acute ischemic stroke. *Cerebrovasc Dis* 1994;4:197–203.

40. Kaste M, Fogelholm R, Erila T, et al. A randomized, double-blind, placebo-controlled trial of nimodipine in acute ischemic hemispheric stroke. *Stroke* 1994;25:1348–1353.

41. Horn J, Limburg M, Vermeulen M. Very Early Nimodipine Use in Stroke (VENUS). *Stroke* 1999;30:243.

42. Buchan AM, Gertler SZ, Li H, et al. A selective N-type Ca(2+)-channel blocker prevents CA1 injury 24 h following severe forebrain ischemia and reduces infarction following focal ischemia. *J Cereb Blood Flow Metab* 1994;14:903–910.

43. Spedding M. Sodium and calcium channel modulators in stroke therapy In: Miller LP, ed. *Stroke therapy: basic, preclinical and clinical directions*. New York: Wiley-Liss, 1999:157–193.

44. http://www.recap.com/alliance

45. http://grimwade.biochem.unimelb.edu.au/~bgl/snx-euro.htm

46. Investigational Drugs weekly highlights Elan progresses with Ziconotide. *NDA* 2000;10:6.

47. Small DL, Morley P, Buchan AM. Preclinical studies of glutamate receptor antagonists for the treatment of cerebral ischemia. In: Miller LP, ed. *Stroke therapy: basic, preclinical and clinical directions*. New York: Wiley-Liss, 199b:65–99.

48. Muir KW, Lees KR. Clinical experience with excitatory amino acid antagonist drugs. *Stroke* 1995;26:503–513.

49. Muir KW. New experimental and clinical data on the efficacy of pharmacological magnesium infusions in cerebral infarcts. *Magnes Res* 1998;11:43–56.

50. Kretschmer BD, Kratzer U, Breithecker K, et al. ACEA 1021, a glycine site antagonist with minor psychotomimetic and amnestic effects in rats. *Eur J Pharmacol* 1997;331:109–116.

51. Dyker AG, Lees KR. Remacemide hydrochloride: a double-blind, placebo-controlled, safety and tolerability study in patients with acute ischemic stroke. *Stroke* 1999a;30:1796–1801.

52. Dyker AG, Lees KR. Safety and tolerability of GV150526 (a glycine site antagonist at the N-methyl-D-aspartate receptor) in patients with acute stroke. *Stroke* 1999b;30:986–992.

53. Harris EW. Subtypes of glutamate receptors: pharmacological classification. In: Stone TW, ed. *CNS neurotransmitters and neuromodulators: glutamate*. New York: CRC Press, 1995:95–125.

54. Small DL, Buchan AM. NMDA antagonists: their role in neuroprotection. *Int Rev Neurobiol* 1997;40:137–172.

55. Ingledine R, Borges K, Bowie D, et al. The glutamate receptor ion channels. *Pharmacol Rev* 1999;51:1–61.

56. Rogawski MA. Therapeutic potential of excitatory amino acid antagonists: channel blockers and 2,3-benzodiazepines. *Trends Pharmacol Sci* 1993;14:325–331.

57. Davis SM, Albers GW, Diener HC, et al. Termination of acute studies involving selfotel treatment. ASSIST Steering Committee. *Lancet* 1997;349:32.

58. Davis SM, Lees KR, Albers GW, et al. Selfotel in acute ischemic stroke. Possible neurotoxic effects of an NMDA antagonist. *Stroke* 2000;31:347–354.

59. Morris GF, Bullock R, Marshall SB, et al. Failure of the competitive N-methyl-D-aspartate antagonist Selfotel (CGS 19755) in the treatment of severe head injury: results of two phase III clinical trials. The Selfotel Investigators. *J Neurosurg* 1999;91:737–743.

60. Lees KR. Cerestat and other NMDA antagonists in ischemic stroke. *Neurology* 1997;49:S66–69.

61. Ryan M, Rhoney DH, Luer MS, et al. New and investigational treatment options for ischemic stroke. *Pharmacotherapy* 1997;17:959–969.

62. http://www.fkpi.com/Release_MHComp/Releases/CamNeuro/12-16-97.html

63. http://www.fkpi.com/Release_MHComp/Releases/CamNeuro/3-9-98.html

64. Seif el Nasr M, Peruche B, Rossberg C, et al. Neuroprotective effect of memantine demonstrated in vivo and in vitro. *Eur J Pharmacol* 1990;185:19–24.

65. *Investigational Drugs Weekly* highlights Phase IIB memantine data released by NTI 4:9, 2000.

66. Palmer GC, Cregan EF, Bialobok P, et al.The low-affinity, use-dependent NMDA receptor antagonist AR-R 15896AR. An update of progress in stroke. *Ann NY Acad Sci* 1999;890:406–420.

67. Biegon A, Joseph AB. Development of HU-211 as a neuroprotectant for ischemic stroke damage. *Neurol Res* 1995;17:275–280.

68. http://www.pharmoscorp.com/pr030700.htm Pharmos announces successful completion of phase II head trauma clinical studies.

69. Bradford AP. IMAGES—Intravenous Magnesium Efficacy in Stroke Trial IMAGES Study Group. *Stroke* 2000;31:248.

70. Muir KW, Lees KR. Dose optimization of intravenous magnesium sulfate after acute stroke. *Stroke* 1998;29:918–923.

71. Strand T, Wester PO. A double blind randomized pilot trial of magnesium therapy in acute cerebral infarction. *7th Nord Meet Cerebrovasc Dis* 1993;37.

72. http://www.npsp.com/LeadProdTech.html#*Stroke* NPS 1506: a novel neuroprotectant.

73. Mueller AL, Artman LD, Balandrin MF, et al. NPS 1506, a novel NMDA receptor antagonist and neuroprotectant. Review of preclinical and clinical studies. *Ann NY Acad Sci* 1999;890:450–457.

74. Synthelabo Reserche. Press Release, Paris, Paris February 6, 1996.

75. Fisher M. Potentially effective therapies for acute ischemic stroke. *Eur Neurol* 1995;35:3–7.

76. Chenard BL, Menniti FS. Antagonists selective for NMDA receptors containing the NR2B subunit. *Curr Pharmacol Des* 1999;5:381–404.

77. Fischer G, Mutel V, Trube G, et al. Ro 25-6981, a highly potent and selective blocker of N-methyl-D-aspartate receptors containing the NR2B subunit. Characterization in vitro. *J Pharmacol Exp Ther* 1997;283:1285–1292.

78. Trube G, Ehrhard P, Malherbe P, et al. The selectivity of Ro 25-6981 for NMDA receptor subtypes expressed in *Xenopus* oocytes. *Soc Neurosci Abst* 1996;22:693–694.

79. Merchant RE, Bullock MR, Carmack CA, et al. A double-blind, placebo-controlled study of the safety, tolerability and pharmacokinetics of CP-101,606 in patients with a mild or moderate traumatic brain injury. *Ann NY Acad Sci* 1999;890:42–50.

80. Zhou ZZL, Cai SX, Whittemore ER, et al. 4-Hydroxy-1-[2-(4-hydroxyphenoxy)ethyl]-4-(4-methylbenzyl)pip. A novel, potent, and selective NR1/2B NMDA receptor antagonist. *J Med Chem* 1999;42:2993–3000.

81. Warner D, Martin D, Ludwig P, et al. In vivo models of cerebral ischemia: effects of parenterally administered NDMA receptor glycine site antagonists. *J Cereb Blood Flow Metab* 1995;15:188–196.

82. Albers GW, Clark WM, Atkinson RP, et al. Dose escalation study of the NMDA glycine-site antagonist licostinel in acute ischemic stroke. *Stroke* 1999;30:508–513.

83. Drugs R D 1:27-8, 1999. Licostinel. ACEA 1021.

84. Xue D, Miyazaki BK, Daniell G, et al. Co-administration of probenecid with licostinel increases the steady state plasma level and reduces the minimum effective dose of neuroprotection in a rat model of transient focal ischemia. *Soc Neurosci Abst* 1999;25:234.6.

85. North American Glycine Antagonist in Neuroprotection (GAIN) Investigators. Phase II studies of the glycine antagonist GV150526 in acute stroke. The North American experience. *Stroke* 2000;31:358–365.

86. Gill R, Lodge D. Pharmacology of AMPA antagonists. *Int Rev Neurobiol* 1997;40:197–232.

87. Levay G, Simo A, Barkoczy J, et al. EGIS-9637, a novel antiischemic drug exerts complex neuroprotective properties. *Soc Neurosci Abst* 1999;25:234.11.

88. Li H, Buchan AM. Treatment with an AMPA antagonist 12 hours following severe normothermic forebrain ischemia prevents CA1 neuronal injury. *J Cereb Blood Flow Metab* 1993;13:933–939.

89. Xue D, Huang ZG, Barnes K, et al. Delayed treatment with AMPA, but not NMDA, antagonists reduces neocortical infarction. *J Cereb Blood Flow Metab* 1994;14:251–261.

90. http://www.neuro.wustl.edu/stroke/therapy/Therapy_1Page44.html

91. Kawasaki-Yatsugi S, Ichiki C, Yatsugi S, et al. Neuroprotective effects of an AMPA receptor antagonist YM872 in a rat transient middle cerebral artery occlusion model. *Neuropharmacology* 2000;39:211–217.

92. Takahashi M, Ni JW, Kawasaki-Yatsugi S, et al. Neuroprotective efficacy of YM872, an alpha-amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptor antagonist, after permanent middle cerebral artery occlusion in rats. *J Pharmacol Exp Ther* 1998;287:559–566.

93. Kawasaki-Yatsugi S, Yatsugi S, Takahashi M, et al. A novel AMPA receptor antagonist, YM872, reduces infarct size after middle cerebral artery occlusion in rats. *Brain Res* 1998;793:39–46.

94. Umemura K, Kondo K, Ikeda Y, et al. Pharmacokinetics and safety of the novel amino-3-hydroxy-5methylisoxazole-4-propionate receptor antagonist YM90K ion healthy men. *J Clin Pharmacol* 1997;37:719–727.

95. Kohara A, Okada M, Ohno K, et al. In vitro characterization of YM872: A selective, potent and highly water soluble AMPA receptor antagonist. *Soc Neurosci Abst* 1996;22:1528.

96. http://www.neuro.wustl.edu/stroke/therapy/Therapy_1Page56.html

97. Wahlgren NG, Ranasinha KW, Rosolacci T, et al. Clomethiazole Acute *Stroke* Study (CLASS): results of a randomised controlled trial of clomethiazole versus placebo in 1360 acute stroke patients. *Stroke* 1999a;30:21–28.

98. Wahlgren NG, Bornhov S, Sharma A, et al. The Clomethizaole Acute *Stroke* Study (CLASS): efficacy results in 545 patients classified as total anterior circulation syndrome (TACS). *J Stroke Cerebravasc Dis* 1999b;8:1–10.

99. Barber PA, Demchuk AM, Zhang J, et al. Hyperacute stroke: the validity and reliability of a novel quantitative CT score in predicting outcome prior to thrombolytic therapy. ASPECTS Study Group. Alberta Stroke Programme Early CT Score. *Lancet* 2000;355:1670–1674.

100. Clark W, Ertag W, Orecchio E, et al. Cervene in acute ischemic stroke: results of a double-blind, placebo-controlled, dose-comparison study. J *Stroke Cerebrovasc Dis* 1999;8:224–230.

101. Haley EC Jr, Kassell NF, Alves WM, et al. Phase II trial of tirilazad in aneurysmal subarchnoid hemorrhage. A report of the Cooperative Aneurysm Study. *J Neurosurg* 1995;82:786–790.

102. Johnston KC, Li JY, Lyden PD, et al. Medical and neurologic complications of ischemic stroke: Experience from the RANTTAS trial. RANTTAS Investigators. *Stroke* 1998;29:1256–1257.

103. Yamaguchi T, Sano K, Takakura K, et al. Ebselen in acute ischemic stroke: a placebo-controlled, double-blind clinical trial. Ebselen Study Group. *Stroke* 1998;29:12–17.

104. Clark WM, Warach SJ, Pettigrew LC, et al. A randomized dose-response trial of citicoline in acute ischemic stroke patients. Citicoline Stroke Study Group. *Neurology* 1997;49:671–768.

105. Gammans RE. ECCO 2000: study of citicoline for treatment of acute ischemic stroke: Final results. *Stroke* 2000;31:278.

106. http://www.cambneuro.com/tech/tech.htm Bayer and CNSI collaborating on GGF2 protein growth factors: a pivotal role glial growth factor 2: numerous applications.

107. Liu J, Solway K, Messing RO, et al. Increased neurogenesis in the dentate gyrus after transient global ischemia in gerbils. *J Neurosci* 1998;18:7768–7778.

108. Chen J, Li Y, Chopp M. Intracerebral transplantation of bone marrow with BDNF after MCAO in rat. *Neuropharmacology* 2000;39:711–716.

109. Yusuf S, Dagenais G, Pogue J, et al. Vitamin E supplementation and cardiovascular events in high-risk patients. The Heart Outcomes Prevention Evaluation Study Investigators. *N Engl J Med* 2000;342:154–160.

110. Toole JF. Vitamin intervention for stroke prevention (VISP). The VISP Investigators. *Stroke* 2000;31:249.

111. Hankey G, Hickling S, Baker R, et al. VITATOPS a study of vitamins to prevent stroke. The VITATOPS study group. *Stroke* 2000;31:267.

112. Kernan WN, Brass LM, Viscoli CM, et al. Estrogen after ischemic stroke: clinical basis and design of the Women's Estrogen for Stroke Trial. *J Stroke Cerebrovasc Dis* 1998;7:63–69.

94

Tau Protein and Tauopathy

Makoto Higuchi

John Q. Trojanowski

Virginia M.-Y. Lee

Makoto Higuchi, John Q. Trojanowski, and Virginia M.-Y. Lee: Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

TAU-POSITIVE FILAMENTOUS LESIONS IN NEURODEGENERATIVE DISEASES

Part of "94 - Tau Protein and Tauopathy "

Neurofibrillary Lesions of Alzheimer's Disease Brains

Although the mechanisms underlying the onset and progression of Alzheimer's disease (AD) have not been fully elucidated, the two diagnostic neuropathologies in the AD brain (i.e., amyloid plaques and neurofibrillary lesions) (1 ,2) have been implicated mechanistically in the degeneration of the AD brain, and they are considered to be plausible targets for the discovery of potential therapeutic agents to treat this common dementing disorder. AD is a genotypically and phenotypically heterogeneous disease. In spite of this genetic heterogeneity, abundant amyloid plaques and neurofibrillary lesions, including neurofibrillary tangles (NFTs), neuropil threads, and plaque neurites are observed consistently in all forms of AD, and both plaques and tangles are required to establish a definite diagnosis of AD in a patient with dementia.

Amyloid plaques are extracellular deposits of fibrils formed by ß-amyloid (Aß) peptides cleaved from APP, but Aß also forms diffuse plaques that contain primarily nonfibrillar deposits of Aß peptides. The neuritic type of amyloid or senile plaque (SP) binds amyloid dyes such as thioflavin-S and Congo red because of the presence of Aß fibrils with a ß-pleated sheet structure. The neurofibrillary AD lesions also contain aggregated filaments, but they are formed by abnormally phosphorylated tau proteins that accumulate as NFTs in neuronal perikarya and as neuropil threads or dystrophic neurites in dendrites and axons. However, NFTs may be released into the extracellular space of the AD brain, following the degeneration of tangle-bearing neurons, and they are referred to as "ghost tangles." Finally, dystrophic neurites are frequently associated with amyloid plaques to form neuritic plaques.

Both amyloid plaques and neurofibrillary lesions are considered to play independent and/or interrelated roles in the mechanisms that underlie the onset and relentless progression of brain degeneration in AD. Indeed, serveral studies have shown that NFTs correlate with the severity of dementia in AD, as do losses of synapses and neurons (10 ,11 ,12 and 13). Although there is a poor correlation between these parameters and the concentration or distribution of amyloid deposits (11 ,13), this could reflect the turnover of these lesions. Furthermore, it has been reported that a small population of AD patients show abundant NFTs but very few amyloid plaques (14), which may signify that there is a causal relationship between the accumulation of NFTs and the clinical manifestations of AD.

Despite heterogeneity in the AD phenotype, the progressive accumulation of NFTs follows a stereotypical pattern as described by Braak and Braak (15), who defined six neuropathologic stages of AD progression determined by the distribution and severity of NFTs. Stage I shows NFTs and neuropil threads confined to pre-α neurons of the transentorhinal cortex, and stage II shows a more remarkable involvement of this area and a mild involvement of the pre-α neurons in the entorhinal cortex. AD brains in stage III show severe neurofibrillary lesions in the above-mentioned regions as well as the emergence of extracellular tangles, and extensive neurofibrillary lesions are found in the deeper layers of entorhinal and transentorhinal cortex in stage IV. Stages III and IV are also characterized by neurofibrillary pathology in layer I of Ammon's horn in the hippocampus and in subcortical nuclei. Finally, increasingly abundant neurofibrillary lesions in isocortical association cortex define stages V and VI.

Neurofibrillary lesions such as NFTs, neuropil threads, and plaque neurites are argyrophilic structures, but they are visualized most effectively using immunohistochemistry and antibodies to phosphorylated tau proteins. In addition to neurofibrillary tau lesions, some neurons show diffuse perikaryal tau immunoreactivity, and this so-called "pretangle"

tau pathology is not stained by amyloid dyes such as thioflavin-S and Congo red, unlike NFTs and other neurofibrillary lesions. Thus, “pretangle” tau pathology may be an early stage in the formation of NFTs prior to the accumulation of abnormal tau filaments.

Tauopathies Other than AD

Neurofibrillary lesions that are positive for thioflavin-S, silver stains, and anti-tau antibodies are also observed as the predominant brain pathology in a group of neurodegenerative disorders other than AD, which are now categorized as tauopathies (Table 94.1). Some of these diseases also show the abundant coexistence of amyloid plaques. For example, neurofibrillary lesions coexist with Aß deposits in AD as well as in Down syndrome (16 ,17), dementia pugilistica (18), and inclusion-body myositis (19 ,20 and 21). Further, in some cases of Gerstmann-Sträussler-Scheinker disease (GSS) (22 ,23), Creutzfeldt-Jakob disease (24), and prion protein cerebral amyloid angiopathy (25), neurofibrillary lesions coexist with prion protein amyloid deposits. On the other hand, amyotrophic lateral sclerosis/parkinsonism-dementia complex (ALS/PDC) found in the Chamorro inhabitants of Guam and Rota in the Mariana Islands shows abundant NFTs but very few amyloid plaques (26 ,27 ,28 and 29). Moreover, neurofibrillary lesions without amyloid plaques are observed in argyrophilic grain dementia (30 ,31), Pick disease (32 ,33 and 34), corticobasal degeneration (CBD) (35 ,36 ,37 and 38), progressive supranuclear palsy (PSP) (39 ,40 and 41), multiple system atrophy (MSA) (42), Niemann-Pick disease type C (43 ,44 and 45), diffuse neurofibrillary tangles with calcification (46), Hallervorden-Spatz disease (47), subacute sclerosing panencephalitis (48), and frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17) (49 ,50). However, some of these disorders, such as MSA, various subtypes of AD, Hallervorden-Spatz disease, and so on also have prominent synuclein brain lesions.

Diseases showing coexistence of tau and amyloid pathologies
Alzheimer's disease
Creutzfeldt-Jakob disease
Dementia pugilistica
Down's syndrome
Gerstmann-Sträussler-Scheinker disease
Inclusion-body myositis
Prion protein cerebral amyloid angiopathy
Diseases without distinct amyloid pathology
Amyotrophic lateral sclerosis/parkinsonism–dementia complex
Argyrophilic grain dementia
Corticobasal degeneration
Diffuse neurofibrillary tangles with calcification
Frontotemporal dementia with parkinsonism linked to chromosome 17
Hallevorden–Spatz disease[a]
Multiple system atrophy[a]
Niemann–Pick disease, type C
Pick's disease
Progressive subcortical gliosis
Progressive supranuclear palsy
Subacute sclerosing panencephalitis
Tangle-predominant Alzheimer's disease

[a]Diseases in which synuclein-positive lesions are the most prominent neuropathologic feature.

TABLE 94.1. DISEASES WITH TAU-POSITIVE NEUROFIBRILLARY LESIONS

The tau pathology of AD is almost limited to neurons, whereas some other tauopathies exhibit both neuronal and glial tau inclusions. Brains of MSA, CBD, PSP, and FTDP-17 contain abundant tau deposits in astrocytes as well as oligodendrocytes (50 ,51 ,52 ,53 ,54 ,55 ,56 ,57 ,58 ,59 ,60 and 61). On the other hand, in familial multiple system tauopathy with presenile dementia (MSTD), affected glial cells are primarily oligodendrocytes (62 ,63 and 64).

Neurofibrillary Lesions with Aging

Abundant amyloid plaques indistinguishable from those in AD brains have been demonstrated in the brains of elderly individuals who are not cognitively impaired (12 ,65 ,66); this indicates that accumulation of amyloid plaques alone is not sufficient to cause dementia. Moreover, nondemented elderly individuals also show sparse neurofibrillary lesions with increasing age, but this occurs in limited brain regions (67). Although extensive analysis by Braak and Braak has suggested that neurofibrillary changes of Braak and Braak stage I/II in elderly people may represent early stages of AD pathology (68), this has yet to be proven in studies of subjects who have been subjected to longitudinal cognitive testing up until the time of death.

Significances of Tau Pathology in Neurodegenerative Disorders

Aggregation of tau into neurofibrillary lesions is a neuropathologic hallmark of many neurodegenerative diseases, and these tauopathies can be subclassified with regard to coexisting amyloid or other brain pathology, the affected cell types and the affected CNS areas. Coexistence of tau and amyloid pathologies in some diseases suggests an interaction between tau and amyloid in mechanisms of brain degeneration. The presence of tau lesions in the brain without amyloid plaques in other tauopathies indicates that these neurofibrillary lesions are not mere consequences of neurotoxicity owing to amyloid deposits, but are direct causes of neurodegeneration. Although it is likely that the diversity of affected cell types and/or CNS regions in tauopathies will be explained more fully based on mechanistic relationships between genetic and biochemical differences among these cells, brain regions, and diseases, the clinicopathologic, biochemical, and genetic aspects of these disorders so far remain unsettled.

Ultrastructure of Filamentous Tau Lesions

According to transmission electron microscopic (EM) and immuno-EM analyses of tau filaments in various neurofibrillary lesions, the filamentous lesions consist of three types of morphologies. Approximately 95% of the neurofibrillary components in AD NFTs are paired helical filaments (PHFs), and the rest consists of straight filaments (SFs) (69 ,70). PHFs have a helical structure consisting of two ribbon-like strands that are paired together in filaments that have a diameter of 8 to 20 nm and a stereotypical periodicity of 80 nm (70 ,71). In Down syndrome, ALS/PDC, prion diseases with tangles, dementia with tangles only, Nieman-Pick disease type C, and the Seattle family A FTDP-17 kindred with the V337M tau gene mutation, the filamentous tau pathology is composed of fibrils that are ultrastructurally indistinguishable from the PHFs in AD tangles (29 ,43 ,44 and 45 ,70 ,72 ,73). Moreover, PSP and Pick disease show tangles composed of numerous SFs and smaller numbers of twisted tau filaments similar to PHFs (63 ,74). Twisted ribbon-like tau filaments that are morphologically different from AD PHFs and SFs are found in the tangles of the familial MSTD FTDP-17 syndrome caused by a G to A mutation in the intron following exon 10 of the tau gene (64), Dutch family 1 FTDP-17 syndrome owing to the P301L mutation in exon 10 of the tau gene (75), and CBD (76). Unlike AD PHFs, these filaments have an irregular periodicity of 90 to 130 nm (64).

Nonetheless, immuno-EM studies have demonstrated that all the filamentous structures in the tangles of all of these tauopathies are composed of aberrantly hyperphosphorylated tau proteins, and possess the same tau epitopes (77 ,78 ,79 ,80 ,81 and 82), although the relative abundance of different pathologic tau isoforms may vary in these tauopathies, as discussed in the following. Currently, there seems to be no association between ultrastructural diversity and biochemical or genetic property in tauopathies. Observation of hybrid filaments suggests a transition from PHF to SF.

BIOCHEMICAL FEATURES OF TAU PROTEINS IN NORMAL AND PATHOLOGIC CONDITIONS

Part of "94 - Tau Protein and Tauopathy "

Localization and Function of Tau Protein

Tau is a low molecular weight component of cytoskeletal structures and is known as one of the microtubule-associated proteins (MAPs). Neuronal MAPs consisting of tau and MAP2 regulate the assembly of microtubules (MTs). Although tau and MAP2 are thought to have similar functions, intracellular localization of tau largely differs from that of MAP2. The mRNAs encoding tau proteins are expressed predominantly in neurons, where these tau proteins are localized mostly to axons of the CNS and PNS under normal physiologic conditions (83 ,84), whereas the neurofibrillary lesions in AD accumulate in the neuronal perikarya, axons, and dendrites. In contrast to the axon-specific distribution of tau in normal states, MAP2 has somatodendritic localization (85 ,86). Although it is likely that the compartment specificity of normal tau and MAP2 in neurons may subserve functional differences such as organization of neuronal polarity and spacing of intermicrotubule distances, or other aspects of axonal and somatodendritic MT distribution and architecture (87 ,88 ,89 ,90 and 91), there is no direct evidence for these different roles for tau and MAP2. Lower expression of tau mRNA and less abundant tau protein have been observed in astrocytes as well as in oligodendrocytes (92 ,93), and this suggests that the formation of glial tau inclusions in several neurodegenerative tauopathies results from the aggregation of tau proteins produced in glial cells themselves.

It has been shown in a number of studies from many laboratories that tau proteins play a major role in regulating neuronal MT assembly and stability (94 ,95 and 96). For example, tau proteins promote the polymerization of tubulin into MTs (97), and tau bound to MTs help stabilize these structures in the polymerized state (98). Moreover, developing neurons treated with antisense oligonucleotides to tau mRNA to block expression of tau fail to extend axon-like processes, suggesting that tau protein also functions in, or is required for, the establishment of neuronal polarity during development (99 ,100). However, mice lacking tau protein present no major phenotypic changes initially, and the only abnormal findings in these mice are decreased MT density and stability in some small-caliber axons (101), but they show motor impairments with age (102). Tau is likely to play an essential role in the development of neurons and even glial cells, but it is also probable that other proteins such as MAP1A can be upregulated to partially compensate for the loss of tau at least early in life, as indicated in the preceding tau-knockout mouse study (101 ,102).

Expression of Multiple Tau Isoforms

As a consequence of alternative mRNA splicing, the single tau gene on the long arm of chromosome 17 gives rise to six brain tau proteins that are normally expressed in the adult human CNS (77 ,103 ,104 and 105) (Fig. 94.1). The differences among these six brain tau isoforms result from the presence of three (3R tau) or four (4R tau) imperfect MT binding repeats of 31 or 32 amino acids in the carboxy-terminal half of each of two sets of these proteins, as well as from the presence of inserts of 29 or 58 amino acids or no insert at all in the amino-terminal region (77 ,103). The tandem repeats in the carboxy-terminal half are encoded by exons 9, 10, 11, and 12, and the alternative splicing of exon 10 (E10) results in the generation of E10+ 4R tau and E10- 3R tau mRNAs and their corresponding 4R and 3R tau isoforms, respectively. This consecutive repeat region

constitutes the MT-binding domain of each tau protein (106 ,107). In each domain, binding affinity to MTs is provided by a binding element that consists of 18 amino acids (107), but the remainder of this motif, known as the interrepeat sequence, also may contribute to the binding of tau to MTs. Indeed, the interrepeat sequence between MT-binding repeats 1 and 2, which is included only in 4R tau isoforms, has a binding affinity for MTs that is more than twofold higher than any MT-binding repeat (108). This may suggests that 4R tau plays a much greater role in regulating the MT-binding than 3R tau, and it is possible that 3R and 4R tau have different MT-binding sites on MTs. The function of the amino-terminal region remains unsettled, but this region is supposed to affect inter-MT distances by forming a bridge between two adjacent MTs. In the PNS, a high molecular weight tau isoform (110 kDa) with one additional exon (exon 4A) is expressed (known as “big tau”) (84 ,110).

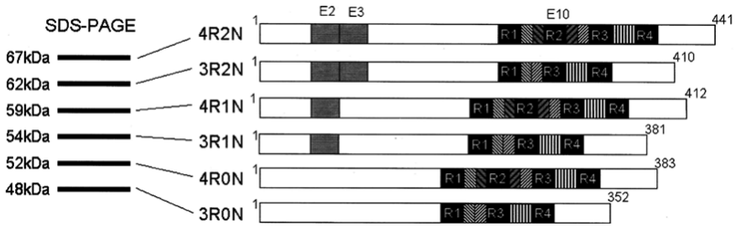


FIGURE 94.1. Six human CNS tau isoforms produced by alternative splicing of the tau gene. The differences among the six isoforms are the number of MT-binding repeat domains *(black boxes)* and the number of amino-terminal inserts. The alternatively spliced exons, exons 2 (E2), 3 (E3), and 10 (E10) are indicated with gray boxes, and hatched boxes indicate the inter-repeat sequences. The recombinant tau proteins run as six bands on SDS-PAGE *(left)*.

The alternative splicing of the six brain tau isoforms is developementally regulated, and only the shortest tau isoform with three repeats and no amino-terminal inserts (i.e. “fetal tau” or 3R0N tau) is present in fetal human brains (102). By analyzing fresh biopsy-derived nomal fragments of human adult brain, it has been demonstrated that the adult CNS contains the following tau isoforms: tau with one amino-terminal insert (1N tau, 50%), tau with no amino-terminal insert (0N tau, 40%) and tau with two amino-terminal inserts (2N tau, 10%) in order of abundance. In the same analysis, the ratio between 4R and 3R tau isoforms has been found to be approximately 1 (111). However, it also is known that the isoform composition of tau protein differs among species. For example, only the three 4R tau isoforms are known to be expressed in the adult rodent brain, while a 3R0N or fetal tau isoform is expressed in the developing CNS of rodents (112). Although the reasons for this difference between the adult rodent and human brains are not known nor is the functional consequence thereof evident at this time, it is possible that the lack of a stem-loop structure in the intron following E10 in rodents versus humans may account for the failure to express 3R tau isoforms in the adult rodent brain.

Phosphorylation of Tau in Normal and AD Brains

Tau is a phosphoprotein, and tau isolated from the developing and adult brain is phosphorylated at multiple sites. PHF-tau extracted from the AD brain shows three major bands (approximately 60, 64, and 68 kDa) and one minor band (approximately 72 kDa) in SDS-PAGE. Enzymatic dephosphorylation of PHF-tau *in vitro* using alkaline phosphatase changes the electrophoretic mobility of these three bands to generate six bands that are identical to the six tau isoforms extracted from normal human brain after dephosphorylation and the six recombinant human tau proteins. This suggests that PHF-tau in AD is composed of all six tau isoforms that are abnormally phosphorylated. Indeed, these PHF-tau bands are detected using antibodies specific for phosphorylated tau epitopes as well as by other phosphorylation-independent anti-tau antibodies.

Approximately 20 serine and threonine residues in tau, some of which are followed by a proline, currently are known to be sites of normal phosphorylation (113 ,114) (Fig. 94.2). Although many of these sites initially were thought to be unique to PHF-tau in AD (114), subsequent studies summarized in the following did not confirm this. Most of these phosphorylation sites are clustered in regions flanking the MT-binding domains, and thus it is presumed that the phosphorylation of these sites influences the binding of tau to MTs. In fact, the conformation of tau is changed by phosphorylation (115), and this reduces binding of tau to MTs (98 ,116 ,117 ,118 and 119), lowers the ability of tau to promote MT assembly (120) and decreases the stability of MTs (121 ,122). It has also been shown that PHF-tau cannot bind to MTs (117 ,118), and that the binding ability is restored after enzymatic dephosphorylation of PHF-tau *in vitro* by phosphatase (118 ,123 ,124). If phosphorylation of a certain site is unique to PHF-tau in AD, elucidation of the mechanism for phosphorylation of such a site could provide much information on the pathogenesis of AD. However, the phosphorylation at multiple sites on tau is considered to be requisite for eliminating the MT-binding ability of PHF-tau (125). In addition, normal human fetal tau is more phosphorylated than tau proteins extracted from

postmortem normal adult brain, but almost all of the phosphorylated sites in normal fetal tau are identical to those in PHF-tau (114 ,117 ,126), and it is known that fetal tau also is less capable of binding to MTs than normal adult tau (117). Furthermore, normal adult human tau isolated from biopsy-derived brain sample is phosphorylated at most of the known phosphorylation sites in PHF-tau, albeit to a much lesser extent (127 ,128). Nevertheless, there still remains a possibility that the difference between PHF-tau and normal tau is not only the extent of phosphorylation, but also the aberrant phosphorylation of some sites that are unique to PHF-tau. In fact, serine212 and serine214, the two phosphorylated residues contained in the epitope that are required for recognition by the anti-phospho-tau antibody AT100/AT10 (129 ,130), are at least two abnormally phosphorylated sites that are known to be unique to PHF-tau (128). This antibody fails to recognize tau isolated from fetal and biopsy-derived normal adult human brain (128), and thus phosphorylation of Ser212/Ser214 appears to be highly specific to PHF-tau in AD, but these sites also may be phosphorylated in the abnormal tau proteins in other tauopathies. Therefore, PHF-tau is likely to be “hyperphosphorylated” (i.e., phosphorylated to a greater extent) as well as “aberrantly phosphorylated” (i.e., phosphorylated at unique sites).

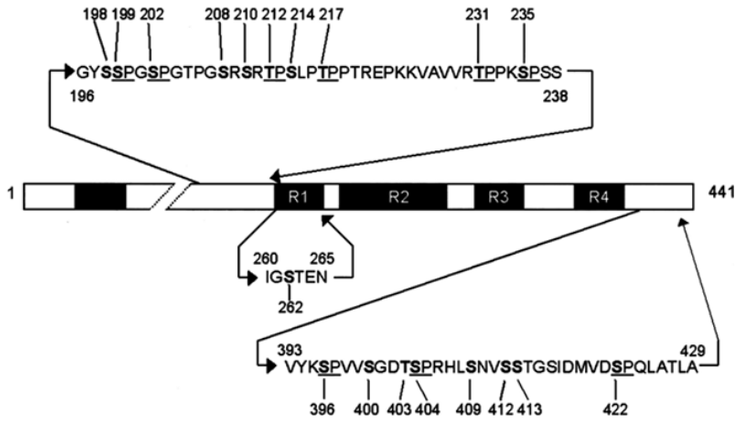


FIGURE 94.2. Phosphorylation sites identified in human PHF-tau. The numbers are based on those in the longest isoform with 441 amino acid residues. The proline-oriented sites are indicated with underlines.

The extent of tau phosphorylation is regulated by the activities of phosphatases and kinases. Therefore, increased activities of kinases and/or decreased activities of phosphatases could presumably lead to the hyperphosphorylation of tau, thereby resulting in the formation of PHF-tau. Although several different kinases have been demonstrated to be capable of phosphorylating tau *in vitro*, the specific kinases that are responsible for the phosphorylation of tau in the living human CNS remain to be identified. Only two kinases, glycogen synthase kinase-3ß (GSK-3ß) and cyclin-dependent kinase 5 (Cdk5), have been copurified with MTs (131 ,132). GSK-3 is known to phosphorylate endogenous tau expressed in neurons (133), and it is abundant in the brain (134). Cdk5 is normally activated by a regulatory protein p35 (135). Cdk5 is supposed to be active in neurons because p35 is expressed primarily in neurons (135 ,136). However, a recent study has shown that a truncated form of p35 (known as p25) accumulates in neurons of the AD brain, and that p25 binds to Cdk5, leading to a deregulation of this kinase (137). Accordingly, hyperactivated Cdk5 may be caused by accumulation of p25, and this may be part of the mechanism that causes the hyperphosphorylation of tau in AD.

Among many phosphatases, protein phosphatase 2A (PP2A) and 2B (PP2B or calcineurin) have been shown to be enzymatically active in biopsy-derived human brain tissue, and *in vitro* studies have shown that these enzymes dephosphorylate several phospho-serine and phospho-threonine residues in tau (70 ,124 ,128 ,129 ,139 ,140). Thus, these phosphatases may be involved in the generation of PHF-tau, but there has been no evidence for decreased phosphatase activity in the AD brain, and thus roles played by these enzymes in formation of PHF-tau are still to be elucidated. Nonetheless, a recent study has demonstrated reduced PP2A subunit mRNAs in the AD hippocampus when NFTs are abundant, so it is plausible that this might contribute to mechanisms underlying PHF-tau formation in AD (140a).

Other Tauopathies

Immunoblot analyses of brains from patients with tauopathies other than AD have demonstrated that insoluble tau fractions are detectable using many different phosphorylation-dependent antibodies to epitopes spanning the tau molecule, suggesting that the filamentous inclusions in these diseases are composed of hyperphosphorylated tau similar

to PHF-tau in AD. Three types of abnormal tau isoform profiles have been found in tauopathies other than AD (Fig. 94.3). For example, similar to AD, three predominant PHF-tau–like bands of 60, 64, and 68 kDa and one minor band of 72 kDa are observed in Down syndrome, GSS disease, ALS/PDC of Guam, Niemann-Pick disease type C, and some FTDP-17 kindreds caused by certain tau mutations (17 ,23 ,29 ,43 ,72). These bands have been shown to contain all six tau isoforms in several studies. On the other hand, biochemical analyses of Pick disease have shown a characteristic pattern of tau isoform composition consisting of two major tau bands of 60 and 64 kDa and one minor tau band of 68 kDa (34 ,141). In some studies, these bands are not positive for anti-phospho-tau antibody 12E8, which recognizes an epitope in E10, but our study demonstrated weak but specific recognition (34). Further, several studies have indicated that these bands include only 3R tau isoforms (142 ,143). The third type of abnormal tau isoform profile is found in CBD (40 ,76), PSP (41), and some FTDP-17 kindreds with specific tau gene mutations (64 ,144), and this profile is characterized by two major tau bands of 64 and 68 kDa as well as one minor tau band of 72 kDa. Notably, several studies have indicated that these tau bands predominantly consist of 4R tau isoforms (64 ,110 ,144 ,145).

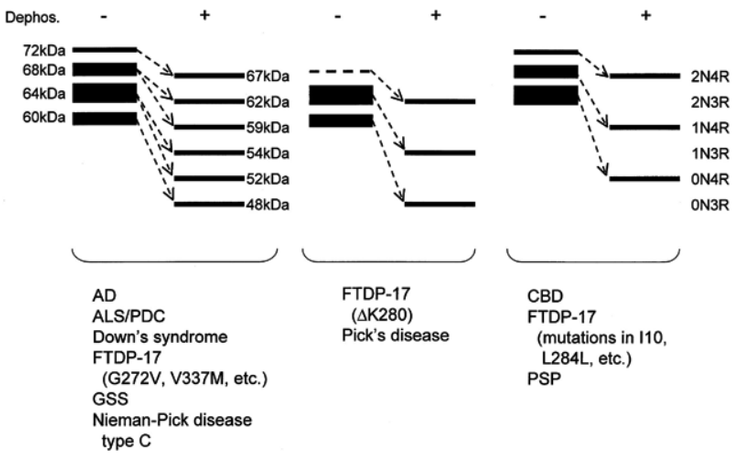


FIGURE 94.3. Schematic representation of sarkosyl-insoluble tau bands from different tauopathies before (-) and after (+) dephosphorylation. The insoluble tau composed of all six isoforms shows three major bands (60, 64, and 68 kDa) and one minor band (72 kDa) before dephosphorylation. In some tauopathies, 3R tau isoforms are major components of the insoluble fraction, running as two major bands (60 and 64 kDa) and one minor band (68 kDa). Tauopathies with insoluble tau consisting of 4R tau isoforms shows two major bands (64 and 68 kDa) and one minor band (72 kDa).

FRONTOTEMPORAL DEMENTIA WITH PARKINSONISM LINKED TO CHROMOSOME 17: CAUSED BY MULTIPLE EXONIC AND INTRONIC TAU GENE MUTATIONS

Part of "94 - Tau Protein and Tauopathy "

FTDP-17 is a group of familial neurodegenerative tauopathies characterized by diverse but overlapping clinical and neuropathologic features (50 ,60 ,146). According to several reports on clinical and neuropathologic features of FTDP-17, three major clinical syndromes have been delineated, and albeit preliminary, they include: disinhibition-dementia-parkinsonism-amyotrophy complex (DDPAC) (55 ,147), pallido-ponto-nigral degeneration (PPND) (148), and MSTD (64). However, it is important to note that more than 20 kindreds caused by diverse tau gene muations and variably characterized phenotypes have been reported so far (50). Clinical characteristics of these FTDP-17 tauopathies variably include memory and language impairments, behavioral and psychiatric abnormalities, extrapyramidal signs, and motor deficits (50), each of which presumably reflects differential degeneration of specific brain regions. However, all FTDP-17 brains from affected patients share a common neuropathology characterized by abundant neuronal and to a lesser extent glial fibrillary lesions composed of hyperphosphorylated tau proteins associated with a remarkable loss of neurons in affected regions (144 ,149 ,150 and 151).

Autosomal dominant inheritance of these FTDP-17 syndromes suggested that one or more genetic mutations might be pathogenic for these disorders, and linkage analyses showed cosegregation of disease with a genetic locus on chromosome 17q21-22 (49 ,50 ,60 ,146 ,152 ,153). Because the pathologic hallmarks of these disorders are tau lesions and the tau gene resides within the disease locus of chromosome 17, the tau gene was an obvious candidate for pathogenic mutations in FTDP-17 kindreds. Thus, as expected, several research groups discovered multiple tau gene mutations in 1998, and these mutations were found to segregate with FTDP-17 patients, but they were not seen in normal individuals (151 ,154 ,155). Further studies have identified at least 20 distinct pathogenic mutations in exons and introns of the tau gene in many FTDP-17 kindreds, many of which were identified for the first time with the identification of a tau gene mutation. Approximately 10 missense mutations were found in exons of the tau gene, and they include K257T, I260V (Hutton M, personal communication),

and G272V (154) in exon 9; N279K (142), P301L (142 ,152 ,155), P301S (156 ,157), and S305N (158 ,159) in E10; V337M (155) in exon 12; and G389R (160) and R406W (154) in exon 13 (numbered according to the longest CNS tau isoform consisting of 441 amino acids). Moreover, two silent mutations have been reported, including L284L (159) and S305S (161) in E10 and a mutation resulting in the deletion of single amino acid ΔK280 (159 ,167) in E10. On the other hand, intronic tau gene mutations in FTDP-17 kindreds are clustered around the 5′ splice site in the intron following E10. They contain E10+3 (151), E10+12 (163), E10+13 (164), E10+14 (145 ,164), E10+16 (164 ,165), and E10+33 (162). The currently known tau gene mutations in FTDP-17 kindreds are listed in Table 94.2 and are depicted schematically in Fig. 94.4 . The increasing number of tau gene mutations that continue to be identified suggests that FTDP-17 is likely to be more frequent than previously recognized.

Mutation	Location (Domain)	Effects on Exon 10 Splicing	Predominant Tau Isoforms	Effects on MT Binding	Disease/Syndrome	Tau-Positive Neuropathology	References
K257T	E9 (R1)	N/A	N/A	N/A	N/A	Tau as well as Pick body-like inclusions	[a]
I260V	E9 (R1)	N/A	N/A	N/A	N/A	N/A	[a]
G272V	E9 (R1)	None	All isoforms	Reduced	HFTD2; "familial Pick's disease"	Inclusions in cortical and subcortical areas	154
N279K	E10 (IR1–2)	Increased	Mainly 4R	None	PPND	Neuronal and glial fibrillary tangles	142
ΔK280	E10 (IR1–2)	Reduced	3R	Reduced	Dutch FTD	N/A	159, 167
L284L (CTT to CTC)	E10 (IR1–2)	Increased	4R	None	"Variant FTD"	Widespread amyloid as well as tau deposits	159
P301L	E10 (R2)	None	All isoforms	Reduced	HFTD1; "Dutch family"	Glial and neuronal tangles	142, 152, 155
P301S	E10, R2	None	All isoforms	Reduced	FTD and CBD-like	Extensive filamentous pathology	156, 157
S305N	E10 (IR2–3)	Increased	Mainly 4R	None	Very early onset presenile dementia, CBD-like	Glial and neuronal inclusions; many unusual ring-shaped NFTs	158, 159
S305S	E10 (IR2–3)	Increased	Mainly 4R	None	PSP-like	Subcortical NFTs	161
V337M	E12 (IR3–4)	None	All isoforms	Reduced	"Seattle family A"	NFTs indistinguishable from AD NFTs	155
G389R	E13 (C-term)	None	All isoforms	Reduced	N/A	Tau as well as Pick body-like inclusions	160
R406W	E13 (C-term)	None	All isoforms	Reduced	"Iowa family," PSP-like	Cortical and subcortical NFTs	154
E10+3 (G to A)	I10 (5′ splice site)	Increased	4R	None	MSTD	Neuronal and glial inclusions	151
E10+12 (C to T)	I10 (5′ splice site)	Increased	4R	None	"FTD Kumamoto"	Neuronal and glial inclusions	163
E10+13 (C to T)	I10 (5′ splice site)	Increased	4R	None	N/A	N/A	164
E10+14 (C to T)	I10 (5′ splice site)	Increased	4R	None	DDPAC	Ballooned neurons with tau staining	145, 164
E10+16 (C to T)	I10 (5′ splice site)	Increased	4R	None	"Australian" pedigree and PSG	Neuronal and glial inclusions	164, 165
E10+33	I10 (5′ splice site)	Increased	4R	ND	N/A	N/A	162

C-term, carboxyterminus; E, exon; I, intron; IR, interrepeat region; N-term, amino-terminus; R, MT-binding repeat.
[a]Hutton M, personal communication.

TABLE 94.2. MUTATIONS ON THE TAU GENE FOUND IN FTDP-17

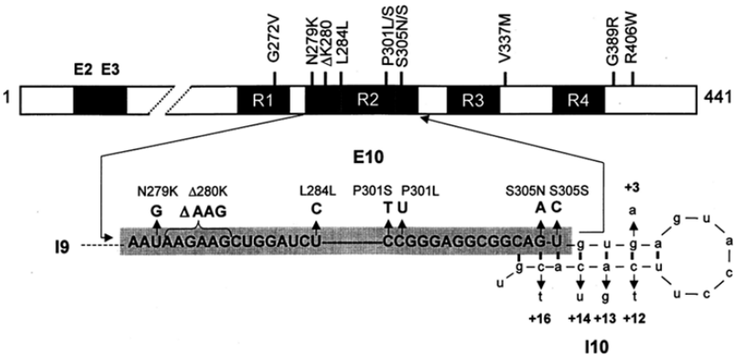


FIGURE 94.4. Mutations identified on tau gene in FTDP-17. The mutation sites are depicted on the longest tau isoform. The alternatively spliced inserts are indicated as gray boxes and MT-binding repeats are shown as black boxes. Sequences in intron 10, which form a stem-loop structure, are presented in lower case.

In FTDP-17, pathogenic mutations in the tau gene may be pathogenic by one or more abnormalities in tau proteins, and at present, two mechanisms have been proposed to mediate the effects of these mutations based on recent molecular and biochemical analyses (111 ,154 ,159 ,166). The first mechanism involves perturbations of the alternative splicing of E10 by mutations in E10 or around the 5′ splice site in the intron following E10, thereby resulting in an altered ratio of 4R tau to 3R tau proteins. The second pathogenic mechanism directly impairs the ability of tau to bind to MTs and to promote the polymerization and stability of MTs.

The ratio of 4R tau to 3R tau isoforms is approximately 1 in the normal adult human brain, but an increase in the 4R/3R ratio of brain tau isoforms has been demonstrated in brains of FTDP-17 patients with mutations clustered around the 5′ splice site in the intron following E10 and with E10 tau gene mutations including: N279K, L284L, S305N, and S305S (111 ,154 ,159). The altered splicing of E10 caused by these mutations results in increased levels of E10+ tau mRNA in FTDP-17 brains presumably owing to greater E10 usage of the E10 5′ splice site as demonstrated by exon-trapping experiments (154 ,159). Biochemical analyses of tau extracted from autopsied brain samples of patients with PPND (N279K), DDPAC (E10+14) and MSTD (E10+3) have shown a predominance of 4R tau isoforms (111 ,145 ,151). It has been suggested that these mutations affect multiple cis-acting elements that enhance or suppress the usage of 5′ splice site of E10 (151 ,154 ,159). A stem-loop structure consisting of sequences around the 5′ splice site in the intron following E10 is thought to inhibit the splicing of E10 presumably by blocking the association of snRNA with the splice site (151 ,154). Accordingly, it is postulated that these intronic mutations and the exonic mutations S305N and S305S may cause a disruption of this stem-loop structure (154). The S305N mutation has also been suggested to increase the strength of the 5′ splice site (167). However, the S305S mutation, which increases the 4R/3R ratio like the S305N mutation, has been demonstrated to weaken the 5′ splice site (161). Another potential regulatory element might be an exon-splicing enhancer (ESE) or exon-splicing silencer (ESS) element within E10 adjucent to the following intron (159). The N279K mutation is thought to augment the ESE and consequently cause an increase in the 4R/3R ratio of tau isoforms because of the fact that it raises the purine content of this purine-rich domain (i.e., by changing TAAGAA to GAAGAA) (168 ,169 ,170 and 171). This mechanistic hypothesis is supported by the observation that the ΔK280 mutation, which produces a deletion of the three adjacent purines (AAG), abolishes the splicing of E10 into tau mRNAs (159). The silent mutation L284L is likely to disturb the ESS (159), but it also is possible that this mutation augments the effects of the ESE.

The binding sites on MTs for 4R tau and 3R tau isoforms have been suggested to be different (109), and increases in the 4R/3R tau isoform ratio is likely to produce an excess

amount of 4R tau isoforms that are not bound to MTs. This abnormal increase of free tau may result in the formation of insoluble tau aggregates and consequently neurodegeneration.

The second hypothetical pathogenic mechanism to account for brain degeneration in FTDP-17 owing to other tau gene mutations suggests that these mutations directly cause deficits in the abilities of tau to bind to MTs and promote assembly and stability of MTs. This disease mechanism has been linked to several tau gene missense mutations including: G272V, ΔK280, P301L, P301S, V337M G389R, and R406W by *in vitro* studies (111 ,159 ,166). On the other hand, the mutations that increase E10 splicing do not have similar effects on the functions of tau (111 ,159). Nonetheless, a loss of the binding ability of tau to MTs may produce an increase in the levels of free tau proteins in the neuronal cytoplasm, and this could promote their fibrillogenesis. Mutant tau proteins are also likely to accelerate the accumulation of insoluble tau filaments within neurons. This notion has been supported by several studies of the *in vitro* assembly of tau filaments, which also demonstrated that tau filament formation is enhanced by heparin using recombinant G272, P301L, V337M, and R406W tau mutant proteins compared to wild-type tau protein (172 ,173). Moreover, mutations in exons other than those in E10 (i.e., V337M and R406W) promote tau aggregation composed of all six isoforms, whereas other E10 mutations (i.e., P301L) increase 4R tau in insoluble FTDP-17 brain fractions (111 ,145). Although there are preliminary data to account for the differential effects of these mutations, additional studies are needed to fully elucidate how they cause diverse FTDP-17 syndromes.

DEVELOPMENT OF EXPERIMENTAL ANIMAL MODELS OF FILAMENTOUS TAU PATHOLOGY

Part of "94 - Tau Protein and Tauopathy "

The discovery of tau gene mutations pathogenic for FTDP-17 indicates that genetic abnormalities directly influence the levels or functions of tau proteins, thereby resulting in the aggregation of insoluble tau and neurodegeneration. However, it is difficult to specify the precise underlying mechanisms whereby these mutations cause distinct biochemical, neuropathologic, and phenotypic abnormalities in FTDP-17 syndromes that vary from patient to patient by analyzing human cases because of the following reasons: (a) limited sample size of kindreds with each mutation; (b) the difficulty of conducting biochemical and pathologic studies in early stages of the disease; and (c) the possibility that several additional but as yet unknown environmental and/or genetic factors might modify the biochemical and clinicopathologic phenotype of FTDP-17. Accordingly, animal models that reproduce tauopathies are required for better understanding of the central roles played by tau abnormalities in neurodegenerative disorders. Such models are also expected to be useful for assessing methods of early diagnosis and the effectiveness of therapeutic agents for the treatment of neurodegenerative tauopathies.

The strategies for making animal models that recapitulate tauopathies are summarized in the following and include:

1. Selection of DNA constructs to be expressed in the CNS of the model, and the use of cDNA or genomic DNA tau constructs is a straightforward strategy to induce accumulations of tau in the CNS of experimental animals. Indeed, tau cDNAs or minigenes have been used to overexpress specific tau isoform(s), and cause an imbalance of the tau isoform profile similar to that seen in human tauopathies. Further, animal models engineered to express human genomic tau DNA using bacterial artificial chromosome (BAC) or a P1-derived artificial chromosome (PAC) containing the entire tau gene may be informative for elucidating the biochemistry (including E10 splicing) and neuropathology (including emergence of tau deposits) in animals with and without a tau gene mutation. Other possible strategies to generate animal models of tauopathies are to express proteins that regulate the phosphorylation of tau proteins, or to express APP, PS-1, and PS-2 in tau transgenic (Tg) animals to investigate the interaction of these molecules with tau to model tauopathies that show coexistence of tau and amyloid pathologies.
2. Overexpression of tau without mutation to assess the effects of excess tau proteins in the cytoplasm of neurons and glia on the formation of tau aggregates. Because animal models expressing a mutant human tau gene may reproduce the biochemical and pathologic abnormalities in FTDP-17, it is necessary to do this by expressing mutations such as ΔK280, P301L, V337M, and R406W, because they impair the ability of tau to bind to MTs and promote assembly and stability of MTs. Further, the generation of tau DNA with mutations that alter the splicing of E10 to produce animal models would be informative using the entire gene or minigene of human tau. However, the phenotypes of these mutant tau animals should be assessed in comparison with animals showing a similar expression level of wild-type human tau.
3. Use of neuron-specific promoters including the Thy-1, 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMG-CoA) and prion protein (PrP) promoters to generate tau pathology in neurons, whereas animals showing glial tau pathology can be developed by using glia-specific promoters including glial fibrillary acid protein (GFAP) promoter for astrocytes and 2′,3′-cyclin-nucleotide phosphodiesterase (CNP) and myelin basic protein (MBP) promoters for oligodendrocytes. Tg animals with genomic DNA are generally driven by endogenous promoters.

The generation of a tau Tg mouse using a cDNA for the longest tau isoform (T40, 4R2N tau) combined with Thy-1 promoter was reported in 1995 (174), followed by a study of tau Tg mice expressing the shortest tau isoform (T44 or fetal tau, 3R0N tau) driven by the HMG-CoA reductase promoter (175). In these studies, somatodendritic overexpression of human tau was observed using anti-phospho-tau antibodies. However, these Tg mice did not show tau aggregates in any of the CNS regions nor other tauopathy-like phenotypic changes, probably owing to the low expression level of the transgene product. Filamentous tau aggregates have been observed in the spinal cord and brainstem of tau Tg mice generated by using a transgene consisting of fetal tau and the PrP promoter (176) (Fig. 94.5). This Tg mouse showed approximately ninefold more tau protein than the wild-type control, and thus the successful development of tau inclusions is presumably owing to the high expression levels of human tau. The tau aggregates in this Tg mouse are spheroidal inclusions in proximal axons, and they showed an increase in number with aging, consistent with an age-dependent increase in the extent of tau phosphorylation and an age-dependent decrease in the solubility of overexpressed human tau. In addition, this tau Tg mouse showed axonal degeneration and reduced axonal transport as well as motor weakness. Hence, this Tg mouse is thought to be a good model for age-related neurodegeneration in tauopathies, and it is useful for studying the time course of CNS degeneration in a human tauopathy. The spheroidal tau inclusions in these tau Tg mice have also been shown to contain neurofilaments (NFs) and tubulin. This colocalization of tau and NFs is found in the inclusions of ALS/PDC spinal cord. Further studies using these tau Tg mice crossed with NF-knockout mice could be informative by determining whether or not NFs promote the formation of tau aggregates, and whether or not tau can form aggregates in the absence of NFs.

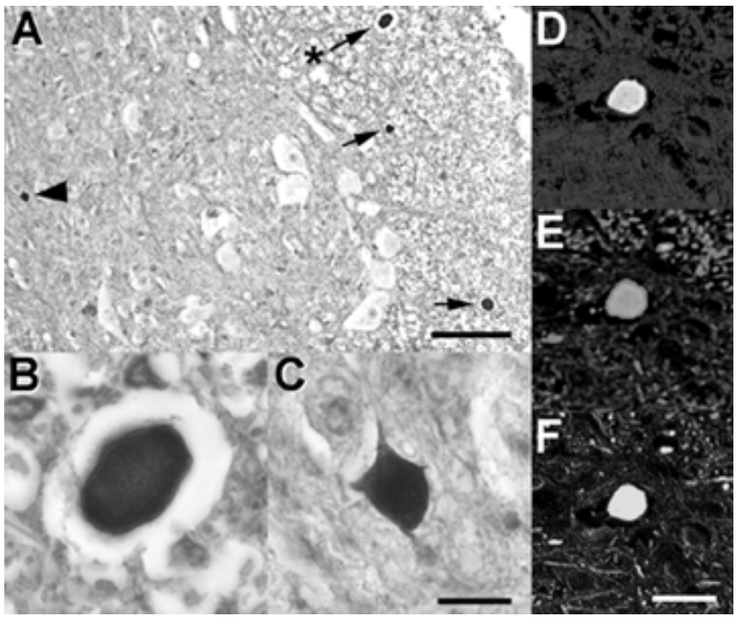


FIGURE 94.5. Spheroidal tau deposits in the spinal cord of tau Tg mice over-expressing the shortest human tau isoform. **A:** Low power field of the spinal cord section of a 6-month-old Tg mouse stained with anti-tau antibody T14. Spheroidal deposits *(arrows)* are observed in axons, and somatodendritic tau stain *(arrowhead)* is found in the neuron. **B:** Higher magnification of the spheroidal tau deposit indicated with asterisk in *(A)*. **C:** Higher magnification of the somatodendritic tau stain shown in *(A)*. **D,E,F:** Triple-labeled indirect immunofluorescence of the spinal cord deposit from a 6-month-old Tg mouse. Blue, anti-tau antibody T14 and AMCA *(D)*; red, anti-low-molecular-mass-neurofilament (NFL) antibody and Text Red *(E)*; and green, anti-high-molecular-mass-neurofilament (NFH) antibody and FITC *(F)*. Note that tau is colocalized with neurofilaments in the spheroid deposits. Scale bar, 50 µm *(A)*; 10 µm *(B,C)*; and 10 µm *(D,E,F)*.

Tau Tg mice with the T40 transgene combined with the Thy-1 promoter have been developed recently (177 ,178). These mice exhibit spheroidal tau inclusions in axons of the spinal cord and brainstem as well as cerebral cortex, and the colocalization of tau and NFs has also observed in these inclusions. Axonal degeneration and corresponding phenotypic changes were found in these Tg mice, and thus they may be regarded as models of neurodegenerative tauopathies with increased 4R tau. In addition, these 4R tau Tg mice showed somatodendritic tau expression to a greater extent than 3R tau Tg mice. This suggests that the difference in affected brain areas between 4R tau and 3R tau Tg mice as well as the effect of predominant tau isoforms on the distribution of pathology should be analyzed by using tau Tg mice with the same promoter and a similar expression level of human tau.

To date, only a few genomic tau Tg mice using PAC and BAC have been generated, and they have shown a somatodendritic pattern of phosphorylated tau expression (179). Although all of the mentioned tau Tg mice have shown a somatodendritic tau expression that resembles the “pretangles” in AD, none of them have developed NFTs containing a ß-pleated sheet structure that can be recognized by thioflavin-S and Congo red. In fact, overexpressed tau in the cytoplasm and processes of neurons is rather diffuse and does not show a filamentous structure by EM (179). One possible method to generate NFTs in tau Tg mice would be to use a mutant tau gene construct to decrease the ability of tau to regulate MTs. Another method would be to follow tau Tg mice showing an age-related increase of tau pathology to vary advanced ages.

One of the major goals in developing animal models is to generate a model for AD, which is the most common neurodegenerative tauopathy. It seems feasible to develop mice with human APP, PS-1, PS-2, and ApoE transgenes to elucidate the mechanism of biochemical and clinicopathologic changes caused by genetic abnormalities in AD. Tg mice overexpressing APP with FAD mutations have shown diffuse and neuritic Aß plaques in the brain, but they have lacked tau-positive NFTs and neuron loss (180 ,181 and 182). Moreover, Tg mice with a PS-1 transgene linked to FAD

with and without a mutant APP transgene have developed no tau pathology (183). Taking together, these studies suggest that the generation of Aß plaques in mice is not sufficient to model AD pathology and to elucidate interactions between tau and Aß pathologies in the pathogenesis of AD. Thus, the generation of double Tg mouse overexpressing tau and mutant APP, and "humanized" genomic tau Tg mouse with tau-knockout mice to over-express only the six human tau isoforms may be better strategies to develop Tg mouse models recapitulating AD pathology.

CONCLUSION

Part of "94 - Tau Protein and Tauopathy "

Genetic, biochemical, and pathologic analyses have indicated that tau proteins play a central role in the pathogenesis of various neurodegenerative tauopathies, including AD. Transgenic experiments have also shown that overexpression of tau can cause neurodegenerative tauopathies in experimental animals. However, genetic abnormalities that influence tau phosphorylation and/or functions of tau in AD remain unsettled. The mutations on APP, PS-1, and PS-2, polymorphism of ApoE and other yet-to-be-identified genetic susceptibilities as well as environmental factors may promote the dysfunction of tau in AD. Among other tauopathies, genetic abnormalities have been confirmed only in FTDP-17 kindreds. The onset of PSP has been associated with an intronic polymorphism in the tau gene (184), but the major genetic factors that cause tau aggregation and neurodegeneration in many tauopathies are still to be elucidated.

REFERENCES

1. The National Institute on Aging, and Reagan Institute Working Group on Diagnostic Criteria for the Neuropathologic Assessment of Alzheimer's Disease: consensus recommendations for the postmortem diagnosis of Alzheimer's disease (review; 68 refs). *Neurobiol Aging* 1997;18:S1–S2.

2. Goedert M, Trojanowski JQ, Lee VMY. The neurofibrillary pathology of Alzheimer's disease. In: Prusiner SB, Rosenberg RN, Di Mauro S, et al, eds. *The molecular and genetic basis of neurological diseases*, second ed. Boston: Butterworth Heineman, 1997:613–627.

3. St. George-Hyslop P, Tanzi RE, Polinsky RJ, et al. The genetic defect causing familial Alzheimer's disease maps on chromosome 21. *Science* 1987;235:885–890.

4. St. George-Hyslop P, Haines J, Rogaev E, et al. Genetic evidence for a novel familial Alzheimer's disease locus on chromosome 14. *Nat Genet* 1992;2:330–334.

5. Levy-Lahad E, Wijsman EM, Nemens E, et al. A familial Alzheimer's disease locus on chromosome 1. *Science* 1995;269:970–973.

6. Levy-Lahad E, Wasco W, Poorkaj P, et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* 1995:269:973–977.

7. Sherrington R, Rogaev EI, Liang Y, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 1995;375:754–760.

8. Strittmatter WJ, Saunders AM, Schmechel D, et al. Apolipoprotein E: high avidity binding to beta-amyloid and increased freqency of type 4 allele in late-onset familial Alzheimer's disease. *Proc Natl Acad Sci USA* 1993;92:4725–4727.

9. Saunders AM, Strittmatter, WJ, Schmechel D, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 1993;43:1467–1472.

10. McKee AC, Kosik KS, Kowall NW. Neuritic pathology and dementia in Alzheimer's disease. *Ann Neurol* 1991;30:156–165.

11. Arriagada PV, Growdon JH, Hedley-Whyte ET, et al. Neurofibrillary tangles but not senile plaques parallel duration and severity of Alzheimer's disease. *Neurology* 1992;42:631–639.

12. Dickson DW, Crystal HA, Mattiace LA, et al. Identification of normal and pathologic aging in prospectively studied nondemented elderly humans. *Neurobiol Aging* 1992;13:179-189.

13. Neve RL, Robakis NK. Alzheimer's disease: a re-examination of the amyloid hypothesis. *Trends Neurosci* 1998;21:15–19.

14. Bancher C, Jellinger KA. Neurofibrillary tangle predominant form of senile dementia of Alzheimer type: a rare subtype in very old subjects. *Acta Neuropathol (Berl)* 1994;88:565–570.

15. Braak H, Braak E. Neuropathologic stageing of Alzheimer-related changes. *Acta Neuropathol (Berl)* 1991;82:239–259.

16. Giaccone G, Tagliavini F, Linoli G, et al. Down patients: extracellular preamyloid deposits precede neuritic degeneration and senile plaques. *Neurosci Lett* 1989;97:232–238.

17. Flament S, Delacourte A, Mann DM. Phosphorylation of tau proteins: a major event during the process of neurofibrillary degeneration. A comparative study between Alzheimer's disease and Down's syndrome. *Brain Res* 1990;516:15–19.

18. Roberts GW, Allsop D, Bruton C. The occult aftermath of boxing. *J Neurol Neurosurg Psychiatry* 1990;53:373–378.

19. Mendell JR, Sahenk Z, Gales T, et al. Amyloid filaments in inclusion body myositis: novel findings provide insight into nature of filaments. *Arch Neurol* 1991;48:1229–1234.

20. Askanas V, Engel WK, Alvarez RB. Light and electron microscopic localization of beta-amyloid protein in muscle biopsies of patients with inclusion-body myositis. *Am J Pathol* 1992;141:31–36.

21. Askanas V, Engel WK, Bilak M, et al. Twisted tubulofilaments of inclusion body myositis muscle resemble paired helical filaments of Alzheimer brain and contain hyperphosphorylated tau. *Am J Pathol* 1994;144:177–187.

22. Ghetti B, Tagliavini F, Masters CL, et al. Gerstmann-Straussler-Scheinkerdisease: II. Neurofibrillary tangles and plaques with PrP-amyloid coexist in an affected family. *Neurology* 1989;39:1453–1461.

23. Tagliavini F, Giaccone G, Prelli F, et al. Aa68 is a component of paired helical filaments of Gerstmann-Straussler-Sheinker disease, Indiana kindred. *Brain Res* 1993;616:325–329.

24. Hsiao K, Dlouhy SR, Farlow MR, et al. Mutant prion proteins in Gerstmann-Straussler-Scheinker disease with neurofibrillary tangles. *Nat Genet* 1992;1:68–71.

25. Ghetti B, Piccardo P, Spillantini MG, et al. Vascular variant of prion protein cerebral amyloidosis with tau-positive neurofibrillary tangles: the phenotype of the stop codon 145 mutation in PRNP. *Proc Natl Acad Sci USA* 1996;93:744–748.

26. Hirano A, Malamud N, Kurland LT. Parkinsonism dementia complex in endemic disease on the island of Guam—pathologic features. *Brain* 1961;84:662.

27. Hof PR, Nimchinsky EA, Buée-Scherrer V, et al. Amyotrophic lateral sclerosis/parkinsonism-dementia complex of Guam: quantitative neuropathology, immunohistochemical analysis of neuronal vulnerability, and comparison with related neurodegenerative disorders. *Acta Neuropathol (Berl)* 1994;88:397–404.

28. Umahara T, Hirano A, Kato S, et al. Demonstration of neurofibrillary tangles and neuropil thread-like structures in spinal cord white matter in parkinsonism-dementia complex on Guam and in Guamanian amyotrophic lateral sclerosis. *Acta Neuropatholo (Berl)* 1994;88:180–184.

29. Buée-Scherrer V, Buee L, Hof PR, et al. Neurofibrillary degeneration in amyotrophic lateral sclerosis/parkinsonism-dementia complex of Guam: immunochemical characterization of tau proteins. *Am J Pathol* 1995;146:924–932.

30. Braak H, Braak E. Argyrophilic grains: characteristic pathology of cerebral cortex in cases of adult onset dementia without Alzheimer changes. *Neurosci Lett* 1987;76:124–127.

31. Itagaki S, McGeer PL, Akiyama H, et al. A case of adult-onset dementia with argyrophilic grains. *Ann Neurol* 1989;26:685–689.

32. Perry G, Stwart D, Friedman R, et al. Filaments of Pick's bodies contain altered cytoskeletal elements. *Am J Pathol* 1987;127:559–568.

33. Murayama S, Mori H, Ihara Y, et al. Immunocytochemical and ultrastructural studies of Pick disease. *Ann Neurol* 1990;27:394–405.

34. Lieberman AP, Trojanowski JQ, Lee VMY, et al. Cognitive, neuroimaging, and pathologic studies in a patient with Pick disease. *Ann Neurol* 1998;43:259–265.

35. Paulus W, Selim M. Corticonigral degeneration with neuronal achromasia and basal neurofibrillary tangles. *Acta Neuropathol (Berl)* 1990;81:89–94.

36. Ksiezak-Reding H, Morgan K, Mattiace LA, et al. Ultrastructure and biochemical composition of paired helical filaments in corticobasal degeneration. *Am J Pathol* 1994;145:1496–1508.

37. Mori H, Nishimura M, Namba Y, et al. Corticobasal degeneration: a disease with widespread appearance of abnormal tau and neurofibrillary tangles, and its relation to progressive supranuclear palsy. *Acta Neuropathol (Berl)* 1994;88:113–121.

38. Wakabayashi K, Oyanagi K, Makifuchi T, et al. Corticobasal degeneration: etiopathologic significance of the cytoskeletal alterations. *Acta Neuropathol (Berl)* 1994;87:545–553.

39. Bancher C, Lassmann H, Budka H, et al. Neurofibrillary tangles in Alzheimer's disease and progressive supranuclear palsy: antigenic similarities and differences. Microtubule-associated protein tau antigenicity is prominent in all types of tangles. *Acta Neuropathol (Berl)* 1987;74:39–46.

40. Flament S, Delacourte A, Verny M, et al. Abnormal tau proteins in progressive supranuclear palsy: similarities and differences with the neurofibrillary degeneration of the Alzheimer type. *Acta Neuropathol (Berl)* 1991;81:591–596.

41. Schmidt ML, Huang R, Martin JA, et al. Neurofibrillary tangles in progressive supranuclear palsy contain the same tau epitopes identified in Alzheimer's disease PHF-tau. *J Neuropathol Exp Neurol* 1996;55:534–539.

42. Papp MI, Kahn JE, Lantos PL. Glial cytoplasmic inclusions in the CNS of patients with multiple system atrophy (striatonigral degeneration, olivopontocerebellar atrophy and Shy-Drager syndrome). *J Neurol Sci* 1989;94:79–100.

43. Auer IA, Schmidt ML, Lee VMY, et al. Paired helical filament tau (PHFtau) in Niemann-Pick type C disease is similar to PHFtau in Alzheimer's disease. *Acta Neuropathol (Berl)* 1995;90:547–551.

44. Love S, Bridges LR, Case CP. Neurofibrillary tangles in Niemann-Pick disease type C. *Brain* 1995;188:119–129.

45. Suzuki K, Parker CC, Pentchev PG, et al. Neurofibrillary tangles in Niemann-Pick disease type C. *Acta Neuropathol (Berl)* 1995;89:227–238.

46. Kosaka K. Diffuse neurofibrillary tangles with calcification: a new presenile dementia. *J Neurol Neurosurg Psychiatry* 1994;57:594–596.

47. Eidelberg D, Sotrel A, Joachim C, et al. Adult onset Hallervorden-Spatz disease with neurofibrillary pathology: a discrete clinico-pathologic entity. *Brain* 1987;110:993–1013.

48. McQuaid S, Allen IV, McMahon J, et al. Association of measles virus with neurofibrillary tangles in subacute sclerosing panencephalitis: a combined in situ hybridization and immunocytochemical investigation. *Neuropathol Appl Neurobiol* 1994;20:103–110.

49. Wilhelmsen KC, Lynch T, Pavlou E, et al. Localization of disinhibition-dementia-parkinsonism-amyotrophy complex to 17q21-22. *Am J Hum Gen* 1994;55:1159–1165.

50. Foster NL, Wilhelmsen K, Sima AA, et al. Frontotemporal dementia and parkinsonism linked to chromosome 17: a consensus conference. Conference participants. *Ann Neurol* 1997;41:706–715.

51. Papp MI, Lantos PL. Accumulation of tubular structures in oligodendroglial and neuronal cells as the basic alteration in multiple system atrophy. *J Neurol Sci* 1992;107:172–182.

52. Nishimura M, Namba Y, Ikeda K, et al. Glial fibrillary tangles with straight tubules in the brains of patients with progressive supranuclear palsy. *Neurosci Lett* 1992;143:35–38.

53. Iwatsubo T, Hasegawa M, Ihara Y. Neuronal and glial tau-positive inclusions in diverse neurologic diseases share common phosphorylation characteristics. *Acta Neuropathol* 1994;88:129–136.

54. Yamaoka LH, Welsh-Bohmer KA, Hulette CM, et al. Linkage of frontotemporal dementia to chromosome 17: clinical and neuropathologic characterization of phenotype. *Am J Hum Gen* 1996;59:1306–1312.

55. Wijker M, Wszolek ZK, Wolters EC, et al. Localization of the gene for rapidly progressive autosomal dominant parkinsonism and dementia with pollido-ponto-nigral degeneration to chromosome 17q21. *Hum Mol Genet* 1996;5:151–154.

56. Barker M, Kwok JB, Kucera S, et al. Localization of frontotemporal dementia with parkinsonism in an Australian kindred to chromosome 17q21-22. *Ann Neurol* 1997;42:794–798.

57. Froelich S, Basun H, Forsell C, et al. Mapping of a disease locus for familial rapidly progressive frontotemporal dementia to chromosome 17q12-21. *Am J Med Genet* 1997;74:380–385.

58. Murrell JR, Koller D, Foroud T, et al. Familial multiple-system tauopathy with presenile dementia is localized to chromosome 17. *Am J Hum Gen* 1997;61:1131–1138.

59. Wilhelmsen KC. Disinhibition-dementia-parkinsonism-amyotrophy complex (DDPAC) is a non-Alzheimer's frontotemporal dementia. *J Neural Transm Suppl* 1997;49:269–275.

60. Heutink P, Stevens M, Rizzu P, et al. Hereditary frontotemporal dementia is linked to chromosome 17q21-q22: a genetic and clinicopathologic study of three Dutch families. *Ann Neurol* 1997;41:150–159.

61. Spillantini MG, Bird TD, Ghetti B. Frontotemporal dementia and parkinsonism linked to chromosome 17: a new group of tauopathies. *Brain Pathol* 1998;8:387–402.

62. Feany MB, Dickson DW. Neurodegenerative disorders with extensive tau pathology: a comparative study and review. *Ann Neurol* 1996;40:139–148.

63. Chin SS-M, Goldman JE. Glial inclusions in CNS degenerative diseases. *J Neuropathol Exp Neurol* 1996;55:499–508.

64. Spillantini MG, Goedert M, Crowther RA, et al. Familial multiple system tauopathy with presenile dementia: a disease with abundant neuronal and glial tau filaments. *Proc Natl Acad Sci USA* 1997;94:4113–4118.

65. Fukumoto H, Asami-Okada A, Suzuki N, et al. Association of A beta 40-positive senile plaques with microglial cells in the brains of patients with Alzheimer's disease and in non-demented aged individuals. *Neurodegeneration* 1996;5:13–17.

66. Fukumoto H, Asami-Okada A, Suzuki N, et al. Amyloid beta protein deposition in normal aging has the same characteristics as that in Alzheimer's disease: predominance of A beta 42(43) and association of A beta 40 with cored plaques. *Am J Pathol* 1996;148:259–265.

67. Wisniewski K, Jervis GA, Moretz RC, Wisniewski HM. Alzheimer neurofibrillary tangles in diseases other than senile and presenile dementia. *Ann Neurol* 1979;5:288–294.

68. Braak H, Braak E. Frequency of stages of Alzheimer-related lesions in different age categories. *Neurobiol Aging* 1997;18:351–357.

69. Kidd M. Paired helical filaments in electron microscopy of Alzheimer's disease. *Nature* 1963;197:192–194.

70. Goedert M, Spillantini MG, Cairns NJ, et al. Tau proteins of Alzheimer paired helical filaments: abnormal phosphorylation of all six brain isoforms. *Neuron* 1992;8:159–168.

71. Crowther RA, Wischik CM. Image reconstruction of the Alzheimer paired helical filament. *EMBO J* 1985;4:3661–3665.

72. Spillantini MG, Crowther RA, Goedert M. Comparison of the neurofibrillary pathology in Alzheimer's disease and familial presenile dementia with tangles. *Acta Neuropathol* 1996;92:42–48.

73. Tagliavini F, Giaccone G, Prelli F, et al. A68 is a component of paired helical filaments of Gerstmann-Straussler-Scheinker disease, Indiana kindred. *Brain Res* 1993;616:325–328.

74. Probst A, Tolnay M, Langui D, et al. Pick disease: hyperphosphorylated tau protein segregates to the somatoaxonal compartment. *Acta Neuropathol* 1996;92:588–596.

75. Spillantini MG, Crowther RA, Kamphorst W et al. Tau pathology in two Dutch families with mutations in the microtubule-binding region of tau. *Am J Pathol* 1998;153:1359–1363.

76. Ksiezak-Reding H, Morgan K, Mattiace LA, et al. Ultrastructure and biochemical composition of paired helical filaments in corticobasal degeneration. *Am J Pathol* 1994;145:1496–1508.

77. Goedert M, Wischik CM, Crowther RA, et al. Cloning and sequencing of the cDNA encoding a core protein of the paired helical filament of Alzheimer disease: identification as the microtubule-associated protein tau. *Proc Natl Acad Sci USA* 1988;85:4051–4055.

78. Kondo J, Honda T, Mori H, et al. The carboxyl third of tau is tightly bound to paired helical filaments. *Neuron* 1988;1:827–834.

79. Wischik CM, Novak M, Thogersen HC, et al. Isolation of a gragment of tau derived from the core of the paired helical filament of Alzheimer disease. *Proc Natl Acad Sci USA* 1988;85:4506–4510.

80. Kosik KS, Orecchio LD, Binder L, et al. Epitopes that span the tau molecule are shared with paired helical filaments. *Neuron* 1988;1:817–825.

81. Lee VMY, Balin BJ, Otvos LJ, et al. A68: a major subunit of paired helical filaments and derivatized forms of normal tau. *Science* 1991;251:675–678.

82. Crowther RA. Straight and paired helical filaments in Alzheimer disease have a common structural unit. *Proc Natl Acad Sci USA* 1991;88:2288–2292.

83. Binder LI, Frankfurter A, Rebhun LI. The distribution of tau in the mammalian central nervous system. *J Cell Biol* 1985;101:1371–1378.

84. Couchie D, Mavilia C, Georgieff IS, et al. Primary structure of high molecular weight tau present in the peripheral nervous system. *Proc Natl Acad Sci USA* 1992;89:4378–4381.

85. Bernhardt R, Matus A. Light and electron microscopic studies of the distribution of microtubule-associated protein 2 in rat brain: a difference between dendritic and axonal cytoskeletons. *J Comp Neurol* 1984;226:203–221.

86. DeCamilli P, Miller PE, Navone F, et al. Distribution of microtubule-associated protein 2 in the nervous system of the rat studied by immunofluorescence. *Neuroscience* 1984;11:817–846.

87. Burton PR, Paige JL. Polarity of axoplasmic microtubules in the olfactory nerve of the frog. *Proc Natl Acad Sci USA* 1981;78:3269–3273.

88. Heidemann SR, Landers JM, Hamborg MA. Polarity orientation of axonal microtubules. *J Cell Biol* 1981;91:661–665.

89. Baas PW, Deitch JS, Black MM, et al. Polarity orientation of microtubules in hippocampal neurons: uniformity in the axon and nonuniformity in the dendrite. *Proc Natl Acad Sci USA* 1988;85:8335–8339.

90. Nadelhaft I. Microtubule densities and total numbers in selected axons of the crayfish abdominal nerve cord. *J Neurocytol* 1974;3:73–86.

91. Chen J, Kanai Y, Cowan NJ, et al. Projection domains of MAP2 and tau determine spacings between microtubules in dendrites and axons. *Nature* 1992;360:674–677.

92. LoPresti P, Szuchet S, Papasozomenos SC, et al. Functional implications for the microtubule-associated protein tau: localization in oligodendrocytes. *Proc Natl Acad Sci USA* 1995;92:10369–10373.

93. Shin RW, Iwaki T, Kitamoto T, et al. Hydrated autoclave pretreatment enhances tau immunoreactivity in formalin-fixed normal and Alzheimer's disease brain tissues. *Lab Invest* 1991;64:693–702.

94. Drubin DG, Feinstein SC, Shooter EM, et al. Nerve growth factor-induced neurite outgrowth in PC12 cells involves the coordinate induction of microtubule assembly and assembly-promoting factors. *J Cell Biol* 1985;101:1799–1807.

95. Drubin D, Kobayashi S, Kellogg D, et al. Regulation of microtubule protein levels during cellular morphogenesis in nerve growth factor-treated PC12 cells. *J Cell Biol* 1988;106:1583–1591.

96. Ferreira A, Busciglio J, Caceres A. Microtubule formation and neurite growth in cerebellar macroneurons which develop *in vitro*: evidence for the involvement of the microtubule-associated proteins, MAP-1a, HMW-MAP2 and Tau. *Brain Res Dev Brain Res* 1989;49:215–228.

97. Weingarten MD, Lockwood AH, Hwo SY, et al. A protein factor essential for microtubule assembly. *Proc Natl Acad Sci USA* 1975;72:1858–1862.

98. Drechsel DN, Hyman AA, Cobb MH, et al. Modulation of the dynamic instability of tubulin assembly by the microtubule-associated protein tau. *Mol Biol Cell* 1992;3:1141–1154.

99. Cáceres A, Kosik KS. Inhibition of neurite polarity by tau antisense oligonucleotides in primary cerebellar neurons. *Nature* 1990;343:461–463.

100. Cáceres A, Potrebic S, Kosik KS. The effect of tau antisence oligonucleotides on neurite formation of cultured cerebellar macroneurons. *J Neurosci* 1991;11:1515–1523.

101. Harada A, Oguchi K, Okabe S, et al. Altered microtubule organization in small-calibre axons of mice lacking tau protein. *Nature* 1994;369:488–491.

102. Ikegami S, Harada A, Hirokawa N. Muscle weakness, hyperactivity, and impairment in fear conditioning in tau-deficient mice. *Neurosci Lett* 2000;279:129–132.

103. Goedert M, Spillantini MG, Jakes R, et al. Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron* 1989;3:519–526.

104. Goedert M, Spillantini MG, Potier MC, et al. Cloning and sequencing of the cDNA encoding an isoform of microtubule-associated protein tau containing four tandem repeats: differential expression of tau protein mRNAs in human brain. *EMBO J* 1989;8:393–399.

105. Andreadis A, Brown WM, Kosik KS. Structure and novel exons of the human tau gene. *Biochemistry* 1992;31:10626–10633.

106. Goedert M, Jakes R. Expression of separate isoforms of human tau protein: correlation with the tau pattern in brain and effects on tubulin polymerization. *EMBO J* 1990;9:4225–4230.

107. Butner KA, Kirschner MW. Tau protein binds to microtubules through a flexible array of distributed weak sites. *J Cell Biol* 1991;115:717–730.

108. Goode BL, Feinstein SC. Identification of a novel microtubule binding and assembly domain in the developmentally regulated inter-repeat region of tau. *J Cell Biol* 1994;124:769–782.

109. Goode BL, Denis PE, Panda D, et al. Functional interactions between the proline-rich and repeat regions of tau enhance microtubule binding and assembly. *Mol Biol Cell* 1997;8:353–365.

110. Goedert M, Spillantini MG, Crowther RA. Cloning of a big tau microtubule-associated protein characteristic of the peripheral nervous system. *Proc Natl Acad Sci USA* 1992;89:1983–1987.

111. Hong M, Zhukareva V, Vogelsberg-Ragaglia V, et al. Mutation-specific functional impairments in distinct tau isoforms of hereditary FTDP-17. *Science* 1998;282:1914–1917.

112. Götz J, Probst A, Spillantini MG, et al. Somatodendritic localization and hyperphosphorylation of tau protein in transgenic mice expressing the longest human brain tau isoform. *EMBO J* 1995;14:1304–1313.

113. Hasegawa M, Morishima-Kawashima M, Takio K, et al. Protein sequence and mass spectrometric analyses of tau in the Alzheimer's disease brain. *J Biol Chem* 1992;267:17047–17054.

114. Morishima-Kawashima M, Hasegawa M, Takio K, et al. Proline-directed and non-proline-directed phosphorylation of PHF-tau. *J Biol Chem* 1995;270:823–829.

115. Hagestedt T, Lichtenberg B, Wille H, et al. Tau protein becomes long and stiff upon phosphorylation: correlation between paracrystalline structure and degree of phosphorylation. *J Cell Biol* 1989;109:1643–1651.

116. Biernat J, Gustke N, Drewes G, et al. Phosphorylation of Ser262 strongly reduces binding of tau to microtubules: distinction between PHF-like immunoreactivity and microtubule binding. *Neuron* 1993;11:153–163.

117. Bramblett GT, Goedert M, Jakes R, et al. Abnormal tau phosphorylation at Ser396 in Alzheimer's disease recapitulates development and contributes to reduced microtubule binding. *Neuron* 1993;10:1089–1099.

118. Yoshida H, Ihara Y. Tau in paired helical filaments is functionally distinct from fetal tau: assembly incompetence of paired helical filament-tau. *J Neurochem* 1993;61:1183–1186.

119. Biernat J, Gustke N, Drewes G, et al. Phosphorylation of Ser262 strongly reduces binding of tau to microtubules: distinction between PHF-like immunoreactivity and microtubule binding. *Neuron* 1993;11:153–163.

120. Lindwall G, Cole RD. Phosphorylation affects the ability of tau protein to promote microtubule assembly. *J Biol Chem* 1984;259:5301–5305.

121. Drechsel DN, Hyman AA, Cobb MH, et al. Modulation of the dynamic instability of tubulin assembly by the microtubule associated protein tau. *Mol Biol Cell* 1992;3:1141–1154.

122. Trinczek B, Biernat J, Baumann K, et al. Domains of tau protein, differential phosphorylation, and dynamic instability of microtubules. *Mol Biol Cell* 1995;6:1887–1902.

123. Drewes G, Lichtenberg-Kraag B, Doring F, et al. Mitogen activated protein (MAP) kinase transforms tau protein into an Alzheimer-like state. *EMBO J* 1992;11:2131–2138.

124. Drewes G, Mandelkow EM, Baumann K, et al. Dephosphorylation of tau protein and Alzheimer paired helical filaments by calcineurin and phosphatase-2A. *FEBS Lett* 1993;336:425–432.

125. Seubert P, Mawal-Dewan M, Barbour R, et al. Detection of phosphorylated Ser262 in fetal tau, adult tau, and paired helical filament tau. *J Biol Chem* 1995;270:18917–18922.

126. Goedert M, Jakes R, Crowther RA, et al. The abnormal phosphrylation of tau protein at Ser-202 in Alzheimer disease recapitulates phosphorylation during development. *Proc Natl Acad Sci USA* 1993;90:5066–5070.

127. Garver TD, Harris KA, Lehman RA, et al. Tau phosphorylation in human, primate, and rat brain: evidence that a pool of tau is highly phosphorylated in vivo and is rapidly dephosphorylated in vitro. *J Neurochem* 1994;63:2279–2287.

128. Matsuo ES, Shin RW, Billingsley ML, et al. Biopsy-derived adult human brain tau is phosphorylated at many of the same sites as Alzheimer's disease paired helical filament tau. *Neuron* 1994;13:989–1002.

129. Hoffmann R, Lee VMY, Leight S, et al. Unique Alzheimer's disease paired helical filament specific epitopes involve double phosphorylation at specific sites. *Biochemistry* 1997;36:8114–8124.

130. Zheng-Fischhofer Q, Biernat J, Mandelkow EM, et al. Sequential phosphorylation of tau by glycogen synthase kinase-3beta and protein kinase A at Thr212 and Ser214 generates the Alzheimer-specific epitope of antibody AT100 and requires a paired-helical filament-like conformation. *Eur J Biochem* 1998;252:542–552.

131. Ishiguro K, Takamatsu M, Tomizawa K, et al. Tau protein kinase I converts normal tau protein into A68-like component of paired helical filaments. *J Biol Chem* 1992;267:10897–10901.

132. Hosoi T, Uchiyama M, Okamura E, et al. Evidence for cdk5 as a major activity phosphorylating tau protein in porcine brain extract. *J Biochem* 1995;117:741–749.

133. Hong M, Lee VMY. Insulin and insulin-like growth factor-1 regulate tau phosphorylation in cultured human neurons. *J Biol Chem* 1997;272:19547–19553.

134. Woodgett JR. Molecular cloning and expression of glycogen synthase kinase-3/factor A. *EMBO J* 1990;9:2431–2438.

135. Tsai LH, Delalle I, Caviness VS Jr, et al. p35 is a neural-specific regulatory subunit of cyclin-dependent kinase 5. *Nature* 1994;371:419–423.

136. Lew J, Huang QQ, Qi Z, et al. A brain-specific activator of cyclin-dependent kinase 5. *Nature* 1994;371:423–426.

137. Patrick GN, Zukerberg L, Nikolic M, et al. Conversion of p35 to p25 deregulates Cdk5 activity and promotes neurodegeneration. *Nature* 1999;402:615–622.

138. Merrick SE, Demoise DC, Lee VMY. Site-specific dephosphorylation of tau protein at Ser202/Thr205 in response to microtubule depolymerization in cultured human neurons involves protein phosphatase 2A. *J Biol Chem* 1996;271:5589–5594.

139. Harris KA, Oyler GA, Doolittle GM, et al. Okadaic acid induces hyperphosphorylated forms of tau protein in human brain slices. *Ann Neurol* 1993;33:77–87.

140. Sontag E, Nunbhakdi-Craig V, Bloom GS, et al. A novel pool of protein phophatase 2A is associated with microtubules and is regulated during the cell cycle. *J Cell Biol* 1995;128:1131–1144.

140a. Vogelsberg-Ragaglia V, Schuck T, Trojanowski JQ, et al. PP2A mRNA expression is quantitatively decreased in Alzheimer's disease hippocampus. *Exp Neurol* 2001;168:402–412.

141. Buee SV, Hof PR, Buee L, et al. Hyperphosphorylated tau proteins differentiate corticobasal degeneration and Pick disease. *Acta Neuropathol (Berl)* 1996;91:351–359.

142. Sergeant N, David JP, Lefranc D, et al. Different distribution of phosphorylated tau protein isoforms in Alzheimer's and Pick diseases. *FEBS Letters*. 1997;412:578–582.

143. Delacourte A, Sergeant N, Wattez A, et al. Vulnerable neuronal subsets in Alzheimer's and Pick disease are distinguished by their tau isoform distribution and phosphorylation. *Ann Neurol* 1998;43:193–204.

144. Reed LA, Grabowski TJ, Schmidt ML, et al. Autosomal dominant dementia with widespread neurofibrillary tangles. *Ann Neurol* 1997;42:564–572.

145. Clark LN, Poorkaj P, Wszolek ZK, et al. Pathogenic implications of mutations in the tau gene in pallido-ponto-nigral degeneration and related chromosome 17-linked neurodegenerative disorders. *Proc Natl Acad Sci USA* 1998;95:13103–13107.

146. Bird TD, Wijsman EM, Nochlin D, et al. Chromosome 17 and hereditary dementia: linkage studies in three non-Alzheimer families and kindreds with late-onset FAD. *Neurology* 1997;48:949–954.

147. Wszolek ZK, Pfeiffer RF, Bhatt MH, et al. Rapidly progressive autosomal dominant parkinsonism and dementia with pallido-ponto-nigral degeneration. *Ann Neurol* 1992;32:312–320.

148. Lynch T, Sano M, Marder KS, et al. Clinical characteristics of a family with chromosome 17-linked disinhibition-dementia-parkinsonism-amyotrophy complex. *Neurology* 1994;44:1878–1884.

149. Spillantini MG, Bird TD, Ghetti B. Frontotemporal dementia and parkinsonism linked to chromosome 17: a new group of tauopathies. *Brain Pathol* 1998;8:387–402.

150. Reed LA, Schmidt ML, Wszolek ZK, et al. The neuropathology of a chromosome 17-linked autosomal dominant parkinsonism and dementia ("pallido-ponto-nigral degeneration"). *J Neuropathol Exp Neurol* 1998;57:588–610.

151. Spillantini MG, Murrell JR, Goedert M, et al. Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proc Natl Acad Sci USA* 1998;95:7737–7741.

152. Murrell JR, Koller D, Foroud T, et al. Familial multiple-system tauopathy with presenile dementia is localized to chromosome 17. *Am J Hum Gen* 1997;61:1131–1138.

153. Lendon CL, Lynch T, Norton J, et al. Hereditary dysphasic disinhibition dementia: a frontotemporal dementia linked to 17q21-22. *Neurology* 1998;50:1546–1555.

154. Hutton M, Lendon CL, Rizzu P, et al. Association of missense and 5′-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* 1998;393:702–705.

155. Poorkaj P, Bird TD, Wijsman E, et al. Tau is a candidate gene for chromosome 17 frontotemporal dementia. *Ann Neurol* 1998;43:815–825.

156. Bugiani O, Murrell JR, Giaccone G, et al. Frontotemporal dementia and corticobasal degeneration in a family with a P301S mutation in tau. *J Neuropathol Exp Neurol* 1999;58:667–677.

157. Sperfeld AD, Collatz MB, Baier H, et al. FTDP-17: an early-onset phenotype with parkinsonism and epileptic seizures caused by a novel mutation. *Ann Neurol* 1999;46:708–715.

158. Iijima M, Tabira T, Poorkaj P, et al. A distinct familial presenile dementia with a novel missense mutation in the tau gene. *Neuroreport* 1999;10:497–501.

159. D'Souza I, Poorkaj P, Hong M, et al. Missense and silent mutations in tau cause FTDP-17 by altering alternative splicing. *Proc Natl Acad Sci USA* 1999;96:55598—55603.

160. Murrell JR, Spillantini MG, Zolo P, et al. Tau gene mutation G389R causes a tauopathy with abundant pick body-like inclusions and axonal deposits. *J Neuropathol Exp Neurol* 1999;58:1207–1226.

161. Stanford PM, Halliday GM, Brooks WS, et al. Progressive supranuclear palsy pathology caused by a novel silent mutation in exon 10 of the tau gene: expansion of the disease phenotype caused by tau gene mutations. *Brain* 2000;123:880–893.

162. Rizzu P, Van Swieten JC, Joosse M, et al. High prevalence of mutations in the microtubule-associated protein tau in a population study of frontotemporal dementia in the Netherlands. *Am J Hum Gen* 1999;64:414–421.

163. Yasuda M, Takamatsu J, D'Souza I, et al. A novel mutation at position +12 in the intron following exon 10 of the tau gene in familial frontotemporal dementia (FTD-Kumamoto). *Ann Neurol* 2000;47:422–429.

164. Morris HR, Perez-Tur J, Janssen JC, et al. Mutation in the tau exon 10 splice site region in familial frontotemporal dementia. *Ann Neurol* 1999;45:270–271.

165. Goedert M, Spillantini MG, Crowther RA, et al. Tau gene mutation in familial progressive subcortical gliosis. *Nat Med* 1999;5:454–457.

166. Hasegawa M, Smith MJ, Goedert M. Tau proteins with FTDP-17 mutations have a reduced ability to promote microtubule assembly. *FEBS Lett* 1998;437:207–210.

167. Senapathy P, Shapiro MB, Harris NL. Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. *Meth Enzynmol* 1990;183:252–278.

168. Xu R, Teng J, Cooper TA. The cardiac troponin T alternative exon contains a novel purine-rich positive splicing element. *Mol Cell Biol* 1993;13:3660–3674.

169. Watakabe A, Tanaka K, Shimura Y. The role of exon sequences in splice site selection. *Genes Dev* 1993;7:407–418.

170. Laviqueur A, La BH, Komblihtt AR, et al. A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Gene Dev* 1993;7:2405–2417.

171. Cooper TA, Mattox W. The regulation of splice-site selection, and its role in human disease. *Am J Hum Gen* 1997;61:259–266.

172. Arrasate M, Perez M, Armas-Portela R, et al. Polymerization of tau peptides into fibrillar structures. The effect of FTDP-17 mutations. *FEBS Lett* 1999;446:199–202.

173. Nacharaju P, Lewis J, Easson C, et al. Accelerated filament formation from tau protein with specific FTDP-17 missense mutations *FEBS Lett* 1999;447:195–199.

174. Götz J, Probst A, Spillantini MG, et al. Somatodendritic localization and hyperphosphorylation of tau protein in transgenic mice expressing the longest human brain tau isoform. *EMBO J* 1995;14:1304–1313.

175. Brion JP, Tremp G, Octave JN. Transgenic expression of the shortest human tau affects its compartmentalization and its phosphorylation as in the pretangle stage of Alzheimer's disease. *Am J Pathol* 1999;154:255–270.

176. Ishihara T, Hong M, Zhang B, et al. Age-dependent emergence and progression of a tauopathy in transgenic mice overexpressing the shortest human tau isoform. *Neuron* 1999;24:751–762.

177. Spittaels K, Van den HC, Van Dorpe J, et al. Prominent axonopathy in the brain and spinal cord of transgenic mice overexpressing four-repeat human tau protein. *Am J Pathol* 1999;155:2153–21659.

178. Probst A, Götz J, Wiederhold KH, et al. Axonopathy and amyotrophy in mice transgenic for human four-repeat tau protein. *Acta Neuropathol (Berl)* 2000;99:469–481.

179. Duff K, Knight H, Refolo LM, et al. Characterization of pathology in transgenic mice over-expressing human genomic and cDNA tau transgenes. *Neurobiol Dis* 2000;7:87–98.

180. Games D, Adams D, Alessandrini R, et al. Alzheimer-type neuropathology in transgenic mice overexpressing V717F beta-amyloid precursor protein. *Nature* 1995;373:523–527.

181. Hsiao K, Chapman P, Nilsen S, et al. Correlative memory deficits, Abeta elevation, and amyloid plaques in transgenic mice. *Science* 1996;274:99–102.

182. Irizarry MC, Soriano F, McNamara M, et al. Abeta deposition is associated with neuropil changes, but not with overt neuronal loss in the human amyloid precursor protein V717F (PDAPP) transgenic mouse. *J Neurosci* 1997;17:7053–7059.

183. Price DL, Sisodia SS. Mutant genes in familial Alzheimer's disease and transgenic models. *Annu Rev Neurosci* 1998;21:479–505.

184. Conrad C, Andreadis A, Trojanowski JQ, et al. Genetic evidence for the involvement of tau in progressive supranuclear palsy. *Ann Neurol* 1997;41:277–281.

Section X
Substance Use Disorders

Charles P. O'Brien

Marian W. Fischman

Substance Use Disorders - Introduction

Although the prevailing view for many decades was that drug dependent patients simply suffered from character weakness, the persuasive data emerging from modern brain imaging techniques and the application of molecular biology methods to animal models of compulsive drug use indicate that this position is no longer tenable. The integration of a number of new technologies has allowed investigators to combine behavioral and neurobiological approaches to more completely evaluate multiple aspects of this difficult problem.

The following 16 chapters detail advances in the biology of substance use disorders, concentrating on those occurring during the 1990s, the decade of the brain. The section concentrates on advances most relevant to neuropsychopharmacology, integrating neurobiology, behavioral biology, and pharmacology. Knowledge of the pathophysiology of drug use disorders has greatly increased with the identification and cloning of receptors for the major drugs of abuse. There is also a much greater understanding of the brain circuits involved, including those common to different classes of drugs. The efficacy of treatment has also increased through the availability of effective medications for alcohol, heroin, and nicotine, as well as behavioral approaches used with cocaine abusers. Also, there is greater acceptance of the chronic disease model, which focuses on functional improvement as the realistic goal of treatment, rather than “cures.”

The terminology used in this section deserves some comment. There is general agreement that there are degrees of severity ranging from occasional drug use to a dangerous but moderately severe state called “abuse” in the American Psychiatric Association Diagnostic and Statistical Manual (DSM), to a severe compulsive state known as “dependence” or “addiction.” There is disagreement, however, on the usefulness of the term “addiction” to denote this severe state that occurs only in the minority of users who lose control and become compulsive drug users with a chronic relapsing clinical course. The DSM-III Revision Committee narrowly voted not to use the term “addiction” because of its prejudicial connotations, opting instead for “dependence.” This was continued in the current version, DSM-IV. The other point of view is that the term “dependence” creates confusion because it is already used to designate the state marked by drug-specific withdrawal symptoms that normally occur when regular drug use is abruptly terminated (“physical” dependence). Dependence also has a long-standing use as a personality disorder descriptor completely unrelated to drug use. Most important, patients with chronic pain receiving opiates often show signs of tolerance and withdrawal symptoms without any behavior that could be categorized as abuse. Physicians who are confused by “dependence” defined as a normal response and “dependence” as a disorder have been known to mistakenly withhold pain medication to “prevent addiction.” We have opted to use the DSM terminology for the title of this section, but the reader will find that there is some inconsistency among the chapters in the use of the terms “addiction” and “dependence” reflecting the current variance in the field over proper terminology.

95

Neurocircuitry of Addiction

Peter W. Kalivas

Peter W. Kalivas: Department of Physiology and Neuroscience, Medical University of South Carolina, Charleston, South Carolina

Addiction can be defined as drug-induced changes in the central nervous system (CNS) that produce maladaptive alterations in spontaneous behavior and in the behavioral response to readministration of that drug. Maladaptive behaviors include those identified as criteria for addiction in the DSM-IV. In general what most psychiatric metrics describe as addiction associated behaviors is the emergence of behaviors to obtain drug reward at the expense of engaging in behaviors to seek natural rewards, ranging from biological rewards such as sex to cultural rewards such as stable personal relationships. The substitution of drug reward for natural reward suggests that the neuropathology of addiction may reside in the same neural systems that mediate the detection and acquisition of natural rewards. This postulate forms a primary premise in the search for the neurobiological basis of addiction, and has revealed a circuit consisting of interconnections among limbic cortex, basal ganglia, and brainstem nuclei that is pathologically modified by repeated drug administration. The drug-induced changes in the structure and function of this circuit are progressive, and to some extent parallel the development of the behavioral characteristics of addiction.

Over the last decade neurobiologists have come to describe the behavioral transition to addiction as a druginduced neuroplastic process (1 ,2 and 3). In parallel with the development and expression of addictive behaviors, the neurobiology of these two components of the transition to addiction can be described as: (a) the sequence of molecular events that establish the neuroplastic changes leading to addiction, and (b) the neuroplastic changes themselves. Accordingly, a number of molecular neuroplastic alterations have been identified in the brain after repeated drug administration, and some of these appear to be important in the development and/or expression of addictive behaviors. However, the process of identifying drug-induced changes is accelerating and producing a deluge of information that is proving increasingly difficult to integrate into a coherent sequence of neuroplastic changes that mediate addiction. In part the difficulty of integration arises because the veracity of each molecular neuroplastic event must be tested in animal models of addiction; a labor-intensive process often employing imprecise tools for selective *in vivo* manipulation of cell biology. Moreover, this problem is compounded by the fact that the neuroadaptations associated with use of a drug of abuse are not entirely predicated on the pharmacology of the drug. It has become increasing clear that in addition to drug pharmacology, the environmental regulation of the drug-induced changes is a critical factor in the development and expression of addiction (4 ,5 and 6).

Although a critical role for environment in the manifestation of addicted behaviors has been obvious for decades in the behavioral evaluation of addiction in humans and animal models, only recently has the drug–environment interface been directly evaluated as critical in the cellular neuroadaptations mediating addiction (7 ,8). The realization that the neuroplasticity defining addiction arises as a collaboration between repeated drug administration and environmental associations is contributing to the construction of a template to help organize and integrate the emerging tide of data that describes the cellular adaptations potentially mediating addiction. Figure 95.1 illustrates how learning to associate environmental stimuli with the molecular action of a drug impinges on brain circuitry to elicit neuroplastic changes producing addiction. Figure 95.1 is used to organize this chapter into the three current arenas of investigation into the neurobiology of addiction: (a) the molecular site of action by a drug and the immediate consequences of drug administration to intracellular signaling and synaptic transmission; (b) the circuitry involved in learning and how this integrates with the molecular action of drugs of abuse; and (c) the overall circuitry of reward that contains both molecular sites of drug action and learning circuits, and forms the site of pathologic change mediating addiction.

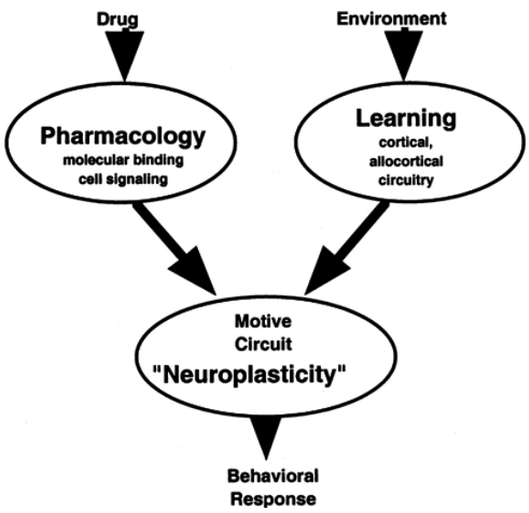


FIGURE 95.1. Collaboration between the pharmacology of a drug of abuse and environmental stimuli. Depiction of the impact that both the molecular binding site of a drug and environmental stimuli have on the neuroplasticity associated with drug addiction.

PHARMACOLOGY

Part of "95 - Neurocircuitry of Addiction "

Molecular Binding Site of the Drug

The molecular site of action and the immediate sequence of cellular events have been successfully decoded for many

drugs of abuse. Ethanol remains a notable exception where a number of candidate binding sites are currently being evaluated for a role in addiction. In contrast, for many drugs of abuse a reasonably accurate portrait of site of action has emerged. Over the last two decades a common direct or indirect action by drugs on dopamine transmission in the projection from the ventral mesencephalon (ventral tegmental area, VTA) to the forebrain has consistently been proposed and experimentally evaluated (8 ,9 ,10 and 11). This historic focus arose from studies employing lesions and the administration of dopamine receptor antagonists that converged on the potential importance of the dopamine projections in mediating natural, electrical, and drug reward. Figure 95.2 illustrates the dopamine projections most frequently evaluated in this regard and includes axon terminations in the amygdala, nucleus accumbens, prefrontal cortex (PFC), and ventral pallidum that arise from cell bodies in the VTA. Examination of the known binding sites for drugs of abuse reveals for some drugs a clear relationship between drug binding site and dopamine transmission (Fig. 95.3). Thus, amphetamine-like psychostimulants inhibit the binding of dopamine to the dopamine transporter and thereby elevate extracellular dopamine (12 ,13). Nicotine binds directly to acetylcholine nicotinic receptors on dopamine cells to increase the firing frequency of the mesocorticolimbic dopamine neurons (14). μ-opioid receptors are found in high density on presynaptic terminals of GABAergic interneurons and GABAergic axonal afferents to the VTA (15 ,16 ,17 ,18 and 19). Stimulation of these receptors inhibits the tonic and stimulated release of GABA, thereby increasing the firing frequency of the dopamine neurons. Alcohol modulates GABA-gated chloride conductance, as well as sodium and calcium conductances gated by glutamate (20 ,21). By promoting chloride conductance and inhibiting glutamate gated sodium and calcium fluxes alcohol is generally found to be inhibitory on neuronal activity. However, electrophysiologic and neurochemical studies reveal that alcohol increases dopamine cell firing and dopamine release in axon terminal fields via an action within the VTA (22 ,23 ,24 and 25). Thus,

it has been hypothesized that alcohol may act preferentially on GABAergic interneurons in the VTA to disinhibit dopamine cells. Moreover, recent evidence demonstrates that the capacity of alcohol to elevate dopamine transmission is blocked by opioid receptor antagonists (26). This indicates that alcohol is directly or indirectly increasing enkephalin release in the VTA, thereby disinhibiting dopamine cell firing. Indeed, this mechanism is likely to contribute to the therapeutic efficacy of opioid antagonist naltrexone in attenuating relapse in alcohol addiction (27 ,28).

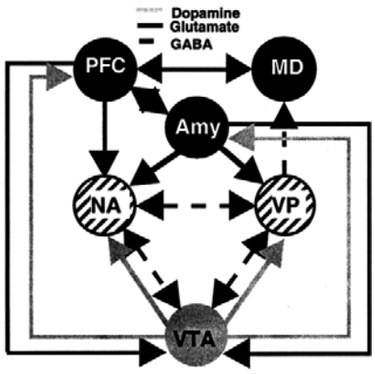


FIGURE 95.2. The circuit thought to be critical for mediating both natural and drug reward behavior. Amy, amygdala; MD, mediodorsal thalamus; NA, nucleus accumbens; PFC, prefrontal cortex; VP, ventral pallidum; VTA, ventral tegmental area.

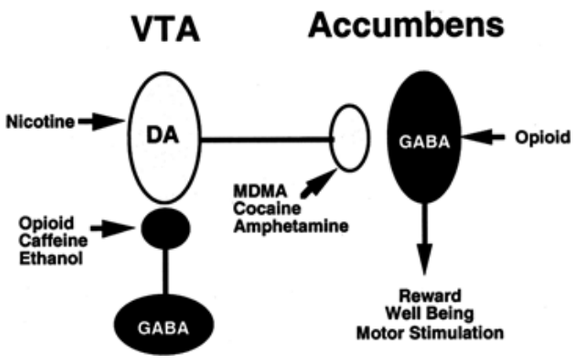


FIGURE 95.3. Sites of action by addictive drugs to augment mesoaccumbens dopamine transmission.

By inhibiting the NMDA subtype of the glutamate receptor, phencyclidine may decrease the activity of GABAergic interneurons in the VTA and thereby activate dopamine cell firing (29). Although these data strongly indicate a common action of drugs of abuse to stimulate dopamine transmission, notable exceptions exist. Of the major classes of drugs of abuse, the allosteric GABAA receptor agonists, including benzodiazepines and barbiturates, do not appear to stimulate dopamine cells or increase dopamine release (30). Also, although systemic administration of μ-opioids clearly activates dopamine transmission by disinhibiting dopamine cell activity, it is equally clear that stimulating μ-opioid receptors located postsynaptic to the dopamine projection in the nucleus accumbens is sufficient to elicit behaviors that are characteristic of addiction (31 ,32 ,33 and 34). Thus, the historic focus on dopamine transmission has identified an initial substrate for most drugs of abuse. However, the exceptions to this rule point clearly to the conclusion that, although an action on dopamine transmission may be a sufficient substrate to initiate drug-induced neuroplasticity relevant to addiction, other substrates are also sufficient.

Evaluating potential nondopaminergic substrates for initiating addiction-related neuroplasticity points most directly to modulating GABA transmission. Figure 95.2 illustrates that dopamine terminals in the nucleus accumbens synapse on GABAergic cells. Moreover, these GABAergic cells innervate GABAergic neurons in the VP. Indeed, Fig. 95.2 shows that there is a topographically organized interconnection among the nucleus accumbens, VP, and VTA that is GABAergic (35 ,36). Given the apparent importance of modulating dopamine transmission in this circuit in the neuroplasticity leading to addiction, it seems reasonable that directly modulating the GABAergic subcircuit with opioids or allosteric GABAergic agonists may also produce neuroplastic changes relevant to the development of addiction. Although substantial evidence exists for μ-opioids that this is true (31 ,32 ,33 and 34), this hypothesis has not yet been evaluated for benzodiazepines or barbiturates.

Pharmacologic Versus Physiologic Release of Dopamine

Studies in animal models measuring the *in vivo* release of dopamine with microdialysis reveal that natural rewarding or aversive stimuli increase dopamine release in the nucleus accumbens (6 ,37 ,38 ,39 ,40 ,41 and 42). However, the physiologic release of dopamine stimulated by environmental stimuli is of substantially less magnitude and duration than the pharmacologic release elicited by most drugs of abuse (22 ,43 ,44 ,45 ,46 and 47). The importance of excessive pharmacologic release of dopamine can be extrapolated from what is known about the behavioral situations that elicit physiologic activation of mesoaccumbens dopamine transmission. Dopamine release and cell firing are increased by the presentation of novel and motivationally relevant environmental stimuli (6 ,48). This has contributed to the notion that enhanced dopamine release in the nucleus accumbens signals the appearance of an important event that requires the creation and engagement of an adaptive behavioral strategy. This signal is supplied to numerous forebrain structures constituting the limbic cortex and basal ganglia (Fig. 95.2) and presumably initiates the recruitment of cortically derived memories and cognitive strategies, as well as motor output. In doing so, dopamine plays a role in initiating and establishing neuroplastic changes associated with developing behavioral strategies necessary to adapt to novel stimuli. One characteristic of the activation of dopamine transmission by environmental stimuli is that the release of dopamine diminishes with repeated exposure to the same stimulus as the organism establishes an adaptive behavioral response (40 ,49 ,50). Thus, dopamine contributes to the establishment of neuroplastic changes that mediate behavioral adaptation to relevant environmental stimuli, but may not be necessary for the expression of those behaviors. Rather, once a behavioral response to a stimulus has been established, the stimulus elicits the behavior via interactions among the limbic cortex, thalamus, and basal ganglia, with less involvement by mesocorticolimbic dopamine transmission (51 ,52). The transition from dopamine-dependent to behaviors less dependent on dopamine is illustrated in Fig. 95.4 as the transition in neural substrates mediating behavioral responding to a novel and familiar stimulus.

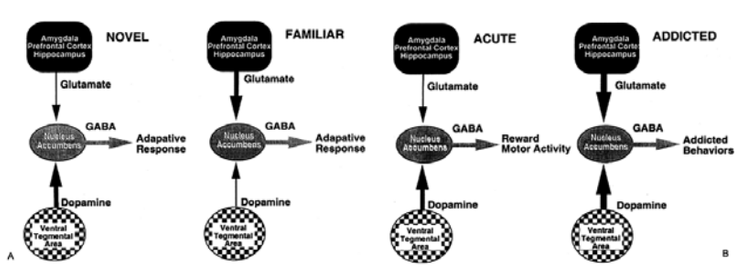


FIGURE 95.4. Changes in the role of mesencephalic versus cortical input to the nucleus accumbens with repeated exposure to natural motivational stimuli **(A)** and drug reward **(B)**. Both a novel stimulus and acute drug administration increase dopamine transmission. After repeated exposure the natural stimulus produces little or no increase in extracellular dopamine and the behavioral response to the stimulus arises primarily from cortical and allocortical interactions with the basal ganglia. Similarly, repeated drug use results in a recruitment of cortico-accumbens circuitry, but is also associated with augmented dopamine transmission. It is postulated that a portion of the pathology of addiction arises from both an overstimulated cortico-accumbens projection and a hyperresponsive mesoaccumbens dopamine projection.

Given the apparent role of dopamine release in establishing adaptive behavioral responses and the magnitude of dopamine release elicited by most drugs of abuse, it is perhaps not surprising that nonphysiologic neuroadaptations ensue, culminating in the maladaptive behaviors characteristic of addiction. Moreover, given that the physiologic release of dopamine is elicited by novel and salient environmental stimuli to facilitate the creation of enduring behavioral strategies, it is reasonable that the pharmacologic magnitude of drug-induced release would result in the addict developing potent learned associations between the dopamine-releasing drug experience and environmental stimuli. This strong association may contribute to the intense cravings that can be induced by environmental stimuli the addict has learned to associate with the drug experience.

LEARNING

Circuitry of Learning

A complete review of the current knowledge of the neural circuitry involved in learning and memory is beyond the scope of this chapter (53 ,54 ,55 and 56); however, it is clear that cortical regions receiving dopaminergic axon terminations, such as the prefrontal and anterior cingulate cortex and allocortical nuclei (including the hippocampus and amygdala) play distinct yet integrated roles in memory and learning. As illustrated in Fig. 95.2 , these cortical and allocortical nuclei have relatively dense glutamatergic projections to the nucleus accumbens and synapse on the same neurons that receive dopaminergic afferents from the VTA (57 ,58). This arrangement provides an anatomic substrate whereby learned information can be integrated and guide adaptive behavioral responding to environmental stimuli. Whether or not these glutamatergic afferents to the nucleus accumbens influence the behavioral effects of drugs of abuse and play a role in the neuroplasticity associated with the development and expression of addictive behaviors is a subject of much recent interest (13 ,59). For example, antagonists of the NMDA subtype of glutamate receptors block the development of behavioral neuroadaptations (e.g., behavioral sensitization of motor activity) to the repeated administration of most drugs of abuse, including alcohol, psychostimulants, and opioids (59 ,60). Moreover, it has been shown that drug craving induced in addicts by environmental stimuli previously paired with the drug experience is associated with metabolic activation of the anterior cingulate and amygdala (4 ,61 ,62). The apparent recruitment of glutamatergic cortical regions by a drug-associated stimulus concurrent with the manifestation of behavioral characteristics of addiction argues for a transition from primarily dopamine-dependent behaviors elicited by acute drug administration to primarily glutamate-dependent behaviors produced by drug or environmental stimuli in drug addicts. Thus, as illustrated in Fig. 95.4 , the development of addiction may consist in part of a transition from dopamine- to glutamate-dependent behaviors akin to what occurs during adaptation to biologically relevant environmental stimuli, such as food, sex, and stress. However, given the initial pharmacologic activation of dopamine transmission produced by most drugs of abuse, the ensuing recruitment of cortical glutamatergic neurons may also contain nonphysiologic constituents that contribute to the pathology of addiction. Moreover, repeated administration of most drugs of abuse, including psychostimulants, opioids, alcohol, and nicotine, increases the capacity of subsequent drug administration to release dopamine in the nucleus accumbens (2 ,63 ,65). This neurochemical reverse-tolerance or sensitization is opposite to what typically occurs with repeated exposure to motivationally relevant environmental stimuli, such as food, sex, and stress, and probably also contributes to the expression of addiction-related behaviors. Of note, the indirect measurement of dopamine release using methylphenidate-induced displacement of radiolabeled D2 receptor antagonist revealed an apparent decrease in dopamine release in cocaine addicts compared with control subjects (66). Assuming the veracity of D2 ligand displacement as an *in vivo*

measure of dopamine release (67), this indicates a potentially important distinction between data in animal models and human cocaine addicts.

NEUROPLASTICITY

Part of "95 - Neurocircuitry of Addiction "

Motivational Circuitry as a Site of Addiction Pathologies

The circuit shown in Fig. 95.2 contains interconnections among mesencephalic dopamine, cortical glutamate, and subcortical (basal ganglia) GABA neurons that can permit the transition from dopamine- to glutamate-dependent behaviors. This circuit has been a focus for research aimed at determining how the neuroplastic changes produced in various nuclei by repeated drug action at a molecular level are integrated with environmental stimuli to form the maladaptive behaviors characteristic of addiction. Studies in this regard are most numerous for psychostimulants; the findings employing this class of addictive drugs provide the primary buttresses of the model. However, emerging data with other drugs of abuse are generally consistent with this view; moreover, there is an abundance of behavioral data supporting the possibility that the neuroplastic changes elicited by repeated administration of one drug impinge on circuitry shared with a different drug of abuse. For example, the repeated administration of some drugs of abuse promotes the acquisition of addiction-related behaviors to another mechanistically similar drug of abuse (68 ,69). Similarly, a number of overlapping cellular neuroadaptations, such as changes in G proteins, tyrosine hydroxylase, and certain immediate early genes, have been identified after repeated exposure to different drugs of abuse (70 ,71 ,72 and 73). Also, a number of studies indicate behavioral and neurochemical overlap between the effects of repeated drug administration and motivationally relevant environmental stimuli (e.g., cross-sensitization and cross-tolerance to stress) (65 ,69 ,74 ,75).

Drug-Induced Changes in Presynaptic and Postsynaptic Dopamine Transmission

As described in the preceding, most addictive drugs produce significant elevations in cortical and subcortical dopamine transmission, and repeated drug-induced dopamine release causes enduring alterations in presynaptic and postsynaptic dopamine transmission. These changes have been characterized using animal models of addiction, including behavioral sensitization studies involving repeated investigator administered drug, as well as studies in which the acquisition, maintenance, and reinstatement of drug self-administration is measured (9 ,76 ,77). The majority of studies have endeavored to identify changes in nucleus accumbens dopamine transmission. However, in light of evidence implicating cortical and allocortical brain regions in addiction, more recent studies have come to focus dopamine innervation of the prefrontal cortex and amygdala. In general, in animal models the repeated administration of drugs of abuse, including amphetamine-like psychostimulants, alcohol, nicotine, and opioids, produces an enduring increase in the capacity of subsequent drug injection or an environmental stimulus, such as stress or sex, to release dopamine in the nucleus accumbens (13 ,37 ,63 ,64 and 65). The enhanced releasability may result in part from increased excitability of dopamine cells (24 ,78 ,79 and 80), but also clearly involves increased presynaptic release, as evidence by the enhanced release of dopamine from synaptosomes or after direct drug administration into the accumbens (81 ,82). Following repeated amphetamine and cocaine administration the elevated presynaptic release of dopamine arises from enhanced presynaptic calcium signaling through calmodulin kinase II (83 ,84 and 85). A number of enduring changes in postsynaptic dopamine signaling have also been identified. Most of these changes can be traced to alterations in gene expression following stimulation of the D1 receptor-signaling cascade (2 ,86). Interestingly, among the genes showing changed expression are protein products, including preprodynorphin, NAC-1, and delta-Fos-B, that appear to dampen enhanced dopamine transmission and/or the expression of behaviors potentially related to addiction, such as motor sensitization and conditioned place preference (87 ,88 and 89). This poses the likelihood that many changes produced by repeated drug administration are not necessarily promoting addictive behaviors but may constitute homeostatic alterations to minimize the impact of pathologic neuroadaptations elicited by repeated drug injection. In addition to changes in gene expression in the nucleus accumbens that affect dopamine transmission, repeated administration of psychostimulants also increases or decreases the synthesis of proteins affecting glutamate transmission that may alter postsynaptic corticofugal neurotransmission (see the following). This includes drug-induced changes in glutamate receptor subunits, as well as proteins involved in the trafficking and signaling of glutamate receptors (13 ,59 ,90 ,91).

In addition to the nucleus accumbens, more recent studies have examined the effect of repeated drug administration on dopamine transmission in the PFC. In general, repeated systemic drug administration reduces the releasability of dopamine in the PFC. Such a blunting of stimulus- or drug-induced dopamine transmission in the PFC has been most clearly shown after repeated cocaine and phencyclidine administration (50 ,92). Further implicating reduced cortical dopamine transmission, destruction of dopamine afferents to or blockade of dopamine receptors in the PFC promotes behavioral sensitization and drug self-administration (93 ,94 and 95). Moreover, the blunting of PFC dopamine transmission may be related to the cognitive dysfunction and increased impulsivity often associated with drug addiction (96).

The effect of repeated drug administration has also been evaluated in the amygdala, and in contrast with the PFC an

augmentation or no change in drug-stimulated dopamine release after repeated drug administration has been observed (97 ,98). Although the effect of repeated drug administration on dopamine release in the PFC and amygdala differs, the changes in dopamine transmission in both structures may contribute to expression of addictive behavior. Although reduced dopamine transmission in the PFC may contribute to cognitive dysfunction in some addicts (see the preceding), enhanced release of dopamine in the amygdala may contribute to the strong learned associations made between drug taking and environmental stimuli that have been repeatedly paired with drug taking. Supporting this possibility, stimulation of dopamine transmission in the amygdala increases stimulus-reward associations, whereas the inhibition of dopamine transmission blocks cue-induced reinstatement of lever pressing for cocaine (99 ,100). Thus, the sensitized release of dopamine in the amygdala may contribute to cue-induced relapse, a conclusion directly supported by imaging studies showing that drug craving in addicts is associated with metabolic activation of the amygdala (4 ,61 ,62), and in experimental animal models where drug-associated cues have been shown to increase glucose uptake or c-fos synthesis in the amygdala (101 ,102).

Prefrontal and Anterior Cingulate Cortex

As outlined above, there is an enduring change in dopamine transmission in the PFC associated with repeated administration of some drugs of abuse. Efforts to deduce the effect of dopamine transmission in the PFC have generally revealed that the actions of dopamine are state-dependent. This is reflected by distinct electrophysiologic effects of dopamine when pyramidal cells in the cortex are relatively active or inactive (e.g., up or down state) and by a biphasic dose–response effect of dopamine agonists and antagonists on cognition (103 ,104). Although the effects of dopamine are state-dependent, dopamine clearly has the capacity to modulate excitatory glutamatergic projections to the nucleus accumbens and VTA (105 ,106 and 107). Thus, the long-term alterations in PFC dopamine transmission by repeated drug use change corticofugal glutamate transmission.

A variety of emerging data indicate that glutamatergic projections to the nucleus accumbens and VTA may be altered by repeated drug administration and that this is consequential in behaviors associated with addiction. Thus, repeated cocaine administration elicits a reduction in basal extracellular glutamate in the nucleus accumbens, whereas subsequent drug exposure results in enhanced glutamate release in both the nucleus accumbens and VTA that depends on dopamine receptor stimulation (108 ,109 and 110). Moreover, the blockade of glutamate receptors in the nucleus accumbens inhibits the expression of behavioral sensitization to cocaine, as well as the capacity of systemic cocaine to elicit drug craving using a reinstatement model of craving in rodents (108 ,111). Also, lesions of the PFC have been shown to block behavioral sensitization to cocaine and morphine, as well as increase extracellular glutamate produced by cocaine in cocaine-pretreated subjects (13 ,112). Taken together with the imaging data in human cocaine addicts showing an association between cocaine craving and metabolic activity in the anterior cingulate cortex (4 ,61), the data to date point toward a critical role in addiction for neuroadaptations in glutamate transmission in the projection from the PFC to the mesoaccumbens dopamine system. Moreover, recent studies may indicate a prepotent role in cocaine addiction for glutamate over dopamine in the nucleus accumbens, although a role for this projection has been more difficult to demonstrate for some drugs of abuse, notably amphetamine (59).

Amygdala-Thalamo-Cortical Circuit

There is substantial evidence for a role by the amygdala in mediating the conditioning of behavior to obtain a natural or drug reward (99 ,113 ,114). Thus, pharmacologically inhibiting neurotransmission within the amygdala prevents the capacity of a cocaine cue to elicit drug-seeking behavior in rodents, and amygdala lesions blunt the expression of conditioned place preference produced by many drugs of abuse. There are two likely connections from the amygdala that are mediating the relationship between learned associations and rewarding stimuli. The first is an interconnected relationship with the amygdala, mediodorsal thalamus, and PFC. Various data suggest that this subcircuit is critical for the formation of learned associations with rewarding stimuli (115). The second connection is among the amygdala, VTA, and nucleus accumbens. Although this subcircuit is important in unconditioned responses to most drugs of abuse (see the preceding), many studies indicate that it is critical for the expression of behaviors elicited by conditioned stimuli to obtain reward, as well as cue-induced drug relapse (114). Although both circuits probably contribute to stimulus-induced drug relapse, the relative role of each may depend on the type of stimulus. For example, relapse precipitated by a low dose of drug, especially a drug known to activate dopamine transmission, is likely to involve the dopamine projection from the VTA to the amygdala, whereas cue-induced relapse may be more involved with the influence the amygdala brings to bear on the thalamocortical projection (99 ,111).

CONCLUSION AND FUTURE DIRECTIONS

Part of "95 - Neurocircuitry of Addiction "

The last decade has witnessed the emergence of a number of seminal developments in our understanding of the circuitry mediating addiction. Primary among these is that there exists substantial overlap in the circuits mediating behavioral responses to non-drug and drug reward (Fig. 95.2 and Fig 95.4), and that the pathology of addiction arises to a great extent

from neuroplasticity within this common circuit. The second critical development is that neuroplasticity to drug reward arises from an interaction between the molecular site of drug action and environmental stimuli associated with the effect of the drug. In this way, cortical and allocortical circuitry important in learning and memory is recruited and undergoes neuroplastic alterations in response to drugs of abuse. The studies outlined in this chapter describe changes in neurotransmission and cell signaling that may be critical in changing the functional status of the circuit illustrated in Fig. 95.2 . Accompanying these intracellular and neurochemical neuroadaptations, it is also important to note the recent studies of Robinson and coworkers (116 ,117), which show that akin to normal learning, repeated exposure to psychostimulants or opioids produces morphologic changes in dendritic branching and synaptic connectivity in many nuclei (Fig. 95.2). Although neurobiologists are well engrossed in identifying neuroplastic changes in various nuclei (Fig. 95.2), there is a lag in understanding how these alterations result in a change in the function of the circuit as a whole. The circuit shown in Fig. 95.2 is highly interconnected and topographically organized (35 ,36). Thus, a change in the function of a cell group in one nucleus is likely to impact on the status of neurons in other nuclei. Predicting and experimentally evaluating these secondary interactions and how they determine the appearance of addictive behavior will be the arena of study over the next decade. This arena can be approached from molecular technologies at one end and imaging studies at the other. The imaging experiments help to define functional subcircuits within the overall reward and learning circuitry, whereas the molecular studies can identify specific cellular neuroadaptations. Using pharmacologic and genetic manipulations these specific cellular neuroadaptations can be evaluated *in vivo* animal models of addiction to determine the relevance of the changes in addictive behaviors and on cellular adaptations in interconnected nuclei within the reward and learning circuits.

As an initial foray into integrative thinking it is becoming clear that there is a distinction between the effects of acute drug and repeated administration on the circuit shown in Fig. 95.2 . In the analysis of the literature outlined in the preceding, this distinction was characterized as a transition from dopamine-dependent effects to behaviors more dependent on cortical glutamate transmission. Although this is an oversimplification based on too few experiments using too few classes of drugs of abuse, it is similar to the physiologic neuroadaptive processes associated with the development of habitual behavioral responses to previously novel motivationally relevant natural stimuli (Fig. 95.4). Moreover, involvement of cortical and allocortical brain regions in addiction imparts a mandate that future researchers to consider the neurobiology of learning and memory as an equal partner with drug neuropharmacology in developing effective therapy for addiction.

ACKNOWLEDGMENT

Part of "95 - Neurocircuitry of Addiction "

This work was supported in part by USPHS grants MH-40817, DA-11809, and DA-03906.

REFERENCES

1. Koob GF, Nestler EJ. The neurobiology of drug addiction. *J Neuropsychiatry Clin Neurosci* 1997;9:482–497.

2. White FJ, Kalivas PW. Neuroadaptations involved in amphetamine and cocaine addiction. *Drug Alcohol Depend* 1998;51:141–154.

3. Nestler EJ, Hope BT, Widnell KL. Drug addiction: a model for the molecular basis of neural plasticity. *Neuron* 1993;11:995–1006.

4. Childress AR, Mozley PD, McElgin W, et al. Limbic activation during cue-induced cocaine craving. *Am J Psychiatry* 1999;156:11–18.

5. O'Brien CP, Childress AR, McLellan AT, et al. Classical conditioning in drug-dependent humans. *Ann NY Acad Sci* 1992;654:400–415.

6. Berridge KC, Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Rev* 1998;28:309–369.

7. Kalivas P, Cornish J, Ghasemzadeh M. Cocaine craving and paranoia: a combination of pharmacology and learning. *Psychiatry Ann* 1998;28:569–576.

8. Berke J, Hyman S. Addiction, dopamine, and the molecular mechanisms of memory. *Neuron* 2000;25:515–532.

9. Robinson TE, Becker JB. Enduring changes in brain and behavior produced by chronic amphetamine administration: a review and evaluation of animal models of amphetamine psychosis. *Brain Res Rev* 1986;11:157–198.

10. Wise RA, Rompre PP. Brain dopamine and reward. *Annu Rev Psychol* 1989;40:191–215.

11. Koob GF, Goeders NE. Neuroanatomical substrates of drug self-administration. In: Cooper SJ, Lieberman JM, eds. *The neuropharmacological basis of reward*. Clarendon, UK: Oxford Press, 1989:214–263.

12. Reith MEA, Meisler BE, Sershen H, et al. Structural requirements for cocaine congeners to interact with dopamine and serotonin uptake sites in mouse brain and to induce stereotyped behavior. *Biochem Pharmacol* 1986;35:1123–1129.

13. Pierce RC, Kalivas PW. A circuitry model of the expression of behavioral sensitization to amphetamine-like psychostimulants. *Brain Res Rev* 1997;25:192–216.

14. Mereu G, Yoon K-WP, Boi V, et al. Preferential stimulation of ventral tegmental area dopaminergic neurons by nicotine. *Eur J Pharmacol* 1987;141:395–399.

15. Gysling K, Wang RY. Morphine-induced activation of A10 dopamine neurons in the rat brain. *Brain Res* 1983;277:119–127.

16. Dilts RP, Kalivas PW. Autoradiographic localization of mu-opioid and neurotensin receptors within the mesolimbic dopamine system. *Brain Res* 1989;488:311–327.

17. Westfall TC, Mereu G, Vickery L, et al. Regulation by nicotine of midbrain dopamine neurons. *Prog Brain Res* 1989;79:173–185.

18. Clarke PBS, Pert A. Autoradiographic evidence for nicotine receptors on nigrostriatal and mesolimbic dopaminergic neurons. *Brain Res* 1985;348:355–358.

19. Nisell M, Nomikos GG, Svensson TH. Systemic nicotineinduced dopamine release in the rat nucleus accumbens is regulated by nicotinic receptors in the ventral tegmental area. *Synapse* 1994;16:36–44.

20. Lovinger DM. Alcohols and neurotransmitter gated ion channels: past, present and future. *Naunyn-Schmiedeberg's Arch Pharmacol* 1997;356:267–282.

21. Fadda F, Rossetti Z. Chronic ethanol consumption: from neuroadaptation to neurodegeneration. *Progr Neurobiol* 1998;56:385–431.

22. Yim HJ, Schallert T, Randall PK, et al. Comparison to local and systemic ethanol effects on extracellular dopamine concentration in rat nucleus accumbens by microdialysis. *Alcohol Clin Exp Res* 1998;22(2):367–374.

23. Bunney E, Appel S, Brodie M. Cocaine potentiates ethanol-induced excitation of dopaminergic reward neurons in the ventral tegmental area. *J Pharmacol Exp Ther* 2000;293(2):383–389.

24. Bailey CP, Manely SJ, Watson S, et al. Chronic ethanol administration alters activity in ventral tegmental area neurons after cessation of withdrawal hyperexcitability. *Brain Res* 1998;803(1,2):144–152.

25. Brodie M, McElvain M, Bradshaw Bunney E, et al. Pharmacological reduction of small conductance calcium-activated potassium current (sK) potentiates the excitatory effect of ethanol on ventral tegmental area dopamine neurons. *J Pharmacol Exp Ther* 1999;290(1):325–333.

26. Gonzales RA, Weiss F. Suppression of ethanol-reinforced behavior by naltrexone is associated with attenuation of the ethanol-induced increase in dialysate dopamine levels in the nucleus accumbens. *J Neurosci* 1998;18(24):10663–10671.

27. Weinrieb R, O'Brien C. Naltrexone in the treatment of alcoholism. *Annu Rev Med* 1997;48:477–487.

28. Swift R. Drug therapy for alcohol dependence. *N Engl J Med* 1999;340:1482–1490.

29. French ED, Ceci A. Non-competitive N-methyl-D-aspartate antagonists are potent activators of ventral tegmental A10 dopamine neurons. *Neurosci Lett* 1990;119:159–162.

30. Zhang H, Weng X, Rosenberg HC. Characterization of substantia nigra pars reticulata neurons based on response to iontophoretically applied GABA and flurazepam. *Life Sci* 1993;53:1911–1919.

31. Olds ME. Reinforcing effects of morphine in the nucleus accumbens. *Brain Res* 1982;237:429–440.

32. Ettenberg A, Pettit HO, Bloom FE, et al. Heroin and cocaine intravenous self-administration in rats: mediation by separate neural systems. *Psychopharmacology* 1982;78:204–209.

33. Svingos AL, Moriwaki A, Wang JB, et al. mu-opioid receptors are localized to extrasynaptic plasma membranes of GABAergic neurons and their targets in the rat nucleus accumbens. *J Neurosci* 1997;17:2585–2594.

34. David V, Cazala P. Anatomical and pharmacological specificity of the rewarding effect elicited by microinjections of morphine into the nucleus accumbens of mice. *Psychopharmacology* 2000;150(1):24–34.

35. Mogenson GJ, Brudzynski SM, Wu M, et al. From motivation to action: A review of dopaminergic regulation of limbic-nucleus accumbens-pedunculopontine nucleus circuitries involved in limbic-motor integration. In: Kalivas PW, Barnes CD, eds. *Limbic motor circuits and neuropsychiatry*. Boca Raton, FL: CRC Press, 1993:193–236.

36. Kalivas PW, Churchill L, Klitenick MA. The circuitry mediating the translation of motivational stimuli into adaptive motor responses. In: Kalivas PW, Barnes CD, eds. *Limbic motor circuits and neuropsychiatry*. Boca Raton, FL: CRC Press, 1993:237–287.

37. Mitchell JB, Gratton A. Opioid modulation and sensitization of dopamine release elicited by sexually relevant stimuli: a high-speed chronoamperometric study in freely behaving rats. *Brain Res* 1990;551:20–27.

38. Tabor MT, Fibiger HC. Feeding-evoked dopamine release in the nucleus accumbens: regulation by glutamatergic mechanisms. *Neuroscience* 1997;76:1105–1112.

39. Mermelstein PG, Becker JB. Increased extracellular dopamine in the nucleus accumbens and striatum of the female rat during paced copulatory behavior. *Behav Neurosci* 1995;109:354–365.

40. Cabib S, Puglishi-Allegra S. Different effects of repeated stressful experiences on mesocortical and mesolimbic dopamine metabolism. *Neuroscience* 1996;73:375–380.

41. Doherty MD, Gratton A. NMDA receptors in nucleus accumbens modulate stress-induced dopamine release in nucleus accumbens and ventral tegmental area. *Synapse* 1997;26:225–234.

42. Kalivas PW, Duffy P. Selective activation of dopamine transmission in the shell of the nucleus accumbens by stress. *Brain Res* 1995;675:325–328.

43. Richardson NR, Gratton A. Behavior-relevant changes in nucleus accumbens dopamine transmission elicited by food reinforcement: an electrochemical study in rat. *J Neurosci* 1996;16:8160–8169.

44. Kuczenski RSD. Dynamic changes in sensitivity occur during the acute response to cocaine and methylphenidate. *Psychopharmacology* 1999;147(1):96–103.

45. Marshall DL, Redfern PH, Wonnacott S. Presynaptic nicotinic modulation of dopamine release in the three ascending pathways studied by in vivo microdialysis: comparison of naive and chronic nicotine-treated rats. *J Neurochem* 1997;68:1511–1519.

46. Kalivas PW, Duffy P, White S. MDMA elicits behavioral and neurochemical sensitization in rats. *Neuropsychopharmacology* 1998;18:469–479.

47. Tanda G, Pontieri FE, Di Chiara G. Cannabinoid and heroin activation of mesolimbic dopamine transmission by a common m1 opioid receptor mechanism. *Science* 1997;276:2048–2050.

48. Schultz W. Predictive reward signal of dopamine neurons. *Am J Physiol* 1998;80:1–27.

49. Deutch AY, Tam S-Y, Roth RH. Footshock and conditioned stress increase in 3,4-dihydroxyphenylacetic acid (DOPAC) in the ventral tegmental area but not substantia nigra. *Brain Res* 1985;333:143–146.

50. Sorg BA, Davidson DL, Kalivas PW, et al. Repeated daily cocaine alters subsequent cocaine-induced increase of extracellular dopamine in the medial prefrontal cortex. *J Pharmacol Exp Ther* 1997;281:54–61.

51. Graybiel A. The basal ganglia and chunking of action repertoires. *Neurobiol Learning Memory* 1998;70:119–136.

52. Jog M, Kubota Y, Connolly C, et al. Building neural representations of habits. *Science* 1999;286:1745–1749.

53. Silva A, Geise K, Fedorov N, et al. Molecular, cellular and anatomical substrates of place learning. *Neurobiol Learning Memory* 1998;70:44–61.

54. McGaugh J. Memory—a century of consolidation. *Science* 2000;287:248–251.

55. Becker J, Morris R. Working memory(s). *Brain Cogn* 1999;41:1–8.

56. Eichenbaum H. The hippocampus and mechanisms of declarative memory. *Behav Brain Res* 1999;103:123–139.

57. Sesack SR, Pickel VM. In the medial nucleus accumbens, hippocampal and catecholaminergic terminals converge on spiny neurons and are in apposition to each other. *Brain Res* 1990;527:266–272.

58. Sesack SR, Pickel VM. Prefrontal cortical efferents in the rat synapse on unlabeled neuronal targets of catecholamine terminal in the nucleus accumbens septi and on dopamine neurons in the ventral tegmental area. *J Comp Neurol* 1992;320:145–160.

59. Wolf ME. The role of excitatory amino acids in behavioral sensitization to psychomotor stimulants. *Progr Neurobiol* 1998;54(6):679–720.

60. Vanderschuren L, Kalivas P. Alterations in dopaminergic and glutamatergic signaling in the induction and expression of behavioral sensitization. *Psychopharmacology* 2000;151:99–120.

61. Grant S, London ED, Newlin DB, et al. Activation of memory circuits during cue-elicited cocaine craving. *Proc Natl Acad Sci USA* 1996;93:12040–12045.

62. Kilts CD, Schweitzer JB, Quinn CK, et al. Functional anatomy of drug craving in human cocaine addiction. *Am J Psychiatry* 1998.

63. Nurmi M, Ashizawa T, Sinclair JD, et al. Effect of prior ethanol experience on dopamine overflow in accumbens of AA and ANA rats. *Eur J Pharmacol* 1996;315(3):277–283.

64. Szumlinski K, Maisonneuve I, Glick S. The potential antiaddictive agent, 18 methoxycoronaridine, blocks the sensitized locomotor and dopamine responses produced by repeated morphine treatment. *Brain Res* 2000;864(1):13–23.

65. George TP, Verrico CD, Roth RH. Effects of repeated nicotine pre-treatment on mesoprefrontal dopaminergic and behavioral responses to acute footshock stress. *Brain Res* 1998;801(1 and 2):36–49.

66. Volkow N, Wang G, Fowler J, et al. Decreased striatal dopaminergic responsiveness in detoxified cocaine-dependent subjects. *Nature* 1997;386:830–833.

67. Wang G, Volkow N, Fowler J, et al. Reproducibility of repeated measures of endogenous dopamine competition with [11C]reclopride in the human brain in response to methylphenidate. *J Nucl Med* 1999;40:1285–1291.

68. Horger BA, Giles MK, Schenk S. Preexposure to amphetamine and nicotine predisposes rats to self-administer a low dose of cocaine. *Psychopharmacology* 1992;107:271–276.

69. Itzhak Y, Martin JL. Effects of cocaine, nicotine, dizocilpine and alcohol on mice locomotor activity: cocaine-alcohol cross-sensitization involves upregulation of striatal dopamine transporter binding sites. *Brain Res* 1999;818(2):204–211.

70. Beitner-Johnson D, Guitart X, Nestler EJ. Common intracellular actions of chronic morphine and cocaine in dopaminergic brain reward regions. In: Kalivas PW, Samson HH, eds. *The neurobiology of drug and alcohol addiction* (Annals of the New York Academy of Sciences, Volume 654). New York: New York Academy of Science, 1992:70–87.

71. Nestler E, Aghajanian G. Molecular and cellular basis addiction. *Science* 1997;278:58–63.

72. Bonci A, Williams JT. A common mechanism mediates long-term changes in synaptic transmission after chronic cocaine and morphine. *Neuron* 1996;16:631–639.

73. Erdtmann-Vourliotis M, Mayer P, Riechert U, et al. Prior experience of morphine application alters the c-fos response to MDMA ('ecstasy') and cocaine in the rat striatum. *Mol Brain Res* 2000;77(1):55–64.

74. Prasad BM, Ulibarri C, Sorg BA. Stress-induced cross-sensitization to cocaine: effect of adrenalectomy and corticosterone after short- and long-term withdrawal. *Psychopharmacology* 1998;136:24–33.

75. Mitchell JB, Stewart J. Facilitation of sexual behaviors in the male rat associated with intra-VTA injections of opiates. *Pharmacol Biochem Behav* 1990;35:643–650.

76. Segal DS, Schuckit MA. Animal models of stimulant-induced psychosis. In: Creese I, ed. *Stimulants: neurochemical, behavioral, and clinical perspectives*. New York: Raven Press, 1983:131–167.

77. Markou A, Weiss F, Gold LH, et al. Animal models of drug craving. *Psychopharmacology* 1993;112:163–182.

78. Zhang X-F, Hu X-T, White FJ, et al. Increased responsiveness of ventral tegmental area dopamine neurons to glutamate after repeated administration of cocaine or amphetamine is transient and selectively involves AMPA receptors. *J Pharmacol Exp Ther* 1997;281:699–706.

79. Corrigal WA, Coen KM, Adamson KL. Self-administered nicotine activates the mesolimbic dopamine system through the ventral tegmental area. *Brain Res* 1994;653:278–284.

80. Gallegoa RA, Lee R-S, Criado JR, et al. Adaptive responses of γ-aminobutyric acid neurons in the ventral tegmental area to chronic ethanol. *J Pharmacol Exp Ther* 1999;291(3):1045–1052.

81. Ohmori T, Abekawa T, Koyama T. Environment modifies the expression of behavioral sensitization produced by methamphetamine: behavioral and neurochemical studies. *Behav Pharmacol* 1995;6:133–142.

82. Pierce RC, Kalivas PW. Amphetamine produces sensitized locomotion and dopamine release preferentially in the nucleus accumbens shell of rats administered repeated cocaine. *J Pharmacol Exp Ther* 1996;275:1019–1029.

83. Pierce RC, Kalivas PW. Repeated cocaine modifies the mechanism by which amphetamine releases dopamine. *J Neurosci* 1997;17:3254–3261.

84. Iwata S-I, Hewlett GHK, Ferrell ST, et al. Enhanced dopamine release and phosphorylation of synapsin I and neuromodulin in striatal synaptosomes after repeated amphetamine. *J Pharmacol Exp Ther* 1997;283:1445–1452.

85. Kantor L, Hewlett Keikilani GH, et al. Enhanced amphetamine- and K+-mediated dopamine release in rat striatum after repeated amphetamine: differential requirements for Ca2+- and calmodulin-dependent phosphorylation and synaptic vesicles. *J Neurosci* 1999;19(10):3801–3808.

86. Nestler EJ, Aghajanian GK. Molecular and cellular basis of addiction. *Science* 1997;278:58–63.

87. Mackler SA, Korutla L, Cha XY, et al. NAC-1 is a brain POZ/BTB protein that can prevent behavioral sensitization in rat. *J Neurosci* 2000,20:6210–6217.

88. Hurd YL, Herkenham M. Molecular alterations in the neostriatum of human cocaine addicts. *Synapse* 1993;13:357–369.

89. Ketz M, Chen, J, Carlezon WA, et al. Expression of the transcription factor dFosB in the brain controls sensitivity to cocaine. *Nature* 1999;401:272–276.

90. Brakeman PR, Lanahan AA, O'Brien R, et al. Homer: a protein that selectively binds metabotropic glutamate receptors. *Nature* 1997;386:221–223.

91. O'Brien R, Xu D, Petralia RS, et al. Synaptic clustering of AMPA receptors by the extracellular immediate early gene product Narp. *Neuron* 1999;23:309–323.

92. Jentsch JD, Taylor JR, Elsworth JD, et al. Altered frontal cortical dopaminergic transmission in monkeys after subchronic phencyclidine exposure: involvement in frontostriatal cognitive deficits. *Neuroscience* 1999;90(3):823–832.

93. Beyer CE, Steketee JD. Dopamine depletion in the medial prefrontal cortex induces sensitized-like behavioral and neurochemical responses to cocaine. *Brain Res* 1999;833(2):133–141.

94. Vezina P, Blanc G, Glowinski J, et al. Blockade of D-1 dopamine receptors in the medial prefrontal cortex produces delayed effects on pre- and postsynaptic indices of dopamine function in the nucleus accumbens. *Synapse* 1994;16:104–112.

95. Schenk S, Horger BA, Peltier R, et al. Supersensitivity to the reinforcing effects of cocaine following 6-hydroxydopamine lesions to the medial prefrontal cortex in rats. *Brain Res* 1991;543:297–235.

96. Jentsch K, Taylor J. Impulsivity resulting form frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology* 1999;146:373–390.

97. Tran-Nguyen LTL, Fuchs RA, Coffey GP, et al. Time-dependent changes in cocaine-seeking behavior and extracellular dopamine levels in the amygdala during cocaine withdrawal. *Neuropsychopharmacology* 1998;19:48–59.

98. Harmer CJ, Hitchcott PK, Morutto SL, et al. Repeated d-amphetamine enhances stimulated mesoamygdaloid dopamine transmission. *Psychopharmacology* 1997;132:247–254.

99. Grimm J, See R. Dissociation of primary and secondary reward-relevant limbic nuclei in an animal model of relapse. *Neuropsychpharmacology* 2000;22(5):473–479.

100. Hitchcott P, Harmer C, Phillips G. Enhanced acquisition of discriminative approach following intra-amygdala d-amphetamine. *Psychopharmacology* 1997;132:237–246.

101. Brown EE, Robertson GS, Fibiger HC. Evidence for conditional neuronal activation following exposure to a cocaine-paired environment: role of forebrain limbic structures. *J Neurosci* 1992;12:4112–4121.

102. Neisewander JL, Baker DA, Fuchs RA, et al. Fos protein expression and cocaine seeking behavior in rats after exposure to a cocaine self-administration environment. *Neuroscience* 2000;20(2):798–805.

103. Yang C, Seamans J, Gorelova N. Developing a neuronal model for the pathology of schizophrenia based on the nature of electrophysiological actions of dopamine in the prefrontal cortex. *Neuropsychopharmacology* 1999;21(2):161–194.

104. Williams GV, Goldman-Rakic PS. Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* 1995;376:572–575.

105. Taber MT, Das S, Fibiger HC. Cortical regulation of subcortical dopamine release: mediation via the ventral tegmental area. *J Neurochem* 1995;65:1407–1410.

106. Karreman M, Moghaddam B. The prefrontal cortex regulates the basal release of dopamine in the limbic striatum: an effect mediated by ventral tegmental area. *J Neurochem* 1996;66:589–598.

107. Brake W, Flores G, Francis D, et al. Enhanced nucleus accumbens dopamine and plasma corticosterone stress responses in adult rats with neonatal excitotoxic lesions to the medial prefrontal cortex. *Neuroscience* 2000;96(4):687–695.

108. Pierce RC, Bell K, Duffy P, et al. Repeated cocaine augments excitatory amino acid transmission in the nucleus accumbens only in rats having developed behavioral sensitization. *J Neurosci* 1996;16:1550–1560.

109. Reid MS, Hsu KJ, Berger SP. Cocaine and amphetamine preferentially stimulate glutamate release in the limbic system: studies on the involvement of dopamine. *Synapse* 1997;27:95–105.

110. Kalivas PW, Duffy T. Repeated cocaine administration alters extracellular glutamate levels in the ventral tegmental area. *J Neurochem* 1998;70:1497–1502.

111. Cornish J, Kalivas P. Glutamate transmission in the nucleus accumbens mediates relapse in cocaine addiction. *J Neurosci* 2000;20:RC89.

112. Tzschentke T, Schmidt W. Differential effects of discrete subarea-specific lesions of the rat medial prefrontal cortex on amphetamine- and cocaine-induced behavioral sensitization. *Cereb Cortex* 2000;10:488–498.

113. Brown EE, Fibiger HC. Differential effects of excitotoxic lesions of the amygdala on cocaine-induced conditioned locomotion and conditioned place preference. *Psychopharmacology* 1993;113:123–130.

114. Everitt BJ, Robbins TW. Amygdala-ventral striatal interactions and reward-related processes. In: Aggleton JP, ed. *The amygdala: neurobiological aspects of emotion, memory and mental dysfunction*. New York: Wiley-Liss, 1992:401–429.

115. Gaffan D, Murray EA. Amygdala interaction with the mediodorsal nucleus of the thalamus and the ventromedial prefrontal cortex in stimulus-reward associative learning in the monkey. *J Neurosci* 1990;8:3144–3150.

116. Robinson TE, Kolb B. Persistent structural modifications in nucleus accumbens and prefrontal cortex neurons produced by previous experience with amphetamine. *J Neurosci* 1997;17:8491–8497.

117. Robinson T, Kolb B. Morphine alters the structure of neurons in the nucleus accumbens and neocortex of rats. *Synapse* 1999;33(2):160–162.

96

Molecular and Cellular Biology of Addiction

Kathy L. Kopnisky

Steven E. Hyman

Kathy L. Kopnisky and Steven E. Hyman: National Institutes of Health, Bethesda, Maryland.

Addiction to alcohol, tobacco, and illegal drugs represents a substantial burden to societies worldwide. In terms of health-related outcomes, addiction results in enormous direct medical costs, premature mortality (tobacco alone may be responsible for 450,000 deaths yearly in the United States), and disability. In terms of broader social costs, addiction results in crime, negative impacts on families, derailed lives, and personal suffering. The major categories of drugs most likely to produce addiction are psychostimulants (including cocaine and amphetamines), opiates, ethanol, nicotine, marijuana, and phencyclidine-like drugs. Understanding the molecular and cellular actions of addictive drugs is obligatory if we are to better understand pathophysiology and develop potent pharmacotherapies to treat addiction. Of course, the molecular and cellular information presented in this chapter cannot be applied directly to the behavioral expression of addiction without putting it into the context of systems level neuroscience described in other chapters.

Acutely, addictive drugs are both rewarding (i.e., interpreted by the brain as intrinsically positive) and reinforcing (i.e., behaviors associated with drug use tend to be repeated). With repeated use, however, addictive drugs produce molecular changes that, within a vulnerable brain, promote continued drug-taking behavior in a manner that becomes increasingly difficult to control. The central feature of addiction is compulsive drug use—the loss of control over the apparently voluntary acts of drug seeking and drug taking. Once it has taken hold, addiction tends to follow a chronic course with periods of abstinence (that may or may not follow treatment), followed by relapse to active drug use. Even after extended periods of drug abstinence, the risk of relapse remains high. From the point of view of developing treatments, a central problem in addiction research includes understanding the molecular processes that lead to compulsive use and the long-term risk of relapse.

Addiction (defined as compulsive use) is not the only long-term effect of addictive drugs. Both addictive and many nonaddictive drugs may produce tolerance and dependence. Tolerance refers to the diminishing effect of a drug after repeated administration at the same dose, or to the need for an increase in dose to produce the same effect. Dependence represents an adaptive state that develops as a homeostatic response to repeated drug administration. Dependence is typically unmasked when drug taking stops, leading to withdrawal symptoms. Withdrawal symptoms may even emerge during active drug use as a result of tolerance, helping to drive increasing dosages or shorter intervals between doses.

Among the addictive drugs, ethanol and opiates produce dependence that has a somatic component, manifested by somatic symptoms during withdrawal, such as hypertension, tremor or seizures for ethanol, and hypertension, lacrimation, and abdominal cramps for opiates. All addictive drugs, including the psychostimulants, can produce an emotional–motivational component of dependence, manifested by symptoms such as dysphoria, anhedonia, and drug craving.

Tolerance and dependence may be prominent features accompanying addiction, but are not required. Indeed, when produced by addictive drugs, tolerance and withdrawal symptoms tend to resolve within days to weeks and therefore cannot account for the persistence of drug addiction (as manifest by the tendency to relapse) for many years. Indeed, both tolerance and dependence can occur with nonaddictive drugs as well. For example, B-adrenergic agonists inhaled for asthma, many antihypertensive drugs, and shorter-acting serotonin selective reuptake inhibitors may produce dependence and withdrawal symptoms on cessation, but do not produce compulsive drug seeking and drug taking. Based on these considerations, the molecular mechanisms underlying tolerance and dependence, and those responsible for addiction may overlap, but cannot be identical.

One other long-term effect of addictive drugs, best documented for psychostimulants, is sensitization, in which repeated administration of a drug elicits escalating effects of a given dose. Sensitization can be operationally defined as a leftward shift in the drug's dose–response curve (1). Because behavioral sensitization to drugs in animal models can be quite long-lived, it has been considered by some to be a model for long-lasting aspects of human drug addiction.

Not every individual who experiments with drugs becomes addicted. Indeed, the likelihood that a person will experiment with drugs, use them repetitively, and progress to addiction, appear to be the product of complex gene–gene and gene–environment interactions, acting together with contextual variables, such as drug availability. Factors related to vulnerability are discussed elsewhere in this volume. This chapter has a dual focus. First, it discusses the initial molecular targets of addictive drugs in the brain; then the molecular and cellular changes induced by drugs in the brain that might be responsible for such clinically significant aspects of drug abuse syndromes as tolerance, dependence, sensitization, and addiction. An enormous number of drug-induced molecular and cellular changes in brain function are already known, not all of which turn out to have clinical relevance. Thus, the chapter does not attempt to produce an exhaustive list of the known molecular effects of addictive drugs, but focuses on a subset of those that illustrate important principles and that can be related to the long-term effects of addictive drugs in humans.

With exceptions from a small number of human postmortem studies, most of what we know about the molecular and cellular actions of addictive drugs comes from animal models. The integration of such information about drug action in the brain with information about human risk factors is in its early stages and will benefit enormously from the eventual discovery of risk-producing alleles from human genetic studies. The discovery of alleles that confer vulnerability to drug use or addiction will help focus molecular and cellular studies of pathophysiology, as well as suggest biochemical pathways that can be exploited for treatment.

MOLECULAR TARGETS OF ADDICTIVE DRUGS

Part of "96 - Molecular and Cellular Biology of Addiction "

The overall effect of each of the addictive drugs depends on the particular neurons and circuits that express their molecular targets, and the nature of those targets. Thus, for example, morphine-like opiates are analgesic and sedating, whereas cocaine is a psychomotor stimulant; these different properties are based on differences in localization and functional properties of the proteins with which they interact, the μ-opioid receptor for morphine and the dopamine reuptake transporter (DAT) for cocaine. However, as described in other chapters in this section, addictive drugs share the ability to activate mesocorticolimbic dopamine projections that are critical substrates for both rewarding and reinforcing effects of natural stimuli. Mesocorticolimbic dopamine projections originate in the ventral tegmental area (VTA) of the ventral midbrain and project to structures that include the nucleus accumbens (NAc) (a complex structure within the ventral striatum that is the best-established substrate for reinforcement), and the prefrontal cerebral cortex. *In vivo* microdialysis studies have indicated that most if not all addictive drugs, including cocaine, amphetamines, opiates, nicotine, and ethanol, cause selective elevation of extracellular dopamine levels in the NAc, and blockade of dopamine neurotransmission in this region attenuates most measurable reinforcing and rewarding effects of addictive drugs (2 ,3). The powerful control over behavior exerted by addictive drugs is thought to result from the brain's inability to distinguish between the activation of reward circuitry by drugs and natural activation of the same circuitry by useful behaviors (e.g., behavioral related to eating or reproduction). Any activity, whether related to drug taking or survival, that activates this circuitry tends to be repeated; however, activation of reward circuitry by addictive drugs can be much more reliable and powerful than activation triggered by natural reinforcers, facilitating repetitive drug use, and with it, the initiation of molecular mechanisms that may produce tolerance, dependence, sensitization, and compulsive use. Although the mesocorticolimbic dopamine system is a site of convergence for the rewarding effects of virtually all major classes of addictive drugs, these drugs act by very different mechanisms.

Psychostimulants

The best-characterized and most widely abused psychostimulants are cocaine and the amphetamines. The details of their mechanisms of actions differ, but both result in increases of extracellular dopamine and other monoamines and produce similar effects on behavior. In humans, psychostimulants increase alertness and produce a sense of well being. In animal studies, psychostimulants produce a dose-dependent increase in locomotor activity at low doses and stereotypies at high doses. If cocaine or amphetamine is used repeatedly, some acute drug effects may diminish (tolerance), whereas others are enhanced (sensitization).

Cocaine and amphetamines produce their effects by potentiating monoaminergic transmission through actions on dopamine, serotonin, and norepinephrine reuptake transporters (4). These proteins normally transport previously released neurotransmitter back into the presynaptic nerve terminal, and thereby terminate transmitter action. Cocaine binds to these transporters and competitively inhibits their functioning, thereby increasing the duration of action of neurotransmitter released into the synaptic cleft. Amphetamines and related drugs increase dopamine, serotonin, and norepinephrine neurotransmission by acting as a substrate for their transporters. Amphetamines are transported into

the presynaptic terminal where they cause neurotransmitter release by reversing the usual direction of transport (i.e., causing transmitter to move into the synapse).

Whereas psychostimulants affect all three transporters, it is their actions at the DAT that are most directly related to the reinforcing effects of psychostimulant drugs. Lesions of the dopamine system or administration of dopamine receptor antagonists, but not similar manipulations of serotonin or noradrenergic systems, markedly attenuate cocaine self-administration. The central role of the DAT in psychostimulant action is highlighted in studies using mice in which the DAT has been genetically inactivated. In the absence of the DAT, animals are insensitive to the locomotor stimulatory effects of cocaine and amphetamine (5). By contrast, animals lacking serotonin (5-HTT$^{-/-}$) or norepinephrine (NET$^{-/-}$) transporters exhibit normal locomotor responses to psychostimulants (6 ,7). Interestingly DAT$^{-/-}$ animals still self-administer cocaine and amphetamine to some degree (8), but the interpretation of this result is complex because DAT$^{-/-}$ animals have very high levels of extracellular dopamine at baseline (lacking a DAT to remove synaptic dopamine), which might magnify a psychostimulant-mediated effect on norepinephrine or serotonin. Much evidence demonstrates that the dopamine system is obligatory for psychostimulant-induced reinforcement, but the serotonin and noradrenergic systems, whose transporters are also inhibited by psychostimulants, may play a role as well.

Opiates

The opiates and their synthetic analogues are the most effective analgesic agents known, and at the same time can produce tolerance, dependence (including somatic dependence), and addiction. Physical dependence on opiates can contribute to addiction, but can also occur independently of it. For example, patients with cancer pain may become physically dependent on these drugs but do not compulsively abuse them.

Opiate drugs bind to receptors for three subtypes of receptors that normally bind endogenous opioid peptides. The three subtypes, denoted μ, κ, and δ, are members of the G protein-coupled receptor family, and all interact with G proteins of the G_i/G_o types. This coupling results in inhibition of adenylyl cyclase, activation of inwardly rectifying K^+ channels, and inhibition of voltage-gated Ca^{2+} channels. Opiate receptors thus typically mediate inhibitory responses that reduce membrane excitability and reduce likelihood of cell firing and neurotransmitter release; however, the different opiate receptor subtypes are expressed on different cells, resulting in different biological effects when stimulated. Morphine-like opiates, including heroin, are both analgesic and addictive, and interact with highest affinity with the μ-receptor.

Morphine-like opiates suppress afferent nociceptive information by acting on opiate receptors contained within a descending pathway extending from the periaqueductal gray matter of the midbrain, to the rostroventral medulla, and then to the dorsal horn of the spinal cord. These same drugs appear to produce both reward and reinforcement by means of at least two mechanisms: (a) activation of the VTA, which results in dopamine release in the NAc; and (b) direct binding to opiate receptors in the NAc, an action that is independent of dopamine. Activation of VTA dopamine neurons by opiates results from disinhibition: morphine-like opiates inhibit GABAergic interneurons in the VTA that tonically inhibit the dopamine projection neurons (9). Increased activity of these dopamine neurons produces increases in extracellular dopamine levels in the NAc. Consistent with this arrangement, the reinforcing effects of intravenous heroin can be partly attenuated by administration of an opioid receptor antagonist directly into the VTA or lesioning the dopaminergic neurons of the VTA. Opiates also produce reinforcement through direct dopamine-independent action on μ, and perhaps δ, receptors expressed by NAc neurons. Consistent with this mechanism, morphine is self-administered even in the presence of dopamine receptor blockade or following a 6-hydroxydopamine lesion that destroys dopamine neurons (10 ,11). Moreover, lesions of the NAc or pharmacologic μ-receptor antagonists applied to the NAc dose dependently reduce the reinforcing effects of heroin and morphine (12 ,13). Thus, opiates and dopamine work through different postsynaptic receptors within the NAc to produce reinforcement.

M and δ-opioid receptor subtypes, both of which are present in the VTA and NAc, may both play a role in opiate reinforcement. In contrast, κ-opioid receptor activation is not reinforcing. Activation of κ receptors can decrease dopamine release in the NAc by both presynaptic mechanisms- there are κ receptors on a subset of dopamine terminals. As a result stimulation of κ opiate receptors may produce aversive responses in both animals and humans. As will be described in the following, κ opiate receptors may play a role in the emotional–motivational aspects of withdrawal from psychostimulants.

Ethanol

Ethanol is a central nervous system depressant that produces behavioral disinhibition, euphoria, reduced anxiety, decreased motor coordination, and sedation. The major mechanisms underlying behavioral effects, including reinforcement, are thought to be facilitation of $GABA_A$ receptors and inhibition of NMDA glutamate receptors (14). At higher doses, ethanol also inhibits the functioning of most voltage-gated ion channels. The molecular mechanisms by which ethanol affects these receptors and channels are not yet certain; two types of mechanisms have been hypothesized. One possible mechanism attributes the effects of ethanol on receptors and channels to its generalized effects on cell membranes, in which it is highly soluble. Certain

ligand-gated and voltage-gated channels may be preferentially affected by ethanol because, as complex multimeric proteins, they may be particularly vulnerable to ethanol-mediated changes in their lipid environment. The alternative hypothesis is that ethanol interacts with specific hydrophobic regions of these proteins to produce allosteric changes in structure, but the convincing demonstration of such interactions is still lacking.

Whether it acts via its general effects on membranes or, more specifically, in interaction with particular regions of proteins, ethanol has been shown to allosterically regulate the $GABA_A$ receptor to enhance GABA-activated Cl^- flux. The anxiolytic and sedative effects of ethanol, like those of barbiturates and benzodiazepines, are believed to result from facilitation of the $GABA_A$ receptor, although the precise mechanism differs for each drug. For example, distinct binding sites on the receptor have been identified for barbiturates and benzodiazepines. The convergence of actions of ethanol, barbiturates, and benzodiazepines on a single receptor result in more than additive effects, which can be responsible for lethal overdoses. In addition, these agents all produce cross-tolerance, thus permitting the use of benzodiazepines in ethanol detoxification protocols.

Not all $GABA_A$ receptors are ethanol sensitive. $GABA_A$ receptor complexes are heteropentamers comprised of combinations of the various members of five distinct subunit families. The subunit combinations vary in different cell types, leading to differences in the sensitivity of $GABA_A$ receptors to ethanol in different brain regions.

Ethanol also acts as an NMDA glutamate receptor antagonist, and allosterically inhibits the passage of glutamate-activated Na^+ and Ca^{2+} currents through the NMDA receptor. Other actions of ethanol that are possibly relevant to its psychotropic effects include potentiation of the action of serotonin at $5\text{-}HT_3$ receptors, which, like NMDA receptors, are excitatory, cation-selective ion channels.

The mechanisms by which ethanol produces reinforcement are not yet known in their entirety. The reinforcing effects of ethanol are partly explained by its ability to activate mesocorticolimbic dopamine circuitry (15), with enhanced release of dopamine in the NAc. It is not known whether this effect is mediated by disinhibition of dopamine neurons at the level of the VTA or whether it occurs at the level of the NAc, nor is it known whether it is caused primarily by facilitation of $GABA_A$ receptors or inhibition of NMDA receptors. Finally, it is not know to what degree opioid, serotonin, and other systems play a role in ethanol-mediated reinforcement. Thus, for example, not only $GABA_A$ receptor antagonists but also opiate antagonists, decrease ethanol self-administration and ethanol-related behavioral effects in rats (16 ,17 and 18). The opiate antagonist naltrexone reduces ethanol self-administration in animals; moreover, naltrexone and other opioid receptor antagonists reduce ethanol consumption, relapse to active drinking, and craving clinically (19 ,20).

Serotonin also appears to be involved in ethanol consumption and reinforcement; ethanol consumption is generally curbed by experimental manipulations that increase serotonergic function, and experiments with rats selectively bred for ethanol preference suggest that strong ethanol preference is associated with reduced serotonergic function. $5\text{-}HT_3$ antagonists such as ondansetron can block both ethanol-induced dopamine release in the NAc and ethanol consumption in rats. Mice lacking $5\text{-}HT_{1B}$ serotonin receptors consume higher levels of ethanol yet demonstrate less ataxia (21).

Nicotine

Nicotine is the main psychoactive ingredient of tobacco and is responsible for the stimulant effects, reinforcement, and dependence that result from tobacco use. Cigarette smoking rapidly delivers nicotine into the bloodstream. Nicotine differs from cocaine and opiates in that it is powerfully reinforcing in the absence of subjective euphoria.

The effects of nicotine are caused by its activation of nicotinic acetylcholine receptors (nAChRs). Nicotinic AChRs are ligand-gated cation channels located both presynaptically and postsynaptically. Presynaptic nAChRs facilitate neurotransmitter release. The reinforcing effects of nicotine depend on an intact mesolimbic dopamine system; nicotine-induced increases in locomotor behavior are also blocked by destruction of mesolimbic dopamine nerve terminals or cell bodies (22). Moreover, nicotine increases dopamine neurotransmission and energy metabolism in the nucleus accumbens (23).

Nicotinic AChRs containing α6 and B2 subunits are highly expressed in VTA dopamine neurons, and seem to be involved in both nicotine-induced dopamine release and reinforcement and in nicotine-induced locomotor activation (24 ,25). Systemic nicotine self-administration is disrupted when nicotinic receptor antagonists are administered directly into the VTA but not when they are administered into the NAc. Nicotine may also have some ability to stimulate dopamine release in the NAc, however, mediated by presynaptic nAChRs located on dopamine terminals within the NAc. Nicotinic AChRs on VTA dopamine neurons are normally activated by cholinergic innervation from the laterodorsal tegmental nucleus or the pedunculopontine nucleus.

Nicotine may also affect reinforcement via the opioid peptide system. Not only dopamine antagonists, but also opiate antagonists, block nicotine-induced behaviors and self-administration (26 ,27). These findings suggest a role for endogenous opioid systems in the reinforcing effects of nicotine, and raise the possibility that such antagonists may be of use in the treatment of nicotine addiction.

Cannabinoids

δ-9-Tetrahydrocannabinol (THC) is the major psychoactive compound contained in marijuana. THC produces effects

in humans that range from mild relaxation, euphoria, analgesia, and hunger to panic attacks. Reinforcing effects of cannabinoids comparable to those of other addictive drugs have not been demonstrated in animals, but cannabinoids have been shown to decrease reward thresholds and promote conditioned place preference in rats (28 ,29). THC increases mesolimbic dopamine transmission in the NAc shell, probably via a μ-opioid receptor-mediated mechanism because μ-receptor antagonists prevent the THCinduced dopamine increases in the brain mesolimbic area (30). Cannabinoids have also been reported to inhibit excitatory glutamatergic neurotransmission in the substantia nigra pars reticulata (31).

THC binds to two cannabinoid receptors denoted CB_1 and CB_2. Of the two, only the CB_1 receptor is expressed in the central nervous system, with high levels in the basal ganglia and limbic system (32). The endogenous ligand for the CB1 receptor appears to be an arachidonic acid derivative, anandamide; however, the nature of anandamide's function in the brain remains speculative. Evidence indicates that other endogenous ligands also may bind at this receptor.

Despite ongoing debates about the addictiveness of cannabinoids in humans, there appear to be many compulsive marijuana users. Withdrawal symptoms typically are not reported with termination of long-term marijuana use, but withdrawal symptoms have been demonstrated in a laboratory setting after four days of marijuana smoking (33). Cannabinoid dependence can be demonstrated experimentally with the use of cannabinoid receptor antagonists, which precipitate profound withdrawal symptoms that are somatic and emotional–motivational. In animals chronically treated with THC, a selective cannabinoid receptor antagonist produced withdrawal symptoms that included head shakes, facial tremors, tongue rolling, biting, wet dog shakes, and ptosis (34). Neurobiologically, withdrawal effects include increases in c-fos expression in the basal ganglia systems and CRF release in the amygdala (35).

Phencyclidine-Like Drugs

Phencyclidine (PCP or angel dust) and ketamine are related drugs classified as dissociative anesthetics. These drugs exhibit psychotomimetic properties, but are distinguished from hallucinogens by their distinct pharmacologic effects, including their reinforcing properties and risks related to compulsive abuse.

The reinforcing properties of PCP and ketamine are mediated by the binding to specific sites in the channel of the NMDA glutamate receptor, where they act as noncompetitive NMDA antagonists. PCP is self-administered directly into the NAc, where its reinforcing effects are believed to result from the blockade of excitatory glutamatergic input to the same medium spiny NAc neurons inhibited by opioids, and also by increases in extracellular dopamine. In contrast, hallucinogens, such as LSD, act at $5\text{-}HT_2$ serotonin receptors.

MOLECULAR AND CELLULAR MECHANISMS OF LONG-LIVED DRUG EFFECTS

Part of "96 - Molecular and Cellular Biology of Addiction "

Homeostasis Versus Associative Learning

Diverse behaviors, symptoms, and signs of substance use disorders coexist clinically, but depending on the drug and on the stage of the disorder, these may involve multiple molecular mechanisms occurring in diverse neural circuits. Heuristically, the types of molecular mechanisms involved in the long-lived effects of addictive drugs may be divided into two major classes: homeostatic adaptations and associative learning. Homeostatic adaptations can be understood as compensatory responses of cells or circuits to excessive bombardment by a drug or to excessive drug-induced neurotransmitter stimulation (e.g., excessive dopamine stimulation). These adaptations tend to dampen drug effects, thus playing a critical role in tolerance and dependence. The adapted state of neurons or neural systems may be unmasked on drug cessation, leading to the production of withdrawal symptoms, as illustrated in the following. Homeostatic adaptations typically occur within reversible bounds, and with removal of the drug, tend to dissipate over days to weeks.

Although clinically significant, homeostatic mechanisms cannot account for the persistent tendency of addicted individuals to relapse, even years after any withdrawal symptoms have subsided. Relapse often occurs on re-exposure to cues associated with drug use, consistent with an important role for associative learning (36). Although homeostatic mechanisms are thought to represent reversible global alterations in the sensitivity of neurons or circuits to neurotransmitters or drugs, associative learning is thought to represent long-lived or permanent alterations in patterns of synaptic connectivity that encode specific information (37). The clear separation between homeostasis and associative learning that has been implied, however, is an oversimplification. For example, there is recent evidence that associative learning mechanisms and compensatory adaptations may interact. Thus, for example, associative learning mechanisms have recently been shown to play a role tolerance to opiate analgesia, that is, the expression of tolerance may be context-dependent (38). Moreover, molecular adaptations that occur as a homeostatic response to drug bombardment may alter the threshold for associative learning involving affected cells.

Recruitment of Different Molecular Mechanisms Over Time

During the earliest periods of drug experimentation, mesocorticolimbic reward circuits are activated via different

mechanisms by different classes of drugs. As noted, a shared property of addictive drugs is to promote dopamine release in multiple forebrain regions, including the NAc, but also including the dorsal striatum, amygdala, and hippocampus, in which dopamine release can act as a reinforcement signal, thus controlling learning processes (39 ,40). As drug use continues, tolerance may occur, leading to dosage escalation. Depending on the drug, somatic dependence and/or emotional–motivational dependence my sustain drug seeking and drug use in attempts to avoid the aversive state of withdrawal. The emotional–motivational aspects of tolerance and dependence may largely occur within the mesocorticolimbic circuitry itself, but molecular adaptations occur in other circuits as well in a drug-specific manner reflecting the location of the target molecules for the given drug. Sensitization to some drug effects may occur, a phenomenon that is especially well documented for psychostimulants. Sensitization may act, *inter alia*, to increase the incentive salience of the drug, and thereby contribute to compulsive drug use (41). At the same time, multiple memory systems are affected by drugs of abuse (42) and, undoubtedly contribute to sustaining active drug use and late relapses (37). What follows are examples of different molecular processes that contribute to different aspects and stages of substance use disorders. These illustrations have been chosen based on the depth of available information, and likely relevance to the clinical situation in humans.

Adaptations That Produce Tolerance and Somatic Dependence to Opiates

Opiates and ethanol produce somatic dependence and withdrawal because their targets are expressed on cells and circuits that regulate bodily functions such as autonomic activity. Tolerance and dependence are generally thought to represent homeostatic adaptations that compensate for overstimulation by a drug or neurotransmitter. During withdrawal, the overcompensated system is suddenly unopposed by the drug it had adapted to counteract. Consequently, withdrawal symptoms appear that generally are opposite to the immediate effects produced by the drug. The molecular adaptations probably responsible for some aspects of tolerance and somatic dependence are best understood for opiates (43).

With repeat administration of mu agonist opiates such as morphine or heroin, both tolerance and dependence emerge. There is a significant somatic component to heroin dependence as manifest by the classic heroin somatic withdrawal syndrome. It had initially been hypothesized that opiate dependence would correlate with significant changes in expression of endogenous opioid peptides or opioid receptors or changes in opioid receptor affinity. This has not turned out to be the case; rather opiate tolerance and dependence appear to be caused by adaptation in postreceptor signaling mechanisms in opiate receptor-bearing cells (Fig. 96.1).

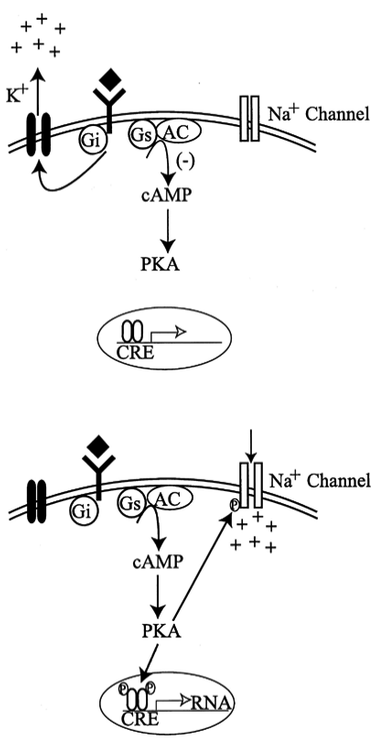


FIGURE 96.1. Mechanism of opiate tolerance and dependence in the locus ceruleus: Acute administration of opiates increases outward K^+ current, thereby hyperpolarizing locus ceruleus cells *(top)*. With chronic opiate use the cAMP signaling system is up-regulated, leading to PKA-dependent phosphorylation of the Na^+ channel. In this state, the channel is more active, allowing Na^+ ions to flow into the cell, increasingly the intrinsic excitability of the cell. Up-regulation of the cAMP system also increases CREB Ser^{133} phosphorylation and CRE-dependent gene transcription. Alterations in CRE-driven genes may contribute to the increased LC neuron excitability as well *(bottom)*.

The locus ceruleus (LC), located in the dorsal pons, is the major noradrenergic nucleus of the brain and regulates arousal, attention, and vigilance. It is involved in responses to stress, and together with other noradrenergic cell groups plays a role in regulation of the autonomic nervous system. Morphine-like opiates acutely inhibit the firing of LC neurons, but tolerance and dependence occur within the LC with continued administration. Thus, despite continued opiate exposure, LC firing rates gradually return to their basal levels. At this point, administration of an opioid receptor

antagonist, such as naloxone or naltrexone, causes a dramatic increase in LC firing rates. In animals, the period of rapid LC firing correlates with the somatic withdrawal syndrome, and drugs, such as the α_2-adrenergic receptor agonist clonidine, which inhibit LC firing, attenuate withdrawal symptoms.

Many of the adaptations that produce LC-mediated tolerance and dependence depend on the cyclic AMP (cAMP) pathway. In the LC, as in most other cell types, μ-opioid receptor activation inhibits the cAMP pathway via Gi activation and stimulates an inwardly rectifying K^+ current by means of direct interactions of the G protein βγ subunits with the channel. M-opiate receptors also inhibit a Na^+ current that is dependent on the cAMP-dependent protein kinase (protein kinase A or PKA) for its activation, and is thus dependent on an active cAMP system. Taken together, the actions of μ agonist opiates on these K+ and Na+ channels expressed by LC neurons decrease the excitability and inhibit the firing of the LC. With long-term opiate administration, however, a homeostatic compensatory response occurs: key components of the cAMP pathway become up-regulated in LC neurons; thus, for example there are increased concentrations of adenylyl cyclase and protein kinase A. This up-regulation increases the intrinsic excitability of LC neurons, by activating the cAMP-dependent Na^+ current. The activation of this current may explain why LC firing rates return to normal despite the continued presence of an opiate (an example of tolerance). These observations may also explain the dramatic increase in LC firing that occurs if an opiate antagonist such as naloxone is administered to precipitate withdrawal (illustrating dependence). The up-regulation of the cAMP pathway has increased the intrinsic excitability of the LC neuron, but as long as μ-opiate agonists continue to be administered, this excitability is counteracted.

Activation of LC neurons during opiate withdrawal owes not only to changes in intrinsic excitability, but also partly to glutamatergic projections to the LC from the nucleus paragigantocellularis of the medulla. Lesions of the paragigantocellularis, or glutamate receptor antagonists administered locally in the LC, attenuate withdrawal-induced increases in LC firing rates by approximately 50%. An up-regulated cAMP pathway also may mediate this effect, as long-term use of opiates causes up-regulation of the cAMP pathway in the paragigantocellularis and its major afferents.

The mechanisms by which the cAMP pathway becomes up-regulated are complex, and may involve both transcriptional and translational mechanisms; however, the importance of transcription factor CREB is supported by experiments in mice with a partial knockout (hypomorphic allele) of CREB. In these animals two of the major CREB isoforms, α and δ, were disrupted. After an opiate administration paradigm that would be expected to produce opiate dependence and naloxone-precipitated withdrawal, these mice exhibited markedly reduced signs of withdrawal including complete absence of sniffing and ptosis (44 ,45).

This is not the whole story, however. Opiate-induced up-regulation of PKA does not involve CREB and may be mediated posttranslationally. The inactive PKA holoenzyme is a heterotetramer composed of two regulatory and two catalytic subunits. When the regulatory subunits are bound by cAMP, the catalytic subunits are free to phosphor late substrate proteins. However, free catalytic subunits of PKA are highly vulnerable to proteolysis, whereas inactive subunits bound to regulatory subunits are proteolysis-resistant. It is currently speculated that PKA subunits accumulate in the LC during long-term opiate treatment because the enzyme is inhibited by the persistent presence of an opiate, keeping it in its inactive holoenzyme form in which subunits would be degradation-resistant. As the number of enzyme molecules increases, the kinase activity can be more readily activated by the low levels of cAMP.

Adaptations That May Produce Tolerance and Somatic Dependence on Ethanol

Like opiates, ethanol produces somatic dependence and withdrawal, although the clinical syndrome is quite distinct, and potentially more dangerous. The molecular mechanisms are less well understood than those underlying opiate tolerance and dependence, but the comparison is instructive. There is some evidence that homeostatic adaptations occur in response to ethanol that decrease $GABA_A$ receptor expression and increase NMDA receptor expression on some neurons. The decrease in receptors for the major inhibitory neurotransmitter and the increase in excitatory receptors would make neurons intrinsically more excitable. With removal of ethanol, a drugs that facilitates $GABA_A$ receptor-mediated Cl- currents and inhibits NMDA receptors, a state of increased neural excitability would be unmasked (Fig. 96.2) leading to withdrawal symptoms such as agitation, tremor, hypertension, and seizures. The decrease in $GABA_A$ receptor function is possibly due in part to decreased $GABA_A$-1 subunit expression in the striatum, cortex, and hippocampus (46). Conversely, chronic ethanol appears to increase the number and function of NMDA receptors (47). During withdrawal, glutamate release is increased for up to 36 hours in the NAc, hippocampus, and striatum (48). Overall, there is increasing evidence to suggest that that the neuronal hyperexcitability evident during ethanol withdrawal result from the combination of reduced $GABA_A$ receptor-mediated inhibition and increased glutamatergic excitation (Fig. 96.2).

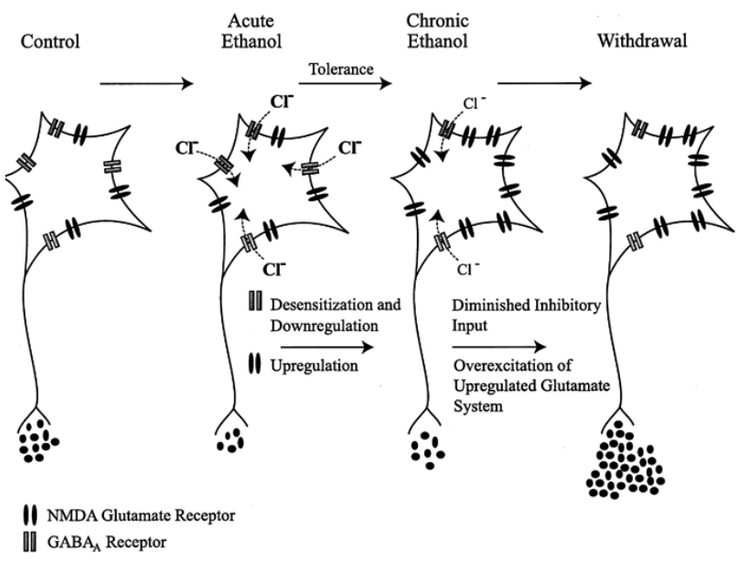


FIGURE 96.2. Hypothetical scheme to explain increased neuronal excitability with ethanol withdrawal: Acute ethanol exposure increases chloride conductance via the $GABA_A$ receptor and inhibits NMDA glutamate receptors, thereby reducing neuronal excitability and glutamate release. With chronic ethanol exposure, there is putative down-regulation of $GABA_A$ receptor subunits and possibly up-regulation of NMDA receptors. On ethanol withdrawal, inhibitory input from ethanol is removed; therefore, excitatory influences are relatively unopposed. The neurons release increased quantities of glutamate, which may act on up-regulated receptors. The unopposed cellular hyperexcitability can promote seizure activity as groups of neurons become overexcited.

Adaptations That Produce Emotional and Motivational Aspects of Dependence

An emotional and motivational component of dependence has been hypothesized to reflect neural adaptations to excessive

dopamine release within brain reward pathways. Emotional and motivational aspects of dependence are inferred in humans from the dysphoria, anhedonia, and, anxiety that may accompany withdrawal from psychostimulants (49) and other addictive drugs. One animal model for emotional–motivational aspects of dependence is elevation of brain stimulation reward thresholds. Direct electrical stimulation of brain reward pathways is both rewarding and reinforcing. Acute administration of addictive drugs decreases the level of stimulation necessary to achieve rewarding levels of stimulation. However, if drugs are repeatedly administered and then withdrawn, there is a marked increase in the threshold for achieving reward, as if the brain reward pathways are in a state of decreased responsiveness (50). Based on this and other models, it has been hypothesized that reduced mesolimbic dopaminergic activity is associated with the emotional–motivational aspects of dependence and withdrawal; however, brain reward circuitry is complex, and it is to be expected that many adaptive processes occur that contribute to dependence. For example, relevant adaptations likely occur both in NAc and in VTA neurons. Indeed there is evidence of up-regulation of the cAMP pathway with chronic administration of addictive drugs both in NAc neurons and in the GABAergic interneurons that innervate dopaminergic neurons of the VTA. Up-regulation of the cAMP pathway in VTA interneurons during withdrawal could lead to increased GABA release and consequently to reduced firing of the dopaminergic cells on which they synapse. Such activity might partially account for the reductions in dopaminergic neurotransmission from the VTA to the NAc observed during early phases of withdrawal and that are believed to contribute to withdrawal symptoms.

One mechanism that could contribute to aversive states that occur with psychostimulant withdrawal is up-regulation of the neuropeptide dynorphin. In the dorsal and ventral striatum, levels of prodynorphin mRNA and dynorphin peptides increase significantly following repeated administration of psychostimulants (51 ,52). A significant increase in prodynorphin mRNA is observed after rats self-administer cocaine (53) and, in postmortem studies of cocaine-dependent human drug abusers, there is a marked induction of prodynorphin, but not other peptide mRNAs in the striatum (54).

Dynorphin peptides are relatively selective for the κ opiate receptor, and exert inhibitory actions in the nervous system via the G protein, Gi. Stimulation of κ opiate receptors on dopamine terminals within the dorsal and ventral striatum appears to decrease dopamine release. Consistent with this, activation of κ receptors is associated behaviorally

with an aversive dysphoric syndrome both in humans (55) and rats (56). Thus, increases in dynorphin peptides produced by chronic cocaine or amphetamine administration may inhibit dopamine release and contribute to emotional–motivational aspects of psychostimulant withdrawal.

Regulation of prodynorphin gene expression by psychostimulants has been shown to be dependent on D1 dopamine receptor stimulation (57) because selective D1 receptor agonists inhibit it. Moreover, the prodynorphin gene is expressed in the striatum in D1 receptor bearing cells (58). D1 receptors are coupled to G_s/G_{olf}, and thus stimulate adenylyl cyclase to produce cAMP, which in turn activates PKA. PKA can then phosphorylate numerous substrates including CREB, a transcription factor that binds cAMP response elements (CREs) in numerous genes. Indeed, cocaine and amphetamine (59) have been shown to induce phosphorylation of CREB in striatal neurons via a D1 receptor-mediated mechanism, and the prodynorphin gene has been shown to be CREB regulated in these same cells (60). Thus, at the same time that D1 receptor stimulation acutely contributes to the acute rewarding effects of cocaine and amphetamine, it also initiating a cascade of homeostatic events that eventually yield compensatory adaptations to excess dopamine stimulation. One of these adaptations is induction of dynorphin peptides (Fig. 96.3).

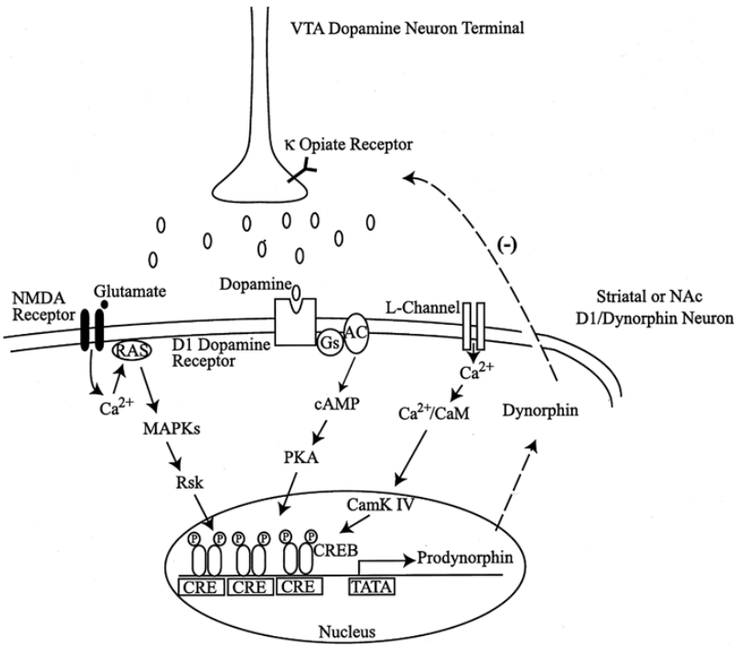


FIGURE 96.3. Dynorphin gene regulation by psychostimulants: implications for central motive states. Cocaine and amphetamine increase levels of dopamine in the nucleus accumbens and dorsal striatum. D1 receptor stimulation by dopamine leads to activation of the cAMP pathway, phosphorylation of CREB and ultimately the transcription of CRE-regulated genes such as c-fos and prodynorphin. Glutamate, via the NMDA receptor, as well as L-type calcium channels similarly contribute to CREB phosphorylation and CRE-driven gene expression in these dopaminoceptive neurons. Release of dynorphin inhibits dopamine release by binding to its presynaptic target, the κ opiate receptor, on DA nerve terminals. The dynorphin negative-feedback mechanism for controlling dopamine levels also contributes to aversive feelings and dysphoria due to its actions at the κ opiate receptor.

Corticotropin Releasing Factor May Also Contribute to Emotional–Motivational Aspects of Withdrawal

One type of adaptation that occurs outside the mesocorticolimbic dopamine system that may contribute to aversive states, and thus drug seeking, is up-regulation of corticotropin releasing factor (CRF). This neuropeptide is expressed in the hypothalamus, central nucleus of the amygdala, and other brain regions. In the hypothalamus CRF has been shown to be critical in the initiation of stress hormone cascades that culminate in the release of cortisol from the adrenal cortex. CRF released in the central nucleus of the amygdala has been implicated in anxiety states. Several studies have implicated CRF systems in the mediation of many of the angiogenic and aversive aspects of drug withdrawal. Increased release of CRF, particularly in the central nucleus of the amygdala, occurs during withdrawal from ethanol, opiates, cocaine, and cannabinoids. CRF antagonists have reversed at least some of the aversive effects of cocaine, ethanol, and opiate withdrawal in laboratory animals.

Alterations in Expression of Transcription Factors May Impact Diverse Neuronal Processes

Drug-induced changes in expression of transcription factors may lead to the altered expression of specific target genes, which in turn may affect both homeostatic adaptations and associative learning. In addition to regulating the peptide precursor gene, prodynorphin, D1 receptor-mediated stimulation of CREB induces a large number of immediate early genes (IEGs) including several that encode transcription factors, including c-fos, fras, junB, and zif268 (61 ,62 ,63 and 64). One interesting finding is that one CREB regulated transcription factor, ΔFosB, a truncated isoform of Fos B has a long half-life compared to all other members of the Fos family of transcription factors. Unlike other members of the Fos family, ΔFosB is only slightly induced by acute stimulation, but because it is long lived, it begins to accumulate with repeated stimulation, including repeated administration of addictive drugs (65). Thus, long-term, but not short-term, administration of cocaine, amphetamine, opiates, nicotine, or PCP induces ΔFosB in the NAc and dorsal striatum. Of all known molecules that exhibit altered levels following drug stimulation, ΔFosB is the longest lived currently known. Accumulation of ΔFosB represents a molecular mechanism by which drug-induced changes in gene expression can persist for weeks—even after drug use has been discontinued. The biological significance of ΔFosB induction will be better understood following identification of the genes that it regulates. Certain AMPA glutamate receptor subunit-encoding genes are among the candidates. Even though ΔFosB is stable, it is ultimately degraded; thus, by itself it cannot mediate the lifelong changes in behavior that accompany addiction.

Associative Learning

Both humans and animals readily learn to self-administer addictive drugs; behaviors that require the specific recognition of drug-associated cues, and the performance of complex action sequences. Cues associated with drug administration acquire motivational significance as illustrated by the conditioned place preference paradigm; for example, rats will choose to spend more time in a location in which they have passively received an injection of psychostimulants than in another location paired with saline injection (66). Associative learning also appears to play a role in psychostimulant sensitization. If, for example, a rat is taken from its home cage to a novel “test” cage for intermittent amphetamine injections, the sensitized locomotor response to a challenge dose is much greater if the challenge is also given in that test cage than if given in a different environment (67 ,68). Such context dependence can dominate the behavioral effects with sensitization expressed in the drug-associated location, no sensitization at all in a different environment (69 ,70). In drug-addicted humans, late relapses appear to involve associative learning, as they often occur after encounters with people, places, or other cues previously associated with drug use (71 ,72). As described, conditioned responses to drug-associated cues persist far longer than withdrawal symptoms (36), and can occur despite years of abstinence from drugs.

At a systems level, context-dependent sensitization in animal models and cue-conditioned relapse in humans suggests that the brain stores specific patterns of drug-related information. Homeostatic responses that increase or decrease the gain on the overall responsiveness of dopaminergic or other neurotransmitter systems in the brain could not mediate selective responsiveness to specific contexts or cues. Thus, general homeostatic mechanisms are not adequate to explain these phenomena. Elsewhere it has been argued that core features of addiction arise from the inappropriate recruitment of molecular mechanisms normally responsible for associative learning (37). On this view, the persistence of drug addiction reflects the persistence of the memory for this learned experience in the form of altered patterns of synaptic connectivity.

At the molecular level, stimulation of dopamine D1 receptors in multiple brain regions, including striatum, promotes activation of the transcription factor CREB (59 ,73) and a transient burst of altered gene expression (74). The induction of multiple transcription factors by this mechanism has already been described. Other psychostimulant induced IEG products that have been described in the striatum include *homer-1a, narp, arc*, and many others (62 ,74 ,75). Some of the genes induced by dopamine and psychostimulants in the striatum have been hypothesized to play

a role in hippocampal LTP, making it tempting to speculate that they may ultimately have a role in synaptic remodeling in the striatum (76 ,77 ,78 and 79). Indeed, D1 receptors have been shown to be required for normal hippocampal long-term potentiation (LTP), an important model of synaptic plasticity. For LTP in the CA1 region of the hippocampus to persist for more than 2 or 3 hours ("late-phase" LTP; L-LTP) there must be increases in postsynaptic cAMP, phosphorylation of CREB, gene transcription, and protein synthesis (64 ,80 ,81 and 82). The requirement for activation of gene expression seems to be transient, because blockers of transcription or translation disrupt hippocampal L-LTP if they are given within a few hours of the LTP-inducing stimulus, but not if given later (83). Activators of the cAMP cascade, including D1 agonists, can induce L-LTP (84 ,85). D1 receptor blockade inhibits hippocampal L-LTP (85 ,86 and 87), and D1-knockout mice do not show L-LTP (88). Therefore, D1 receptor activation in the hippocampus may act to gate synaptic plasticity, helping to determine whether changes in synaptic strength are long lasting or merely transient.

A role for dopamine receptors in the modification of synaptic strength fits well with the idea that increases in extracellular dopamine can act as a reinforcement learning signal in striatum (89). LTP (and also LTD, long-term depression) is found at corticostriatal synapses *in vivo* (90) and *in vitro* (91 ,92). Some groups have found that striatal LTP can be modified by dopamine receptor stimulation (91 ,93 ,94). Moreover, based on genetic manipulations, CREB has been implicated in both invertebrate and vertebrate models of synaptic plasticity and long-term memory (80 ,81 and 82 ,95). Moreover, changes in striatal synaptic physiology and synaptic structure occur in response to psychostimulant administration (96). At the systems level, dorsal regions of striatum appear to be involved in the learning and execution of complex automatized behavioral sequences, particularly in response to external cues. Ventral striatal areas are involved in acting on the motivational significance of such cues. Thus drug-induced synaptic plasticity in both regions may contribute to drug use through consolidation of drug-taking and -seeking behaviors. Many questions remain, but the central outstanding issue is the identification of genes transiently induced by addictive drugs, the products of which produce stable remodeling of synapses.

CONCLUSION

Part of "96 - Molecular and Cellular Biology of Addiction "

All of the initial molecular targets of drugs of abuse have been characterized and cloned. However, the molecular biology of processes relevant to tolerance, dependence, sensitization, and most important, compulsive drug use, and late relapse, are in their relatively early stages. Striking progress has been made in identifying large numbers of molecular changes initiated by drugs of abuse, but coherent biological implications of these changes can currently be described for only a few situations, such as somatic dependence on opiates. Even for more difficult problems, however, powerful tools are on the horizon. It is imperative, for example, to investigate the mechanisms by which dopamine excess might produce long-lived pathological associative memories that could underlie compulsive drug use and late relapse. Fortunately, in the very near future, a complete set of mammalian genes will be available in arrays, and similar collections of proteins will follow, albeit with some delay. Given these reagents, we will be limited only by our neurobiological imaginations.

REFERENCES

1. Altman J, Everitt BJ, Glautier S, et al. The biological, social and clinical bases of drug addiction: commentary and debate. *Psychopharmacology (Berl)* 1996;125:285–345.

2. Di Chiara G, Imperato A. Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proc Natl Acad Sci USA* 1988;85:5274–5278.

3. Wise RA, Mendrek A, Carlezon WA Jr. MK-801 (dizocilpine): synergist and conditioned stimulus in bromocriptine-induced psychomotor sensitization. *Synapse* 1996;22:362–368.

4. Rudnick G, Clark J. From synapse to vesicle: the reuptake and storage of biogenic amine neurotransmitters. *Biochim Biophys Acta* 1993;1144:249–263.

5. Giros B, Jaber M, Jones SR, et al. Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. *Nature* 1996;379:606–612.

6. Bengel D, Murphy DL, Andrews AM. Altered brain serotonin homeostasis and locomotor insensitivity to 3,4-methylenedioxymethamphetamine ("Ecstasy") in serotonin transporter-deficient mice. *Mol Pharmacol* 1998;53:649–655.

7. Xu F, Gainetdinov RR, Wetsel WC, et al. Mice lacking the norepinephrine transporter are supersensitive to psychostimulants. *Nat Neurosci* 2000;3:465–471.

8. Rocha BA, Fumagalli F, Gainetdinov RR, et al. Cocaine self-administration in dopamine-transporter knockout mice [see comments] [published erratum appears in *Nat Neurosci* 1998;1(4):330]. *Nat Neurosci* 1998;1:132–137.

9. Johnson SW, North RA. Opioids excite dopamine neurons by hyperpolarization of local interneurons. *J Neurosci* 1992;12:483–488.

10. Pettit HO, Ettenberg A, Bloom FE, et al. Destruction of dopamine in the nucleus accumbens selectively attenuates cocaine but not heroin self-administration in rats. *Psychopharmacology (Berl)* 1984;84:167–173.

11. Ettenberg A, Pettit HO, Bloom FE, et al. Heroin and cocaine intravenous self-administration in rats: mediation by separate neural systems. *Psychopharmacology (Berl)* 1982;78:204–209

12. Zito KA, Vickers G, Roberts DC. Disruption of cocaine and heroin self-administration following kainic acid lesions of the nucleus accumbens. *Pharmacol Biochem Behav* 1985;23:1029–1036.

13. Vaccarino FJ, Bloom FE, Koob GF. Blockade of nucleus accumbens opiate receptors attenuates intravenous heroin reward in the rat. *Psychopharmacology (Berl)* 1985;86:37–42.

14. Buck KJ. Molecular genetic analysis of the role of GABAergic systems in the behavioral and cellular actions of alcohol. *Behav Genet* 1996;26:313–323.

15. Yoshimoto K, McBride WJ, Lumeng L, et al. Ethanol enhances the release of dopamine and serotonin in the nucleus accumbens of HAD and LAD lines of rats. *Alcohol Clin Exp Res* 1992;16:781–785.

16. Hyytia P, Koob GF. GABAA receptor antagonism in the extended amygdala decreases ethanol self-administration in rats. *Eur J Pharmacol* 1995;283:151–159.

17. Liljequist S, Engel J. Effects of GABAergic agonists and antagonists on various ethanol- induced behavioral changes. *Psychopharmacology (Berl)* 1982;78:71–75.

18. Tabakoff B, Hoffman PL. Alcohol addiction: an enigma among us. *Neuron* 1996;16:909–912.

19. Weinrieb RM, O'Brien CP. Naltrexone in the treatment of alcoholism. *Annu Rev Med* 1997;48:477–487.

20. Spanagel R, Zieglgansberger W. Anti-craving compounds for ethanol: new pharmacologic tools to study addictive processes. *Trends Pharmacol Sci* 1997;18:54–59.

21. Crabbe JC, Phillips TJ, Feller DJ, et al. Elevated alcohol consumption in null mutant mice lacking 5-HT1B serotonin receptors. *Nat Genet* 1996;14:98–101.

22. Louis M, Clarke PB. Effect of ventral tegmental 6-hydroxydopamine lesions on the locomotor stimulant action of nicotine in rats. *Neuropharmacology* 1998;37:1503–1513.

23. Pontieri FE, Tanda G, Orzi F, et al. Effects of nicotine on the nucleus accumbens and similarity to those of addictive drugs [see comments]. *Nature* 1996;382:255–257.

24. Picciotto MR, Zoli M, Rimondini R, et al. Acetylcholine receptors containing the beta2 subunit are involved in the reinforcing properties of nicotine. *Nature* 1998;391:173–177.

25. le Novere N, Zoli M, Lena C, et al. Involvement of alpha6 nicotinic receptor subunit in nicotine-elicited locomotion, demonstrated by in vivo antisense oligonucleotide infusion. Neuroreport 1999;10:2497–2501.

26. Zarrindast MR, Pazouki M, Nassiri-Rad S. Involvement of cholinergic and opioid receptor mechanisms in nicotine- induced antinociception. *Pharmacol Toxicol* 1997;81:209–213.

27. Koob GF, Nestler EJ. The neurobiology of drug addiction. *J Neuropsychiatry Clin Neurosci* 1997;9:482–497.

28. Gardner EL, Paredes W, Smith D, et al. Facilitation of brain stimulation reward by δ9- tetrahydrocannabinol. *Psychopharmacology (Berl)* 1988;96:142–144.

29. Lepore M, Vorel SR, Lowinson J, et al. Conditioned place preference *Life Sci* 1995;56:2073–2080.

30. Tanda G, Pontieri FE, Di Chiara G. Cannabinoid and heroin activation of mesolimbic dopamine transmission by a common mu1 opioid receptor mechanism [see comments]. *Science* 1997;276:2048–2050.

31. Szabo B, Wallmichrath I, Mathonia P, et al. Cannabinoids inhibit excitatory neurotransmission in the substantia nigra pars reticulata. *Neuroscience* 2000;97:89–97.

32. Glass M, Dragunow M, Faull RL. Cannabinoid receptors in the human brain: a detailed anatomical and quantitative autoradiographic study in the fetal, neonatal and adult human brain. *Neuroscience* 1997;77:299–318.

33. Haney M, Ward AS, Comer SD, et al. Abstinence symptoms following smoked marijuana in humans. *Psychopharmacology* 1999;141:395–404.

34. Aceto MD, Scates SM, Lowe JA, et al. Cannabinoid precipitated withdrawal by the selective cannabinoid receptor antagonist, SR 141716A. *Eur J Pharmacol* 1995;282:R1–R2.

35. Rodriguez dF, Carrera MRA, Navarro M, et al. Activation of corticotropin-releasing factor in the limbic system during cannabinoid withdrawal [see comments]. *Science* 1997;276:2050–2054.

36. O'Brien CP, Childress AR, McLellan AT, et al. Classical conditioning in drug-dependent humans. *Ann NY Acad Sci* 1992;654:400–415.

37. Berke JD, Hyman SE. Addiction, dopamine, and the molecular mechanisms of memory. *Neuron* 2000;25:515–532.

38. Cepeda-Benito A, Tiffany ST, Cox LS. Context-specific morphine tolerance on the paw-pressure and tail-shock vocalization tests: evidence of associative tolerance without conditioned compensatory responding. *Psychopharmacology (Berl)* 1999;145:426–432.

39. Schultz W. The phasic reward signal of primate dopamine neurons. *Adv Pharmacol* 1998;42:686–690.

40. Schultz W. Predictive reward signal of dopamine neurons. *J Neurophysiol* 1998;80:1–27.

41. Robinson TE, Berridge KC. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Brain Res Rev* 1993;18:247–291.

42. White NM. Addictive drugs as reinforcers: multiple partial actions on memory systems. *Addiction* 1996;91:921–949.

43. Nestler EJ, Aghajanian GK. Molecular and cellular basis of addiction. *Science* 1997;278:58–63.

44. Maldonado R, Blendy JA, Tzavara E, et al. Reduction of morphine abstinence in mice with a mutation in the gene encoding CREB [see comments]. *Science* 1996;273:657–659.

45. Blendy JA, Maldonado R. Genetic analysis of drug addiction: the role of cAMP response element binding protein. *J Mol Med* 1998;76:104–110.

46. Hirouchi M, Hashimoto T, Kuriyama K. Alteration of GABAA receptor alpha 1-subunit mRNA in mouse brain following continuous ethanol inhalation. *Eur J Pharmacol* 1993;247:127–130.

47. Faingold CL, N'Gouemo P, Riaz A. Ethanol and neurotransmitter interactions—from molecular to integrative effects. *Prog Neurobiol* 1998;55:509–535.

48. Rossetti ZL, Carboni S. Ethanol withdrawal is associated with increased extracellular glutamate in the rat striatum. *Eur J Pharmacol* 1995;283:177–183.

49. Gawin FH. Cocaine addiction: psychology and neurophysiology [published erratum appears in *Science* 1991;253:494]. *Science* 1991;251:1580–1586.

50. Koob GF, Le Moal M. Drug abuse: hedonic homeostatic dysregulation. *Science* 1997;278:52–58.

51. Steiner H, Gerfen CR. Cocaine-induced c-fos messenger RNA is inversely related to dynorphin expression in striatum. *J Neurosci* 1993;13:5066–5081.

52. Hanson GR, Merchant KM, Letter AA, et al. Characterization of methamphetamine effects on the striatal-nigral dynorphin system. *Eur J Pharmacol* 1988;155:11–18.

53. Hurd YL, Brown EE, Finlay JM, et al. Cocaine self-administration differentially alters mRNA expression of striatal peptides. *Brain Res Mol Brain Res* 1992;13:165–170.

54. Hurd YL, Herkenham M. Molecular alterations in the neostriatum of human cocaine addicts. *Synapse* 1993;13:357–369.

55. Pfeiffer A, Knepel W, Braun S, et al. Effects of a kappa-opioid agonist on adrenocorticotropic and diuretic function in man. *Horm Metab Res* 1986;18:842–848.

56. Bals-Kubik R, Ableitner A, Herz A, et al. Neuroanatomical sites mediating the motivational effects of opioids as mapped by the conditioned place preference paradigm in rats. *J Pharmacol Exp Ther* 1993;264:489–495.

57. Shippenberg TS, Rea W. Sensitization to the behavioral effects of cocaine: modulation by dynorphin and kappa-opioid receptor agonists. *Pharmacol Biochem Behav* 1997;57:449–455.

58. Gerfen CR, Engber TM, Mahan LC, et al. D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons [see comments]. *Science* 1990;250:1429–1432.

59. Konradi C, Cole RL, Heckers S, et al. Amphetamine regulates gene expression in rat striatum via transcription factor CREB. *J Neurosci* 1994;14:5623–5634.

60. Cole RL, Konradi C, Douglass J, et al. Neuronal adaptation to amphetamine and dopamine: molecular mechanisms of prodynorphin gene regulation in rat striatum. *Neuron* 1995;14:813–823.

61. Graybiel AM. Basal ganglia—input, neural activity, and relation to the cortex. *Curr Opin Neurobiol* 1991;1:644–651.

62. Cole AJ, Bhat RV, Patt C, et al. D1 dopamine receptor activation of multiple transcription factor genes in rat striatum. *J Neurochem* 1992;58:1420–1426.

63. Hope B, Kosofsky B, Hyman SE, et al. Regulation of immediate early gene expression and AP-1 binding in the rat nucleus accumbens by chronic cocaine. *Proc Natl Acad Sci USA* 1992;89:5764–5768.

64. Nguyen TV, Kosofsky BE, Birnbaum R, et al. Differential expression of c-fos and zif268 in rat striatum after haloperidol, clozapine, and amphetamine. *Proc Natl Acad Sci USA* 1992;89:4270–4274.

65. Hope BT, Nye HE, Kelz MB, et al. Induction of a long-lasting AP-1 complex composed of altered Fos-like proteins in brain by chronic cocaine and other chronic treatments. *Neuron* 1994;13:1235–1244.

66. Tzschentke TM, Schmidt WJ. Blockade of morphine- and amphetamine-induced conditioned place preference in the rat by riluzole. *Neurosci Lett* 1998;242:114–116.

67. Badiani A, Browman KE, Robinson TE. Influence of novel versus home environments on sensitization to the psychomotor stimulant effects of cocaine and amphetamine. *Brain Res* 1995;674:291–298.

68. Hinson RE, Poulos CX. Sensitization to the behavioral effects of cocaine: modification by Pavlovian conditioning. *Pharmacol Biochem Behav* 1981;15:559–562.

69. Pert A, Post R, Weiss SR. Conditioning as a critical determinant of sensitization induced by psychomotor stimulants. *NIDA Res Monogr* 1990 3:208–241.

70. Anagnostaras SG, Robinson TE. Sensitization to the psychomotor stimulant effects of amphetamine: modulation by associative learning. *Behav Neurosci* 1996;110:1397–1414.

71. Childress AR, McLellan AT, O'Brien CP. Role of conditioning factors in the development of drug dependence. *Psychiatr Clin North Ami* 1986;9:413–425.

72. Shiffman S. “Addiction versus stages of change models” vs. “addiction and stages of change models.” *Addiction* 1996;91:1289–1290.

73. Hyman SE. Shaking out the cause of addiction [comment]. *Science* 1996;273:611–612.

74. Berke JD, Paletzki RF, Aronson GJ, et al. A complex program of striatal gene expression induced by dopaminergic stimulation. *J Neurosci* 1998;18:5301–5310.

75. Fosnaugh JS, Bhat RV, Yamagata K, et al. Activation of arc, a putative “effector” immediate early gene, by cocaine in rat brain. *J Neurochem* 1995;64:2377–2380.

76. Cole AJ, Saffen DW, Baraban JM, et al. Rapid increase of an immediate early gene messenger RNA in hippocampal neurons by synaptic NMDA receptor activation. *Nature* 1989;340:474–476.

77. Lyford GL, Yamagata K, Kaufmann WE, et al. Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. *Neuron* 1995;14:433–445.

78. Brakeman PR, Lanahan AA, O'Brien R, et al. Homer: a protein that selectively binds metabotropic glutamate receptors [see comments]. *Nature* 1997;386:284–288.

79. O'Brien RJ, Xu D, Petralia RS, et al. Synaptic clustering of AMPA receptors by the extracellular immediate- early gene product Narp [see comments]. *Neuron* 1999;23:309–323.

80. Bourtchuladze R, Frenguelli B, Blendy J, et al. Deficient long-term memory in mice with a targeted mutation of the cAMP- responsive element-binding protein. *Cell* 1994;79:59–68.

81. Frey U, Frey S, Schollmeier F, et al. Influence of actinomycin D, a RNA synthesis inhibitor, on long-term potentiation in rat hippocampal neurons in vivo and in vitro. *J Physiol (Lond)* 1996;490(Pt 3):703–711.

82. Nguyen PV, Kandel ER. A macromolecular synthesis-dependent late phase of long-term potentiation requiring cAMP in the medial perforant pathway of rat hippocampal slices. *J Neurosci* 1996;16:3189–3198.

83. Frey U, Morris RG. Synaptic tagging and long-term potentiation [see comments]. *Nature* 1997;385:533–536.

84. Frey U, Huang YY, Kandel ER. Effects of cAMP simulate a late stage of LTP in hippocampal CA1 neurons. *Science* 1993;260:1661–1664.

85. Huang YY, Kandel ER. D1/D5 receptor agonists induce a protein synthesis-dependent late potentiation in the CA1 region of the hippocampus [see comments]. *Proc Natl Acad Sci USA* 1995;92:2446–2450.

86. Frey U, Matthies H, Reymann KG, et al. The effect of dopaminergic D1 receptor blockade during tetanization on the expression of long-term potentiation in the rat CA1 region in vitro. *Neurosci Lett* 1991;129:111–114.

87. Frey U, Schroeder H, Matthies H. Dopaminergic antagonists prevent long-term maintenance of posttetanic LTP in the CA1 region of rat hippocampal slices. *Brain Res* 1990;522:69–75.

88. Matthies H, Becker A, Schroeder H, et al. Dopamine D1-deficient mutant mice do not express the late phase of hippocampal long-term potentiation. *Neuroreport* 1997;8:3533–3535.

89. Wickens JR, Kotter R. Cellular models of reinforcement. In: Houk JC, Davis JL, Beiser DG. *Models of information processing in the basal ganglia*. Cambridge, MA: MIT Press, 1995:187–204.

90. Charpier S, Deniau JM, et al. In vivo activity-dependent plasticity at cortico-striatal connections: evidence for physiological long-term potentiation. *Proc Natl Acad Sci USA* 1997;94:7036–7040.

91. Calabresi P, Saiardi A, Pisani A, et al. Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. *J Neurosci* 1997;17:4536–4544.

92. Kombian SB, Malenka RC. Simultaneous LTP of non-. *Nature* 1994;368:242–246.

93. Wickens JR, Begg AJ, Arbuthnott GW. Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex in vitro. *Neuroscience* 1996;70:1–5.

94. Pennartz CM, Ameerun RF, Groenewegen HJ, et al. Synaptic plasticity in an in vitro slice preparation of the rat nucleus accumbens. *Eur J Neurosci* 1993;5:107–117.

95. Nguyen PV, Abel T, Kandel ER. Requirement of a critical period of transcription for induction of a late phase of LTP. *Science* 1994;265:1104–1107.

96. Robinson TE, Kolb B. Persistent structural modifications in nucleus accumbens and prefrontal cortex neurons produced by previous experience with amphetamine. *J Neurosci* 1997;17:8491–8497.

97

Recent Advances in Animal Models of Drug Addiction

Toni S. Shippenberg

George F. Koob

Toni S. Shippenberg: National Institutes of Health, National Institute on Drug Abuse, Bethesda, Maryland.

George F. Koob: Department of Neuropharmacology, The Scripps Research Institute, La Jolla, California.

DEFINITIONS AND VALIDATION OF ANIMAL MODELS

Part of "97 - Recent Advances in Animal Models of Drug Addiction "

Definitions of Drug Addiction

Drug addiction is defined as a compulsion to take a drug with loss of control in limiting intake (33). It is considered a chronic disorder because the risk of relapse remains high even after completion of treatment and prolonged abstinence. In 1968, the term drug dependence replaced that of addiction in the nomenclature of the World Health Organization and the American Psychiatric Association. Defined as a cluster of cognitive, behavioral, and physiologic symptoms indicative of an individual continuing substance use despite significant substance related problems, this term has become the accepted diagnostic term for compulsive use of a psychoactive substance. When defined as described, it is analogous to the term addiction. However, this term should not be confused with physical or psychic dependence, conditions in which the cessation or reduction of drug usage results in a withdrawal syndrome. Withdrawal and tolerance often are associated with compulsive drug use; however, they are not required for drug addiction. Although individuals suffering from chronic pain may develop tolerance to the analgesic effects of an opiate and experience withdrawal symptoms, they do not exhibit signs of compulsive drug-seeking behavior.

The concept of reinforcement has provided the cornerstone for current theories and animal models of drug addiction. A reinforcer is defined operationally as “any event that increases the probability of a response” and often is used interchangeably with “reward.” In general, drugs function as positive or conditioned reinforcers by virtue of their rewarding effects, and reward often connotes additional attributes of a drug (e.g., pleasure) that cannot be easily defined operationally.

This chapter reviews animal models currently used to examine the neurobiological basis of drug addiction and the role of reinforcement processes in its initiation, maintenance, and reinstatement. Emphasis is place on more recently developed models, and where possible, the models are evaluated in terms of reliability and predictability to the human condition. Potential pitfalls to consider when interpreting data also are discussed.

ANIMAL MODELS OF THE POSITIVE REINFORCING EFFECTS OF DRUGS

Part of "97 - Recent Advances in Animal Models of Drug Addiction "

Drugs of abuse function as positive reinforcing stimuli; this action has provided the framework for currently used animal models of addiction. It is also clear that humans and experimental animals will readily self-administer these agents in the absence of a withdrawal state. Earlier models of drug reinforcement used operant paradigms in nonhuman primates; however, many of these same paradigms now are utilized in rodents. The use of these rodent models, together with the development of modern neurobiological techniques, has provided important information regarding the neurobiology of addiction (11 ,15 ,36 ,45 ,51).

Operant Intravenous Drug Self-Administration

Drugs of abuse are readily self-administered intravenously by experimental animals, and, in general, drugs that are self-administered correspond to those that have high abuse potential. However, that not all drugs abused by humans are self-administered by experimental animals (e.g., hallucinogens). Furthermore, there are species- and strain-related differences in the degree to which a drug is self-administered (48 ,62). A detailed review of intravenous self-administration was presented in the previous edition (46); therefore, only key points are presented here.

For intravenous self-administration studies, the subject acquires a drug infusion by performing a discrete response. The number and pattern of responding required for each infusion is determined by the schedule of reinforcement imposed by the experimenter. Drug availability typically is signaled by an environmental stimulus. The dependent variables are the number of infusions obtained or the rate of responding during a session. In addition to the intravenous route of administration, the intragastric or oral route can be employed (see the following).

Simple Schedules

In fixed-ratio schedules, the number of responses required for drug infusion is set at a fixed number. In rodents, these schedules generally will not maintain stable responding below a certain unit dose and, within the range of doses that maintains stable responding, self-administration rate is inversely related to dose. Within the range of doses that maintains stable responding, animals increase their selfadministration rate as the unit dose is decreased. Conversely, animals reduce their rate of self-administration as the unit dose is increased. In a fixed-interval schedule, the frequency of injections is determined primarily by the interval schedule imposed and not the response rate. Therefore, the use of these two schedules can provide information regarding both the motivational effects of a drug and potential nonspecific motor effects that can confound data interpretation.

Progressive-Ratio Schedules

Progressive-ratio schedules are used to evaluate the reinforcing efficacy of a self-administered drug. In this procedure the response requirements for each successive drug reinforcement are increased and the *breaking point* (the point at which the animal will no longer respond) is determined (89). Breaking points are defined either as the largest ratio requirement that the subject completes or as the number of ratios completed by the subject per session. This value represents the maximum work a subject will perform to receive an infusion of a drug. Because the dependent measure in progressive-ratio schedules is not directly related to the rate of responding, interpretive problems associated with using rate of responding as a measure of reinforcement efficacy are avoided.

This schedule has been used to study the relative reinforcing efficacy of compounds both within and across drug classes (61) as well as the neural basis of drug reinforcement (58 ,61 ,78 ,89). Drug craving has been conceptualized as the incentive motivation to self-administer a drug that has been previously consumed; therefore, this schedule can also provide an animal model of drug craving. However, the breaking point reflects both the unconditioned (reinforcing) and conditioned incentive effects of drugs and does not allow for assessment of drug seeking in the absence of drug administration.

Multiple Schedules

Clinical definitions of drug addiction and dependence typically refer to the disruptive effects of addiction on non—drug-related activities. The use of multiple schedules of reinforcement enables the application of concepts of behavioral economics (e.g., commodities, consumption, price, and demand) to operant behavior. It also can provide a control for nonselective effects of drug reinforcement. In these procedures, self-administration of a drug is incorporated into a multiple component schedule with other reinforcers. Studies using these procedures have shown that the contingencies for concurrent reinforcers can affect behavior asymmetrically and that the response requirement for reinforcers can affect drug self-administration (12). Drugs also can function as substitutes for, complements of, or be independent from, the “price” of one another and can be interpreted in economic terms. In addition to evaluating the selectivity of manipulations that apparently reduce the reinforcing efficacy of a drug, these procedures can provide information regarding those factors affecting “loss of control,” as well as behavioral and/or pharmacologic therapies for the treatment of addiction. Behavior maintained by alternate presentation of natural reinforcers (e.g., food and water) and drugs of abuse in the same session and with the same reinforcement requirements has been reported for several species (16 ,20 ,97 ,98).

Second-Order Schedules

In a second-order schedule, completion of an individual component (or part) of the schedule produces the terminal event (drug infusion) according to another overall schedule. Typically, completion of a unit schedule results in the presentation of a brief stimulus, and completion of the overall schedule results in the delivery of the stimulus and the drug. Second-order schedules have the advantage that they maintain high rates of responding and extended sequences of behavior before any drug infusion occurs. Therefore, acute drug effects on response rates are minimized. High response rates are maintained even for doses that decrease rates when several injections are self-administered during a session (43). In addition, this schedule requires extended sequences of behavior, thus, modeling the human condition in which drug taking is preceded by a series of behaviors (e.g., procurement, preparation). Although second-order schedules are more typically used in nonhuman primate studies of addiction, their use in rodents is increasing (8 ,55 ,73).

Oral Drug-Self Administration

Oral self-administration has focused largely on alcohol because of the obvious face validity of oral alcohol self-administration

and because intravenous self-administration of alcohol is difficult to sustain in rodents (39).

Home Cage Drinking and Preference

A simple approach to studying the motivation to consume a drug is to measure the volume consumed when a drinking bottle is available in the home cage. These procedures have been particularly useful for characterizing genetic differences in drug preference, most often alcohol preference (53), and for initial studies on the effects of pharmacologic treatments on drug intake and preference. Usually a choice is offered between a drug solution and alternative solutions, one of which is often water, and the proportion of drug intake relative to total intake is calculated as a preference ratio. For two-bottle choice testing of alcohol in mice or rats, animals are singly housed and a bottle containing 10% alcohol and one containing water are placed on each cage. Most commonly, animals are allowed free choice of these drinking solutions for successive 24-hour periods with simultaneous free access to food. However, limited access to the drug can induce high drug intakes in short periods of time. Although alcohol is most often studied with these procedures, similar studies have been done with cocaine (40).

Operant Conditioning

A more “motivational” approach is to have animals work to obtain drugs for oral consumption using operant procedures. The advantages of such an approach are numerous. The effort to obtain the substance can be separated from the consummatory response (e.g., drinking) and intake easily can be charted over time. In addition, different schedules of reinforcement can be used to change baseline parameters.

For operant self-administration of alcohol, rats can be trained to lever press for alcohol using a variety of techniques all designed to overcome the aversive taste and after effects of initial exposure to alcohol. One approach involves using a sweetened solution fading procedure (80). Alcohol concentrations are increased to a final concentration of 10% over 20 days, with each concentration being mixed first with saccharin or sucrose and then presented alone. Using this approach, animals can be trained to lever press for concentrations of alcohol up to 40% (80). They will perform on fixed-ratio schedules and progressive-ratio schedules and obtain significant blood alcohol levels in a 30-minute session.

Operant self-administration of oral alcohol has also been validated as a measure of the reinforcing effects of alcohol in primates (93). Similar studies have been published with other drugs of abuse (60 ,90 ,93).

Reliability and Predictability of Self-Administration Procedures

Drug self-administration has both reliability and predictive validity. The dependent variable provides a reliable measure of the motivation to obtain drugs (e.g., the amount of work an animal will perform to obtain drug) or, in an alternative framework, in demonstrating that drugs function as powerful reinforcers. Responding maintained by drugs as reinforcers is stable across sessions and can be altered predictably by neurotransmitter antagonists. Intravenous drug selfadministration also has predictive validity. Drugs having high reinforcement potential in experimental animals have reinforcing effects in humans as measured by both operant and subjective reports (49).

Potential Pitfalls

As discussed previously, self-administration of a drug can vary as a function of the dose available, species or strain tested and the duration of self-administration sessions. It also is influenced by the availability of alternate reinforcers, the presence or absence of environmental stimuli that signal drug infusions, post-reinforcement interval, and prior history of the subject (54). In the progressive-ratio paradigm, breaking point is influenced by the size of the increment by which each ratio increases as well as the initial response requirement that starts a session. Because self-administration typically results in an inverted U-shaped curve, both leftward and rightward shifts in the dose–effect function will decrease self-administration of some unit doses but simultaneously increase self-administration of other doses. The interpretation of downward shifts in the dose–effect function also is sometimes problematic. Therefore, construction of full dose–effect functions is essential in self-administration studies. Since a given pretreatment may decrease self-administration by having nonspecific effects on behavior (e.g., sedation), the influence of a pretreatment on non-drug reinforcers should be assessed.

Conditioned Place Preference

Conditioned place preference is a classical conditioning procedure in which administration of a drug is paired with one distinct environment and administration of placebo with another. After several environmental pairings, allowing noninjected animals access to both environments and measuring the time spent in each assesses the time spent in each environment. The animal’s choice to spend more time in either environment provides a direct measure of the conditioned reinforcing effect of a drug. Animals exhibit a conditioned preference for an environment associated with drugs that function as positive reinforcers (e.g., spend more time in the drug-paired compared to placebo-paired environment) and avoid those that induce aversive states (e.g., conditioned

place aversion). This procedure permits assessment of the conditioning of drug reinforcement and can provide indirect information regarding the positive and negative reinforcing effects of drugs. Place conditioning has been used in conjunction with gene transfer and homologous recombination techniques to delineate the neural basis of druginduced reinforcement (15 ,74).

The apparatus used in conditioning experiments consists of two environments that are differentiated from each other on the basis of color, texture, and/or lighting. The distinctiveness of the environments is essential for the development of conditioning. In the unbiased design, the environments are manipulated so that animals differentiate one from the other but do not exhibit an innate preference for either of the place cues. Pairing of drug with a particular environment is counterbalanced and change in the time spent in the drug-paired environment can be directly attributed to the conditioned reinforcing effects of a drug. Although quality control experiments confirming the unbiased nature of the procedure are conducted periodically, experiments do not require a preconditioning phase to assess pretest preferences, thus preventing the potential confound of latent inhibition and decreasing the time necessary for a particular experiment.

In the biased design, animals exhibit a preference for one of the place cues prior to conditioning. A preconditioning phase, in which animals are allowed access to both environments, is necessary to determine the innate preference of each animal. The drug then is paired with the preferred or nonpreferred environment depending on whether the drug is assumed to produce aversive or positive reinforcing effects, respectively. Although this design is used often, data interpretation can be problematic because place preferences may indicate incentive motivational effects of a drug or a decrease in the aversive properties of the least-preferred environment.

Reliability and Predictability of Conditioned Place Preference Procedures

The conditioned place preference paradigm has reliability and validity. Drugs that produce conditioned preferences for the drug-associated environment are those that function as positive reinforcers in other paradigms. Conditioned aversions also are observed in response to drugs that are negative reinforcers or produce aversive or dysphoric states in human subjects (34 ,66).

Potential Pitfalls

In place conditioning studies, the drug is administered noncontingently and there is evidence that the behavioral and neurochemical effects of abused drugs differ depending on whether drug administration is controlled by the subject.

Route of drug administration, number of environmental pairings, and duration of conditioning sessions (18 ,23) can profoundly affect place conditioning and should be controlled for. Because tests of conditioning are conducted in the absence of drug, the issue of state-dependency also must be addressed.

Place conditioning now is used in many studies assessing genotype-dependent differences in drug sensitivity. However, a lack of a conditioned response may indicate a loss of the reinforcing effects of a drug or a generalized impairment of learning or memory processes required for the acquisition or performance of a conditioned response. In addition, genotype-dependent differences in the saliency of environmental cues used for conditioning may occur. Finally, issues of interpretation and latent inhibition limit the utility of biased place conditioning procedures.

Brain Stimulation Reward Thresholds

Electrical self-stimulation of certain brain areas is rewarding for animals and humans as demonstrated by the fact that subjects will readily self-administer the stimulation (69). The powerful nature of the reward effect produced by intracranial self-stimulation (ICSS) is indicated by the behavioral characteristics of the ICSS response, which include rapid learning and vigorous execution of the stimulation-producing behavior. (See ref. 28 for review.) The high reward value of ICSS has led to the hypothesis that ICSS directly activates neuronal circuits that are activated by conventional reinforcers (e.g., food, water, and sex). In bypassing much of the input side of these neuronal circuit(s), ICSS provides a unique tool to investigate the influence of various substances on reward and reinforcement processes. ICSS differs significantly from drug self-administration in that, in the ICSS procedure, the animal is working to directly stimulate presumed reinforcement circuits in the brain, and the effects of the drugs are assessed on these reward thresholds. Drugs of abuse decrease thresholds for ICSS, and there is a good correspondence between the ability of drugs to decrease ICSS thresholds and their abuse potential (47).

Many ICSS procedures have been developed over the years, but an important methodologic advance has been the development of procedures that provide a measure of reward threshold that is unconfounded by influences on motor and performance capability. These are the rate-frequency curve-shift procedure, and the discrete-trial, current-intensity procedure (28 ,47 ,64). These have been reviewed in detail previously (46) and are not discussed here.

Potential Pitfalls

Brain stimulation reward has the advantage of directly interfacing with brain reward circuits and as such eliminates any interference with consummatory-like behaviors. In addition, it is a validated and reliable measure of brain reward. Potential pitfalls, however, include the requirement for surgery

(e.g., the implantation of electrodes). The surgery itself is routine but does require specialized equipment. Another variable in this domain is the brain site selected. Some brain regions support higher rates of brain stimulation reward than others and there may be different circuits activated by different sites.

In addition, animals need to be trained for several weeks to obtain stable rates of responding or stable thresholds. This training requirement and the extensive surgical requirements virtually force the use of within-subject designs. As a result, steps must be taken to address order-effects and analyze such potential confounds.

Animal Models of the Subjective Effects of Drugs: Drug Discrimination

The use of the drug discrimination paradigm in studies of drug addiction is based on two hypotheses. First, the same components of a drug's actions subserve discriminative stimulus effects in animals and subjective effects in humans. Second, discriminative stimulus effects of drugs may contribute to drug taking in intermittent users and to relapse of addiction in former drug addicts. In this latter view, discriminative stimuli signal the availability of a reinforcer and therefore set the occasion to engage in those behaviors that enable consumption of the reinforcing drug. Evidence has been obtained that stimuli predictive of drug administration elicit drug-seeking and -taking behavior and can retard the extinction of responding for psychostimulants (24 ,87 ,97) suggesting that the discriminative stimulus effects of a drug contribute to the genesis of these behaviors.

In a typical experiment, animals are trained to emit a particular response following administration of a fixed drug dose (e.g., depression of one lever designated the drug-associated lever) and to press another lever (saline designated lever) following administration of saline under a fixed schedule of reinforcement. Most commonly, an appetitively motivated operant procedure is used in which animals are food or water deprived. Responding on the training-condition appropriate lever results in the delivery of food or water. Training is continued until the animal reliably selects the appropriate lever after drug or saline administration. Once trained, tests of stimulus generalization or antagonism are implemented to determine whether other doses of the training drug or a specific drug treatment produces stimulus effects qualitatively similar to or different from that of the training drug.

As with other operant paradigms, various reinforcement schedules (e.g., fixed-ratio, fixed-interval, differential reinforcement of low response rate) and response measures (e.g., nose poking, maze running) can be used. Dose 1 versus dose 2 and drug 1 versus drug 2 versus saline discriminations also can be employed. Details can be found in the following references (13 ,29 ,46 ,72).

Reliability and Predictability of Drug Discrimination Procedures

Drug discrimination offers both reliability and predictive validity. The dependent variable is very reliable as a measure of the interoceptive effects of drugs. Stimulus generalization gradients are stable once drug discrimination is acquired and neurotransmitter antagonists alter the stimulus effects of various drugs predictably. Drug discrimination also has predictive validity in that drugs that produce discriminative stimulus effects that generalize to known drugs of abuse have been shown to have abuse liability.

Potential Pitfalls

Generalization gradients are dependent on the dose of drug used for training. Certain neurotransmitter antagonists attenuate the discriminative stimulus effects of a drug when a low training dose is employed. However, these same antagonists fail to modify the discriminative stimulus effects of the same drug when a higher training is employed (41). Similarly, generalization to partial agonists or mixed agonists/antagonists can differ depending on the training dose employed (19); therefore, the use of multiple training doses is essential.

Different test procedures (extinction versus reinforced responding on the lever on which the first schedule requirement is completed) may yield different results depending on the variable used to measure generalization. As with all animal models, species and strain differences as well as the experimental history of an animal can alter the discriminative stimulus effects of a drug (94). Finally, subtle differences in the discriminative stimulus effects of a drug may occur depending on whether appetitive or aversively maintained responding is employed.

ANIMAL MODELS OF THE NEGATIVE REINFORCING EFFECTS OF DRUG WITHDRAWAL

Part of "97 - Recent Advances in Animal Models of Drug Addiction "

Withdrawal from chronic drug administration usually is characterized by responses opposite to the acute initial actions of the drug. Many of the overt physical signs associated with withdrawal from drugs (e.g., alcohol and opiates) can be quantified easily. However, motivational measures of abstinence have proven to be more sensitive measures of drug withdrawal and powerful tools for exploring the neurobiological bases for the motivational aspects of drug dependence.

Operant Drug Self-Administration

Drug self-administration can be conducted under conditions in which animals are rendered physically dependent

on the drug (e.g., abstinence from drug use results in a withdrawal syndrome), and the procedures are similar to those discussed in the preceding. Although it is clear that animals will self-administer drugs in the absence of withdrawal, some evidence suggests that physical dependence can increase the reinforcing efficacy of a drug. Monkeys made physically dependent on morphine show increases in their progressive-ratio performance compared to animals that do not exhibit withdrawal symptomology (101). Also, baboons in a discrete-trials choice procedure for food and heroin showed significant behavioral plasticity when allowed periodic access to heroin or food (20). In the withdrawal state, one would hypothesize that the animals would be much less likely to respond for food, even if the cost of heroin in terms of response requirements was dramatically increased. Thus, the reinforcing value of drugs may change depending on the presence or absence of a withdrawal state. The neurobiological basis for such a change is only beginning to be investigated, but much evidence has been generated to show that drug withdrawal can produce an aversive or negative motivational state that is manifested by changes in a number of behavioral measures including response disruption, increased drug intake, changes in reward thresholds, and place aversions.

Recent studies with alcohol have shown that rats with a history of self-administration of alcohol will self-administer alcohol during withdrawal in sufficient quantities to prevent withdrawal symptoms and maintain blood alcohol levels above 100 mg% (75). To assess the relationship of withdrawal severity, blood alcohol levels, and alcohol self-administration in dependent and nondependent rats, rats were trained to lever press for 10% alcohol versus water using the saccharin fadeout procedure and subjected to induction of dependence on alcohol (75). Dependent animals allowed to respond for alcohol during a second 12-hour test period showed sustained alcohol intake that maintained blood alcohol levels above 100 mg% throughout the 12-hour period, and a virtual elimination of alcohol withdrawal scores (75) (Fig. 97.1). Animals not allowed access to alcohol during withdrawal on a third test showed a precipitous drop in blood alcohol levels and a dramatic increase in withdrawal scores (75) (Fig. 97.1). These results show that rats will maintain and sustain lever pressing for alcohol during dependence if the animals have a history of lever pressing for alcohol to the point of suppressing alcohol withdrawal and maintaining blood alcohol levels.

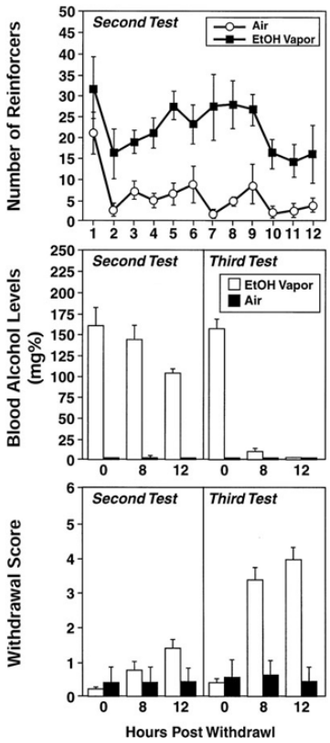


FIGURE 97.1. Operant responding for alcohol across a 12-hour test period by air-exposed and alcohol vapor-exposed rats *(top)*. In addition, blood alcohol levels *(middle)* and alcohol withdrawal severity *(bottom)* obtained during test 2 (while rats were allowed access to alcohol in the operant boxes) and test 3 (while in home cages) are shown. The animals were divided into two groups. One group of animals was assigned to 2 weeks of alcohol exposure in alcohol vapor chambers. The second group was exposed to control air. Rats then were tested in the operant boxes with access to 10% alcohol and water across two 12-hour periods separated by 4 days of vapor exposure. A third and final withdrawal phase was included after another 4 days of vapor exposure; however, animals were kept in their home cages and not allowed to respond for alcohol. Blood was collected for blood alcohol determinations, and observational withdrawal signs were rated during tests 2 and 3 at 0, 8, and 12 hours post withdrawal. Data are expressed as means ± SEM. Taken with permission from Roberts AJ, Cole M, Koob GF. Intra-amygdala muscimol decreases operant ethanol self-administration in dependent rats. *Alcohol Clin Exp Res* 1996;20:1289–1298.

Responding for Non-Drug Reinforcers

Several operant schedules have been used to characterize the response-disruptive effects of drug withdrawal (37 ,84). However, response disruption can be caused by any number of variables from motor problems to malaise and decreases in appetite, and thus other measures must be used to rule out nonspecific effects (see the following).

Conditioned Place Aversion

The conditioned place preference paradigm can also be used to characterize the conditioned aversive effects of drug withdrawal. Rodents are exposed to one environment while undergoing withdrawal and to another in the absence of a withdrawal state. During tests of conditioning, animals are allowed access to both environments and the time spent in each is determined. To date, this procedure has been used almost exclusively to study withdrawal from opiate drugs. Administration of opioid receptor antagonists to animals rendered physically dependent on morphine via implantation of morphine pellets or repeated injections of an opiate produces dose-related conditioned place aversions, an effect that can be observed after only a single conditioning session with the antagonist (27 ,34). In contrast, the administration of the same doses of antagonist to opiate-naive animals fails to produce a conditioned response. Interestingly, the minimum effective dose of an antagonist that produces conditioned place aversions in animals physically dependent on morphine is less than that producing quantifiable somatic withdrawal signs suggesting that this technique is a particularly sensitive model for evaluating the affective component of drug withdrawal. Although place conditioning typically has been used to characterize antagonist-precipitated withdrawal, more recent work indicates its utility for studies of spontaneous withdrawal (10).

Brain Stimulation Reward

ICSS thresholds have been used to assess changes in systems mediating reinforcement processes during the course of drug dependence. Although no actual negative reinforcement is measured using this technique, it is included in this section because it constitutes a model of the aversive motivational state associated with the negative reinforcement of drug abstinence in dependent animals. Acute administration of psychostimulant drugs lowers ICSS threshold (i.e., increases ICSS reward) (47), and withdrawal from chronic administration of these same drugs elevates ICSS thresholds (i.e., decreases ICSS reward) (22 ,44 ,56) (Fig. 97.2). Similar results have been observed with precipitated withdrawal in opiate-dependent rats (82). Rats showed dramatic increases in ICSS thresholds to naloxone injections that occurred in a dose-related manner and at doses below which obvious physical signs of opiate withdrawal were manifest. These doses of naloxone had no effect on reward thresholds in nondependent animals.

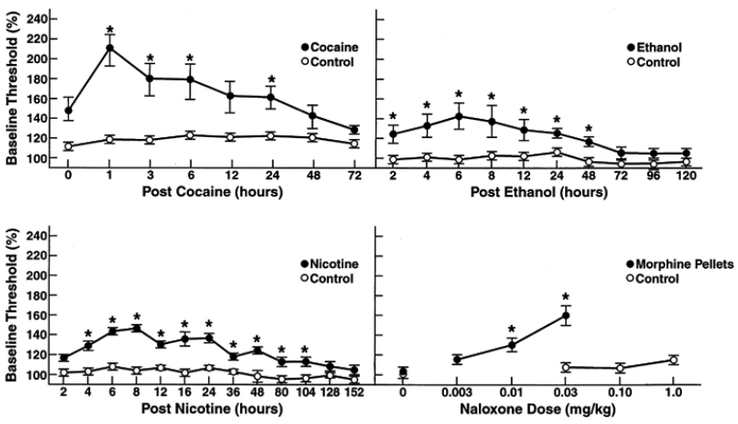


FIGURE 97.2. Changes in reward threshold associated with chronic administration of four major drugs of abuse. Reward thresholds were determined using a rate-independent discrete-trials threshold procedure for intracranial self-stimulation (ICSS) of the medial forebrain bundle. **A:** Rats equipped with intravenous catheters were allowed to self-administer cocaine for 12 straight hours prior to withdrawal and reward threshold determinations. Elevations in threshold were dose-dependent with longer bouts of cocaine self-administration yielding larger and longer-lasting elevations in reward thresholds. Taken with permission from Markou A, Koob GF. Postcocaine anhedonia: an animal model of cocaine withdrawal. *Neuropsychopharmacology* 1991;4:17–26. **B:** Elevations in reward thresholds with the same ICSS technique following chronic exposure to alcohol of approximately 200 mg% in alcohol vapor chambers. Taken with permission from Schulteis G, Markou A, Cole M, et al. Decreased brain reward produced by ethanol withdrawal. *Proc Natl Acad Sci USA* 1995;92:5880–5884. **C:** Elevations in reward thresholds during spontaneous withdrawal after termination of chronic administration of nicotine hydrogen tartrate (9.0 mg/kg per day for 7 days; *n* = 8) or saline (*n* = 6). Taken with permission from Epping-Jordan MP, Watkins SS, Koob GF, Markou A. Dramatic decreases in brain reward function during nicotine withdrawal. *Nature* 1998;393:76–79. **D:** Elevations in reward thresholds following administration of very low doses of the opiate antagonist naloxone to animals made dependent on morphine using two, 75-mg morphine (base) pellets implanted subcutaneously. Taken with permission from Schulteis G, Markou A, Gold LH, et al. Relative sensitivity to naloxone of multiple indices of opiate withdrawal: A quantitative dose-response analysis. *J Pharmacol Exp Ther* 1994;271:1391–1398. Asterisks (*) refer to significant differences between treatment and control values. Values are mean ±SEM

Drug Discrimination

Drug discrimination can be used to characterize both specific and nonspecific aspects of withdrawal. Generalization to an opiate antagonist provides a more general nonspecific measure of opiate withdrawal intensity and time course (26 ,30). Examples of a more specific aspect of withdrawal are animals that have been trained to discriminate pentylenetetrazol, an anxiogenic-like substance, from saline in alcohol-, diazepam-, and opiate-dependent animals. During withdrawal, generalization to the pentylenetetrazol cue has suggested an anxiogenic-like component to the withdrawal syndrome (14 ,21).

Ethological Measures

Animals models of withdrawal that illustrate aversive stimulus effects can be extended to observational measures, some of which may be common to withdrawal from many different drugs of abuse. Increased anxiety-like responses are observed following abstinence from cocaine, opiates, benzodiazepines, and alcohol (9 ,25 ,79 ,81 ,85). Measures used to assess anxiety-like responses during include validated animal models of anxiety such as the elevated plus-maze, light-dark test, defensive withdrawal, and defensive burying.

Advantages and Disadvantages of Animal Models of the Negative Reinforcing Effects of Drugs

The advantages and disadvantages of models used to evaluate drug withdrawal are similar to those described for the positive reinforcing effects of drugs. Clearly, each of the paradigms described has weaknesses, but when combined can provide powerful insights into the motivational effects of drug abstinence.

ANIMAL MODELS OF ESCALATION IN DRUG INTAKE

Part of "97 - Recent Advances in Animal Models of Drug Addiction "

Animal Models of Sensitization to the Reinforcing Effects of Drugs

Preclinical studies have shown that the repeated intermittent administration of psychostimulants, opiates, and alcohol can result in a long-lasting enhancement of their behavioral effects (92). This phenomenon, referred to as sensitization, has been implicated in the psychosis that occurs in some individuals following repeated psychostimulant use. A role of sensitization in both vulnerability to drug addiction and drug craving has been hypothesized (77). Both self-administration and conditioned place preference procedures have been used to evaluate sensitization to the reinforcing effects of drugs in experimental animals.

Intravenous Self-Administration

In self-administration studies, sensitization to the positive reinforcing effects of drugs is assessed. Typically, animals receive daily, noncontingent injections of a drug or placebo.

Self-administration sessions are then initiated. The rate of acquisition of self-administration and/or the number of animals acquiring stable drug self-administration then is determined. Because sensitization is defined as an increase in the potency and/or efficacy of a drug in producing a particular response following its repeated administration, the rate of acquisition of drug self-administration should be increased and the threshold dose effective in supporting self-administration should be decreased. Several laboratories have shown that the rate of acquisition of psychostimulant self-administration is increased in animals that have received noncontingent injections of these agents indicating the development of sensitization (38 ,71). The prior administration of amphetamine also increases the acquisition rate of cocaine self-administration (and, conversely, the prior administration of cocaine increases the acquisition rate of amphetamine self-administration), suggesting that cross-sensitization develops to the positive reinforcing effects of psychostimulants.

Conditioned Place Preference

The conditioning procedure used to study sensitization is identical to that described above except that the dose of conditioning drug or the number of environmental pairings used typically are those that are ineffective in producing a conditioned response in previously drug-naive animals. Animals receive repeated noncontingent administration of a drug, and place conditioning can be initiated at various time points following the cessation of drug administration. An increase in the potency of a drug provides a direct measure of sensitization (Fig. 97.3). Alternatively, by employing doses that are subthreshold and threshold for producing a conditioned response, changes in drug potency and efficacy following prior drug exposure can be determined. Using these procedures, long-lasting sensitization and cross-sensitization to the conditioned reinforcing effects of opiates and psychostimulants has been shown (50 ,52 ,88).

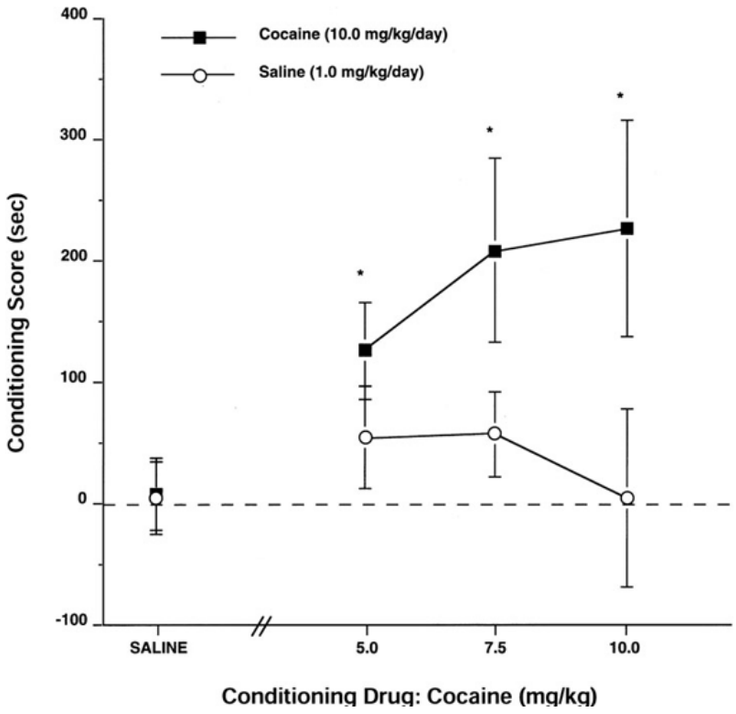


FIGURE 97.3. Sensitization to the conditioned reinforcing effects of cocaine. Rats received once daily home cage injections of cocaine or saline for 5 days. Place conditioning (two cocaine; two saline) commenced 3 days later. Cocaine was ineffective in producing a conditioned response after two environmental pairings. In animals with a prior history of cocaine, doses of cocaine as low as 5.0 mg/kg produced significant conditioned place preferences. Ordinate: Conditioning score defined as time in drug-paired environment minus time in saline-paired environment. Asterisks (*) denote significant place conditioning. Taken with permission from Shippenberg TS, Heidbreder C. Sensitization to the conditioned rewarding effects of cocaine: pharmacologic and temporal characteristics. *J Pharmacol Exp Ther* 1995;273:808–815.

Escalation in Drug Self-Administration Produced by a History of Drug Intake

A critical issue for the study of the neurobiology of addiction is to develop animal models for the transition between controlled/moderate drug intake and uncontrolled/excessive drug intake. Animal models of increased drug intake based on prolonged exposure to drug now have been described in rats for cocaine, heroin, and alcohol (1 ,2 and 3 ,76).

The pattern of drug self-administration dramatically differs depending on the duration of access. With 1 hour of access to cocaine, drug intake remained at the level of training intake and was stable over time. In contrast, with 6 hours of access per session, cocaine intake gradually escalated to levels significantly above the training baseline (Fig. 97.4). The dose–effect function was shifted up and not to the right or left (2). Abstinence of a month returned the escalated intake to pre-escalation baseline, but escalation was reinstated rapidly at a level higher than that seen before abstinence.

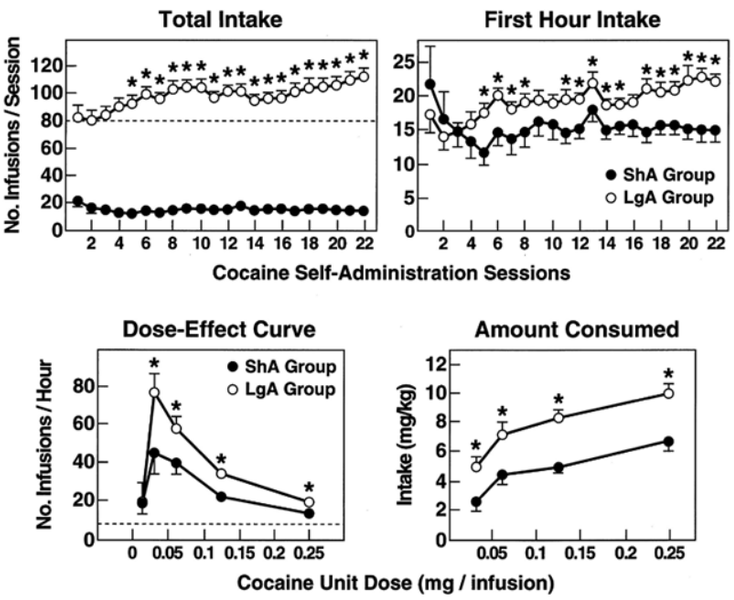


FIGURE 97.4. Reproduction of escalation of cocaine use. **A:** In Long-Access (LgA) rats (*n* = 12) but not in Short-Access (ShA) rats (*n* = 12), mean total cocaine intake (±SEM) started to increase significantly from session 5 ($p < .05$; sessions 5 to 22 compared to session 1) and continued to increase thereafter ($p < .05$; session 5 compared to sessions 8 to 10, 12, 13, and 17 to 22). **B:** During the first hour, LgA rats self-administered more infusions than ShA rats during sessions 5 to 8, 11, 12, 14, 15, and 17 to 22 ($p < .05$). **C:** Mean infusion (±SEM) per cocaine dose tested. LgA rats took significantly more infusions than ShA rats at doses of 31.25, 62.5, 125, and 250 µg per infusion ($p < .05$). **D:** After escalation, LgA rats took more cocaine than ShA rats regardless of the dose ($p < .05$). *($p < 0.05$ (Student's *t* test after appropriate one-way and two-way analysis of variance). Taken with permission from Ahmed SH, Koob GF. Transition from moderate to excessive drug intake: change in hedonic set point. *Science* 1998;282:298–300.

Similar results have been observed in rats trained to self-administer heroin intravenously. Two groups of rats were trained on 1-hour continuous access to intravenous heroin self-administration and then one group was allowed access for 11 hours continuously. In the animals with 11-hour access, intake gradually increased over time, whereas in the animals continued on 1-hour access there was no change in intake over time. The animals with 11-hour access to heroin were slower to extinguish heroin-seeking behavior.

Animals with a history of alcohol exposure sufficient to produce dependence show a similar increase in baseline alcohol intake long after acute withdrawal (76). Operant oral alcohol self-administration was established in rats and then animals were exposed to alcohol vapor chambers for a sufficient period to produce physical dependence on alcohol, detoxified, and then allowed a 2-week period of protracted

abstinence. Operant responding was enhanced during protracted abstinence by 30% to 100% and remained elevated for 4 to 8 weeks post acute withdrawal.

ANIMAL MODELS OF RELAPSE: CONDITIONED REINFORCING EFFECTS OF DRUGS

Part of "97 - Recent Advances in Animal Models of Drug Addiction "

The role of environmental stimuli in the control of drug-taking behavior is a major focus of addiction research. This interest stems from the view that any account of drug abuse must address those factors that precede and motivate drug taking, as well as those that underlie the reinforcing consequences of drug delivery. Environmental cues repeatedly paired with primary reinforcers can acquire incentive properties via classical conditioning processes (57 ,87 ,97). It has been postulated that these conditioned reinforcing effects contribute to drug craving and relapse to addiction. Indeed, human studies have shown that the presentation of stimuli previously associated with drug delivery increases the likelihood of relapse as well as self-reports of craving and motivation to engage in drug taking (17 ,68).

Positive Reinforcing Effects of Stimuli Associated with Drug Self-Administration: Conditioned Reinforcement Paradigm

The conditioned reinforcement paradigm allows characterization of the incentive value imparted on formerly neutral environmental stimuli that have been repeatedly associated with drug self-administration. In this paradigm, subjects usually are trained in an operant chamber containing two levers. Responses on one lever result in the presentation of a brief stimulus followed by a drug injection (active lever)

whereas responses on the other lever have no consequences (inactive lever) (86). The ability of the previous neutral, drug-paired stimulus to maintain responding in the absence of drug injections provides a measure of the reinforcing value of these stimuli. This procedure provides a stringent test for the conditioned incentive effects of drugs because responding for drug-associated stimuli occurs under extinction conditions (e.g., in the absence of drug). It also provides an animal model of drug craving because the incentive motivational effects of a stimulus are examined in the absence of drug taking.

Second-Order Schedules

Second-order schedules also can be used to evaluate the conditioned reinforcing effects of drugs. To assess the effects of conditioned reinforcement, the number of responses with the paired stimulus can be compared to the number of responses with a nonpaired stimulus. For example, substitution of drug-paired stimuli with nondrug-paired stimuli actually can decrease response rates (43). This maintenance of performance with drug-paired stimuli appears to be analogous to the maintenance and reinstatement of drug seeking in humans with the presentation of drug-paired stimuli (17). In rats, a decrease in responding and an increase in the latency to initiate responding occurs in response to withholding a stimulus paired with cocaine self-administration (8). The schedule can be repeated several times during a test session, resulting in multiple infusions of drug. However, drug craving in the absence of drug can be assessed by terminating sessions immediately after the first drug infusion that occurs after completion of the terminal schedule.

Extinction with and without Cues Associated with Intravenous Drug Self-Administration

Extinction procedures provide measures of the incentive or motivational effects of drugs by assessing the persistence of drug-seeking behavior in the absence of response–contingent drug availability. In this paradigm, subjects first are trained to self-administer a drug until stable self-administration patterns are exhibited. Extinction sessions are identical to training sessions except that no drug is delivered after the completion of the response requirement.

Measures provided by an extinction paradigm reflect the degree of resistance to extinction and include the duration of extinction responding and the total number of responses emitted during the entire extinction session. The probability of reinstating responding under extinction conditions with drug-paired stimuli or even stimuli previously paired with drug withdrawal can be examined.

Both stimulant and opiate self-administration have been consistently reinstated following priming injections of drug (31 ,55). Responding during extinction is greater in the presence of the conditioned stimulus than in its absence (73). Similar results have been obtained in an operant runway task (57). It is also apparent that environmental stimuli predictive of cocaine self-administration reliably elicit drug-seeking behavior in experimental animals and that responding for these stimuli is highly resistant to extinction (39 ,87 ,97).

Reinstatement of Extinguished Drug-Seeking Behavior in an Animal Model of Relapse: Use of Discriminative Stimuli

Rat models of “relapse” induced by drug-related stimuli also can involve the use of a drug-predictive discriminative stimulus (S?). This stimulus is paired with response-contingent presentation of a stimulus that has been contiguously paired with drug presentations (i.e., a conditioned stimulus, or CS) to elicit recovery of responding at a previously active lever after prior extinction of alcohol-seeking behavior. Discriminative stimuli signal the availability of a reinforcer, and thereby provide motivation to engage in behavior that brings the organism into contact with the reinforcer. A condition often associated with drug craving in humans is cognitive awareness of drug availability (63). Discriminative stimuli, therefore, may have a prominent role in craving and the resumption of drug-seeking behavior in abstinent individuals. Moreover, the response-contingent CS, acting as a conditioned reinforcer, may contribute to the maintenance of subsequent drug-seeking behavior once initiated. In fact, these contingencies can be conceptualized to resemble those associated with the relapse process in humans in that certain drug-related cues may provide the initial central motivational state to engage in drug-seeking behavior, whereas others may maintain this behavior until the primary reinforcer is obtained.

To investigate the role of drug-associated stimuli in the motivational effects of a history of cocaine self-administration, rats were trained to associate discriminative stimuli (SΔ) with response-contingent availability of intravenous cocaine versus saline (97) (Fig. 97.5). The rats then were subjected to repeated extinction sessions during which cocaine, saline, and the respective SΔ were withheld until the rats reached extinction. Subsequent re-exposure to the cocaine SΔ, but not the nonreward SΔ, produced strong recovery of responding at the previously active lever in the absence of any further drug availability. The behavioral significance of the cocaine SΔ was further confirmed by the fact that the rats initially tested in the presence of the nonreward SΔ showed complete recovery of responding when subsequently presented with the cocaine SΔ, but rats that had shown robust reinstatement ceased responding when later tested under nonreward SΔ conditions. These results support the hypothesis that learned responses to drugrelated environmental stimuli can be important factors in the reinstatement of drug-seeking in animals and provide

a powerful model for elucidating the neuropharmacologic basis for such effects that are related to the human concepts of relapse and craving (97).

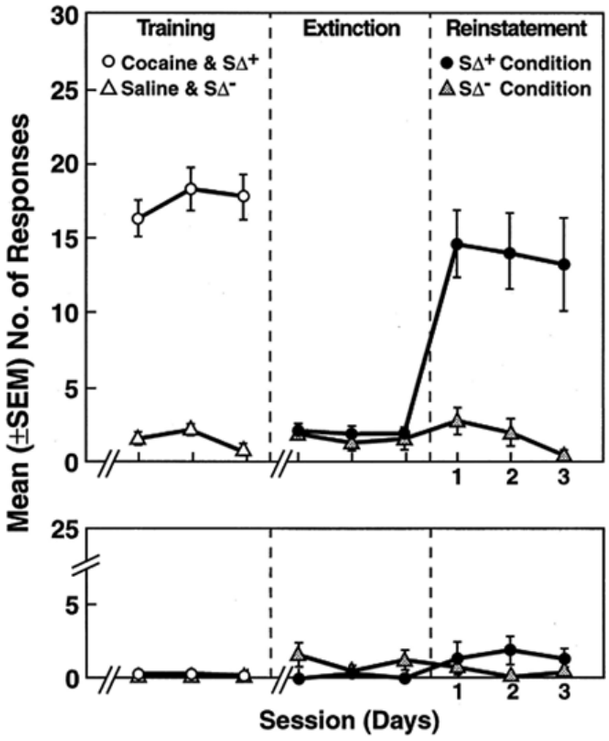


FIGURE 97.5. Lever-press responses during self-administration training, extinction, and reinstatement sessions at the active *(A)* and inactive *(B)* lever. Training phase: cocaine-reinforced responses during the final 3 days of the self-administration phase in rats (n = 15) trained to associate S?s with the availability of intravenous cocaine (S^+) versus saline (S^-). No differences were observed between responses during the first and second daily hour of cocaine availability, and responses for cocaine or saline between rats designated for testing under S^+ versus S^- conditions during the initial 3 days of the reinstatement phase. The data were, therefore, collapsed across groups and daily cocaine sessions for the purpose of this illustration. Extinction phase: extinction responses at criterion. The extinction criterion (= 4 responses per session over 3 consecutive days) was reached within 16.4 ± 3.8 days (averaged across rats designated for testing under the S^+ versus S^- condition during the reinstatement phase). Reinstatement phase: responses under the S^+ (n = 7) and S^- (n = 8) reinstatement conditions. Exposure to the S^+ elicited significant recovery of responding in the absence of further drug availability. Responding in the presence of the S^- remained at extinction levels. Taken with permission from Weiss F, Maldonado-Vlaar CS, Parsons LH, et al. Control of cocaine-seeking behavior by drug-associated stimuli in rats: effects on recovery of extinguished operant-responding and extracellular dopamine levels in amygdala and nucleus accumbens. *Proc Natl Acad Sci USA* 2000;97:4321-4326.

Cues associated with oral self-administration and availability of alcohol also can reinstate responding in the absence of the primary reinforcer (42 ,96). In addition, and consistent with the well-established conditioned cue reactivity in human alcoholics, the motivating effects of alcohol-related stimuli are highly resistant to extinction in that they retain their efficacy in eliciting alcohol-seeking behavior over more than 1 month of repeated testing (96).

Place Conditioning

Place conditioning procedures can be modified to serve as a model of relapse. Place aversions to opiate withdrawal last for over 8 weeks (94) and are resistant to extinction. Attempts to modify such conditioned effects could hypothetically contribute to knowledge of the factors that contribute to relapse or "craving." One also could envisage the use of cue- and drug-induced reinstatement of an extinguished place conditioning response as a measure of relapse (67).

Reliability and Predictability

Each of the techniques described has reliability and predictive validity. Presentation of stimuli associated with drug injection induces drug craving in humans and maintains responding in the conditioned reinforcement, second-order schedule, and extinction paradigms. The presence or absence of cues associated with drug administration alters the reinstatement of extinguished drug-seeking behavior in predictable ways.

CONCLUSIONS AND FUTURE RESEARCH

Part of "97 - Recent Advances in Animal Models of Drug Addiction "

Although it is very difficult to find an animal model of any psychiatric disorder that mimics the entire syndrome, one can reasonably validate animal models for different symptoms of mental disorders (32). In the realm of addiction research, the observation that animals readily self-administer drugs has led to arguments of face validity. Although intravenous drug self-administration meets the criteria of reliability, predictability, and face validity, it does not represent the whole syndrome of addiction (see the following). Other aspects of the addiction syndrome can indeed be modeled, but again, it is incorrect to consider any one of these an animal model of addiction. The DSM-IV criteria for substance dependence and animal models relevant to their study are summarized in Table 97.1 .

TABLE 97.1. ANIMAL MODELS FOR THE CRITERIA OF DSM-IV

DSM-IV	Animal Models
A. A maladaptive pattern of substance use, leading to clinically significant impairment or distress as occurring at any time in the same 12-month period:	
(1) Need for markedly increased amounts of substance to achieve intoxication or desired effect; or markedly diminished effect with continued use of the same amount of substance	(1) Tolerance to reinforcing effects: Cocaine Opiates
(2) The characteristic withdrawal syndrome for substance; or substance (or closely related substance) is taken to relieve or avoid withdrawal symptoms	(2) Increased: Reward thresholds: Cocaine, Opiates, Alcohol, Nicotine, Tetrahydrocannabinol Anxiety-like responses: Cocaine, Opiates, Alcohol, Tetrahydrocannabinol
(3) Persistent desire or one or more unsuccessful efforts to cut down or control substance use	(3) Conditional positive reinforcing effects: Cocaine Opiates Alcohol
(4) Substance used in larger amounts or over a longer period than the person intended	(4) Cocaine binge Opiate intake (dependent animals) Alcohol intake (dependent animals) Alcohol Deprivation Effect
(5) Important social, occupational, or recreational activities given up or reduced because of substance use	(5) Choice paradigms Behavioral economics—loss of elasticity
(6) A great deal of time spent in activities necessary to obtain substance, to use substance, or to recover from its effect	(6) Opiate self-administration during withdrawal Alcohol self-administration during withdrawal Progressive-ratio responding
(7) Continued substance use despite knowledge of having a persistent problem that is likely to be caused or exacerbated by substance use	(7) Cocaine binge toxicity

From: American Psychiatric Association, 1994.

Tolerance (criterion 1) and withdrawal (criterion 2) no longer define addiction, as illustrated by the change in criteria outlined in DSM-III versus DSM-IIIR and DSM-IV (5 ,6 and 7); however, evidence is accumulating to suggest that a common element associated with addiction is a motivational form of withdrawal that is reflected in a compromised brain reward system (see the preceding). This not only reaffirms the importance of withdrawal in addiction (e.g., criterion 2: "characteristic withdrawal syndrome") but also adds the dimension of a persistent motivational change that may be reflected in criteria 7 of the DSM-IV: "continued use

despite knowledge of having had a persistent psychological problem that is likely to be exacerbated by the substance" (Table 97.1).

The two DSM-IV criteria that are best modeled by drug self-administration are criteria 3 and 4, respectively: "the persistent desire to cut down or control substance use" and "the substance taken in larger amounts than intended." Drugs of abuse are readily self-administered by animals and, in general, drugs that are self-administered correspond to those that have high abuse potential in humans (60 ,61).

The chronic relapsing nature of drug addiction is perhaps best illustrated by criterion 3 of Table 97.1 : "persistent desire or one or more unsuccessful attempts to cut down or control substance use." Two difficult states to define that are related to relapse are craving and protracted abstinence. Presumably, such states reflect some prolonged post acute withdrawal perturbation or vulnerability to reinstatement of drug-seeking behavior and ultimately compulsive use. A residual deficit state in the reward system or sensitization of the reward system to stimuli that predict drug effects, or some combination of both, could be responsible for this vulnerability (see the preceding).

Animal models of drug craving and relapse continue to be developed but to date have reflected secondary sources of reinforcement such as conditioned reinforcement (91) or residual changes in motivational state or a combination of the two. Second-order schedules can be used as a measure of the conditioned reinforcing properties of drugs (43). Recent work suggests that reliable responding for cocaine can be obtained with a second-order schedule in rats (99). The conditioned place preference paradigm also provides a measure of conditioned reinforcement that is conceptually similar to the measures provided by the operant paradigms. More recently, stimuli that predict drug availability have been shown to be powerful cues for reinstating drug-seeking behavior (97), and a history of drug intake produces escalation in drug intake (1 ,3).

The remaining criterion for substance dependence as defined by DSM-IV can be linked to animal models by extension to the models described in the preceding. "Substance taken in large amounts or over a longer period of time than the person intended" (criterion 4) clearly is reflected in animal models of self-administration with unlimited access, or in situations of limited access where reinforcement value

is challenged by dose–effect functions or progressive-ratio procedures. “Important social, occupational, or recreational activities given up because of substance use” (criterion 5) has been demonstrated in animal models involving choice procedures (16) and involving behavioral economics paradigms (20). “A great deal of time spent in activities necessary to obtain the substance” (criterion 6) is reflected in animal models of drug self-administration during withdrawal (see the preceding). Finally, “continued substance use despite knowledge of having a persistent or recurrent physical or psychological problem that is exacerbated by the drug” (criterion 7) may be reflected in animal models of toxicity associated with chronic drug self-administration such as with cocaine, or prolonged changes in reward thresholds following chronic drug exposure (83).

Drug addiction in humans has been characterized as occurring in several stages, although progress from one stage to the next is not inevitable. The first stage is initiation or acquisition, which may lead to habitual use, physical or psychic dependence, and loss of control. An individual may stop taking a drug at any stage; however, relapse to drug taking after a period of abstinence is common. The extent to which the procedures discussed here model the human condition to the point of reliability and predictive validity requires further assessment.

Animal models of addiction are critical for advances in the study of addiction. Addiction is a chronic relapsing disorder comprised of multiple stages and multiple sources of reinforcement. As discussed, the motivating factors for the development, maintenance, and persistence of drug addiction can be broken down into four major sources of reinforcement: positive reinforcement, negative reinforcement, conditioned positive reinforcement, and conditioned negative reinforcement (100). Much progress has been made in identifying the neuronal substrates for the acute positive reinforcing effects of drugs of abuse. A more recent focus has been on the neuronal substrates for negative reinforcement and the conditioned reinforcing effects that contribute to relapse. The future challenge will be to explore the mechanisms involved in animal models of craving and relapse and to relate these mechanisms to vulnerability to addiction. A major advantage of animal models is in the translation of the human condition to the animal model (face validity) and the translation of the neurobiological measures back to the human condition in order to predict vulnerability (predictive validity).

Perhaps the best example of the translation value of animal models of addiction is the development of medications for the treatment of drug abuse (4 ,35). The opiate antagonist naltrexone long has been known to block self-administration of alcohol (4 ,95), and preclinical studies eventually led to the use of naltrexone to successfully prevent relapse in detoxified alcoholics (70).

Animal models are critical for the delineation of genetic and environmental factors that lead to and predict vulnerability to addiction. Context-independent drug administration (e.g., experimenter-administered drugs) can provide information about brain changes associated with a history of drug administration and that alter sensitivity to the effects of a drug. However, drug administration in the context of sensitive and validated animal models provides a much more powerful means of linking drug actions and sensitivities to biological and environmental perturbations. The successful implementation of procedures designed to assess functional genomic activity (e.g., screening for changes in the expression of gene activity) to the study of addiction, will require animal models that are reliable and have predictive validity if they are to contribute to our understanding of the neurobiology of drug addiction.

ACKNOWLEDGMENTS

Part of "97 - Recent Advances in Animal Models of Drug Addiction "

The authors would like to thank Mike Arends for his help with the preparation of this manuscript. This is publication number 13454-NP from The Scripps Research Institute. Research was supported by National Institutes of Health grants AA06420 and AA08459 from the National Institute on Alcohol Abuse and Alcoholism, DK26741 from the National Institute of Diabetes and Digestive and Kidney Diseases, and DA04043, DA04398 and the Intramural Research Program from the National Institute on Drug Abuse.

REFERENCES

1. Ahmed SH, Koob GF. Transition from moderate to excessive drug intake: change in hedonic set point. *Science* 1998;282:298–300.

2. Ahmed SH, Koob GF. Long-lasting increase in the set point for cocaine self-administration after escalation in rats. *Psychopharmacology* 1999;146:303–312.

3. Ahmed SH, Walker JR, Koob GF. Persistent increase in the motivation to take heroin in rats with a history of drug escalation. *Neuropsychopharmacology* 2000;22:413–421.

4. Altshuler HL, Phillips PE, Feinhandler DA. Alteration of ethanol self-administration by naltrexone. *Life Sci* 1980;26:679–688.

5. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, third ed. Washington, DC: American Psychiatric Association, 1980.

6. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, third ed (rev). Washington, DC: American Psychiatric Association, 1987.

7. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, fourth ed (rev). Washington, DC: American Psychiatric Association, 1994.

8. Arroyo M, Markou A, Robbins TW, et al. Acquisition, maintenance and reinstatement of intravenous cocaine self-administration under a second-order schedule of reinforcement in rats: effects of conditioned cues and continuous access to cocaine. *Psychopharmacology* 1998;140:331–344.

9. Basso AM, Spina M, Rivier J, et al. Corticotropin-releasing factor antagonist attenuates the “anxiogenic-like” effect in the defensive burying paradigm but not in the elevated plus-maze following chronic cocaine in rats. *Psychopharmacology* 1999;145:21–30.

10. Bechara A, Nader K, van der Kooy D. Neurobiology of withdrawal motivation: evidence for two separate aversive effects produced in morphine-naive versus morphine-dependent rats by both naloxone and spontaneous withdrawal. *Behav Neurosci* 1995;109:91–105.

11. Berke JD, Hyman SE. Addiction, dopamine, and the molecular mechanisms of memory. *Neuron* 2000;25:515–532.

12. Bickel WK, DeGrandpre RJ, Higgins ST. The behavioral economics of concurrent drug reinforcement: a review and reanalysis of drug self-administration research. *Psychopharmacology* 1995;118:250–259.

13. Bowen CA, Gatto GJ, Grant KA. Assessment of the multiple discriminative stimulus effects of ethanol using an ethanol-pentobarbital-water discrimination in rats. *Behav Pharmacol* 1997;8:339–352.

14. Brandt MR, France CP. Chronic l-alpha acetylmethadol in rhesus monkeys: discriminative stimulus and other behavioral measures of dependence and withdrawal. *J Pharmacol Exp Ther* 1998;287:1029–1037.

15. Carlezon WA Jr, Thome J, Olson VG, et al. Regulation of cocaine reward by CREB. *Science* 1998;282:2272–2275.

16. Carroll ME, Lac ST, Nygaard SL. A concurrently available nondrug reinforcer prevents the acquisition or decreases the maintenance of cocaine-reinforced behavior. *Psychopharmacology* 1989;97:23–29.

17. Childress AR, McLellan AT, Ehrman R, et al. Classically conditioned responses in opioid and cocaine dependence: A role in relapse? *NIDA Res Monogr* 1988;84:25–43.

18. Cunningham CL, Dickinson SD, Grahame NJ, et al. Genetic differences in cocaine-induced conditioned place preference in mice depend on conditioning trial duration. *Psychopharmacology* 1999;146:73–80.

19. Dykstra LA, Preston KL, Bigelow GE. Discriminative stimulus and subjective effects of opioids with mu and kappa activity: data from laboratory animals and human subjects. *Psychopharmacology* 1997;130:14–27.

20. Elsmore TF, Fletcher GV, Conrad DG, et al. Reduction of heroin intake in baboons by an economic constraint. *Pharmacol Biochem Behav* 1980;13:729–731.

21. Emmett-Oglesby MW, Mathis DA, Moon RT, et al. Animal models of drug withdrawal symptoms. *Psychopharmacology* 1990;101:292–309.

22. Epping-Jordan MP, Watkins SS, Koob GF, et al. Dramatic decreases in brain reward function during nicotine withdrawal. *Nature* 1998;393:76–79.

23. Ettenberg A, Raven MA, Danluck DA, et al. Evidence for opponent-process actions of intravenous cocaine. *Pharmacol Biochem Behav* 1999;64:507–512.

24. Falk JL, Lau CE. Stimulus control of addictive behavior: persistence in the presence and absence of a drug. *Pharmacol Biochem Behav* 1995;50:71–75.

25. File SE, Baldwin HA, Aranko K. Anxiogenic effects in benzodiazepine withdrawal are linked to the development of tolerance. *Brain Res Bull* 1987;19:607–610.

26. France CP, Woods JH. Discriminative stimulus effects of naltrexone in morphine-treated rhesus monkeys. *J Pharmacol Exp Ther* 1989;250:937–943.

27. Funada M, Shippenberg TS. Differential involvement of D1 and D2 dopamine receptors in the expression of morphine withdrawal in rats. *Behav Pharmacol* 1996;7:448–453.

28. Gallistel CR. Self-stimulation. In: Deutsch JA, ed. *The physiologic basis of memory*, second ed. New York: Academic Press, 1983:73–77.

29. Gauvin DV, Holloway FA. Cue dimensionality in the 3-choice pentylenetetrazole-saline-chlordiazepoxide discrimination task. *Behav Pharmacol* 1991;2:417–428.

30. Gellert VF, Holtzman SG. Discriminative stimulus effects of naltrexone in the morphine-dependent rat. *J Pharmacol Exp Ther* 1979;211:596–605.

31. Gerber GJ, Stretch R. Drug-induced reinstatement of extinguished self-administration behavior in monkeys. *Pharmacol Biochem Behav* 1975;3:1055–1061.

32. Geyer MA, Markou A. Animal models of psychiatric disorders. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:787–798.

33. Goldstein A. *Addiction: from biology to drug policy*. New York: WH Freeman, 1994.

34. Hand TH, Koob GF, Stinus L, et al. Aversive properties of opiate receptor blockade: evidence for exclusively central mediation in naive and morphine-dependent rats. *Brain Res* 1988;474:364–368.

35. Heyser CJ, Schulteis G, Durbin P, et al. Chronic acamprosate eliminates the alcohol deprivation effect while having limited effects on baseline responding for ethanol in rats. *Neuropsychopharmacology* 1998;18:125–133.

36. Hiroi N, Brown JR, Haile CN, et al. FosB mutant mice: loss of chronic cocaine induction of Fos-related proteins and heightened sensitivity to cocaine’s psychomotor and rewarding effects. *Proc Natl Acad Sci USA* 1997;94:10397–10402.

37. Holtzman SG, Villarreal JE. Operant behavior in the morphine-dependent rhesus monkey. *J Pharmacol Exp Ther* 1973;184:528–541.

38. Horger BA, Giles MK, Schenk S. Preexposure to amphetamine and nicotine predisposes rats to self-administer a low dose of cocaine. *Psychopharmacology* 1992;107:271–276.

39. Hyytia P, Schulteis G, Koob GF. Intravenous heroin and ethanol self-administration by alcohol-preferring AA and alcohol-avoiding ANA rats. *Psychopharmacology* 1996;125:248–254.

40. Jentsch JD, Henry PJ, Mason PA, et al. Establishing orally self-administered cocaine as a reinforcer in rats using home-cage pre-exposure. *Prog Neuropsychopharmacol Biol Psychiatry* 1998;22:229–239.

41. Kantak KM, Riberdy A, Spealman RD. Cocaine-opioid interactions in groups of rats trained to discriminate different doses of cocaine. *Psychopharmacology* 1999;147:257–265.

42. Katner SN, Magalong JG, Weiss F. Reinstatement of alcohol-seeking behavior by drug-associated discriminative stimuli after prolonged extinction in the rat. *Neuropsychopharmacology* 1999;20:471–479.

43. Katz JL, Goldberg SR. Second-order schedules of drug injection: implications for understanding reinforcing effects of abused drugs. *Adv Substance Abuse* 1991;4:205–223.

44. Kokkinidis L, McCarter BD. Postcocaine depression and sensitization of brain-stimulation reward: Analysis of reinforcement and performance effects. *Pharmacol Biochem Behav* 1990;36:463–471.

45. Koob GF, Le Moal M. Drug abuse: hedonic homeostatic dysregulation. *Science* 1997;278:52–58.

46. Koob GF. Animal models of drug addiction. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York, 1995:759–772.

47. Kornetsky C, Esposito RU. Euphorigenic drugs: effects on the reward pathways of the brain. *Federation Proc* 1979;38:2473–2476.

48. Kuzmin A, Johansson B. Reinforcing and neurochemical effects of cocaine: differences among C57, DBA and 129 mice. *Pharmacol Biochem Behav* 2000;65:399–406.

49. Lamb RJ, Preston KL, Schindler CW, et al. The reinforcing and subjective effects of morphine in post-addicts: a dose-response study. *J Pharmacol Exp Ther* 1991;259:1165–1173.

50. Le Pen G, Duterte-Boucher D, Costentin J. Sensitization to the rewarding effects of the specific dopamine uptake inhibitor GBR12783. *J Pharmacol Exp Ther* 1998;286:688–696.

51. Leshner AI, Koob GF. Drugs of abuse and the brain. *Proc Assoc Am Phys* 1999;111:99–108.

52. Lett BT. Repeated exposures intensify rather than diminish the rewarding effects of amphetamine, morphine, and cocaine. *Psychopharmacology* 1989;98:357–362.

53. Li TK. Pharmacogenetics of responses to alcohol and genes that influence alcohol drinking. *J Stud Alcohol* 2000;61:5–12.

54. Macenski MJ, Meisch RA. Ratio size and cocaine concentration effects on oral cocaine-reinforced behavior. *J Exp Anal Behav* 1998;70:185–201.

55. Markou A, Arroyo M, Everitt BJ. Effects of contingent and noncontingent cocaine on drug-seeking behavior measured using a second-order schedule of cocaine reinforcement in rats. *Neuropsychopharmacology* 1999;20:542–555.

56. Markou A, Koob GF. Postcocaine anhedonia: an animal model of cocaine withdrawal. *Neuropsychopharmacology* 1991;4:17–26.

57. McFarland K, Ettenberg A. Reinstatement of drug-seeking behavior produced by heroin-predictive environmental stimuli. *Psychopharmacology* 1997;131:86–92.

58. McGregor A, Roberts DCS. Effect of medial prefrontal cortex injections of SCH23390 on intravenous cocaine self-administration under both a fixed and progressive ratio schedule of reinforcement. *Behav Brain Res* 1995;67:75–80.

59. Meil WM, See RE. Conditioned cued recovery of responding following prolonged withdrawal from self-administered cocaine in rats: an animal model of relapse. *Behav Pharmacol* 1996;7:754–763.

60. Meisch RA, Lemaire GA, Cutrell EB. Oral self-administration of pentobarbital by rhesus monkeys: relative reinforcing effects under concurrent signalled differential-reinforcement-of-low-rates schedules. *Drug Alcohol Depend* 1992;30:215–225.

61. Mello NK, Lukas SE, Bree MP, et al. Progressive ratio performance maintained by buprenorphine, heroin and methadone in Macaque monkeys. *Drug Alcohol Depend* 1988;21:81–97.

62. Mello NK, Negus SS. Preclinical evaluation of pharmacotherapies for treatment of cocaine and opioid abuse using drug self-administration procedures. *Neuropsychopharmacology* 1996;14:375–424.

63. Meyer RE, Mirin SM. *The heroin stimulus: implications for a theory of addiction*. New York: Plenum Press, 1979.

64. Miliaressis E, Rompre PP, Laviolette P, et al. The curve-shift paradigm in self-stimulation. *Physiol Behav* 1986;37:85–91.

65. Morgan D, Cook CD, Picker MJ. Sensitivity to the discriminative stimulus and antinociceptive effects of mu opioids: role of strain of rat, stimulus intensity, and intrinsic efficacy at the mu opioid receptor. *J Pharmacol Exp Ther* 1999;289:965–975.

66. Mucha RF, Herz A. Motivational properties of kappa and mu opioid receptor agonists studies with place and taste preference conditioning. *Psychopharmacology* 1985;86:274–280.

67. Mueller D, Stewart J. Cocaine conditioned place preference: Endurance, extinction and cocaine-induced reinstatement. *Society Neurosci Abstr* 1999;25:1574.

68. O'Brien CP, Childress AR, McLellan AT, et al. Classical conditioning in drug-dependent humans. *Ann NY Acad Sci* 1992;654:400–415.

69. Olds J, Milner P. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J Comp Physiol Psychol* 1954;47:419–427.

70. O'Malley SS, Jaffe AJ, Chang G, et al. Naltrexone and coping skills therapy for alcohol dependence: a controlled study. *Arch Gen Psychiatry* 1992;49:881–887.

71. Piazza PV, Deminiere JM, LeMoal M, et al. Factors that predict individual vulnerability to amphetamine self-administration. *Science* 1989;245:1511–1513.

72. Picker MJ, Cook CD. Discriminative stimulus effects of opioids in pigeons trained to discriminate fentanyl, bremazocine and water: Evidence of pharmacologic selectivity. *Behav Pharmacol* 1997;8:160–173.

73. Ranaldi R, Roberts DCS. Initiation, maintenance and extinction of cocaine self-administration with and without conditioned reward. *Psychopharmacology* 1996;128:89–96.

74. Risinger FO, Bormann NM, Oakes RA. Reduced sensitivity to ethanol reward, but not ethanol aversion, in mice lacking 5-HT1B receptors. *Alcohol Clin Exp Res* 1996;20:1401–1405.

75. Roberts AJ, Cole M, Koob GF. Intra-amygdala muscimol decreases operant ethanol self-administration in dependent rats. *Alcohol Clin Exp Res* 1996;20:1289–1298.

76. Roberts AJ, Heyser CJ, Cole M, et al. Excessive ethanol drinking following a history of dependence: animal model of allostasis. *Neuropsychopharmacology* 2000;22:581–594.

77. Robinson TE, Berridge KC. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Rev* 1993;18:247–291.

78. Rocha BA, Fumagalli F, Gainetdinov RR, et al. Cocaine self-administration in dopamine-transporter knockout mice. *Nature Neurosci* 1998;1:132–137.

79. Rodriguez de Fonseca F, Carrera MRA, Navarro M, et al. Activation of corticotropin-releasing factor in the limbic system during cannabinoid withdrawal. *Science* 1997;276:2050–2054.

80. Samson HH. Initiation of ethanol reinforcement using a sucrose-substitution procedure in food- and water-sated rats. *Alcohol Clin Exp Res* 1986;10:436–442.

81. Sarnyai Z, Biro E, Gardi J, et al. Brain corticotropin-releasing factor mediates "anxiety-like" behavior induced by cocaine withdrawal in rats. *Brain Res* 1995;675:89–97.

82. Schulteis G, Carrera R, Markou A, et al. Motivational consequences of naloxone-precipitated opiate withdrawal: A dose-response analysis. *Society Neurosci Abstr* 1993;19:1247.

83. Schulteis G, Markou A, Cole M, et al. Decreased brain reward produced by ethanol withdrawal. *Proc Natl Acad Sci USA* 1995;92:5880–5884.

84. Schulteis G, Markou A, Gold LH, et al. Relative sensitivity to naloxone of multiple indices of opiate withdrawal: A quantitative dose-response analysis. *J Pharmacol Exp Ther* 1994;271:1391–1398.

85. Schulteis G, Yackey M, Risbrough V, et al. Anxiogenic-like effects of spontaneous and naloxone-precipitated opiate withdrawal in the elevated plus-maze. *Pharmacol Biochem Behav* 1998;60:727–731.

86. Schuster CR, Woods JH. The conditioned reinforcing effects of stimuli associated with morphine reinforcement. *Int J Addict* 1968;3:223–230.

87. See RE, Grimm JW, Kruzich PJ, et al. The importance of a compound stimulus in conditioned drug-seeking behavior following one week of extinction from self-administered cocaine in rats. *Drug Alcohol Depend* 1999;57:41–49.

88. Shippenberg TS, Heidbreder C. Sensitization to the conditioned rewarding effects of cocaine: pharmacologic and temporal characteristics. *J Pharmacol Exp Ther* 1995;273:808–815.

89. Stafford D, LeSage MG, Glowa JR. Progressive-ratio schedules of drug delivery in the analysis of drug self-administration: a review. *Psychopharmacology* 1998;139:169–184.

90. Stewart BS, Lemaire GA, Roache JD, et al. Establishing benzodiazepines as oral reinforcers: Midazolam and diazepam self-administration in rhesus monkeys. *J Pharmacol Exp Ther* 1994;271:200–211.

91. Stewart J. Conditioned stimulus control of the expression of sensitization of the behavioral activating effects of opiate and stimulant drugs. In: Gormezano I, Wasserman EA, eds. *Learning and memory: the behavioral and biological substrates*. Hillsdale, NJ: LEA, 1992:129–151.

92. Stewart J, Badiani A. Tolerance and sensitization to the behavioral effects of drugs. *Behav Pharmacol* 1993;4:289–313.

93. Stewart RB, Bass AA, Wang NS, et al. Ethanol as an oral reinforcer in normal weight rhesus monkeys: dose-response functions. *Alcohol* 1996;13:341–346.

94. Stinus L, Caille S, Koob GF. Opiate withdrawal-induced place aversion lasts for up to 16 weeks. *Psychopharmacology* 2000;149:115–120.

95. Volpicelli JR, Alterman AI, Hayashida M, et al. Naltrexone in the treatment of alcohol dependence. *Arch Gen Psychiatry* 1992;49:876–880.

96. Weiss F. Ethanol-associated environmental stimuli potently reinstate ethanol-seeking behavior following extinction and abstinence. In satellite symposium entitled: CNS Mechanisms in Alcohol Relapse. Research Society on Alcoholism Annual Meeting, Denver, CO, 2000.

97. Weiss F, Maldonado-Vlaar CS, Parsons LH, et al. Control of cocaine-seeking behavior by drug-associated stimuli in rats: effects on recovery of extinguished operant-responding and extracellular dopamine levels in amygdala and nucleus accumbens. *Proc Natl Acad Sci USA* 2000;97:4321–4326.

98. Weissenborn R, Yackey M, Koob GF, et al. Measures of cocaine-seeking behavior using a multiple schedule of food and drug self-administration in rats. *Drug Alcohol Depend* 1995;38:237–246.

99. Whitelaw RB, Markou A, Robbins TW, et al. Excitotoxic lesions of the basolateral amygdala impair the acquisition of cocaine-seeking behaviour under a second-order schedule of reinforcement. *Psychopharmacology* 1996;127:213–224.

100. Wikler A. Dynamics of drug dependence: implications of a conditioning theory for research and treatment. *Arch Gen Psychiatry* 1973;28:611–616.

101. Yanagita T. An experimental framework for evaluations of dependence liability of various types of drugs in monkeys. *Bull Narc* 1973;25:57–64.

98

Vulnerability Factors for Alcoholism

Marc A. Schuckit

M. A. Schuckit: Department of Psychiatry, Veterans Affairs San Diego Healthcare System, San Diego, California.

This is an exciting and challenging time in the search for genes that have impact on the risk for alcohol abuse and dependence (alcoholism). Family, twin, and adoption studies offer solid evidence that genetic factors contribute to the risk for severe and repetitive alcohol-related life problems, and at least 30 genetically influenced characteristics are being evaluated for their possible impact on the alcoholism risk (1 ,2 ,3 and 4). Similar to other complex genetic disorders, these risk factors are heterogeneous, and combine to explain an estimated 60% of the variance, often interacting with environmental forces that contribute to the remaining 40% (1 ,2 and 3). Complicating the search for specific genetic influences even further is the probability that most of the characteristics, or phenotypes, are themselves influenced by multiple genes, and the lack of precision of the definition of the broad phenotype, alcohol abuse, or dependence (1 ,4).

This chapter reviews the progress made during the 7 years since the publication of the previous edition (5) in the search for genes that influence the risk for alcoholism. Most of the genetic studies have used variations in two approaches described in more detail in Chapter 99 , as well as in additional recent review articles (1 ,3 ,6). *Candidate gene studies* (also called case-control, association, or forward genetic investigations) begin with the hypothesis that a specific gene is related to the characteristic under study (7). If a limited number of multiple forms, or polymorphisms, of the gene are known, the proportion of individuals with and without the characteristic who have the specified polymorphism can be determined. Although this approach is an important step in identifying specific genes related to a phenotype, it is prone to false-positive results that occur if there are additional differences between groups with and without the characteristic (8). This problem, called population stratification, can be overcome by two variations of case-control studies where the relationship between the genetic marker and the characteristic, or phenotype, is evaluated among related individuals using a transmission disequilibrium or a haplotype relative risk approach (3 ,7).

The second and usually more labor-intensive technique is the *genetic linkage study*, or genome scan. This approach requires determining the presence of the phenotype and gathering blood for genotyping from either multiple generations of a large number of families or a large number of sibling pairs. The relationship between the phenotype and genetic signposts, or markers, across the 23 chromosomes is then evaluated. Unfortunately, genome scans are likely to identify only relatively powerful genes that explain a substantial proportion of the risk, and the data are best analyzed only when the mode of inheritance (e.g., dominant or recessive) is known. The potential linkage of a particular characteristic to a specific signpost helps identify areas of chromosomes of interest, after which more focused candidate gene analyses can be used to test individual genes near that marker.

This chapter describes some of the more promising results of the application of such approaches to alcohol use disorders, and briefly synthesizes the wide range of phenotypic and genotypic markers into a framework focusing on the several possible themes of potentially related characteristics described in Table 98.1 . Reflecting space limitations, emphasis is placed on genetic factors relating to alcoholism, rather than to drugs in general; most studies focus on human rather than animal work; and review articles are often highlighted rather than a series of primary data sources.

Level of response (LR)
P3/disinhibition/ASPD/type 2/B
Independent axis I disorders
Opioids
Alcohol metabolizing enzymes

ASPD, antisocial personality disorder.

TABLE 98.1. POSSIBLE BROAD FAMILIES OF RISK FACTORS

SOME SPECIFIC PHENOTYPIC AND GENETIC MARKERS OF INTEREST

Alcohol has impact on many neurochemical systems, whereas the use and consequences associated with this drug relate to many additional factors including personality traits, reinforcement, craving, and withdrawal symptoms (9 ,10). This panoply of phenomena has produced a wide range of hypotheses regarding how different phenotypes and candidate genes might contribute to the acquisition of drinking behaviors and the development of associated problems. Thus, judgment was needed to decide which potential phenotypic or genetic markers to highlight, and the reader is encouraged to turn to additional reviews (1 ,6 ,11 ,12 and 13). The information offered below is, somewhat arbitrarily, divided into results relevant to more broadly based phenotypes and those related to specific enzymes or genes.

Some Potentially Interesting Broad Phenotypes

A Low Level of Response (LR) to Alcohol

This characteristic, first proposed in 1975, reflects the need for higher doses of alcohol to produce an effect (1 ,10 ,14). The low LR to alcohol might enhance the probability of heavy drinking, encourage the formation of peer groups with similar drinking habits, and facilitate the acquisition of tolerance. LR is classically measured as the level of change in subjective feelings of intoxication, motor performance, hormone levels, and/or electrophysiologic measures observed at specific blood alcohol concentrations, but can also be evaluated by a self-report of the number of drinks required for specific effects (10 ,15 ,16). Low LRs are seen in about 40% of the estimated 700 children of alcoholics evaluated in the majority of the studies, and has been reported to characterize Native Americans, while high LRs and lower alcoholism risks have been noted for Jews and some Asian groups (17 ,18 ,19 ,20 ,21 ,22 ,23 ,24 and 25).

A longitudinal study of 453 20-year-old sons of alcoholics and controls showed that a low LR was a significant predictor of later alcohol abuse or dependence, explaining a significant proportion of the relationship between family history and alcoholism and alcoholic outcome in that sample (14 ,15). A 10-year follow-up of 64 sons of alcoholics and controls in Denmark found a significant relationship between less alcohol-induced physiologic changes and a high risk for alcohol abuse or dependence (26), whereas a 4-year follow-up of about 100 men and women found a relationship between a lower intensity of alcohol-induced incoordination and future alcohol problems for the men (27). Finally, a 10-year follow-up of almost 400 members of Australian twin pairs confirmed the relationship of the intensity of response to the alcoholic outcome for men (28).

The LR to alcohol is genetically influenced (29 ,30 ,31 ,32 and 33). Studies in animals have identified quantitative trait loci (QTLs) associated with LR, and lower reactions to alcohol are often associated with higher alcohol intake (29 ,30). In humans, the LR is more similar in identical than fraternal twin pairs, and among first-degree relatives compared to unrelated individuals (28 ,32 ,33).

A candidate gene study reported additive relationships to LR for alleles for the γ-aminobutyric acid receptor Aα6 ($GABA_{A\alpha6}$) and the serotonin transporter gene, whereas a separate sample revealed areas of potential linkages on chromosomes 1 and 21 (32 ,33). Additional candidate genes potentially tied to LRs include neuropeptide Y (NPY), neurotensin, adenylyl cyclase (AC) systems, protein kinase, aspects of serotonin (5-HT), and the hypothalamic-pituitary-adrenal (HPA) axis. However, a low LR appears to operate relatively independently of the event-related potential (ERP) and personality variables discussed in below, and is generally independent of alcohol metabolizing enzymes, but might crossover with some EEG characteristics (see below) (19 ,24 ,34).

Low Voltage or Low Alpha and EEGs

An overall low-voltage EEG pattern and a relative paucity of synchronized EEG waveforms appear to characterize a variety of conditions including anxiety and depression (6 ,34 ,35). The importance of the EEG results reflects reports that alcohol increases alpha power, and that alcoholics might show lower amounts of alpha, higher beta activity, and/or lower voltage on the EEG overall (24 ,36). Such EEG patterns might have impact on the risk for alcoholism via several different mechanisms, including higher levels of anxiety.

Some children of alcoholics exhibit lower amounts of slower alpha and/or higher levels of faster alpha or beta EEG activity, and alcohol increases the proportion of slower alpha, whereas other offspring demonstrate lower overall EEG voltage (6 ,24 ,34 ,35 ,36 and 37). Finally, children of alcoholics demonstrate less intense, or shorter lasting EEG changes with alcohol, and there is a significant correlation between a low LR and EEG findings in the same subjects (24 ,34).

Evidence supporting genetic influences in EEG patterns includes greater similarities in identical versus fraternal twins, and a report of mendelian inheritance for slow alpha (38 ,39). Some relevant EEG characteristics might relate to genes on chromosome 20, although not all reports agree (40).

Impulsivity, the Antisocial Personality Disorder (ASPD), Types 2 and B Alcoholism

Conduct disorder (CD) in childhood and subsequent ASPD represent persisting patterns of impulsivity and difficulty benefiting from punishment, often associated with aggressive and criminal behavior. These disorders are potentially

related to personality characteristics of novelty- or sensation seeking, extroversion, neuroticism, and the absence of harm avoidance, attributes that might increase the risk for alcohol abuse or dependence (41). Related concepts that overlap with ASPD include type 2 and type B subgroups of alcoholics whose early onset and more severe course is often associated with criminality and dependence on other drugs (42 ,43).

Almost two-thirds of people with ASPD are alcoholics, although only about 20% of alcoholic men fulfill criteria for this disorder (41 ,42 and 43). There is impressive familial crossover between CD/ASPD and alcoholism, and offspring of ASPD parents are likely to have a more severe alcoholic course (44 ,45). Genetic factors are important in personality traits and some personality disorders, and ASPD and alcohol dependence are often inherited together (44 ,45 ,46 and 47).

ASPD (or CD) and substance use disorders might share the characteristic of neuronal disinhibition (41 ,48). The high levels of impulsivity and sensation seeking in CD and ASPD might, in turn, relate to aspects of 5-HT functioning, and/or to subtypes of GABA receptors (49 ,50). ASPD, or neuronal disinhibition, also appear to correlate with the ERP characteristic described in the next section (41), but neither ASPD nor the relevant ERP values correlate with a low LR to alcohol (19 ,51).

The P300 (P3) Wave of the ERP

The P300, or P3, is a positive polarity brain wave in the ERP paradigm observed approximately 300 msec after the presentation of an infrequent but anticipated stimulus (41 ,48). Diminished amplitudes or shorter latencies of this wave reflect problems in attending and interpreting subtle environmental events (48 ,52 ,53). Low P3 amplitudes are seen in schizophrenia, major depressive disorder, attention-deficit/hyperactivity disorder (ADHD), and Alzheimer's disease (6 ,52 ,53).

Adult alcoholics demonstrate small P3 waves even after extended periods of abstinence, although this might diminish with time (6 ,48 ,52). Although most investigators consider this a robust marker for adult alcoholism, a low amplitude P3 might more closely reflect states of temporary depression, a developmental delay, or aspects of CD or ASPD (41 ,52 ,53).

A relatively low amplitude P3 is seen in 20% to 35% of the offspring of alcoholics (6 ,48). A 4-year follow-up of 13 sons of alcoholics, 11 sons of nonalcoholics with family histories of alcoholism, and 12 family-history–negative controls revealed a relationship between a low P3 amplitude and delinquent behavior as well as subsequent substance use, although alcohol abuse and dependence were not directly tested (54). Another 8-year follow-up of 11 children of alcoholics and nine controls revealed lower P3 amplitudes for those four subjects who went on to develop an alcohol use disorder (55). Although neither of these investigations controlled for the potential impact of CD and ASPD, the results are consistent with the relationship between P3 amplitude and the alcoholism risk.

As is true of many electrophysiologic measures, P3 amplitudes are genetically influenced (56 ,57). Data supporting these conclusions come from twin investigations, family studies, and a genetic linkage study that highlighted findings on chromosomes 2, 5, 6, and/or 13.

There are several interesting potential crossovers between P3 amplitude and other phenotypic or genetic markers of the alcoholism risk (19 ,52 ,53). Some investigators believe that most of the variance explained by P3 relates to CD or ASPD, or that finding might reflect activities of several dopamine (DA) receptors, which, in turn, might be related to impulsiveness or CD (41 ,52 ,53 ,58). At the same time, groups demonstrating low LRs to alcohol do not appear to have smaller P3 waves (19 ,51).

Independent Major Psychiatric Disorders

Alcoholics have an elevated risk for several major psychiatric disorders (59). For some people the psychiatric symptoms are likely to be temporary manifestations of intoxication and withdrawal, and for others alcohol problems might develop in the context of poor judgment and loss of control during major psychiatric syndromes (60 ,61). With schizophrenia, bipolar manic depressive disease, and ADHD, poor judgment, feelings of alienation, and high impulsivity might increase the chance of repeated heavy drinking (62 ,63 and 64).

The anxiety disorders are a series of syndromes with different clinical symptoms (in addition to anxiety), divergent treatments, and potentially different etiologies (65 ,66). Even after controlling for the potential impact of temporary substance-induced disorders, at least two anxiety syndromes—panic disorder and social phobia—appear to be tied to alcohol dependence, whereas some others, including obsessive-compulsive disorder, are not (65).

Each of the independent major psychiatric disorders listed above is likely to be genetically influenced (66). Most have abnormalities related to the HPA axis and to neurochemical systems including DA (especially relevant to schizophrenia), 5-HT (most closely tied to schizophrenia, anxiety, and mood disorders), norepinephrine (NE) and GABA (each tied to panic and mood disorders), and NPY (in anxiety) (66). Thus, as discussed below, genes related to the relevant psychiatric disorders might indirectly increase the risks for alcoholism through the clinical manifestations of the syndrome. However, these disorders might act independently of LR (14).

The Hypothalamic-Pituitary-Adrenal (HPA) Axis

Cortisol, adrenocorticotropin hormone (ACTH), corticotropin-releasing factor (CRF), prolactin, and related hormones are important in physiologic well-being, mood, anxiety,

aggression, and reactions to stress, including alcohol withdrawal (66 ,67 ,68 ,69 and 70). Aspects of HPA functioning might relate to an alcoholic predisposition in any of several ways. These include a possible contribution to a low LR, reflections of the actions of neurotransmitters such as 5-HT or DA in the course of some psychiatric disorders, and via the opioid system.

Some alcohol-dependent men and women have high cortisol, and children of alcoholics have a differential hormonal response to alcohol, opioid antagonists, and other challenges when compared to controls with some persisting after abstinence (67 ,68). Although there is some evidence that recovering alcoholics and children of alcoholics show lower levels of HPA response to alcohol, these same groups might show higher levels of response to naltrexone, an opioid antagonist (67 ,68 and 69).

The functioning of the HPA axis is at least partially under genetic control (70). Thus, activity of this system might contribute to the alcoholism risk, perhaps indirectly through several channels described in other sections.

Other Potentially Interesting Broad Phenotypes

Additional phenomena, including other personality disorders (e.g., borderline personality disorder), additional clinical characteristics (e.g., deficiencies in neuropsychological functioning), and other psychiatric syndromes [e.g., posttraumatic stress disorder (PTSD) or generalized anxiety disorder] have also been hypothesized to be related to the alcoholism risk (6 ,11 ,71).

Thus, reflecting space constraints, this discussion is not exhaustive and the reader should consider additional review papers.

Specific Proteins and Candidate Genes Potentially Related to the Alcoholism Risk

This subsection highlights some specific genes, enzymes, and other proteins that might relate to an alcoholism vulnerability. The distinction between broad potential phenotypic markers of alcoholism (discussed above) and the more specific markers noted here is somewhat arbitrary, but might be heuristically useful.

Adenylyl Cyclase (AC) and G Proteins

These proteins include three membrane-bound components of receptors, G (or guanine nucleotide binding) proteins, and the AC enzyme that are part of a complex second messenger system that translates the impact of alcohol, neurotransmitters, and other substances on the cell membrane or receptors into changes within the cell (72 ,73). The G proteins facilitate the coupling of at least nine different isoforms of AC to cell membrane-bound receptors, which in turn affects production of 3′,5′-cyclic adenosine-monophosphate (cAMP), with the latter often used as a measure of the system's activity. This complex is affected by several neurotransmitters, and has impact on a variety of actions within the cell, including gene expression, and might impact on a variety of psychiatric disorders, including depression. G proteins come in several forms that either stimulate (G_s) or inhibit (G_i) the process.

Although results can differ when platelets or lymphocytes from patients or cells *in vitro* are used, alcohol appears to stimulate G_s. Recently detoxified alcoholics and their nonalcoholic relatives might have lower cAMP production following chemical stimulation of platelets or white blood cells, results that have been hypothesized to be related to a reduced G_s binding (72 ,74). However, AC activity is also affected by drinking and by withdrawal, and differences from controls can change depending on periods of abstinence or following the production of multiple generations of lymphocytes in cell culture (73 ,75). In the final analysis, differences between alcoholics and controls are probably both state and trait phenomena.

Regarding the latter, activity of the system appears to be genetically influenced, and it is possible the divergence in AC functioning is especially prominent in alcoholics with family histories of this disorder (75). The production of cAMP in chemically stimulated cells has also been investigated in children of alcoholics who might share lower levels of G_s-protein–stimulated cAMP production, especially if they have multiple alcoholic relatives (72 ,76).

One theory attempting to integrate these findings is that an avenue of alcoholism risk might be a low innate activity of the G_s system, with acute alcohol causing a temporary stimulation, after which abstinence from alcohol produces the opposite effect, which might lead to more alcohol intake in an attempt to compensate (77). The underlying difference has been hypothesized to involve a reduction of gene expression of the α subunit of the G_s protein (72). The process might have an impact on the development of tolerance, and, perhaps, the need for higher levels of alcohol to have an effect. Another mechanism is suggested by findings regarding the role of alcohol in decreasing neuronal excitability through enhancing G-protein–coupled inwardly rectifying potassium channels (GIRKs) (78). Perhaps these results might also tie into the actions of protein kinase and the effects of other markers such as NPY, as well as through additional neurochemical systems such as 5-HT (79 ,80).

Protein Kinase C (PKC)

These proteins encompass at least three families of enzymes that, similar to AC, have important functions in translating the effects of neurotransmitters on receptors into the cell. These calcium-activated, phospholipid-dependent proteins are widely distributed in the body, and function by phosphorylating

target proteins, including G proteins, and thus, mediating changes in intracellular signaling (81 ,82).

The direction of the impact of alcohol on PKC activity can be different with acute versus chronic administration, but in general ethanol affects the movement of this protein from the area around the nucleus to the cytoplasm (83). The changes in PKC subsequently have impact on the actions of several neurotransmitter receptors, including 5-HT and $GABA_A$, and thus, are likely to affect alcohol intoxication and tolerance.

Alcohol-dependent individuals may have higher amounts of PKC-ε, a form that might inhibit the actions of $GABA_A$ receptors, possibly tying PKC-ε to a low LR to alcohol (82). Although no data are yet available in children of alcoholics, mice genetically engineered for an absence of PKC-ε have both a high sensitivity to alcohol and lower self-administration of this drug (82). Such animals also show a decreased reaction to pain, perhaps reflecting changes in opioid activity (84). There is additional evidence that PKC-ε knockout mice have a lower intensity of reaction to alcohol, and less ability to develop tolerance to at least some effects of the drug (81).

Neuropeptide Y (NPY)

NPY is a widely distributed neurotransmitter that affects multiple receptor subtypes including Y1 (in the amygdala where NPY decreases feelings of anxiety), and Y5 (in the hypothalamus where NPY might increase appetitive behaviors) (79 ,85). NPY appears to act through G proteins, producing an inhibition of AC production, and this transmitter can facilitate the release of DA in the nucleus accumbens (86). It has been hypothesized to play a role in eating disorders, depression, anxiety, and the actions of opioids (87).

Acute alcohol intake has impact on NPY release, which in turn affects the release of DA, possibly contributing to some rewarding effects of alcohol or adding to some psychiatric symptoms (85 ,86 ,88). Chronic alcohol intake and withdrawal are associated with increased NPY in the hypothalamus, and increased responsiveness of CRF to NPY (85).

Alcohol-preferring rats have a QTL on chromosome 4 [logarithm of odds (LOD) = 8.6], which explains about a third of the enhanced alcohol intake, and which is located in an area where NPY has been mapped (89). In addition, rats bred to consume high levels of alcohol have increased NPY activity in the amygdala (perhaps reflecting levels of anxiety), along with decreased NPY in the frontal cortex and hippocampus (perhaps reflecting a lower level of satiety) (85 ,88). Mice genetically engineered for an absence of NPY drink more alcohol and have a lower intensity of response compared to wild-type mice, whereas transgenetic mice with increased NPY have less alcohol consumption and higher responses to alcohol (79 ,90). Studies have not yet been carried out in alcoholics or their offspring.

Opioid-Like Substances, Including ß-Endorphin

The opioids are endogenous proteins, including endorphins and enkephalins, as well as most of the prescription pain medications, methadone, heroin, codeine, and morphine, each of which bind to opioid receptors. Their actions diminish pain, decrease respirations, cause euphoria, and produce a decreased motility in the gut. There are a variety of opioid receptor subtypes including μ, κ, and δ, with μ most closely tied to analgesic and reinforcing effects (91). Opioids have impact on DA, 5-HT, and NPY activity, and relate to some psychiatric disorders, such as depression (92).

Acute alcohol intake results in the release of endogenous opioids, and stimulates relevant receptors (6 ,93). Aspects of tolerance and withdrawal from alcohol might relate to changes in functioning of the μ receptors, and alcohol-preferring animals have an increase in these receptors in the ventral tegmentum, along with a greater increase in ß-endorphin following alcohol (94 ,95 and 96). An exaggerated HPA response to naltrexone (an opioid antagonist) has been reported in alcoholics and their relatives, perhaps reflecting less baseline opioid functioning (97). Decreased ß-endorphin has been noted in the cerebrospinal fluid (CSF) of abstinent alcoholics, and their relatives demonstrate more release of ß-endorphin following alcohol (94 ,96). Opioid antagonists, such as naltrexone and nalmephene, can decrease the self-administration of alcohol in animals and humans, perhaps by blunting the stimulatory effect of alcohol, enhancing the sedative effects of this drug, and/or through decreased levels of reinforcement from alcohol (93 ,97 ,98). A μ opioid receptor gene might be located near a QTL for alcohol preference in mice (99), and there is a possible association between alcoholism and some of the six more known alleles of the μ opioid receptor (OPRM1), although not all studies agree (6 ,94 ,100).

The Serotonin (5-HT) Systems

The actions of this neurotransmitter, which are mediated through the 5-HT transporter (5-HTT) and more than 14 different receptors, affect HPA functioning, anxiety, impulsivity, eating behaviors, depression, and other conditions (6). Numerous drugs of abuse have impact on 5-HT systems, including alcohol, and 5-HT, in turn, also interacts with other neurotransmitters, especially DA (6). The following data suggest that different genes affecting 5-HT levels could increase the alcoholism risk through several different mechanisms.

5-HT agonists can simulate signs of intoxication, and, perhaps, feelings of craving (101). Alcoholics, especially those with aggressiveness or an early onset of their substance use disorder, may have lower levels of platelet and brain 5-HT, diminished responses to 5-HT boosting drugs, and lower levels of 5-HT metabolites in the CSF (6 ,102). Treatment with drugs that boost 5-HT in the synapse [e.g., selective

serotonin reuptake inhibitors (SSRIs)], produces a modest decrease in voluntary alcohol intake in animals and humans (6).

Alcohol preference in animals is associated with a QTL near the genes encoding for the 5-HTT (103). The s-allele may relate to nervousness, harm avoidance, and other forms of anxiety that might tie in to axis I anxiety disorders and more severe alcohol withdrawal, although not all authors agree (49 ,104). The l-allele, which might produce a protein that more rapidly takes up 5-HT from the synapse, has been tied to a low LR to alcohol and an enhanced alcoholism risk (32). Another gene that controls the production of the rate limiting enzyme in the synthesis of 5-HT, *tryptophan hydroxylase*, might also relate to lower 5-HT levels and more impulsiveness, suicidality, as well as a predisposition toward alcoholism (105).

Genes for some 5-HT receptors might be associated with higher alcohol intake either directly or through ASPD, depressive disorders, schizophrenia, or anxiety disorders. Findings include a high receptor density for 5-HT_{1A} or a decrease in 5-HT_{1B} activity in alcohol-preferring rats, with 5-HT_{1B} knockout mice demonstrating higher levels of alcohol intake (106 ,107). Turning to a second family of receptors, there is evidence of a decrease in 5-HT_{2C} receptor sensitivity in alcoholics, along with a potential increase in the density of these proteins in the hippocampus in alcohol-preferring rats (108). A third family has also been implicated through the actions of the 5-HT_3 receptor, which promotes the release of dopamine in the nucleus accumbens in the context of alcohol (6 ,109).

Children of alcoholics might have more rapid uptake of 5-HT in platelets, perhaps indicating a lower level of 5-HT in the synapse that might relate to LR (110). This is consistent with lower LR to alcohol in the offspring who have the l-allele of the 5-HTT (32). Finally, a drug that antagonizes activity of the 5-HT_3receptor, ondansetron, both decreases subjective feelings of intoxication with alcohol and decreases alcohol intake in alcoholics and their relatives (109).

The Potential Importance of Dopamine (DA)

This neurotransmitter has broad effects in the brain, including in the mesolimbic system where it functions as a mediator of reward or pleasure (6 ,111). DA impacts on the risk for heavy drinking and alcoholism through potentially diverse mechanisms including the reinforcing effects of the drug, personality characteristics, and via several psychiatric disorders.

Ethanol causes the release of DA in the mesolimbic system, affects DA neurons in the ventral tegmentum, and the reinforcement from alcohol decreases when DA antagonists are given (6 ,111 ,112). There might be a general decrease in overall DA functioning among more violent alcoholics, as evidenced by lower levels of DA metabolites in the CSF, and DA activity might also relate to personality characteristics such as novelty seeking and to ASPD (113 ,114).

These findings have led to a search for specific DA markers possibly tied to a vulnerability toward alcoholism. A decrease in the D2 receptor density has been reported in the brain of alcohol-preferring rodents and some alcoholics, as has a blunted hormonal response to D2 agonists, at least soon after withdrawal (115 ,116). Lower levels of DA in the synapse might result from a higher density of DA uptake as seen in alcohol-preferring primates, although possibly reflecting withdrawal, the opposite was reported in the striatum in a small sample of nonviolent alcoholics (116 ,117).

Although it is not clear whether they are functional, several alleles of the Taq1A marker for the D2 DA receptor (DRD2) have been reported to be linked to alcoholism, especially severe early-onset problems, and thus possibly to ASPD (6 ,118). However, results relating to this candidate have not been replicated in genome scans, and there are as many nonconfirmatory studies as there are positive ones (6 ,119). Additional interest has been expressed regarding the D4 receptor and several alleles of the DA transporter, but with conflicting results (120 ,121).

GABA, Norepinephrine (NE), and Monoamine Oxidase (MAO)

This subsection briefly reviews several markers that might relate to the alcoholism risk. GABA, a ubiquitous inhibitory neurotransmitter, has an important role in several conditions possibly related to the alcoholism risk including anxiety, mood disorders, schizophrenia, and aggressive behaviors (32 ,66). There are multiple GABA receptors, with special interest for alcohol intoxication or withdrawal for the estimated 13 or more subunits for the $GABA_A$ receptor complex (6 ,32 ,122). Alcohol-dependent men and women have a decreased density of $GABA_A$ receptors, and might show decreased responses to lorazepam in frontal brain regions and in the basal ganglia, while demonstrating abnormal responses to a benzodiazepine antagonist flumazenil (123 ,124). A diminished response to brain depressants might occur with a common mutation of the $GABA_{A\alpha 6}$ receptor, which might also reflect a low LR to alcohol (32 ,122). In addition, a possible predisposition toward alcohol dependence might link to an area of chromosome 4 near genes noted to have an impact on GABA functioning (32 ,125).

Monoamines, including 5-HT, NE, and DA, are metabolized in part by MAO. Alcoholics, especially those with concomitant ASPD, might demonstrate low MAO activities, perhaps reflecting alternate forms of genes, although this finding might be an artifact of recent smoking (126).

Finally, alcoholics, especially those with multiple alcoholic relatives, might have a blunted hormonal response to drugs that have impact on NE, especially during withdrawal and early abstinence (127). Thus, NE might also increase the alcoholism risk through vulnerability for panic and other anxiety disorders.

Alcohol-Metabolizing Enzymes

The best-documented genetic factors related to alcoholism involve genes controlling the two prominent alcohol-metabolizing enzymes, alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH). As discussed in Chapter 99 , Asian men and women who lack the low k_m, mitochondrial ALDH (i.e., have the homozygous ALDH2-2, 2-2 genotype) and who, thus, produce very high levels of the first breakdown product of alcohol (acetaldehyde), have almost no risk for alcohol dependence (11 ,12 ,128). Heterozygotes with the ALDH2-1, 2-2 genotype produce higher acetaldehyde levels than Asians with ALDH2-1, 2-1, and have an enhanced level of response to alcohol and lower risk for alcoholism. For ADH, individuals carrying the genotypes that are associated with more rapid metabolism of alcohol (e.g., those with ADH2-2, 2-3, or 3-1 alleles) more rapidly produce acetaldehyde and have a diminished alcoholism risk, especially if these genotypes are associated with the relevant ALDH markers described above (129 ,130). The ADH enzyme forms appear to exert less influence on the alcoholism risk, but are applicable to a broader range of ethnic and racial groups.

Thus, ADH and ALDH enzyme patterns have several potential mechanisms associated with an altered risk for alcohol dependence. These include an aversive reaction to alcohol related to very high levels of acetaldehyde in ALDH2-2 homozygotes, whereas individuals who are heterozygotes for ALDH2-2 and those with more rapid metabolism of alcohol through ADH forms might decrease their risk through an increased intensity of reaction to alcohol, not necessarily an overall more aversive response (25).

AN ATTEMPT TO SYNTHESIZE THESE DATA

Part of "98 - Vulnerability Factors for Alcoholism "

Background

This chapter has briefly reviewed many genetically influenced characteristics that may be relevant to the alcoholism risk. However, it is unlikely that there are 30 or so independent genetically influenced trait markers for alcoholism, and thus the findings are likely to represent a more limited number of overarching phenomena, or families of risk factors.

As in the selection of markers of interest, the identification of possible families of findings, or domains, requires subjective judgment. Initially, I was tempted to highlight a separate domain for 5-HT and another for DA markers, and I recognize that it is possible that the functioning of the HPA axis might be a core mediator of risk by itself. However, I believe that most of these markers might function as correlates of several different mechanisms through which broader domains of influence operate. Therefore, I propose that the majority of the genetically related markers of the alcoholism risk might fall into about five relatively independent overarching categories (Table 98.1). The specific markers are summarized in Table 98.2 .

	LR	Disinhibition	Axis II	Opioids	ALDH/ADH
Broad markers					
EEG alpha	X		X		
Voltage	X		X		
HPA	X		X	X	X
5-HT levels		X	X	X	
DA levels		X	X	X	
Neuropsychiatric		X	X		
Genes/proteins					
AC	X		X		
G protein	X		X		
PKC	X			X	
NPY	X		X	X	
$5\text{-HT}_{1A/1B/2C}$	?	?	?		
5-HT_3	X				
5-HTT	X	X	X		
TOH		X	X		
DRD2		X			
D4		X			
DAT		X			
$GABA_A$	X	X	X		

AC, adenylyl cyclase; ADH, alcohol dehydrogenase; ALDH, aldehyde dehydrogenase; DA, dopamine; DAT, dopamine transporter; DRD2, dopamine receptor D2; GABA, γ-aminobutyric acid; HPA, hypothalemic-pituitary-adrenal axis; 5-HT, serotonin; LR, level of response; NPY, neuropeptide Y; PKC, protein kinase C.

TABLE 98.2. POSSIBLE FAMILIES OF RISK FACTORS

A Domain Related to a Low LR to Alcohol

A low LR is a useful place to begin to demonstrate how some of the literature might be synthesized. Although the central aspect of this family of findings might be altered 5-HT or HPA axis functioning, it is equally plausible that a low LR to alcohol is the characteristic that pulls together the other findings.

Theoretically, a low LR to alcohol might reflect the final phenotypic path through which many different markers operate. Dampened responses to alcohol in children of alcoholics have been observed for several hormones, electrophysiologic measures, subjective feelings of intoxication, and motor performance. This array of findings might reflect altered second messenger and intracellular signaling actions that are less sensitive to alcohol-induced neurochemical changes, including genes that affect G proteins, AC, and PKC. Other mechanisms that might independently affect the LR to alcohol might include genes that have impact on NPY, GABA, 5-HT, and DA. The subsequent diminished effect of alcohol on neurons might also function to produce less change in the EEG and HPA axis responses, thereby helping to mediate the reaction to alcohol overall (24 ,67).

Other components of some of these same systems, including DA, 5-HT, DA, and HPA functioning, are also likely to function through additional mechanisms as mediators of several other domains of risk factors. For example, alternative aspects of these systems might have impact on disinhibition, independent psychiatric disorders, and opioid system functioning. As discussed further below, most of those additional attributes are likely to be distinct from those that relate to a low LR. It is hoped that the identification of specific genes linked to LR will help pinpoint which of these neurochemical systems contribute most to this domain.

A Domain Encompassing Disinhibition, a Low P3 Amplitude, ASPD, and Early-Onset and More Severe Subtypes of Alcoholism

A low P3 amplitude of the ERP and aspects of CD and/or ASPD characterize a substantial minority of young children of alcoholics; aspects of this domain are genetically influenced, and these phenomena relate to the future alcoholism risk. Although low LR appears to be relatively specific for the alcoholism risk (19 ,24), the disinhibition domain might enhance the risk for all substance use disorders (41 ,48 ,52). Theoretically, these markers might be associated with poor judgment and impulsive behavior, which both increase the risk for ingesting substances and for problems learning how to control their use.

This potential domain appears to operate independently of a low LR and alcohol-metabolizing enzymes, and there are few data that would link these phenomena to the opioids and major axis I psychiatric disorders (14 ,53). It is possible that one childhood psychiatric disorder, ADHD, might also fit into the disinhibition domain, and other investigators might believe that some anxiety disorders might fit here as well.

A number of neurochemical markers described above might fit together under a general heading of disinhibition. These include several DA-related genes (e.g., DRD2, and DRD4 receptors), the DA transporter, and aspects of the DA function tied to sensation seeking, novelty seeking, as well as early-onset and severe alcoholism (6 ,58). Aspects of 5-HT might also be relevant, with the major finding here (as opposed to the LR domain where only the 5-HTT might contribute) being low 5-HT functioning overall, and genes having impact on tryptophan hydroxylase (6 ,102). Several additional findings that are discussed in more detail in other publications might also apply to this domain, including cognitive difficulties with executive cognitive functioning, which might cross over with CD, ASPD, and the substance use disorders.

Independent Axis I Major Psychiatric Disorders as a Potential Domain of Risk

The central hypothesis for this domain is that genes that contribute to the development of some psychiatric disorders might indirectly increase the risk for heavy drinking and alcohol-related problems. The axis I disorders most closely tied to an elevated alcoholism risk are schizophrenia and bipolar manic depressive disorder as described above (60 ,61 ,64). An enhanced alcoholism risk might also be associated with panic disorder and social phobia, and possibly PTSD or generalized anxiety disorder. Each relevant condition is itself a complex genetic disorder, with separate, but perhaps overlapping, sets of genes. Furthermore, it is possible that different environmental events add to or detract from the risk for each of these conditions.

For this discussion, it is not essential to determine if the individual with schizophrenia, for example, is drinking to decrease the symptoms of their underlying and independent disorder (although this contention is not well supported) (63 ,131), or if the problems were a result of the combination of poor judgment, a large amount of free time, and living in a heavy drinking environment. In either case, the search for genetic factors in alcoholism might be more efficient if the potential impact of genes related to these independent disorders is considered. Once genetic markers for an additional characteristic (such as LR) have been found, they can be tested in these more complex subjects to determine whether it adds further to the risk.

The core characteristics of this domain might only indirectly relate to independent psychiatric disorders. It is possible that the predispositions toward both alcoholism and a

psychiatric disorder might operate through the same genes that have impact on the 5-HT, DA, or the HPA systems. Or the relationships could reflect transmission disequilibrium for the genes having impact on the alcoholism risk and those related to some of the psychiatric disorders. The answer to these questions might be more easily addressed once specific genes linked to the alcoholism risk and those linked to the relevant independent psychiatric disorders have been identified.

A Possible Domain Related to the Opioid System

A low level of activity of endogenous opioids could alter the intensity of reinforcement from alcohol (93 ,94). The markers reported above, as well as characteristics potentially related to the opioid systems, might form a domain that appears to be relatively independent of other risk factors. The decision to place opioids in a separate family of findings rests with evidence of high levels of alcohol dependence among opioid-dependent individuals, the closer than expected relationship between alcohol and opioid consumption in animals, and my subjective judgment that opioid systems are likely to function primarily as a characteristic subsumed under a separate global domain.

There is crossover between this hypothesized domain and the functioning of the HPA axis, and 5-HT, DA systems. Thus, it is possible this is not a separate domain of influence, but the possibility of a relatively unique impact is worth considering.

The Importance of Alcohol-Metabolizing Enzymes

Both the genetic control and the impact on drinking behaviors for alcohol-metabolizing enzymes have been well established (18 ,128). The risk for alcohol dependence among individuals with the ALDH2-2, 2-2 genotype is close to zero. ALDH2-2, 2-1 heterozygotes have significantly lower levels of risk as, apparently, do some people who have the more efficient ADH2-2, 2-3, and 3-1 alleles. The mechanisms through which the relevant genes are likely to operate include an aversive effect of alcohol at high acetaldehyde levels (as seen with ALDH2-2 homozygotes), and possibly through an enhanced LR to alcohol (for ALDH2-2, 2-1 heterozygotes and the relevant ADH alleles).

Despite some crossover with LR for ALDH heterozygotes, it is likely that the alleles controlling these alcohol-metabolizing enzymes operate as a relatively separate domain of risk. Few, if any, data tie these alleles to disinhibition or axis II major psychiatric disorders, and a strong link to opioid systems seems unlikely. However, it is possible that some of the impact of acetaldehyde might operate through elevations in HPA hormones, and the accompanying neurochemical changes might have impact on second messenger or other intracellular signaling systems.

CONCLUDING REMARKS

Part of "98 - Vulnerability Factors for Alcoholism "

During the 7 years since publication of the previous edition, there have been exciting developments regarding the study of complex genetically influenced disorders. Progress in the studies of the genetic factors in alcohol dependence has been important because this is one of the most prevalent major psychiatric conditions, and is associated with substantial morbidity and mortality.

Some readers will be most interested in the citations from publications over the prior decade regarding the broad phenotypic or more focused genotypic markers described here. It is also hoped that some will also benefit from considering the hypothesis that many of these markers of risk can be grouped together into overarching phenomena including a low LR disinhibition, independent axis II psychiatric disorders, the opioid system, or alcohol-metabolizing enzymes.

The idea of searching for central themes among the markers is potentially more important than whether any of the hypothesized domains survive the test of time. It is possible that several of the families of findings discussed here would have been better subsumed under changes in the HPA axis, that findings related to specific neurochemical systems might each represent separate domains of influence, and that some of the hypothesized domains might be epiphenomena operating under the umbrella of genes that affect the levels of functioning in specific neurochemical systems. However, as is true of all fields of science, it is important to outline possible examples of a generic approach, in the hope of stimulating additional research to determine the most appropriate domains of influence.

ACKNOWLEDGMENTS

Part of "98 - Vulnerability Factors for Alcoholism "

This research was supported by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) grants 05526 and 08403, the Veterans Affairs Research Service, and funds provided by the State of California for medical research on alcohol and substance abuse through the University of California, San Francisco.

REFERENCES

1. Schuckit MA. New findings in the genetics of alcoholism. *JAMA* 1999;281:1875–1876.

2. Kendler KS, Walters EE, Neale MC, et al. The structure of the genetic and environmental risk factors for six major psychiatric disorders in women: phobia, generalized anxiety disorder, panic disorder, bulimia, major depression, and alcoholism. *Arch Gen Psychiatry* 1995;52:374–383.

3. Hyman SE. Introduction to the complex genetics of mental disorders. *Biol Psychiatry* 1999;45:518–521.

4. McGue M. The behavioral genetics of alcoholism. *Curr Directions Psychol Sci* 1999;8:109–115.

5. Bloom GE, Kupfer DJ. *Psychopharmacology: the fourth generation of progress: an official publication of the American College of Neuropsychopharmacology*. New York: Raven Press, 1994.

6. Hill SY. Biologic phenotypes associated with individuals at high risk for developing alcohol-related disorders, Part 1. *Addiction Biol* 2000;5:5–22.

7. Malhotra AK, Goldman D. Benefits and pitfalls encountered in psychiatric genetic association studies. *Biol Psychiatry* 1999;45:544–550.

8. Wihelmsen KC. Does tumor necrosis factor play a role in alcoholic steatohepatitis? The potential pitfalls of a case-controlled allelic association analysis. *Hepatology* 1997;26:232–233.

9. Nutt D. Alcohol and the brain: pharmacological insights for psychiatrists. *Br J Psychiatry* 1999;175:114–119.

10. Schuckit MA. Biological, psychological and environmental predictors of the alcoholism risk: a longitudinal study. *J Stud Alcohol* 1998;59:485–494.

11. Foroud T, Li T-K. Genetics of alcoholism: a review of recent studies in human and animal models. *Am J Addict* 1999;8:261–278.

12. Li T-K. Pharmacogenetics of responses to alcohol and genes that influence alcohol drinking. *J Stud Alcohol* 2000;61:5–12.

13. Goldman D. General and specific inheritance of substance abuse and alcoholism (commentary). *Arch Gen Psychiatry* 1998;55:964–965.

14. Schuckit MA, Smith TL. An 8-year follow-up of 450 sons of alcoholic and control subjects. *Arch Gen Psychiatry* 1996;53:202–210.

15. Schuckit MA, Smith TL. The relationships of a family history of alcohol dependence, a low level of response to alcohol, and six domains of life functioning to the development of alcohol use disorders. *J Stud Alcohol* 2000;61:827–835.

16. Schuckit MA, Tipp JE, Smith TL, et al. The relationship between self-rating of the effects of alcohol and alcohol challenge results in ninety-eight young men. *J Stud Alcohol* 1997;58:397–404.

17. Pollock VE. Meta-analysis of subjective sensitivity to alcohol in sons of alcoholics. *Am J Psychiatry* 1992;149:1534–1538.

18. Schuckit MA, Smith TL, Kalmijn J, et al. Response to alcohol in daughters of alcoholics: a pilot study and a comparison with sons of alcoholics. *Alcohol Alcohol* 2000;35:242–248.

19. Schuckit MA, Smith TL, Kalmijn J, et al. The relationship of a low response to alcohol to ERP and personality measures. *Alcohol Clin Exp Res* 2000;24:27A.

20. Schuckit MA, Tsuang JW, Anthenelli RM, et al. Alcohol challenges in young men from alcoholic pedigrees and control families: a report from the COGA project. *J Stud Alcohol* 1996;57:368–377.

21. Erblich J, Earleywine M. Children of alcoholics exhibit attenuated cognitive impairment during an ethanol challenge. *Alcohol Clin Exp Res* 1999;23:476–482.

22. Newlin DB, Thomson JB. Chronic tolerance and sensitization to alcohol in sons of alcoholics. *Alcohol Clin Exp Res* 1991;15:399–405.

23. Monteiro MG, Klein JL, Schuckit MA. High levels of sensitivity to alcohol in young adult Jewish men: a pilot study. *J Stud Alcohol* 1991;52:474–469.

24. Ehlers CL, Garcia-Andrade C, Wall TL, et al. Electroencephalographic responses to alcohol challenge in Native American Mission Indians. *Biol Psychiatry* 1999;45:776–787.

25. Wall TL, Johnson ML, Horn SM, et al. Evaluation of the self-rating of the effects of alcohol form in Asian Americans with aldehyde dehydrogenase polymorphisms. *J Stud Alcohol* 1999;60:784–789.

26. Volavka J, Czobor P, Goodwin D, The electroencephalogram after alcohol administration in high-risk men and the development of alcohol use disorders 10 years later. *Arch Gen Psychiatry* 1996;53:258–263.

27. Rodriguez LA, Wilson JR, Nagoshi CT. Does psychomotor sensitivity to alcohol predict subsequent alcohol use? *Alcohol Clin Exp Res* 1993;17:155–161.

28. Heath A, Madden PAF, Bucholz KK, et al. Genetic differences in alcohol sensitivity and the inheritance of alcoholism risk. *Psychol Med* 1999;29:1069–1081.

29. Baldwin HA, Wall TL, Schuckit MA, et al. Differential effects of ethanol on punished responding in the P and NP rats. *Alcohol Clin Exp Res* 1991;15:700–704.

30. Browman KE, Crabbe JC. Quantitative trait loci affecting ethanol sensitivity in BXD recombinant inbred mice. *Alcohol Clin Exp Res* 2000;24:17–23.

31. Moore MS, DeZazzo J, Lub AY, et al. Ethanol intoxication in *Drosophila*: genetic and pharmacological evidence for regulation by the cAMP signaling pathway. *Cell* 1998;93:997–1007.

32. Schuckit MA, Mazzanti C, Smith TL, et al. Selective genotyping for the role of 5-HT_{2A}, 5-HT_{2C}, and $GABA_{A6}$ receptors and the serotonin transporter in the level of response to alcohol: a pilot study. *Biol Psychiatry* 1999;45:647–651.

33. Schuckit MA, Edenberg HJ, Kalmijn J, et al. A genome-wide search for genes that relate to a low level of response to alcohol. *Alcohol Clin Exp Res* 2001;25:323–329.

34. Ehlers CL, Schuckit MA. Evaluation of EEG alpha activity in sons of alcoholics. *Neuropsychopharmacology* 1991;4:199–205.

35. Enoch MA, White KV, Harris CR, et al. Association of low-voltage alpha EEG with a subtype of alcohol use disorders. *Alcohol Clin Exp Res* 1999;23:1312–1319.

36. Pollock VE, Volavka J, Goodwin DW, et al. The EEG after alcohol administration in men at risk for alcoholism. *Arch Gen Psychiatry* 1983;40:857–861.

37. Bauer LO, Hesselbrock VM. EEG, autonomic and subjective correlates of the risk for alcoholism. *J Stud Alcohol* 1993;54:577–589.

38. van Beijsterveldt CEM, Boomsma DI. Genetics of the human electroencephalogram (EEG) and event-related brain potentials (ERPs): a review. *Hum Genet* 1994;94:319–330.

39. O'Connor S, Sorbel J, Morzorati S, et al. A twin study of genetic influences on the acute adaptation of the EEG to alcohol. *Alcohol Clin Exp Res* 1999;23:494–501.

40. Mazzanti CM, Bergen A, Enoch MA, et al. Identification of a Ser857-Asn857 substitution in DRK1 (KCNB1), population frequencies and lack of association to the low voltage alpha EEG trait. *Hum Genet* 1996;2:134–137.

41. Bauer LO, Hesselbrock VM. P300 decrements in teenagers with conduct problems: implications for substance abuse risk and brain development. *Biol Psychiatry* 1999;46:263–272.

42. Schuckit M A, Tipp JE, Smith TL, et al. An evaluation of type A and B alcoholics. *Addiction* 1995;90:1189–1203.

43. Irwin M, Schuckit M, Smith TL. Clinical importance of age at onset in type 1 and type 2 primary alcoholics. *Arch Gen Psychiatry* 1990;47:320–324.

44. Slutske WS, Heath AC, Dinwiddie SH, et al. Common genetic risk factors for conduct disorder and alcohol dependence. *J Abnorm Psychol* 1998;107:363–374.

45. Puttler LI, Zucker RA, Fitzgerald HE, et al. Behavioral outcomes among children of alcoholics: during the early and middle childhood years: familial subtype variations. *Alcohol Clin Exp Res* 1998;22:1962–1972.

46. Cadoret RJ, Yates WR, Troughton E, et al. Genetic-environmental interaction in the genesis of aggressivity and conduct disorders. *Arch Gen Psychiatry* 1995;52:916–924.

47. Benjamin J, Ebstein RP, Lesch K-P. Genes for personality traits: implications for psychopathology. *Int J Neuropsychopharmacol* 1998;1:153–168.

48. Begleiter H, Porjesz B. What is inherited in the predisposition toward alcoholism? A proposed model. *Alcohol Clin Exp Res* 1999;23:1125–1135.

49. Sander T, Harms H, Dufeu P, et al. Serotonin transporter gene variants in alcohol-dependent subjects with dissocial personality disorder. *Biol Psychiatry* 1998;43:908–912.

50. Sander T, Ball D, Murray R, et al. Association analysis of sequence variants of the GABA(A), alpha6, beta2, and gamma2 gene cluster and alcohol dependence. *Alcohol Clin Exp Res* 1999;23:427–431.

51. Ehlers CL, Garcia-Andrade C, Wall TL, et al. Determinants of P3 amplitude and response to alcohol in Native American Mission Indians. *Neuropsycholopharmacology* 1998;18:282–292.

52. Hill SY, Shen S, Locke J, et al. Developmental delay in P300 production in children at high risk for developing alcoholrelated disorders. *Biol Psychiatry* 1999;46:970–981.

53. Hill SY, Locke J, Steinhauer SR. Absence of visual and auditory P300 reduction in nondepressed male and female alcoholics. *Biol Psychiatry* 1999;46:982–989.

54. Berman SM, Whipple SC, Fitch RJ, et al. P3 in young boys as a predictor of adolescent substance abuse. *Alcohol* 1993;10:69–76.

55. Hill SW, Steinhauer S, Lowers L, et al. Eight-year longitudinal follow-up of P300 and clinical outcome in children from high-risk for alcoholism families. *Biol Psychiatry* 1995;37:823–827.

56. O'Connor SJ, Morzorati S, Christian JC, et al. Heritable features of the auditory oddball event-related potential: peaks, latencies, morphology and topography. *Electroencephalogr Clin Neurophysiol* 1994;92:115–125.

57. Begleiter H, Porjesz B, Reich T, et al. Quantitative trait loci analysis of human event-related brain potentials: P3 voltage. *Electroencephalogr Clin Neurophysiol* 1998;108:244–250.

58. Hill SY, Locke J, Zezza N, et al. Genetic association between reduced P300 amplitude and the DRD2 dopamine receptor A1 allele in children at high risk for alcoholism. *Biol Psychiatry* 1998;43:40–51.

59. Kessler RC, Crum RM, Warner LA, et al. Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Arch Gen Psychiatry* 1997;54:313–321.

60. Schuckit MA, Tipp JE, Bergman M, et al. Comparison of induced and independent major depressive disorders in 2,945 alcoholics. *Am J Psychiatry* 1997;154:948–957.

61. Raimo EB, Schuckit MA. Alcohol dependence and mood disorders. *Addict Behav* 1998;23:933–946.

62. Disney ER, Elkins IJ, McGue M, et al. Effects of ADHD, conduct disorder, and gender on substance use and abuse in adolescence. *Am J Psychiatry* 1999;156:1515–1521.

63. Tsuang JW, Lohr JB. Effects of alcohol on symptoms in alcoholic and nonalcoholic patients with schizophrenia. *Hosp Community Psychiatry* 1994;45:1229–1230.

64. Winokur G, Coryell W, Endicott J, et al. Familial alcoholism in manic-depressive (bipolar) disease. *Am J Med Genet* 1996;67:197–201.

65. Schuckit MA, Tipp JE, Bucholz KK, et al. The lifetime rates of three major mood disorders and four major anxiety disorders in alcoholics and controls. *Addiction* 1997;92:1289–1304.

66. Goodwin DW, Guze SB. *Psychiatric diagnosis*, fifth ed. New York: Oxford University Press, 1996.

67. Wand GS, Mangold D, Ali M. Adrenocorticotropin responses to naloxone in sons of alcohol-dependent men. *J Clin Endocrinol Metab* 1999;84:64–68.

68. Adinoff B, Martin PR, Bone GHA, et al. Hypothalamic pituitary adrenal axis functioning and cerebrospinal fluid corticotrophin releasing hormone and corticotrophin levels in alcoholics after recent and long term abstinence. *Arch Gen Psychiatry* 1990;47:325–330.

69. Wand GS, Mangold D, Ali M, et al. Adrenocortical responses and family history of alcoholism. *Alcohol Clin Exp Res* 1999;23:1185–1190.

70. Rivier C, Lee S. Acute alcohol administration stimulates the activity of hypothalamic neurons that express corticotropin-releasing factor and vasopressin. *Brain Research* 1996;726:1–10.

71. Corral MM, Holguín SR, Cadaveira F. Neuropsychological characteristics in children of alcoholics: familial density. *J Stud Alcohol* 1999;60:509–513.

72. Ratsma JE, Gunning WB, Leurs R, et al. Platelet adenylyl cyclase activity as a biochemical trait marker for predisposition to alcoholism. *Alcohol Clin Exp Res* 1999;23:600–604.

73. Wand GS, Waltman C, Martin CS, et al. Differential expression of guanosine triphosphate binding proteins in men at high and low risk for the future development of alcoholism. *J Clin Invest* 1994;94:1004–1011.

74. Parsian A, Todd RD, Cloninger CR, et al. Platelet adenylyl cyclase activity in alcoholics and subtypes of alcoholics. *Alcohol Clin Exp Res* 1996;20:745–751.

75. Menninger JA, Barón AE, Tabakoff B. Effects of abstinence and family history for alcoholism on platelet adenylyl cyclase activity. *Alcohol Clin Exp Res* 1998;22:1955–1961.

76. Lex BW, Ellingboe J, Larosa K, et al. Platelet adenylate cyclase and monoamine oxidase in women with alcoholism or a family history of alcoholism. *Harvard Rev Psychiatry* 1993;1:229–237.

77. Hoffman PL, Tabakoff B. A commentary on mechanisms. *Alcohol Alcohol* 1996;31:333–340.

78. Lewohl JM., Wilson WR, Mayfield RD, et al. G-protein-coupled inwardly rectifying potassium channels are targets of alcohol action. *Nature Neurosci* 1999;2:1084–1090.

79. Tecott LH, Heberlein U. Y do we drink? *Cell* 1998;95:733–735.

80. Pandey SC, Mittal N, Lumeng L, et al. Involvement of the cyclic AMP-responsive element binding protein gene transcription factor in genetic preference for alcohol drinking behavior. *Alcohol Clin Exp Res* 1999;23:1425–1434.

81. Bowers BJ, Owen EH, Collins AC, et al. Decreased ethanol sensitivity and tolerance development in gamma-protein kinase C null mutant mice is dependent on genetic background. *Alcohol Clin Exp Res* 1999;23:387–397.

82. Hodge CW, Mehmert KK, Kelley SP, et al. Supersensitivity to allosteric $GABA_A$ receptor modulators and alcohol in mice lacking PKC. *Nature Neurosci* 1999;2:997–1002.

83. Stubbs CD, Slater SJ. Ethanol and protein kinase C. *Alcohol Clin Exp Res* 1999;23:1552–1560.

84. Khasar SG, Lin Y-H, Martin A, et al. A novel nociceptor signaling pathway revealed in protein kinase C mutant mice. *Neuron* 1999;24:253–260.

85. Ehlers CL, Li T-K, Lumeng L, et al. Neuropeptide Y (NPY) levels in ethanol-naive alcohol preferring and nonpreferring rats and in Wistar rats following ethanol exposure. *Alcohol Clin Exp Res* 1998;22:1778–1782.

86. Ault DT, Radeff JM, Werling LL. Modulation of [3H] dopamine release from rat nucleus accumbens by neuropeptide Y may involve a signal-like receptor. *J Pharmacol Exp Ther* 1998;284:553–560.

87. Heilig M, Widerlov E. Neurobiology and clinical aspects of neuropeptide Y. *Crit Rev Neurobiol* 1995;9:114–136.

88. Hwang BH, Zhang J-K, Ehlers CL, et al. Innate differences of neuropeptide Y (NPY) in hypothalamic nuclei and central nucleus of the amygdala between selectively bred rats with high and low alcohol preference. *Alcohol Clin Exp Res* 1999;23:1023–1030.

89. Carr LG, Foroud T, Bice P, et al. A quantitative trait locus for alcohol consumption in selectively bred rat lines. *Alcohol Clin Exp Res* 1998;22:884–887.

90. Thiele TE, Marsh DJ, Ste Marie L, et al. Ethanol consumption and resistance are inversely related to neuropeptide Y levels. *Nature* 1998;396:366–369.

91. Kieffer BL. Opioids: first lessons from knockout mice. *TIPS* 1999;20:19–26.

92. Krausz M, Degkwitz P, Kühne A, et al. Comorbidity of opiate dependence and mental disorders. *Addict Behav* 1998;23:767–783.

93. Heyser CJ, Roberts AJ, Schulteis G, et al. Central administration of an opiate antagonist decreases oral ethanol self-administration in rats. *Alcohol Clin Exp Res* 1999;23:1468–1476.

94. Sander T, Gscheidel N, Wendel B, et al. Human mu-opioid receptor variation and alcohol dependence. *Alcohol Clin Exp Res* 1998;22:2108–2110.

95. De Waele JP, Kiianmaa K, Gianoulakis C. Distribution of the mu and delta opioid binding sites in the brain of the alcohol-preferring AA and alcohol-avoiding ANA lines of rats. *J Pharmacol Exp Ther* 1995;275:518–527.

96. Gianoulakis C, Krishnan B, Thavundayl J. Enhanced sensitivity of pituitary ß-endorphin to ethanol in subjects at high risk of alcoholism. *Arch Gen Psychiatry* 1996;53:250–257.

97. Wand GS, Mangold D, El Deiry S, et al. Family history of alcoholism and hypothalamic opioidergic activity. *Arch Gen Psychiatry* 1998;55:1114–1119.

98. Swift RM, Whelihan W, Kuznetsov O, et al. Naltrexone-induced alterations in human ethanol intoxication. *Am J Psychiatry* 1994;151:1463–1467.

99. Alexander RC, Heydt D, Ferraro TN, et al. Further evidence for a quantitative trait loci on murine chromosome 10 controlling morphine preference in inbred mice. *Psychiatr Genet* 1996;6:29–31.

100. Bergen AW, Kokoszka J, Peterson R, et al. Mu opioid receptor gene variants: Lack of association with alcohol dependence. *Mol Psychiatry* 1997;2:490–494.

101. Buydens-Branchey L, Branchey M, Fergueson P, et al. Hormonal, psychological and alcohol craving changes after m-chlorophenylpiperazine administration in alcoholics. *Alcohol Clin Exp Res* 1997;21:220–227.

102. Fils-Aime ML, Eckhardt MJ, George DT, et al. Early-onset alcoholics have lower cerebrospinal fluid 5-hydroxyindoleacetic acid levels than late-onset alcoholics. *Arch Gen Psychiatry* 1996;53:211–216.

103. Melo JA, Shendure J, Pociask K, et al. Identification of sex-specific quantitative trait loci controlling alcohol preference in C57BL/6 mice. *Nature Genet* 1996;13:147–152.

104. Katsuragi S, Kunugi H, Sano A, et al. Association between serotonin transporter gene polymorphism and anxiety-related traits. *Biol Psychiatry* 1999;45:368–370.

105. Nielsen DA, Virkkunen M, Lappalainen J, et al. A tryptophan hydroxylase gene marker for suicidality and alcoholism. *Arch Gen Psychiatry* 1998;55:593–602.

106. Wong DT, Reid LR, Li T-K. Greater abundance of serotonin$_{1A}$ receptor in some brain areas of alcohol-preferring (P) rats compared to -nonpreferring (NP) rats. *Pharmacol Biochem Behav* 1993;46:173–177.

107. Crabbe JC, Phillips TJ, Feller DJ, et al. Elevated alcohol consumption in null mutant mice lacking 5-HT$_{1B}$ serotonin receptors. *Nat Genet* 1996;14:98–101.

108. Pandey SC, Lumeng L, Li T-K. Serotonin$_{2C}$ receptors and serotonin$_{2C}$ receptor-mediated phosphoinositide hydrolysis in the brain of alcohol-preferring (P) and alcohol-nonpreferring (NP) rats. *Alcohol Clin Exp Res* 1996;20:1038–1042.

109. Johnson BA, Roache JD, Javors MA, et al. Ondansetron for reduction of drinking among biologically predisposed alcoholic patients: a randomized controlled trial. *JAMA* 2000;284:963–971.

110. Ernouf D, Compagnon P, Lothion P, et al. Platelet 3H 5-HT uptake in descendants from alcoholic patients: a potential risk factor for alcohol dependence. *Life Sci* 1993;52:989–995.

111. Koob GF, Le Moal M. Drug abuse: hedonic homeostatic dysregulation. *Science* 1998;278:52–58.

112. Pfeffer AO, Samson HH. Haloperidol and apomorphine effects on ethanol reinforcement in free feeding rats. *Pharmacol Biochem Behav* 1988;29:343–350.

113. Limson R, Goldman D, Roy A, et al. Personality and cerebrospinal fluid monoamine metabolites in alcoholics and controls. *Arch Gen Psychiatry* 1991;48:437–441.

114. Cloninger CR, Adolfsson R, Svrakic NM. Mapping genes for human personality. *Nature Genet* 1996;12:3–4.

115. Heinz A, Dettling M, Kuhn S, et al. Blunted growth hormone response is associated with early relapse in alcohol-dependent patients. *Alcohol Clin Exp Res* 1995;19:62–65.

116. Hietala J, West C, Syvalahti E, et al. Striatal D2 dopamine receptor binding characteristics in vivo in patients with alcohol dependence. *Psychopharmacology* 1994;116:285–290.

117. Mash DC, Staley JK, Doepel FM, et al. Altered dopamine transporter densities in alcohol-preferring vervet monkeys. *NeuroReport* 1996;7:457–462.

118. Cook CCH, Palsson G, Turner A, et al. A genetic linkage study of the D_2 dopamine receptor locus in heavy drinking and alcoholism. *Br J Psychiatry* 1996;169:243–248.

119. Edenberg HJ, Foroud T, Koller DL, et al. A family-based analysis of the association of the dopamine D2 receptor (DRD2) with alcoholism. *Alcohol Clin Exp Res* 1998;22:505–512.

120. Franke P, Schwab SG, Knapp M, et al. DAT1 gene polymorphism in alcoholism: a family-based association study. *Biol Psychiatry* 1999;45:652–654.

121. Parsian A, Chakraverty S, Fisher L, et al. No association between polymorphisms in the human dopamine D3 and D4 receptors genes and alcoholism. *Am J Med Genet* 1997;74:281–285.

122. Iwata N, Cowley DS, Radel M, et al. Relationship between a GABA$_{A6}$ Pro385Ser substitution and benzodiazepine sensitivity. *Am J Psychiatry* 1999;156:1447–1449.

123. Gilman S, Koeppe RA, Adams K, et al. Positron emission tomographic studies of cerebral benzodiazepine-receptor binding in chronic alcoholics. *Ann Neurol* 1996;40:163–171.

124. Volkow ND, Wang G-J, Begleiter H, et al. Regional brain metabolic response to lorazepam in subjects at risk for alcoholism. *Alcohol Clin Exp Res* 1995;19:510–516.

125. Long JC, Knowler WC, Hanson RL, et al. Evidence for genetic linkage to alcohol dependence on chromosomes 4 and 11 from autosome-wide scan in an American Indian population. *Am J Med Genet* 1998;81:216–221.

126. Anthenelli RM, Tipp J, Li T-K, et al. Platelet monoamine oxidase activity in subgroups of alcoholics and controls: results from the Collaborative Study on the Genetics of Alcoholism. *Alcohol Clin Exp Res* 1998;22:598–604.

127. Krystal JH, Webb E, Cooney NL, et al. Serotonergic and noradrenergic dysregulation in alcoholism: m-chlorophenylpiperazine and yohimbine effects in recently detoxified alcoholics and healthy comparison subjects. *Am J Psychiatry* 1996;153:83–92.

128. Chen C-C, Lu R-B, Chen Y-C, et al. Interaction between the functional polymorphisms of the alcohol-metabolism genes in protection against alcoholism. *Am J Hum Genet* 1999;65:795–807.

129. Neumark YD, Friedlander Y, Thomasson HR, et al. Association of the ADH2*2 allele with reduced ethanol consumption in Jewish men in Israel: a pilot study. *J Stud Alcohol* 1998;59:133–139.

130. Whitfield JB. Meta-analysis of the effects of alcohol dehydrogenase genotype on alcohol dependence and alcoholic liver disease. *Alcohol Alcohol* 1997;32:613–619.

131. Atkinson JH, Slater MA, Patterson TL, et al. Prevalence, onset, and risk of psychiatric disorders in men with chronic low back pain: a controlled study. *Pain* 1991;45:111–121.

99

Molecular and Cellular Genetics of Alcohol Addiction

Mary-Anne Enoch

David Goldman

Mary-Anne Enoch and David Goldman: Laboratory of Neurogenetics, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, Maryland.

Worldwide, in societies where moderate alcohol consumption is accepted as a pleasurable pastime and even an enhancer of health and well-being, it has historically been observed that a sizable minority is unable to keep within safe limits of consumption. Such individuals may abuse alcohol or become dependent. Alcoholism is today among the most pervasive psychiatric disorders. In the United States, the lifetime prevalence of alcohol dependence, the severe form of alcoholism, is 8% to 14% (1). The ratio of alcohol dependence to abuse is 1.5:1. Individuals often maintain a pattern of alcohol abuse without dependence for many years (2). Serious drinking frequently begins in adolescence, and approximately 40% of alcoholics develop their first symptoms of addiction between the ages of 15 and 19 years (3). As discussed in this chapter, heritability studies suggest that an individual's genotype confers a particular level of vulnerability, or risk. Comparative studies across populations suggest that sociocultural factors determine differences in thresholds above which an individual is likely to go beyond social drinking and slip into abuse or addiction. Is the development of alcoholism due to a unique set of biochemical and neurobiological determinants or are the causes of addiction common to many substances? Are there preexisting behavioral traits that predispose to alcoholism? These are some of the questions addressed by this chapter.

GENETIC INFLUENCES ON ALCOHOLISM: HERITABILITY STUDIES

Part of "99 - Molecular and Cellular Genetics of Alcohol Addiction "

Alcoholism is a heterogeneous disease in which the expression of genetic vulnerability is modified by environmental factors. Some of the environmental influences are uniquely experienced by the individual (nonshared) and some are shared among different individuals within the family. Numerous studies have shown that alcoholism is familial. In the National Comorbidity Survey of 5,877 individuals, it was found that alcohol use disorders aggregate significantly in families with an odds ratio of 1.93 (4). To have an alcoholic parent is a significant risk factor for the development of the disease; children of alcoholics are five times more likely to develop alcohol-related problems than children of nonalcoholics (5). It has been shown that the transmission of the vulnerability to alcoholism from parents to their daughters is due largely or entirely to genetic factors (6).

Studies of heritability, a measure of the genetic component of variance in interindividual vulnerability, indicate that genetic influences are substantially responsible for the observed patterns of familiality. Adoption studies have shown that alcoholism in biological parents predicts alcoholism in children even when the child is reared by unrelated adoptive parents (7 ,8). Large, well-constructed twin studies (8 ,9 and 10) have demonstrated that genetic factors are important in determining vulnerability to alcoholism, particularly in the more severe forms of the disease (11).

Over the past few years, several heritability analyses have been performed using the population-based Virginia Twin Registry. A study of 1,030 U.S. Caucasian female twin pairs demonstrated that whether using narrow, intermediate, or broad definitions, the concordance for alcoholism was consistently higher in monozygotic (MZ) than dizygotic (DZ) twin pairs, and the heritability of alcoholism in women is 0.50 to 0.60 (6 ,9). In a study of 3,516 U.S. male twins, it was found that 0.48 to 0.58 of the variation in liability for both alcohol abuse and dependence was attributable to additive genetic factors with the remainder attributable to nonshared environmental factors, which is most accurately labeled as "other," including as it does other sources of variance such as measurement error (12). The results of these two studies were confirmed in a recent analysis of 5,091 U.S. male twins and 4,168 female twins from the same registry (13). This study also added the information that the genetic sources of variability are partially, but not

completely, overlapping in men and women. Compared with the expected genetic correlations of 0.5 for same-sex DZ twin pairs, heritability was 0.20 to 0.24 [95% confidence interval (CI) = 4% to 45%] for opposite-sex pairs. Although the heritability value of approximately 0.5 for alcohol dependence in men and women was replicated in an analysis of 2,685 male and female twin pairs from the Australian twin registry (10), no evidence was found for sex differences in sources of genetic influence. It may be that the relatively small number of opposite sex pairs—592—in the Heath et al. (10) study compared with the larger number—1,428—in the Prescott et al. (13) study limited the power to find a difference.

GENETIC HETEROGENEITY IN ALCOHOLISM

Part of "99 - Molecular and Cellular Genetics of Alcohol Addiction "

Although there is good evidence for substantial heritability for alcoholism, individual differences in clinical presentation suggest variation in origins of vulnerability. Alcoholics vary in their drinking patterns, the severity of their symptoms, and in behavioral, physical, and psychiatric sequelae. Vulnerability may reside in personality or psychiatric traits that predispose to alcohol-seeking behavior, differential response to the effects of alcohol, or differential predisposition to addiction. Vulnerability factors found in some, but by no means all, alcoholics include attentional deficits reflected by low amplitude of the P300 event-related potential (14), anxiety reflected by the low-voltage alpha EEG trait (15), and diminished subjective response to alcohol (16).

COMORBIDITY OF ALCOHOLISM AND OTHER PSYCHIATRIC DISORDERS: GENETIC DIATHESES?

Part of "99 - Molecular and Cellular Genetics of Alcohol Addiction "

Alcohol dependence is often comorbid with other psychiatric disorders, including drug abuse, major depression (MD), anxiety disorders (ADs), and bulimia nervosa (BN), or antisocial personality disorder (ASPD) (17 ,18). Lifetime co-occurrence is more common among women than men and is positively associated with the persistence of alcohol dependence in both men and women (18). Population comorbidity, in which two disorders may be observed to co-occur in excess, may be due to shared causation. A strength of genetic epidemiologic studies is their ability to detect evidence of shared genetic and familial environmental causation.

Severe alcoholism, suicidality, and impulsivity tend to coexist in the same individuals, usually male. The relative risk of alcoholism is significantly increased in males with either ASPD, attention-deficit/hyperactivity disorder, or childhood conduct disorder, and there is evidence of co-inheritance of ASPD and alcoholism. Alcoholism in women is associated with anxiety and affective disorders (18) and increased neuroticism (10). These data notwithstanding, a recent evaluation of the co-inheritance of alcoholism and other disorders revealed that the inheritance of alcoholism is remarkably distinct. Female twins from the Virginia Twin Registry were evaluated for alcoholism, MD, BN, phobia, generalized anxiety disorder (GAD), and panic. Alcoholism emerged as the one disorder with a large disease-specific genetic component: approximately 75% of the genetic variance. In addition, smaller components of the genetic liability to alcoholism also loaded onto a factor common to MD and GAD as well as a factor common to phobia, panic, and BN (19).

COMORBIDITY OF ALCOHOLISM WITH OTHER SUBSTANCE ABUSE: GENETIC DIATHESES?

Part of "99 - Molecular and Cellular Genetics of Alcohol Addiction "

Alcohol, cocaine, opiate, and tobacco (nicotine) dependency co-occur more often in the population than would be expected from their frequencies (20). This raises the possibility that there may be substance-general, as well as substance-specific, components to the heritability of alcoholism (21).

Both alcoholism and drug disorders are familial; two large studies have evaluated the familial aggregation of alcohol and drug dependence (22 ,23). Both studies found that relatives of drug-disorder probands across a wide range of substances, including opioids, cocaine, and cannabis, had a greater rate of drug disorders themselves than relatives of controls. However, this comorbidity occurred largely independently from cotransmission of alcoholism, suggesting that the transmission of alcoholism and other drug disorders is largely independent.

The strongest evidence of a shared, as well as a specific, addictive tendency is between alcohol and nicotine. It has long been observed that there is a relationship between smoking and alcoholism. More than 80% of alcoholics smoke cigarettes and 70% are heavy smokers, compared with 30% of the general population who smoke and 10% who smoke heavily (24). In a multivariate genetic analysis of the use of tobacco and alcohol in 774 MZ and 809 DZ male and female twin pairs from the Virginia Twin Registry, the univariate heritability of alcohol consumption was 0.60 in men and 0.47 in women, and the heritability of tobacco use was 0.49 in men and 0.51 in women. Tobacco use had a stronger loading on this shared genetic factor (0.56 in males, 0.49 in females), than alcohol consumption (0.12 in males, 0.19 in females) (25). However, the precise level of the co-inheritance is less certain than the existence of co-inheritance, and the level of genetic sharing may depend on how the phenotypes are determined. In an analysis of 2,220 MZ and 2,373 DZ U.S. male twin pairs from the National Academy of Sciences–National Research Council’s World War II Twin Registry, a heavy smoking, heavy alcohol genetic factor accounted for 0.45 of the heritable

variance in heavy drinking and 0.35 of the heritable variance in heavy smoking (26), and substance-specific genetic influences contributed 0.22 of the total heritable variance in heavy smoking and 0.55 for heavy alcohol use (26). A study of 3,356 male twin pairs from the U.S. Vietnam Era Twin Registry found a substantial genetic correlation (*r* = 0.68) between alcohol and nicotine dependence. Of the total variance in risk for alcohol dependence, 0.26 was common with the genetic influence on nicotine dependence (27).

THE NEUROBIOLOGY OF ALCOHOL ADDICTION

Part of "99 - Molecular and Cellular Genetics of Alcohol Addiction "

The essential features of addiction are loss of control over consumption, compulsion to obtain the next stimulus, and continuation of abuse despite knowledge of negative health and social consequences. Tolerance and dependence are due to neuroadaptations. Processes of reward and reinforcement play their most crucial role at the start of the path to addiction, after which long-lasting or permanent neuroadaptations occur. It is likely that genetic variation in this neurobiology predisposes some individuals to a pattern of increased craving and loss of control.

Addictive substances affect a range of neurotransmitter systems in different regions of the brain. However, a pathway that appears to be crucial to the action of all addictive drugs is the mesolimbic dopamine system, which originates in the ventral tegmental area (VTA) of the midbrain and projects to the nucleus accumbens (NAC), with projections also to the limbic system and the orbitofrontal cortex (28). The amygdala, hippocampus, and medial prefrontal cortex send excitatory projections to the NAC. The mesolimbic dopamine pathway is associated with the ability to feel pleasure. Serotoninergic neurons originating in the dorsal and median raphe nuclei project to mesolimbic structures, including the VTA and NAC, and may exert inhibitory control on mesolimbic dopamine neuron activity (29) (see Chapter 95).

Alcohol, psychostimulants, opiates, and nicotine (as well as tetrahydrocannabinol and phencyclidine) exert their primary reinforcing or reward effects by releasing dopamine (DA) in the NAC. The acute reinforcing actions of psychostimulant drugs is mediated both by the blockade of DA binding to its transporter, preventing the reuptake of DA from the synaptic cleft (20), and by interaction with multiple DA receptors including D1, D2, and D3 (28). Cocaine blocks the reuptake of serotonin (5-HT) and norepinephrine in a similar fashion. A functional down-regulation of 5-HT_3 receptors in the NAC may contribute to cocaine tolerance (30), whereas chronic alcohol intake increases the sensitivity of 5-HT_3 receptors (31). Chronic cocaine and alcohol administration also disrupts the endogenous opioid system (20). Nicotine's reinforcing effect is through activation of nicotinic receptors in the VTA, ultimately leading to release of dopamine in the NA (32), but the rewarding effects are also mediated by the cholinergic and serotoninergic neurotransmitter systems. The acute reward effects of opioids are enhanced by activation of μ (and possibly also δ) receptors in the VTA.

Enhanced γ-aminobutyric acid (GABA), glutamate, dopaminergic, opioid peptide, and serotoninergic neurotransmission have been associated with acute ethanol administration, and potentially mediate some of alcohol's reinforcing effects (33). In contrast to opioids, which bind to specific receptors, ethanol appears to act on a variety of targets within the cell membrane in a less specific manner, inducing effects on neurotransmitter and neurohormone membrane receptors and receptor-gated and voltage-activated ion channels as well as modulating neurotransmitter release (34). Alterations in calcium channels may be a major component of the changes that occur in the physical dependence on ethanol (35).

GABA is the major inhibitory neurotransmitter in brain. The development of tolerance may be related to ethanol-induced adaptive changes in the $GABA_A$ receptor system. $GABA_A$ receptors are the primary site of action of benzodiazepines and barbiturates, which are used in the treatment of alcohol withdrawal symptoms.

Increasing evidence suggests that ethanol's inhibition of the glutamatergic neurotransmitter pathways, especially at the level of the postsynaptic *N*-methyl-D-aspartate (NMDA) receptor, may be an important cause of its neurotoxic effects (36). Glutamate is the major excitatory brain neurotransmitter with up to 40% of all synapses being glutamatergic (36). Inhibition of GABAergic interneurons mediated via ethanol's effect on the NMDA receptor may result in disinhibition of forebrain glutamatergic neurons that augment dopamine release (37). Changes in the NMDA receptor system may underlie intoxication and withdrawal symptoms (38) as well as "blackouts" (36). Homotarrine (Acamprosate), a structural analogue of glutamate and an NMDA-receptor antagonist, has been shown to almost double the abstinence rate in recovering alcoholics (39).

Some of the rewarding effects of ethanol result from activation of μ opioid receptors in the VTA and/or δ receptors in the NAC by the ethanol-induced release of endogenous β-endorphin from terminals of the VTA and NAC as well as release of enkephalins from intrinsic enkephalin neurons in the NAC (34). The success of the opioid antagonist naltrexone in modifying drinking behavior, controlling craving and preventing relapse in alcoholics indicates that opioid receptor-mediated mechanisms are activated by alcohol (34 ,40).

The co-inheritance of nicotine and alcohol dependence (27) may relate to the finding that ethanol enhancement of DA release in the NAC appears to require activation of nicotinic receptors in the VTA. Neuronal nicotinic acetylcholine receptors are structurally related to $GABA_A$receptors (41).

Activation of these nicotinic receptors may enhance ethanol reinforcement and voluntary intake (42).

Some studies have implicated involvement of the serotonin receptors 5-HT_3 and possibly 5-HT_4 in DA release in the NAC (31 ,43). 5-HT_3 is the only known serotonin receptor that directly gates an ion channel and hence ethanol directly potentiates the effects of serotonin at this site (44). Chronic alcohol intake increases the sensitivity of 5-HT_3 receptors (31) and results in enhanced GABAergic activity. The 5-HT_3 antagonist ondansetron has been shown to reduce alcohol consumption in alcoholics (45).

GENETIC STUDIES OF ALCOHOLISM IN HUMANS

Part of "99 - Molecular and Cellular Genetics of Alcohol Addiction "

In a heterogeneous disease such as alcoholism it is likely that there are different environment-related thresholds and different genes and gene variants. In addition there could be additive or nonadditive (epistatic) interactions between variants of multiple genes. Classic genetic analyses in rodents show that certain alcohol-related phenotypes (e.g., alcohol sensitivity) can be polygenic. On the other hand, human data showing an approximately 2:1 MZ/DZ twin ratio of concordance for alcoholism and high recurrence rates in first-degree relatives of alcoholics followed by a progressively decreasing risk in proportion to relatedness are compatible with additivity of inheritance, and do not favor multiple genes. The alcoholism genes identified so far—*ALDH2* and *ADH2*—were discovered individually but act additively when they co-occur (vide infra).

To dissect the multiple genetic influences on alcoholism vulnerability, it may be necessary or useful to consider several phenotypes representing different aspects of the disease. Attempts have been made to classify alcoholics into more homogeneous clinical groups by severity (dependence or abuse), withdrawal signs, tolerance, medical sequelae, or latent class phenotypes. Cloninger (46) divided alcoholics on a clinical and genetic epidemiologic basis into type 1 (milieu-limited, later onset) and type 2 (early onset, male dominated, associated with ASPD), thus linking premorbid personality with alcoholism vulnerability and identifying an alcoholism subtype, type II, with a stronger genetic predisposition. This classification is supported by a study involving the National Comorbidity Survey of 5,877 people in which Kendler et al. (4) demonstrated that there are two underlying dimensions of liability for alcoholism, drug disorders, ASPD, MD, and GAD that are familially transmitted with moderate specificity: (a) chronic dysphoric symptoms of anxiety and depression, and (b) acting-out behaviors and harmful substance use.

Genetics Of Alcohol Metabolism

At the present time, the genes for alcohol metabolism are the only genes that are known to have a major impact on the development of alcoholism. One gene variant (allele) is protective and the other is a vulnerability allele. Alcohol dehydrogenase (ADH) metabolizes ethanol to acetaldehyde, a toxic intermediate, which is in turn converted to acetate by aldehyde dehydrogenase (ALDH). Approximately half the population of Southeast (SE) Asian countries such as China, Japan, and Korea have functional polymorphisms at four different genes: *ADH2, ADH3, ALDH1*, and *ALDH2*. Across populations, the ALDH2-2 variant appears on a similar genetic background (haplotype) and thus has probably had the same evolutionary origin (47). The most important variants are ALDH2-2 (Glu_{487}-Lys_{487}) and ADH2-2 (Arg_{47}-His_{47}). ALDH2-2 dominantly inactivates *ALDH2*, the ALDH that is mitochondrially localized and responsible for most acetaldehyde metabolism in cells. ADH2-2 is a superactive variant. Allelic variation at ADH3 apparently exerts no independent effect on the risk for alcoholism; however, ADH3-1 is in linkage disequilibrium with ADH2-2 (48 ,49) and is thus also predictive of vulnerability. ADH2-2 and ALDH2-2 raise the levels of acetaldehyde by increasing the rate of synthesis, by decreasing the rate of metabolism, and by interacting additively, but not synergistically (50). The result is that ingestion of even small amounts of ethanol produces an unpleasant reaction characterized by facial flushing, headache, hypotension, palpitations, tachycardia, nausea, and vomiting (51). In an analogous fashion, disulfiram, used in the treatment of alcoholism, and some antiprotozoal drugs such as metronidazole, inhibit *ALDH2* and thereby cause a flushing reaction after alcohol consumption. Therefore, the protective effect of *ALDH2* genotypes can be regarded as analogous to protection with disulfiram, as this flushing reaction, severe in homozygotes but milder in heterozygotes, deters individuals with the protective alleles from becoming alcoholic. The allele frequency of the dominantly acting ALDH2-2 is 0.3 in Japanese and Chinese, and hence about one in two individuals experience flushing after alcohol consumption. Their risk of alcoholism is reduced about four- to tenfold. Approximately 10% of Japanese are ALDH2-2/ALDH2-2 homozygotes. Thus far, only one alcoholic ALDH2-2/ALDH2-2 homozygote has been observed across a series of studies in which several hundred alcoholics have been genotyped, and that individual is the focus of a report (52).

The genetic influence of *ALDH2* variants on alcohol consumption is modified by environment. Tu and Israel (53) found that acculturation accounted for some of the variance (7–11%) in alcohol consumption for SE Asian males born in North America, although the *ALDH2* polymorphism predicted two-thirds of the alcohol consumption and excessive alcohol use. Also, there are large differences in the prevalence of alcohol dependence in populations that have similar *ALDH2* allele frequencies. The frequencies of the *ADH2* and *ALDH2* variants are similar, but the prevalence of alcoholism is 2.9% in Taiwan and and 17.2% in

Korea, suggesting that there are interactions with other genetic and environmental factors (54).

The *ADH2* genotype has been shown to be an independent factor contributing to the risk of alcoholism (50) and acts additively with the ALDH2-2 variant (Table 99.1). A pilot study found that the ADH2-2 allele accounts for 20% to 30% of the alcohol intake variance between two groups of light-drinking and heavy-drinking Israeli Jews, and suggests that the relatively high frequency of the ADH2-2 allele might contribute to the generally perceived lower levels of alcohol consumption and increased sensitivity to alcohol observed among Jews (55 ,56).

ALDH2	*ADH2*	Protective Factor
*2/*2	*2/*2, *1/*2, or *1/*1	High
*2/*1	*2/*2 or *2/*1	
	*1/*1	
*1/*1	*2/*2	
	*2/*1	
	*1/*1	None

ADH, alcohol dehydrogenase; ALDH, aldehyde dehydrogenase.

TABLE 99.1. THE RELATIONSHIP OF ALDH2 AND ADH2 GENOTYPES AND THE RISK FOR ALCOHOLISM IN SOUTHEAST ASIANS (48)

The ADH2-3 allele is a high activity variant identified in approximately 25% of African-Americans, which increases the rate of ethanol metabolism (57). In one report, the presence of ADH2-3 in African-American mothers drinking during pregnancy was associated with a lower rate of alcohol-related birth defects (58).

DETERMINING THE GENETIC BASIS OF VULNERABILITY TO ALCOHOLISM

Part of "99 - Molecular and Cellular Genetics of Alcohol Addiction "

Genetic analysis of complex disorders is complicated by the fact that any single gene is likely to account for only a small part of the variance. To detect subtle genetic effects, large samples are needed. The four methods (59) most widely used are (a) linkage analysis: the inheritance pattern of phenotypes and genotypes are elucidated in pedigrees; (b) allele sharing methods: affected relatives are compared to detect excess genotype sharing; (c) association (case-control) studies: unrelated affected and unaffected individuals are compared; (d) analysis of inbred, transgenic, and gene-knockout animals (principally mice and rats).

Linkage analysis has been successfully employed in finding major genes in diseases such as Huntington's disease, but it has limited power to detect genes of modest effect (60). An alternative approach is to employ an endophenotype as a trait specific marker, e.g., the low-voltage alpha EEG trait that is associated with alcoholism and anxiety disorders (15). Such endophenotypes may be influenced by variation at fewer genes. The brain is relatively inaccessible, so it has been more difficult to obtain biochemical and physiologic measures that identify more specific genetic subtypes as was done decades ago for certain other common, broadly defined diseases (e.g., anemia). Association studies of candidate genes, although laborious, have far greater power for untangling the genetics of complex diseases than linkage analysis (60). New approaches, including TDT (transmission disequilibrium test) analysis (61 ,62) and ethnic matching using informative markers (63), have been used to avoid the problems of population stratification, i.e., ethnic mismatching of cases and controls.

Rodent Models: Quantitative Trait Locus Analyses

Rodent strains are inbred to produce large numbers of genetically identical animals that can be maintained under controlled environmental conditions and intercrossed when required. The neurobiology of reinforcement and reward was elucidated largely through behavioral and anatomic studies in rodents. Several behavioral traits in rodents are continuous and polygenic. Each of the multiple genes responsible for such quantitative traits is termed a quantitative trait locus (QTL). Several QTLs may influence one trait, one QTL may affect several traits, and these QTLs can be individually mapped, with the ultimate goal being to identify genes that play a role in human addiction to alcohol. The knockout of an individual gene in the mouse can reveal a potential role for the equivalent (homologous) gene in the human.

Several QTLs for alcohol-associated behaviors have been identified in mice by using recombinant inbred strains that differ widely with respect to many alcohol-related traits, and by follow-up studies using interstrain crosses and congenics. The behaviors for which QTLs have been mapped include acute and chronic alcohol withdrawal sensitivity, alcohol consumption, and alcohol-associated hypothermia. Buck et al. (64) have shown that 68% of the genetic variability for genes influencing alcohol withdrawal severity can be assigned to QTLs on mouse chromosomes 1, 4, and 11. The locus on chromosome 11 accounted for 12% of the genetic variability in withdrawal liability and was near the genes for the γ_2, α_1, and α_6 subunits of $GABA_A$ receptors. Furthermore, a γ_2 subunit polymorphism has now been found to be genetically correlated with alcohol withdrawal severity in mice (65). QTLs for alcohol-induced hypothermia, alcohol consumption, and certain responses to amphetamine and morphine are located on chromosome 1 (66) and also on chromosome 9 near the serotonin $5\text{-}HT_{1B}$ receptor and dopamine D2 receptor genes (67). These genetic interrelationships between different phenotypes indicate that the same genes influence different alcohol-associated behaviors, that several genes may affect one phenotype, and that some loci

for drug abuse are not drug specific. This is also evident in studies of mice in which specific genes that map to the QTL regions or candidate genes located elsewhere in the mouse genome have been knocked out. 5-HT_{1B} knockout mice drink twice as much ethanol, are less intoxicated, and show enhanced aggression compared with normal mice (68). On chronic exposure to alcohol they show less evidence of tolerance. These mice also work harder to self-administer cocaine and show an increased locomotor response, behaving as if already sensitized to the drug (69).

The dopamine-related genes that have been knocked out in mice are the DRD4 dopamine receptor, which is located at the site of one of the alcohol QTLs, the D1 and D2 dopamine receptors, the dopamine transporter, and VMAT2 (the vesicular transporter). The DRD4 knockout mice appear to be supersensitive to ethanol, cocaine, and amphetamine (70). Mice lacking the D2 receptor consume less alcohol than normal mice (71). VMAT2 knockout mice have a pronounced supersensitivity to cocaine, amphetamine, and ethanol (72).

For morphine preference, three loci identified on murine chromosomes 1, 6, and 10 are apparently responsible for nearly 85% of the genetic variance in this trait (73). The μ opioid receptor gene is located at the site of the largest QTL, and this QTL also affects consumption of alcohol and cocaine (73).

QTL analysis in rodents has limitations. Frequently the result is a large genomic region of interest rather than a gene. There may be functional compensation in knockout mice during development. Mice and humans may not share the same functional variants at the same allele; for example, the ALDH2-2 allele is not even present in all human populations and is not found in mice. Another problem is that behaviors in mice cannot be freely extrapolated to humans; for example alcohol preference in mice may well be taste preference (74). However, QTL analyses in mice are useful for the identification of candidate genes and gene regions.

GENETICS OF REWARD NEUROCIRCUITS, AND NEUROCIRCUITS REGULATING IMPULSE CONTROL

Part of "99 - Molecular and Cellular Genetics of Alcohol Addiction "

Candidate Gene Approach: Case-Control Association Studies

A logical approach to understanding vulnerability to alcohol addiction is to look directly for variants in genes involved in neurotransmitter pathways implicated in ethanol use. Of particular interest is the reward pathway, incorporating serotoninergic, GABAergic, dopaminergic, opioid, and glutamatergic neurotransmission, and the largely serotoninergic impulse-control pathway. Genes for neurotransmitter metabolizing enzymes, transporters, and receptors are good candidates. Because of the complexity of causation of alcoholism, any genetic determinants of vulnerability to alcoholism and sensitivity to alcohol's effects are likely to be subtle.

Serotonin

Serotonin is involved in behavioral inhibition and is a target for the pharmacologic treatment of alcoholism. Selective serotonin reuptake inhibitors play a limited role in modifying craving for alcohol and also modify other comorbid behaviors such as depression and anxiety.

Pathologically low levels of serotonin may contribute to impulsivity and ASPD; for example, a group of criminal, alcoholic Finns was shown to have low cerebrospinal fluid (CSF) 5-hydroxyindolacetic acid (5-HIAA), the lowest levels being found in those who had committed impulsive crimes (75). These are the alcoholics who would be most likely to have a serotonin gene variant affecting function.

Serotonin Transporter

The availability of brainstem serotonin transporter, measured by (I-123) β-CIT and single photon emission computed tomography, has been found to be significantly reduced in alcoholics, and correlated with ratings of depression and anxiety during withdrawal (76). A functional polymorphism, 5-HTTLPR, in the serotonin transporter promoter region (77) has been associated fairly consistently with anxiety/dysphoria (77 ,78). Several association analyses have shown that the s-allele, which reduces transcriptional efficiency, is increased in French alcoholics (79), severely affected German alcoholics (80), and early-onset, violent Finnish alcoholics with ASPD (81). However, neither linkage nor association for the s-allele was found in a family-based TDT analysis of U.S. alcoholics from the COGA (Collaborative Study on the Genetics of Alcoholism) data set, some with withdrawal symptoms (82). A Japanese association study of alcoholics with withdrawal seizures was also negative (83), and in a small study the long allele was found to be associated with reduced sensitivity to alcohol, i.e., with individuals who may be more vulnerable to developing alcoholism (84). Population stratification may be a problem with these association studies as allele frequencies have been shown to vary in European-American, African-American, and Japanese populations (85).

Serotonin-Metabolizing Enzymes

A tryptophan hydroxylase (TPH) intron variant that affects splicing is associated with reduced 5-HIAA and suicidality in impulsive alcoholics (86 ,87).

Serotonin Receptors

Several serotonin receptors are known to be abundant in the NAC: 5-HT_{1B}, 5-HT_{2C}, 5-HT_3, 5-HT_4, and 5-HT_6. There are as yet few published studies in which these serotonin receptors have been genotyped in humans.

Studies in rats suggest that activation of 5-HT_{1B} receptors

in the NAC may be inhibitory on the behavioral effects of elevated mesolimbic dopamine transmission (88) by primarily modulating the activity of glutamatergic hippocampo-accumbens pathways and only secondarily altering NAC DA levels (89). In a large sib-pair linkage analysis of Finnish alcoholic criminal offenders, significant evidence of linkage and association of antisocial alcoholism to *HTR1B G861C* was found, and this was also observed in a Southwest American Indian tribe, suggesting that a locus predisposing to antisocial alcoholism may be linked to *HTR1B* at 6q13-15 (90).

Activation of 5-HT_{2C} receptors inhibits DA release in the NAC (91). However, the functional Cys23Ser polymorphism does not appear to be associated with alcohol dependence (92).

5-HT_3 receptors may be involved in several facets of alcohol-seeking behavior, alcohol intoxication, and addiction (93); however, at the present time there are no published studies on the role of 5-HT_3 variants in alcoholics.

GABA Receptors

Cross-tolerance of benzodiazepines with ethanol and their effectiveness in treating alcohol withdrawal suggests that $GABA_A$ receptors play a key role in alcohol's effects. The $GABA_A$ receptor exists as a number of subtypes that are composed of combinations of at least 14 different subunits. Preliminary studies indicate that the Pro385Ser substitution in $GABA_A\alpha6$ may play a role in benzodiazepine sensitivity (94) and may be associated with a lower level of response to the acute effects of alcohol (84). Several studies have found associations between $GABA_A\alpha6$ and alcohol dependence (84 ,95) and antisocial alcoholism (96). Differences in allele frequencies between alcoholics and controls have been found in $GABA_A\alpha3$ but not in a $GABA_A\alpha1$ (97). There are positive (95) and negative (96) association studies for $GABA_A\beta2$ and alcohol dependence.

Dopamine

Dopamine is involved in arousal, reward, and motivation. Structural variants, some altering function or level of expression of gene product, have been found in the dopamine transporter (DAT) and in several dopamine receptor genes (*DRD2, DRD3*, and *DRD4*). At the present time, no role for variation in dopamine-related genes in alcoholism has been consistently demonstrated. The controversial association of a *DRD2* dopamine receptor polymorphism with alcoholism has been replicated in some case-control studies but not in numerous others (98), nor was it supported in two very large family linkage studies (99), one of which used the functionally impaired 311Cys variant. These family studies were not subject to the potential problems of ethnic stratification inherent in some of the *DRD2* case-control association studies (100). Two recently discovered transcriptionally significant promoter polymorphisms offer promising tools for understanding the roles of *DRD2* (101) and DAT (102) in alcoholism.

Opioid Receptors

Three endogenous opioid receptors (μ, δ, and κ) are the targets of the major opioid peptides (β-endorphin, enkephalins, and dynorphins, respectively). The rewarding properties of μ- and δ-receptor ligands are brought about by activation of the mesolimbic dopamine system. Activation of κ receptors is dysphoric. Human and animal studies implicate the opioid system, particularly the μ opioid receptor, in both initial sensitivity or response to alcohol, and in the rewarding or reinforcing effects of alcohol. Subjects at high risk for alcoholism have been shown to have lower basal plasma β-endorphin levels but a more pronounced release after exposure to ethanol (103). Some studies have found associations of the gene encoding the μ opioid receptor, OPRM1, with some form of drug dependence (104). However, association and sib-pair linkage analyses of Asn40Asp, a μ opioid receptor polymorphism, in 100 U.S. Caucasians, 324 Finnish Caucasians, and 367 American Indians showed no significant association, even though the study had 80% power to detect a small to moderate effect of OPRM1 variation on alcohol dependence (105). Findings were also negative from a large study of German alcoholics (106) and a study of 891 drug- or alcohol-dependent subjects from African-American, European-American, and Hispanic origins (107).

NMDA Receptors

At the present time there are no published studies on the role of NMDA variants in alcoholism. Such studies would be of interest because of the role of NMDA receptors in reward, intoxication, and withdrawal, and potentially for the pharmacogenetics of acamprosate.

Nicotinic Receptors

Two classes of neuronal nicotinic acetylcholine receptor (nAchR) subunits (eight α and three β) have been identified (108). The most abundant receptor subtype in brain is composed of β_2 and α_4 subunits (109). Several lines of evidence, including drug preference in knockout mice (110), suggest that the nAchR β_2 subunit gene (*CHRNB2*) is involved in the reinforcing properties of nicotine. However, none of the *CHRNB2* variants found so far in humans has been associated with nicotine dependence (109). This gene has yet to be genotyped in alcoholics.

Whole Genome Linkage Scans

The power of the genetic linkage analysis approach has been greatly improved by the recent collection of very large, carefully

phenotyped, family and population data sets such as that of COGA, a multicenter, family genetic study and the National Institute on Alcohol Abuse and Alcoholism (NIAAA) Southwestern Indian family sample. Two studies utilizing these data sets have detected evidence of linkage of alcoholism to several chromosomal regions, with some convergent findings (111 ,112). In the Southwestern American Indian population isolate, suggestive evidence was found for linkage of alcohol dependence to the ADH region on chromosome 4q and to two regions harboring neurogenetic candidate genes. Those locations were chromosome 11p, in close proximity to the *DRD4* dopamine receptor and tyrosine hydroxylase gene (the rate-limiting enzyme in dopamine biosynthesis), and chromosome 4p, near a GABA receptor gene cluster (111). In the COGA families, which derive from the cosmopolitan, diverse population of the United States, modest evidence was also found of linkage to the ADH region on 4q. There was also evidence of linkage to chromosomes 1 and 7, and to chromosome 2 at the location of the opioid receptor gene (112). In addition, there was evidence of linkage of the P300 event-related potential alcoholism-associated trait to chromosome 6q in the region of a glutamate receptor (GRIK2), and to chromosome 2q near the location of two acetylcholine receptors (113).

DISCUSSION

Part of "99 - Molecular and Cellular Genetics of Alcohol Addiction "

There is abundant evidence of substantial heritability (0.5 to 0.6) of both broad and narrow definitions of alcoholism in men and women. Although the quantitative role of genetic risk factors is approximately equal in both sexes, the lower concordance of opposite-sex pairs suggests some gender-specific action of genes.

Genetic vulnerability to alcoholism may originate in personality or psychiatric traits that predispose to alcohol-seeking behavior, differential response to the effects of alcohol, or differential predisposition to addiction. Studies of the co-inheritance of alcoholism and other psychiatric disorders are beginning to emerge. There is evidence of co-inheritance of ASPD and alcoholism in men. Although alcoholism is associated with anxiety and affective disorders in women, it has been shown that 75% of the genetic variance of alcoholism is disease specific. Alcohol, nicotine, and other substance abuse disorders have been noted to co-occur, yet recent studies have shown that the transmission of alcoholism is largely independent of that of other drugs of abuse with the exception of nicotine, with which there is a substantial (0.68) genetic correlation.

The mesolimbic dopamine system is fundamental to the neurobiology of addiction. We are only at the very beginning stages of understanding the complexities of ethanol's interactions with this system. Enhanced GABA, glutamate, dopaminergic, opioid peptide, and serotoninergic neurotransmission have been associated with acute ethanol administration and potentially mediate some of alcohol's reinforcing effects. Genetic variants in the serotoninergic system—5-HT_{1B}, TPH, and possibly 5-HTTLPR—have been associated with alcoholism, particularly in males with antisocial, impulsive features. Several studies have found associations between $GABA_A\alpha 6$ and alcohol dependence. There have been no conclusive genetic findings in the dopaminergic or opioid systems to date.

Future studies are likely to focus on finding genetic variants in the neuroreceptors and ion channels that have been demonstrated to be affected by ethanol, including $GABA_A$ receptors, NMDA receptors, non-NMDA glutamate receptors, 5-HT_3 receptors, voltage-gated calcium channels, and neuronal nACh receptors. Of particular interest will be functional genetic variants that are directly capable of altering reward, tolerance, and withdrawal, thereby predisposing individuals to addiction to alcohol.

REFERENCES

1. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington DC: American Psychiatric Press, 1994.

2. Schuckit MA, Smith TL, Danko GP, et al. Five-year clinical course associated with DSM-IV alcohol abuse or dependence in a large group of men and women. *Am J Psychiatry* 2001;158(7):1084–1090.

3. Helzer JE, Burnam A, McEvoy LT. Alcohol abuse and dependence. In: Robins L, Reiger D, eds. Psychiatric disorders in America: the Epidemiological Catchment Area Study. New York: Macmillan, 1991:81–115.

4. Kendler KS, Davis CG, Kessler RC. The familial aggregation of common psychiatric and substance abuse disorders in the National Comorbidity Survey: a family history study. *Br J Psychiatry* 1997;170:541–548.

5. Midanik L. Familial alcoholism and problem drinking in a national drinking survey. *Addict Behav* 1983;8:133–141.

6. Kendler KS, Neale MC, Heath AC, et al. A twin-family study of alcoholism in women. *Am J Psychiatry* 1994;151(5):707–715.

7. Goodwin DW. Alcoholism and heredity. *Arch Gen Psychiatry* 1979;36:57–61.

8. Heath AC. Genetic influences on alcoholism risk. A review on adoption and twin studies. *Alcohol Health Res World* 1995;19(3):166–171.

9. Kendler KS, Heath AC, Neale MC, et al. A population-based twin study of alcoholism in women. *JAMA* 1992;268(14):1877–1882.

10. Heath AC, Bucholz KK, Madden PAF, et al. Genetic and environmental contributions to alcohol dependence risk in a national twin sample: consistency of findings in women and men. *Psychol Med* 1997;27:1381–1396.

11. Cloninger CR, Bohman M, Sigvardsson S. Inheritance of alcohol abuse: cross-fostering analysis of adopted men. *Arch Gen Psychiatry* 1981;38:861–868.

12. Prescott CA, Kendler KS. Genetic and environmental contributions to alcohol abuse and dependence in a population-based sample of male twins. *Am J Psychiatry* 1999;156(1):34–40.

13. Prescott CA, Aggen SH, Kendler KS. Sex differences in the source of genetic liability to alcohol abuse and dependence in a population-based sample of U.S. twins. *Alcohol Clin Exp Res* 1999;23(7):1136–1144.

14. Begleiter H, Porjesz B, Bihari B, et al. Event-related brain potentials in boys at risk for alcoholism. *Science* 1984;225:1493–1495.

15. Enoch M-A, White KV, Harris CR, et al. Association of low-voltage alpha EEG with a subtype of alcohol use disorders. *Alcohol Clin Exp Res* 1999;23(8):1312–1319.

16. Schuckit MA. Low level of response to alcohol as a predictor of future alcoholism. *Am J Psychiatry* 1994;151(2):184–189.

17. Regier DA, Farmer ME, Rae DS, et al. Comorbidity of mental disorders with alcohol and other drug abuse: results from the Epidemiologic Catchment Area (ECA) Study. *JAMA* 1990;264:2511–2518.

18. Kessler RC, Crum RM, Warner LA, et al. Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Arch Gen Psychiatry* 1997;54:313–321.

19. Kendler KS, Walters EE, Neale MC, et al. The structure of the genetic and environmental risk factors for six major psychiatric disorders in women. *Arch Gen Psychiatry* 1995;52:374–383.

20. Kreek MJ. Cocaine, dopamine and the endogenous opioid system. *J Addict Dis* 1996;15(4):73–96.

21. Goldman D, Bergen A. General and specific inheritance of substance abuse and alcoholism. *Arch Gen Psychiatry* 1998;55(11):964–965.

22. Bierut LJ, Dinwiddie SH, Begleiter H, et al. Familial transmission of substance dependence: alcohol, marijuana, cocaine and habitual smoking. *Arch Gen Psychiatry* 1998;55:982–988.

23. Merikangas KR, Stolar M, Stevens DE, et al. Familial transmission of substance use disorders. *Arch Gen Psychiatry* 1998;55:973–979.

24. NIAAA. Alcohol and tobacco. In: *Alcohol alert*. Washington, DC: NIAAA, 1998;39:1–4.

25. Hettema JM, Corey LA, Kendler KS. A multivariate genetic analysis of the use of tobacco, alcohol and caffeine in a population-based sample of male and female twins. *Drug Alcohol Depend* 1999;57(1):69–78.

26. Swan GE, Carmelli D, Cardon LR. Heavy consumption of cigarettes, alcohol and coffee in male twins. *J Stud Alcohol* 1997;58:182–190.

27. True WR, Xian H, Scherrer JF, et al. Common genetic vulnerability for nicotine and alcohol dependence in men. *Arch Gen Psychiatry* 1999;56:655–661.

28. Koob GF, Le Moal M. Drug abuse: hedonic homeostatic dysregulation. *Science* 1997;278:52–58.

29. Di Matteo V, Di Giovanni G, Di Mascio M, et al. SB 242084, a selective serotonin2C receptor antagonist increases dopaminergic transmission in the mesolimbic system. *Neuropharmacology* 1999;38(8):1195–205.

30. King GR, Xiong Z, Ellinwood EH. Blockade of accumbens 5-HT3 receptor down-regulation by ondansetron administered during continuous cocaine administration. *Eur J Pharmacol* 1999;364(2 and 3):79–87.

31. Yoshimoto K, Yayama K, Sorimachi Y, et al. Possibility of 5-HT3 receptor involvement in alcohol dependence: a microdialysis study of nucleus accumbens dopamine and serotonin release in rats with chronic alcohol consumption. *Alcohol Clin Exp Res* 1996;20(9 suppl):311A-319A.

32. Jarvik ME, Schneider NG. Nicotine. In: Lowinson JH, Ruiz P, Millman RB, eds. *Substance abuse: a comprehensive textbook*, second ed. Baltimore: Williams & Wilkins, 1992:334–356.

33. Markou A, Kosten TR, Koob GF. Neurobiological similarities in depression and drug dependence: a self medication hypothesis. *Neuropsychopharmacology* 1998;18(3):135–174.

34. Herz A. Endogenous opioid systems and alcohol addiction. *Psychopharmacology* 1997;129:99–111.

35. Little HJ. The role of calcium channels in drug dependence. *Drug Alcohol Depend* 1995;38:173–194.

36. Tsai G, Coyle JT. The role of glutamatergic neurotransmission in the pathophysiology of alcoholism. *Annu Rev Med* 1998;49:173–184.

37. Nestler EJ, Hope BT, Widnell KL. Drug addiction: a model for the molecular basis of neural plasticity. *Neuron* 1993;11:995–1005.

38. Koob GF, Roberts AJ, Schulteis G, et al. Neurocircuitry targets in ethanol reward and dependence. *Alcohol Clin Exp Res* 1998;22(1):3–9.

39. Sass H, Soyka M, Mann K, et al. Relapse prevention by acamprosate. Results from a placebo-controlled study on alcohol dependence. *Arch Gen Psychiatry* 1996;53:673–680.

40. O'Malley SS. Opioid antagonists in the treatment of alcohol dependence: clinical efficacy and prevention of relapse. *Alcohol Alcohol* 1996;31(Suppl. 1):77–81.

41. Narahashi T, Aistrup GL, Marszalec W, et al. Neuronal nicotinic acetylcholine receptors: a new target site of ethanol. *Neurochem Int* 1999;35(2):131–141.

42. Harris RA. Ethanol actions on multiple ion channels: which are important? *Alcohol Clin Exp Res* 1999;23(10):1563–1570.

43. Mylecharane EJ. Ventral tegmental area 5-HT receptors: mesolimbic dopamine release and behavioral studies. *Behav Brain Res* 1996;73(1 and 2):1–5.

44. Grant KA. The role of 5-HT3 receptors in drug dependence. *Drug Alcohol Depend* 1995;38:155–171.

45. Sellers EM, Tonetto T, Romach MK, et al. Clinical efficacy of the 5-HT3 antagonist ondansetron in alcohol abuse and dependence. *Alcohol Clin Exp Res* 1994;18:879–885.

46. Cloninger CR. Neurogenetic adaptive mechanisms in alcoholism. *Science* 1987;236:410–416.

47. Peterson RJ, Goldman D, Long JC. Nucleotide sequence diversity in non-coding regions of ALDH2 as revealed by restriction enzyme and SSCP analysis. *Hum Genet* 1999;104(2):177–87.

48. Chen CC, Lu RB, Chen YC, et al. Interaction between the functional polymorphisms of the alcohol-metabolism genes in protection against alcoholism. *Am J Hum Genet* 1999;65(3):795–807.

49. Osier M, Pakstis AJ, Kidd JR el al. Linkage disequilibrium at the ADH2 and ADH3 loci and risk of alcoholism. *Am J Hum Genet* 1999;64(4):1147–1157.

50. Thomasson HR, Edenburg HJ, Crabb DW, et al. Alcohol and aldehyde dehydrogenase genotypes and alcoholism in Chinese men. *Am J Hum Genet* 1991;48:677–681.

51. Harada S, Agarwal DP, Goedde HW. Aldehyde dehydrogenase deficiency as cause of facial flushing reaction to alcohol in Japanese. *Lancet* 1982;2(8253):982.

52. Chen Y-C, Lu R-B, Peng G-S, et al. Alcohol metabolism and cardiovascular response in an alcoholic patient homozygous for the ALDH2-2 variant gene allele. *Alcohol Clin Exp Res* 1999;23(12):1853–1860.

53. Tu GC, Israel Y. Alcohol consumption by orientals in North America is predicted largely by a single gene. *Behav Genet* 1995;25(1):59–65.

54. Goldman D. Genetic transmission. In:Galanter M, ed. *Recent developments in alcoholism, vol 11: ten years of progress*. New York: Plenum Press, 1993;14:231–248.

55. Neumark YD, Friedlander Y, Thomasson HR, Li TK. Association of the ADH2-2 allele with reduced ethanol consumption in Jewish men in Israel: a pilot study. *J Stud Alcohol* 1998;59(2):133–139.

56. Monteiro MG, Klein JL, Schuckit MA. High levels of sensitivity to alcohol in young adult Jewish men: a pilot study. *J Stud Alcohol* 1991;52(5):464–469.

57. Thomasson HR, Beard JD, Li TK. ADH2 gene polymorphisms are determinants of alcohol pharmacokinetics. *Alcohol Clin Exp Res* 1995;19(6):1494–1499.

58. McCarver DG, Thomasson HR, Martier SS, et al. Alcohol dehydrogenase-2-3 allele protects against alcohol-related birth defects among African Americans. *J Pharmacol Exp Ther* 1997;283(3):1095–1101.

59. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994;265:2037–2048.

60. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;273:1516–1517.

61. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993;52:506–516.

62. Ewens WJ, Spielman RS. The transmission/disequilibrium test: history, subdivision, and admixture. *Am J Hum Genet* 1995;57:455–464.

63. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 1999;65:220–228.

64. Buck KJ, Metten P, Belknap JK, et al. Quantitative trait loci involved in genetic predisposition to acute alcohol withdrawal in mice. J Neurosci 1997;17(10):3946–3955.

65. Buck KJ, Hood HM. Genetic association of a GABA(A) receptor gamma2 subunit variant with severity of acute physiological dependence on alcohol. *Mamm Genome* 1998;9(12):975–978.

66. Crabbe JC, Belknap JK, Mitchell SR, et al. Quantitative trait loci mapping of genes that influence the sensitivity and tolerance to ethanol-induced hypothermia in BXD recombinant inbred mice. *J Pharmacol Exp Ther* 1994;269(1):184–192.

67. Crabbe JC, Buck KJ, Metten P, et al. Strategies for identifying genes underlying drug abuse susceptibility. In: Lee TNH, ed. *Molecular approaches to drug abuse research*, vol 3. *NIDA Res Monogr* 1996;161:201–219.

68. Crabbe JC, Phillips TJ, Feller DJ, et al. Elevated alcohol consumption in null mutant mice lacking 5-HT1B serotonin receptors. *Nat Genet* 1996;14(1):98–101.

69. Rocha BA, Scearce-Levie K, Hen R. Increased vulnerability to cocaine in mice lacking the serotonin-1B receptor. *Nature* 1998;393(6681):175–178.

70. Rubinstein M, Phillips TJ, Bunzow JR, et al. Mice lacking dopamine D4 receptors are supersensitive to ethanol, cocaine and methamphetamine. *Cell* 1997;90:991–1001.

71. Phillips TJ, Brown KJ, Burkhart-Kasch S, et al. Alcohol preference and sensitivity are markedly reduced in mice lacking dopamine D2 receptors. *Nat Neurosci* 1998;1(7):610–615.

72. Wang YM, Gainetdinov RR, Fumagalli F, et al. Knockout of the vesicular monoamine transporter 2 gene results in neonatal death and supersensitivity to cocaine and amphetamine. *Neuron* 1997;19(6):1285–1296.

73. Berrettini WH, Ferraro TN, Alexander RC, et al. Quantitative trait loci mapping of three loci controlling morphine preference using inbred mouse strains. *Nat Genet* 1994;7:54.

74. Grisel JE, Metten P, Hughes AM, et al. DBA/2J Mice like beer. RSA abstract 52. *Alcohol Clin Exp Res* 1998;suppl 22(3):12A.

75. Linnoila M, Virkkunen M, Roy A. Biochemical aspects of aggression in man. *Clin Neuropharmacol* 1986;9(suppl 4):377–379.

76. Heinz A, Ragan P, Jones DW, et al. Reduced central serotonin transporters in alcoholism. *Am J Psychiatry* 1998;155(11):1544–1549.

77. Lesch KP, Bengel D, Heils A, et al. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 1996;274(5292):1527–1531.

78. Mazzanti CM, Lappalainen J, Long JC, et al. Role of the serotonin transporter promoter polymorphism in anxiety-related traits. *Arch Gen Psychiatry* 1998;55:936–940.

79. Hammoumi S, Payen A, Favre JD, et al. Does the short variant of the serotonin transporter linked polymorphic region constitute a marker for alcohol dependence? *Alcohol* 1999;17(2):107–112.

80. Sander T, Harms H, Lesch KP, et al. Association analysis of a regulatory variation of the serotonin transporter gene with severe alcohol dependence. *Alcohol Clin Exp Res* 1997;21(8):1356–1359.

81. Hallikainen T, Saito T, Lachman HM, et al. Association between low activity serotonin transporter promoter genotype and early onset alcoholism with habitual impulsive violent behavior. *Mol Psychiatry* 1999;4(4):385–388.

82. Edenberg HJ, Reynolds J, Koller DL, et al. A family-based analysis of whether the functional promoter alleles of the serotonin transporter gene HTT affect the risk for alcohol dependence. *Alcohol Clin Exp Res* 1998;22(5):1080–1085.

83. Ishiguro H, Saito T, Akazawa S, et al. Association between drinking-related antisocial behavior and a polymorphism in the serotonin transporter gene in a Japanese population. *Alcohol Clin Exp Res* 1999;23(7):1281–1284.

84. Schuckit MA, Mazzanti C, Smith TL, et al. Selective genotyping for the role of 5-HT2A, 5-HT2C, and $GABA_{A6}$ receptors and the serotonin transporter in the level of response to alcohol: a pilot study. *Biol Psychiatry* 1999;45:647–651.

85. Gelernter J, Kranzler H, Cubells JF. Serotonin transporter protein (SLC6A4) allele and haplotype frequencies and linkage disequilibria in African- and European-American and Japanese populations and in alcohol-dependent subjects. *Hum Genet* 1997;101(2):243–246.

86. Nielsen DA, Goldman D, Virkkunen M, et al. Suicidality and 5-hydroxyindole acetic acid concentration associated with a tryptophan hydroxylase polymorphism. *Arch Gen Psychiatry* 1994;51(1):34–38.

87. Nielsen DA, Virkkunen M, Lappalainen J, et al. A tryptophan hydroxylase gene marker for suicidality and alcoholism. *Arch Gen Psychiatry* 1998;55:593–602.

88. Fletcher PJ, Korth KM. Activation of 5-HT1B receptors in the nucleus accumbens reduces amphetamine-induced enhancement of responding for conditioned reward. *Psychopharmacol (Berl)* 1999;142(2):165–174.

89. Boulenguez P, Rawlins JN, Chauveau J, et al. Modulation of dopamine release in the nucleus accumbens by 5-HT1B agonists: involvement of the hippocampo-accumbens pathway. *Neuropharmacology* 1996;35(11):1521–1529.

90. Lappalainen J, Long JC, Eggert M, et al. Linkage of antisocial alcoholism to the serotonin 5-HT1B receptor gene in 2 populations. *Arch Gen Psychiatry* 1998;55(11):989–994.

91. Willins DL, Meltzer HY. Serotonin 5-HT2C agonists selectively inhibit morphine-induced dopamine efflux in the nucleus accumbens. *Brain Res* 1998;781(1 and 2):291–299.

92. Lappalainen J, Long JC, Virkkunen M, et al. HTR2C Cys23Ser polymorphism in relation to CSF monoamine metabolite concentrations and DSM-III-R psychiatric diagnoses. *Biol Psychiatry* 1999;46:821–826.

93. Lovinger DM. 5-HT3 receptors and the neural actions of alcohols: an increasingly exciting topic. *Neurochem Int* 1999;35:125–130.

94. Iwata N, Cowley DS, Radel M, et al. Relationship between a $GABA_{A\alpha6}$ Pro385Ser substitution and benzodiazepine sensitivity. *Am J Psychiatry* 1999;156:1447–1449.

95. Loh EW, Smith I, Murray R, et al. Association between variants at the GABAAbeta2, GABAAalpha6 and GABAAgamma2 gene cluster and alcohol dependence in a Scottish population. *Mol Psychiatry* 1999;4(6):539–544.

96. Sander T, Ball D, Murray R, et al. Association analysis of sequence variants of GABA(A) alpha 6, beta 2, and gamma2 gene cluster and alcohol dependence. *Alcohol Clin Exp Res* 1999;23(3):427–431.

97. Parsian A, Cloninger CR. Human GABAA receptor alpha 1 and alpha 3 subunits genes and alcoholism. *Alcohol Clin Exp Res* 1997;21(3):430–433.

98. Paterson AD, Sunohara GA, Kennedy JL. Dopamine D4 receptor gene: novelty or nonsense? *Neuropsychopharmacology* 1999;21(1):3–16.

99. Goldman D, Urbanek M, Guenther D, et al. Linkage and association of a functional DRD2 variant (Ser311Cys) and DRD2 markers to alcoholism, substance abuse and schizophrenia in Southwestern American Indians. *Am J Med Genet (Neuropsychiatr Genet)* 1997;74:386–394.

100. Goldman D, Urbanek M, Guenther D, et al. A functionally deficient DRD2 variant (Ser311Cys) is not linked to alcoholism and substance abuse. *Alcohol* 1998;16(1):47–52.

101. Ishiguro H, Arinami T, Saito T, et al. Association study between the -141C Ins/Del and TaqI A polymorphisms of the dopamine D2 receptor gene and alcoholism. *Alcohol Clin Exp Res* 1998:22(4):845–848.

102. Ueno S, Nakamura M, Mikami M, et al. Identification of a novel polymorphism of the human dopamine transporter (DAT1) gene and the significant association with alcoholism. *Mol Psychiatry* 1999;4(6):552–557.

103. Gianoulakis C. Implications of endogenous opioids and dopamine in alcoholism: human and basic science studies. *Alcohol Alcohol Suppl* 1996;1:33–42.

104. Town T, Abdullah L, Crawford F, et al. Association of a functional mu-opioid receptor allele (+118A) with alcohol dependency. *Am J Med Genet* 1999;88(5):458–461.

105. Bergen AW, Kokoszka J, Peterson R, et al. Mu opioid receptor gene variants: lack of association with alcohol dependence. *Mol Psychiatry* 1997;2:490–494.

106. Sander T, Gscheidel N, Wendel B, et al. Human mu-opioid receptor variation and alcohol dependence. *Alcohol Clin Exp Res* 1998;22(9):2108–2110.

107. Gelernter J, Kranzler H, Cubells JF. Genetics of two mu opioid receptor gene (OPRM1) exon 1 polymorphisms: population studies, and allele frequencies in alcohol- and drug-dependent subjects. *Mol Psychiatry* 1999;4(5):476–483.

108. Boyd RT. The molecular biology of neuronal nicotinic acetylcholine receptors. *Crit Rev Toxicol* 1997;27:299–318.

109. Silverman MA, Neale MC, Sullivan PF, et al. Haplotypes of four novel single nucleotide polymorphisms in the nicotinic acetylcholine receptor ß2-subunit (CHRNB2) gene show no association with smoking initiation or nicotine dependence. *Am J Med Genet* 2000;96:646–653.

110. Picciotto MR, Zoli M, Rimondini R, et al. Acetylcholine receptors containing the beta2 subunit are involved in the reinforcing properties of nicotine. *Nature* 1998;391(6663):173–177.

111. Long JC, Knowler WC, Hanson RL, et al. Evidence for genetic linkage to alcohol dependence on chromosomes 4 and 11 from an autosome-wide scan in an American Indian population. *Am J Med Genet (Neuropsychiatr Genet)* 1998;81:216–221.

112. Reich T, Edenberg HJ, Goate A, et al. Genome-wide search for genes affecting the risk for alcohol dependence. *Am J Med Genet* 1998;81(3):207–215.

113. Begleiter H, Porjesz B, Reich T, et al. Quantitative trait loci analysis of human event-related brain potentials: P3 voltage. *Electroencephalogr Clin Neurophysiol* 1998;108:244–250.

100

Ethanol Abuse, Dependence, and Withdrawal: Neurobiology and Clinical Implications

John H. Krystal

Boris Tabakoff

John H. Krystal: Department of Psychiatry, Yale University School of Medicine, VA Alcohol Research Center, VA Connecticut Healthcare System, West Haven, Connecticut.

Boris Tabakoff: Department of Pharmacology, University of Colorado School of Medicine, Denver, Colorado.

Ethanol produces a striking array of behavioral effects in humans that are dependent on the dose of ethanol administered, whether ethanol levels are rising or falling, and the rate of change of ethanol levels (1). The euphoric, stimulant, and anxiolytic effects of ethanol contribute to its wide recreational use, whereas its sedative effects contribute to its consumption as a nonprescription hypnotic. Yet, at intoxicating doses, ethanol clouds memory and judgment (2), slows information processing and reaction time (3), impairs coordination (4), and disinhibits impulsive or aggressive behavior (5). Perhaps, then, it is not surprising that some of the most damaging consequences of ethanol abuse reflect the impact of ethanol intoxication on behaviors, such as driving, that place a high demand on these cognitive functions (6).

With chronic administration, ethanol produces both adaptation and neurotoxicity in the brain, accounting for tolerance and dependence. The ethanol withdrawal syndrome (7) includes anxiety, insomnia, and symptoms of sympathetic nervous system arousal. With more severe levels of dependence and repeated episodes of withdrawal, abstinence may be associated with substantial sympathetic arousal, agitation, perceptual changes, confusion, and seizures. These symptoms may emerge together in the context of delirium tremens, a potentially life-threatening complication of ethanol dependence that generally develops within the initial week of sobriety. Another syndrome, alcoholic hallucinosis, may develop during any phase of the cycle of intoxication and withdrawal. It is associated with persisting hallucinations that may or may not remit with extended sobriety.

Acute and protracted abstinence are important contexts for treatment to ensure the medical safety of recovering patients and to prevent their relapse to ethanol use. Although the most severe consequences of withdrawal appear during the initial week of sobriety, protracted components of withdrawal persist for many months (8). Protracted withdrawal symptoms include insomnia, anergia, and depressed mood. The effort to rid oneself of withdrawal symptoms may be an important motivator for relapse to ethanol use. As will be reviewed later, this phase of recovery is also associated with gains in cognitive function, cortical activity, and brain structure.

Nutritional deficits complicate the natural history of alcoholism (9). If thiamine-deficient individuals ingest glucose before thiamine repletion, the resulting demand on thiamine pyrophosphate-dependent metabolic processes compromises neuronal metabolic functions and may produce cell death associated with the Wernicke-Korsakoff syndrome (10). This eponym refers to a constellation of learning and memory impairment, psychosis, and motor findings.

This chapter provides a brief and selective overview of the basic and clinical neuroscience of alcoholism. Ethanol is now known to have multiple specific actions in the brain that contrast with the historical focus on its nonspecific perturbation of neuronal membrane bulk lipids (11 ,12). This chapter discusses the acute and chronic effects of ethanol at its protein targets in the brain, and the neural circuitry of human alcoholism that has become the focus of structural and functional neuroimaging studies.

MOLECULAR TARGETS OF THE ACTION OF ETHANOL: RELEVANCE TO INTOXICATION AND DEPENDENCE

Part of "100 - Ethanol Abuse, Dependence, and Withdrawal: Neurobiology and Clinical Implications "

Amino Acid Neurotransmission

Glutamate

Glutamate Receptors as an Ethanol Target in the Brain

Glutamate is the major excitatory neurotransmitter in the central nervous system (CNS) and glutamate utilizes both

ionotropic and metabotropic receptors to transduce its signal. The ionotropic receptors, receptor-gated ion channels, include the *N*-methyl-D-aspartate (NMDA) α-amino-3-hydroxy-5-methyl-4-isoxazoleproprionate (AMPA), and kainate receptors (13). NMDA receptors are gated by membrane potential and the simultaneous binding of both glycine and glutamate (14). Ethanol at concentrations as low as 5 mM (20 mg%) inhibits ion flux through the NMDA receptor-gated ion channels, making NMDA receptors one of the highest affinity ethanol targets in the brain (Fig. 100.1) (see ref. 15 for review). It shows lower affinity for AMPA and kainate glutamate receptors (15).

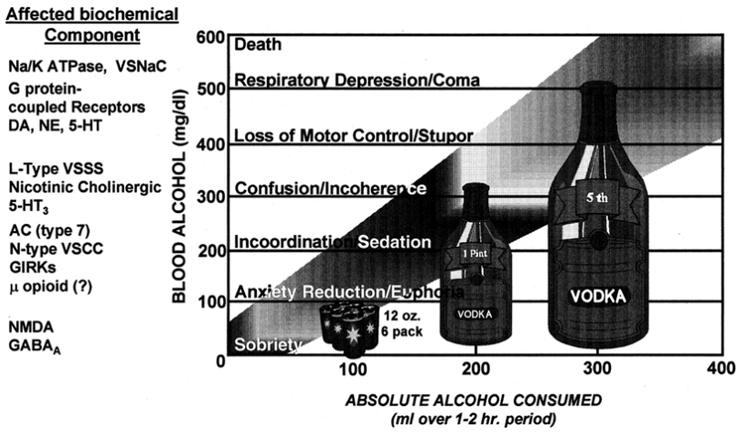


FIGURE 100.1. The relative potency of ethanol for its protein targets in the brain. Shown is the relationship between amounts of alcohol consumed and neurotransmission, neuroexcitatory components, and behavioral actions.

Ethanol effects on the NMDA receptor function are dependent on the concentration of glycine and the phosphorylation status of the receptor. In cultured cerebellar granule cells, ethanol lowered the NMDA receptor affinity for glycine, and ethanol effects were partially reversed by raising glycine levels (16). A protein kinase C (PKC)-mediated phosphorylation event may gate the effects of ethanol on glycine affinity at the NMDA receptor (16). Other protein kinases (e.g., tyrosine kinases such as Fyn and Src) may also play a role in ethanol's actions on the NMDA receptor (17).

Regional variations in NMDA receptor subunit composition contribute to distinctions in ethanol effects on NMDA receptor function across brain regions (see ref. 14 for review). Functional NMDA receptors are composed of an NR1 subunit, with at least eight splice variants, and one NR2 subunit from among the four known subtypes (NR2 A–D). The relative affinity of ethanol for NMDA receptor subtypes may be important to its dose-related effects in the brain (18), but this issue remains under intensive investigation (19).

The inhibition of NMDA receptor function by ethanol has direct neuroprotective consequences (20). NMDA antagonist effects of ethanol may influence its modulation of the release of other neurotransmitters (21 ,22), perhaps reflecting the capacity of low-dose NMDA antagonism to preferentially attenuate the activation of local inhibitory circuits (23).

The subjective effects of ethanol are tested in animals by measuring their ability to discriminate the effects of ethanol and other drugs. NMDA antagonists substitute for ethanol in these experiments, where they resemble the effects of higher doses of ethanol than ethanol doses that are most similar to the effects of γ-aminobutyric acid (GABA) agonists (24). Supporting the importance of the NMDA site, WSP (withdrawal seizure prone) and WSR (withdrawal seizure resistant) mice differ in their NMDA receptor density in several brain regions (25) and in their capacity to discriminate ethanol from other substances (26).

NMDA Receptor Adaptations with Ethanol Tolerance and Dependence

Multiple lines of evidence implicate NMDA receptor up-regulation as a mechanism contributing to acute ethanol withdrawal. Chronic ethanol administration up-regulates

NMDA receptor number, particularly in the cerebral cortex and hippocampus (27). During acute ethanol withdrawal, NMDA receptor increases are associated with tremors, anxiety, ataxia, and convulsions (27). Similarly, the increased hippocampal NMDA receptor density in WSP mice is related to their increased expression of withdrawal seizures relative to WSR mice (25). Additionally, NMDA antagonists given during withdrawal suppress withdrawal seizures (28). Lastly coadministration of the ganglioside GM_1 and ethanol prevents NMDA receptor up-regulation and the display of withdrawal seizures (29).

NMDA receptor up-regulation is subunit specific. In cultured cells and whole animals, chronic ethanol administration increases the levels of the NR2A and NR2B protein subunits and their subunit messenger RNA (mRNA) levels (30). Some, but not all, studies also suggest that NR1 subunit protein level increases may be accompanied by increases in NR1 mRNA levels (31). In cultured cells, the increases in NMDA receptor subunit proteins are associated with increased NMDA receptor function (32).

The consequences of NMDA receptor up-regulation during dependence are compounded by increases in glutamate release associated with the initiation of abstinence (33). Perhaps as a result, ethanol withdrawal is associated with seizures and neurotoxicity (see ref. 34 for review).

Glutamate: Clinical Correlates

Glutamate and the Complex Discriminative Stimulus Effects of Ethanol

As is seen in Fig. 100.2 , the NMDA antagonist, ketamine, produced dose-related ethanol-like effects in recently detoxified alcoholic patients (35). As with the preclinical studies (24), both the intensity and the degree of similarity of the ethanol-like effects of ketamine were greater at a higher subanesthetic dose (0.5 mg/kg) than at 0.1 mg/kg. Ketamine did not stimulate ethanol craving in patients, although craving was associated with the ethanol-like effects of another NMDA antagonist dextromethorphan (36).

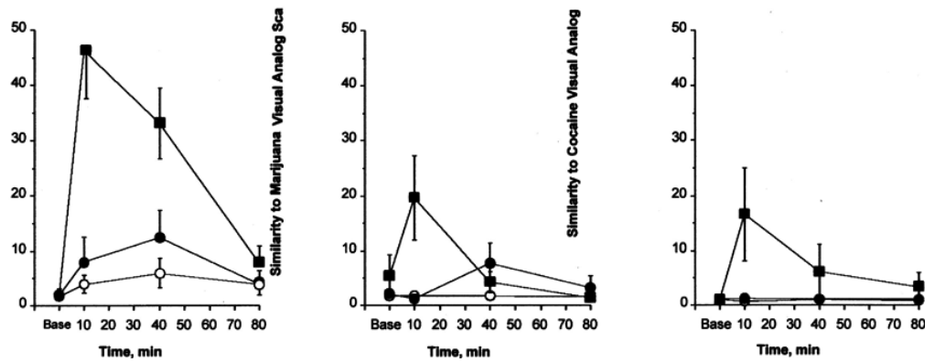


FIGURE 100.2. Similarity of the effects of placebo (*open circles*), ketamine 0.1 mg/kg (*solid circles*), and ketamine 0.5 mg/kg (*solid squares*) to ethanol (**left**), marijuana (**middle**), and cocaine (**right**) in alcoholic patients (n = 20). Values are expressed as mean ± standard error of the mean (SEM). Ketamine effects were significantly more similar to ethanol than both marijuana and cocaine by post-hoc contrast (F1 = 6.7, p = .02). (Data are from ref. 70.)

Clinical studies examining the interactive effects of ketamine and other drugs may provide insights into the NMDA antagonist component of ethanol effects in the brain. The NMDA antagonist-induced euphoria does not yet appear dopamine dependent. For example, the euphoric properties of ketamine are not blocked by haloperidol pretreatment (37) or markedly potentiated by amphetamine pretreatment (38). These findings parallel clinical findings describing the lack of interaction of ethanol and amphetamine (39). In contrast, the euphoric effects of ketamine (40), like ethanol (41), are attenuated by pretreatment with the μ opiate receptor antagonist naltrexone. Ethanol may possess actions at other brain targets that attenuate the dysphoric properties arising from its blockade of NMDA receptors including the facilitation of $GABA_A$ receptor function (42) or blockade of voltage-gated cation channels (43 ,44).

Glutamatergic Dysregulation in Ethanol Dependent Patients: Relationship to the Familial Risk to Develop Alcoholism

Postmortem studies of brain tissue from ethanol-dependent individuals suggest that the B_{max} or K_d of NMDA receptors are increased in cortical structures alcoholics (45 ,46). *In vivo*, ethanol withdrawal increases cerebrospinal fluid (CSF) glutamate levels. The combination of increased glutamate release during withdrawal and NMDA receptor up-regulation promotes withdrawal-related neural plasticity and neurotoxicity (47). With repeated episodes of withdrawal, patients show increased seizure risk (48) and hyperreflexia (49).

NMDA receptor function in recovering ethanol-dependent

patients also may shift the reward valence of the NMDA antagonist component of ethanol response from dysphoria to euphoria, promoting further alcohol use. Recently detoxified ethanol-dependent patients show reductions in their sensitivity to the perceptual, mood, and cognitive effects of ketamine (50) and the glycine-B partial agonist D-cycloserine (51). In contrast, preliminary data suggest that these patients exhibited preserved euphoric responses to ketamine relative to a healthy comparison group (Krystal, unpublished data). Thus, NMDA receptor alterations associated with ethanol dependence might contribute to relapse to ethanol use in two ways: (a) by contributing to the signs and symptoms of ethanol withdrawal, and (b) by enhancing the rewarding properties or reducing the dysphoric properties of ethanol during the early phases of relapse to ethanol use.

Healthy individuals at increased familial risk for developing alcoholism, relative to a “family history negative” group, show reductions in the dysphoric effects of NMDA receptor antagonists resembling the changes seen in ethanol dependent patients (52). Thus, inherited differences in NMDA receptor function may contribute to alterations in the set point for sensitivity to ethanol effects that promote the development of the abuse of ethanol. Further research is needed to clarify the impact of ethanol dependence and alcoholism vulnerability on glutamatergic function.

The mechanism through which ketamine sensitivity is altered in individuals at increased familial risk for alcoholism is not yet clear. In individuals without a family history of alcoholism, antagonism of voltage-gated cation channels clearly reduces the dysphoric effects of ketamine and may enhance its euphoric effects (43 ,44); i.e., these pretreatments may produce changes in the reward valence of ketamine effects that are similar to the alterations associated with a family history of ethanol dependence. The genes underlying altered ketamine response in individuals at risk for alcoholism are not yet known.

Acamprosate: A Putative Glutamatergic Pharmacotherapy for Alcoholism

Acamprosate is a homotaurine derivative without ethanol-like behavioral effects that reduces alcohol consumption in animals (53). This drug is a promising agent for treating alcoholism (54 ,55). It reduces NMDA receptor function, contributing to its capacity to suppress ethanol withdrawal (56). However, it also has promotes some NMDA receptor-mediated effects (57). The site of action of acamprosate is not known and the mechanisms related to its efficacy are not fully understood.

GABA

GABA Receptors as Targets for Ethanol Action

Ethanol produces sedative-hypnotic effects that resemble other drugs that facilitate $GABA_A$ receptor function, particularly muscimol, benzodiazepines, and barbiturates (15). The similarity between the behavioral effects of GABA agonists and ethanol is dose-dependent, with the greatest similarity between these drug classes observed with relatively low training doses of ethanol (15).

Ethanol has effects at the $GABA_A$ receptor at concentrations of 10 to 100 mM. In one study, ethanol acted similarly to benzodiazepines by potentiating the effects of GABA (58), whereas in another study ethanol resembled barbiturates by increasing the entry of chloride without the addition of GABA (59). The microsac preparation employed in these studies, however, would be expected to contain certain sufficient amounts of endogenous GABA to influence their interpretation.

Ethanol actions at $GABA_A$ receptors vary across brain regions and may be related to the differential expression of $GABA_A$ receptor subunits (60). This receptor is composed of five subunits that associate to form a Cl^- channel. Subunit families (e.g., α, ß, γ, δ) and isoforms within each family (e.g., α_1–α_6) have been identified (61). Variation in subunit composition imparts functional distinctions in $GABA_A$ receptor subtypes relevant to ethanol action (61). For example, isoforms of the γ subunit [i.e., α_{2S} or α_{2L} (62)] that differ in their sensitivity to PKC isoforms differ in their sensitivity to ethanol (63). However, studies now question the importance of these particular $GABA_A$ receptor subunits (64). Some of these differences between studies may reflect the importance of a particular PKC isoform, PKC-ε.

The adenylyl cyclase signaling system and protein kinase A (PKA) also modulate ethanol’s actions at $GABA_A$ receptors. Ethanol potentiates GABA inhibition of cerebellar Purkinje cell firing, but only when there is concomitant stimulation of ß-adrenergic receptors (65) and activation of PKA (66).

Chronic Actions of Ethanol on $GABA_A$ Receptor Function

Changes in the levels of $GABA_A$ receptor subunits occur in animals treated chronically with ethanol (67). It consistently decreases the mRNA and protein for the α_1 subunit of the $GABA_A$ receptor, whereas other subunits may show no change or even increases (67). *In vitro* studies form ethanol-treated animals demonstrated a reduced ability of ethanol to potentiate GABA-mediated chloride flux, suggestive of the development of ethanol tolerance (58). *In vivo*, hyperexcitability, fear behaviors, and convulsions during acute ethanol withdrawal may reflect decreases in GABAergic neurotransmission (68). WSP and WSR mice, noted earlier to differ in NMDA receptor regulation, also differ in predicted directions with respect to their $GABA_A$ receptor characteristics and their vulnerability to withdrawal seizures (69).

GABA: Clinical Correlates

GABA Systems and Alcoholism Vulnerability

To date, the genes underlying the GABA-related vulnerability for

alcoholism are unknown. A haplotype relative-risk study did not find evidence associating either the α_1- or α_3-subunit genes with alcoholism, although suggestive data supporting further study was obtained for the latter gene in an association study (70). The strongest tie between $GABA_A$ receptor involvement and the vulnerability to alcoholism has come from studies evaluating ethanol and benzodiazepine effects in populations at high-risk for developing alcoholism. Most (71 ,72), but not all (73), studies found that male offspring of alcoholics have reduced sensitivity to the cognitive, behavioral, and motor effects of both benzodiazepines and ethanol. Similarly, individuals at risk for alcoholism exhibited blunted cerebellar metabolic inhibition, but not cortical metabolic inhibition, following a dose of lorazepam (74). However, there may be greater sensitivity to or preference for the euphoric effects of benzodiazepines in this population (71), although these effects are not uniformly replicated (75). The rewarding effects of alcohol and benzodiazepines in humans may be increased in anxious individuals, who tend to experience more pronounced anxiolytic effects of these drugs (76).

Clinical Evidence of GABA Dysregulation in Alcoholism

GABAergic regulation differs in alcoholic and nonalcoholic populations. GABA levels are reduced in plasma, CSF, and brain in recently detoxified alcoholics during the first month of detoxification (77 ,78). In patients, low plasma GABA levels predicted relapse (79), perhaps because low GABA levels were associated with protracted abstinence symptoms. These changes may in part reflect ethanol dependence and withdrawal-related effects on the levels or function of enzymes regulating GABA synthesis and degradation (80).

Reductions in $GABA_A$ receptor binding in postmortem and antemortem neuroimaging studies may reflect the combined effects of vulnerability, ethanol dependence, and alcoholism-related neurotoxicity. Some postmortem studies found reductions in $GABA_A$/benzodiazepine (BZ) binding, but other studies had conflicting results (81 ,82 and 83). Variability between studies may reflect a compensatory up-regulation in BZ/$GABA_A$ binding that follows alcoholism-related neurotoxicity (84) or the failure to employ ligands that differentiate between subtypes of $GABA_A$ receptors. *In vivo*, positron emission tomography (PET) and single photon emission computed tomography (SPECT) neuroreceptor imaging has provided evidence of reductions in BZ binding (Fig. 100.3) (85 ,86). Neurotoxicity contributed to, but did not account for, reductions in BZ receptor binding in patients (87). Reductions in BZ binding may be consistent with evidence of reduced regional brain metabolic sensitivity to lorazepam in ethanol-dependent individuals (88). However, metabolic blunting did not rapidly recover with sobriety (89), raising the possibility that this deficit reflected genetic vulnerability or irreversible toxicity.

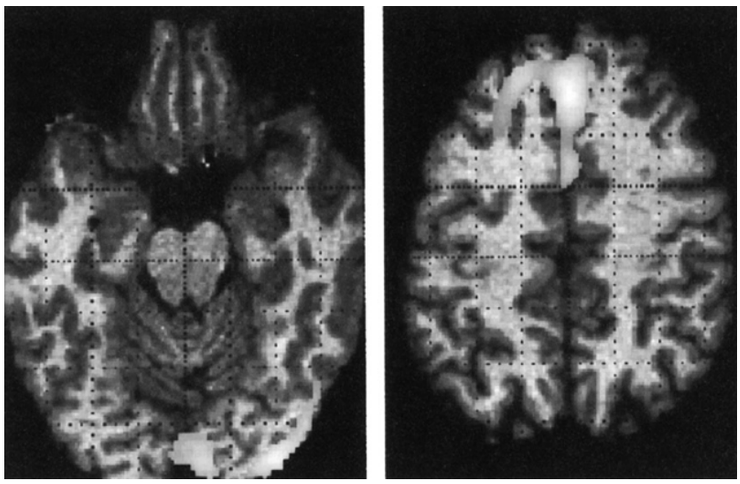


FIGURE 100.3. Transaxial slice of regions where benzodiazepine receptor distribution volume in alcoholic patients was significantly lower than in comparison subjects based on a statistical parametric mapping analysis (*yellow*). Single photon emission computed tomography (SPECT) data were superimposed on an MRI template. (From Abi-Dargham A, Krystal JH, Anjivel S, et al. Alterations of benzodiazepine receptors in type II alcoholics measured with SPECT and [123I]iomazenil. *Am J Psychiatry* 1998;155:1550–1555, with permission.) See color version of figure.

Psychopharmacologic studies also implicate GABA systems in withdrawal and relapse. Clearly, drugs facilitating $GABA_A$ receptor function including BZs, barbiturates, and anticonvulsants suppress acute ethanol withdrawal (90).

The BZ antagonist flumazenil may reduce ethanol intoxication in humans at doses greater than required to antagonize BZ effects (91). However, this drug does not produce ethanol withdrawal symptoms in ethanol-dependent individuals (92). It is not yet clear whether flumazenil exposure during ethanol withdrawal alters the course of recovery from ethanol dependence.

Neurosteroids

There is growing interest in the role of neuroactive intermediates in sex steroid synthesis and metabolism in alcoholism (93). Allopregnanolone, a coagonist of the steroid anesthetic site of the $GABA_A$ receptor, shares discriminative stimulus properties with ethanol and suppresses the ethanol abstinence syndrome in animals (94 ,95). Allopregnanolone also may potentiate the effects of ethanol (96). Allopregnanolone levels, like those of its precursor progesterone, vary markedly during the menstrual cycle. In the late luteal phase, drops in allopregnanolone levels may contribute to premenstrual mood disturbances (97) and increase the intensity of the discriminative stimulus properties of ethanol (98). These factors may increase ethanol consumption during this phase of the menstrual cycle (99). Supporting this view, higher allopregnanolone levels in the late luteal phase are associated with reduced triazolam self-administration in women (100). Thus, cyclical variation in neurosteroid levels may be a focus for future pharmacotherapies for female alcoholics.

Voltage-Sensitive Calcium Channels (VSCCs)

Preclinical Studies

In the brain, VSCCs play a major role in gating synaptic calcium influx and thereby modulating a range of calcium-dependent intracellular processes, membrane potential, and neurotransmitter release (101 ,102). There are six known VSCC classes: L-type (dihydropyridine-sensitive), N-type (“neuronal”), P-type (“Purkinje”), Q-type, R-type, and T-type (“transient”) channels (103 ,104).

Ethanol blocks L-type channels. Supporting this hypothesis, L-type VSCC antagonists show some ethanol-like effects in rats (105). Ethanol, at concentrations over 50 mM, inhibited [$^{45}Ca^{2+}$] influx in *in vitro* preparations (106). These studies suggested a wide range of L-type channel sensitivity to inhibition by ethanol, ranging from 10 to >200 mM, perhaps reflecting differences in channel subunit composition. The variability in L-type channel sensitivity to ethanol may depend on the characteristics of its subunits, post-translational modifications, and other regulatory mechanisms (107 ,108).

Chronic exposure to ethanol *in vivo* or cultured cells up-regulates L-type channels via a PKC-dependent mechanism (108). The up-regulation of L-type channels may contribute to signs of ethanol withdrawal (109 ,110), perhaps in part, by partially depolarizing cell membranes and recruiting NMDA receptors.

Electrophysiologic studies with transformed cells or pituitary neuron terminals also indicated that both N and T channels could be inhibited by an ethanol concentration of 50 mM (111 ,112). Recent evidence (113) also indicates that chronic administration of ethanol to mice up-regulates the number and function of N-type calcium channels.

VSCCs: Clinical Correlates

Ethanol actions at VSCCs may modulate its behavioral effects in humans. Despite the apparent similarities between the effects of ketamine and ethanol, ketamine produces more profound perceptual effects than ethanol at doses studied to date (114). However, the perceptual effects of ketamine are attenuated in humans by both the L-type VSCC antagonist nimodipine (44) and by lamotrigine, a drug with multiple effects on cation channels, including an antagonist action at P- and N-VSCCs (43). These studies suggest that the combination of NMDA and VSCC antagonist properties of ethanol enhance its tolerability.

The relevance of adaptations in VSCCs for the clinical phenomenology of ethanol withdrawal is not yet clear. L-channel antagonists may reduce the severity of some withdrawal symptoms, but they do not clearly suppress withdrawal seizures (115). However, the existing studies are limited by shortcomings in their study design and the selection of agents with limited CNS penetration.

Serotonin (5-HT)

5-HT systems contribute to the discriminative properties of ethanol in animals and humans. Ethanol facilitates that activity of $5\text{-}HT_{1B}$, $5\text{-}HT_{2C}$, and $5\text{-}HT_3$ receptors, and it shares discriminative stimulus properties with drugs acting at these sites (15 ,116).

The administration of the 5-HT partial agonist, *m*-chlorophenylpiperazine (mCPP), produces a euphoric effect that is perceived as ethanol-like in early-onset alcoholic patients (117 ,118 and 119). However, mCPP effects were not specifically similar to ethanol, i.e., the effects were similar to several substances of abuse (117). The two mCPP studies that administered mCPP intravenously also reported the induction of craving (117 ,119), whereas the study administering this drug orally found the opposite (120). mCPP also produced anxiety and irritability (117). The induction of dysphoria by this drug may have contributed to the elicitation of craving. Several studies report that the cortisol or prolactin response to mCPP is reduced in early onset patients (118 ,119 ,121). Further, the cerebral metabolic response to mCPP was reduced in early-onset alcoholics (122). In contrast, the euphoric responses to mCPP were enhanced in early-onset

patients relative to patients with a later onset of alcoholism (119).

The site of action of the ethanol-like effects of mCPP is not clear. Its partial 5-HT_{2C} agonist action appears to figure most prominently in its general discriminative stimulus effects (123). mCPP also has activity at 5-HT_3 (124) and 5-HT_7 receptors (125). Preliminary data suggested that ritanserin, a drug that blocks 5-HT (5-HT_{2A}, 5-HT_{2C}, 5-HT_6, 5-HT_7, 5-HT_{1D}) and dopamine (D2) receptors, reduced mCPP effects in healthy humans (125 ,126 ,127 ,128 and 129). The lack of specificity regarding the site of action of both mCPP and ritanserin limits the interpretation of the mechanisms underlying the ethanol-like and craving effects of mCPP. Also, the failure of ritanserin as an alcoholism pharmacotherapy (130) further raises concerns about the therapeutic applicability of the mCPP studies.

Molecular genetic studies further increased interest in genetic variation associated with the function of the 5-HT transporter and the regulation of central 5-HT turnover. Although the findings have not been replicated (131 ,132), two groups have associated 5-HT transporter alleles with reduced central 5-HT function or poor impulse control in alcoholic individuals (133 ,134). However, one study failed to find that alleles of the 5-HT transporter were associated with alcoholism (131). One hypothesis guiding these studies was that reduction in the efficacy or availability of synaptic 5-HT resulting from enhanced density or function of the 5-HT transporter would contribute to the constellation of behaviors associated with early-onset alcoholism (135 ,136). Alternatively, reduced density of 5-HT transporter binding in the brain might reflect reductions in the density of 5-HT terminals that might contribute to reduced central 5-HT function (137 ,138). Studies suggest that 5-HT transporter antagonists have limited efficacy for alcoholism and may make some early-onset patients worse (139 ,140). However, 5-HT–related vulnerability may be reflected in comorbid conditions. Ethanol-dependent patients with depression may benefit from treatment with 5-HT transporter antagonists (141), and patients with comorbid anxiety may benefit from the addition of the 5-HT_{1A} agonist buspirone (142).

Preclinical research suggests that ethanol stimulates 5-HT_3 receptors and that 5-HT_3 antagonists may attenuate the discriminative stimulus properties of ethanol and ethanol self-administration in animals (143). In humans, the 5-HT_3 antagonist odansetron did not robustly attenuate the discriminative stimulus effects of ethanol (144). However, a clinical trial employing odansetron found evidence of efficacy in early-onset ethanol-dependent patients (145).

Catecholamines

Dopamine

Dopamine-mediated neurotransmission has received much attention due to the involvement of dopamine in rewarding properties of ethanol and other drugs (146). Doses of ethanol that produce motor stimulation or ataxia increase the firing rate of midbrain dopamine neurons of unanesthetized rats (147). *In vitro*, ethanol added to brain slices in concentrations of 20 to 320 mM also stimulated the activity of ventral tegmental dopamine neurons (148).

Ethanol increases dopamine release in brain regions involved in the reinforcing effect of ethanol, such as the ventral tegmental area and nucleus accumbens (21). Further, rats bred to drink ethanol, compared to ethanol nonpreferring animals, show increased dopamine release associated with ethanol consumption (149). In addition, dopaminergic drugs alter ethanol self-administration in animals (150). Ethanol also has effects on dopamine release that may be mediated by opioid and nicotinic cholinergic systems (151 ,152). During ethanol withdrawal, there are reductions in dopamine release in the ventral striatum and in the nucleus accumbens (153). These decreases may contribute to withdrawal-related dysphoria. Ethanol, NMDA receptor antagonists, and L-type VSCC antagonists attenuated these dopamine deficits (153 ,154 and 155).

Norepinephrine

The locus coeruleus (LC) contains the cell bodies for the brain dorsal noradrenergic system (156). LC basal activity and activation are reduced by ethanol, an action that may contribute to sedative effects of ethanol (157 ,158). Ethanol also produced biphasic effects on norepinephrine turnover in the brain, with low doses increasing turnover and higher doses depressing turnover. The sensitivity of noradrenergic systems to ethanol effects varies among brain regions (159).

Catecholamines: Clinical Correlates

Modulation of catecholamine function modulates the stimulant and intoxicating effects of ethanol. Catecholamine synthesis inhibition, produced by α-methyl-para-tyrosine, modestly reduced ethanol intoxication in healthy humans (160). Similarly, dexamphetamine and methamphetamine pretreatment either had no effect or modestly potentiated ethanol intoxication in humans (39). In contrast, the α_2-adrenergic antagonist yohimbine potentiated the intoxicating effects of ethanol, but did not substantially alter the subjective sense of euphoria associated with intoxication (161). These data may conflict with other studies suggesting that ß-adrenergic stimulation attenuates ethanol intoxication, whereas ß-adrenergic blockade enhances intoxication (162 ,163). In recently detoxified early-onset alcohol-dependent patients, yohimbine effects have a low degree of similarity to ethanol effects and it reduced ethanol craving levels below baseline (117).

Several studies document reduced sensitivity of both dopamine and noradrenergic receptors in recently detoxified patients. Reduced sensitivity of dopamine receptors is suggested

by blunted growth hormone responses to dopamine agonists (164). These data are consistent with neuroimaging data, suggesting that the density of dopamine transporter binding is preserved but striatal D2 receptor density is decreased in ethanol-dependent patients (137 ,165). Down-regulation of postsynaptic α_2-adrenergic receptors is suggested by blunted growth hormone responses to clonidine and increased cortisol responses to yohimbine (121 ,166). In contrast, presynaptic noradrenergic activity appears to normalize rapidly following withdrawal, as measured by the CSF levels of the norepinephrine metabolite 3-methoxy-4-hydroxyphenethyleneglycol (MHPG) (167). Similarly, the presynaptic component of the noradrenergic response to yohimbine, reflected by plasma MHPG levels, is normal in patients sober for approximately 1 month (121).

Genetic studies relating catecholamine receptor alleles to the vulnerability to alcoholism have been a promising but controversial research strategy that has not yet born fruit. Initial studies suggested that D2 receptor alleles were associated with alcoholism (168 ,169). However, more definitive subsequent studies were negative (170). Reanalysis of published studies suggested that the positive findings reflected ethnic differences between the patient and control population (171). A subsequent study also suggested that D2 receptor alleles predicted an anticraving response to bromocriptine (172). However, studies using dopamine agonists to treat alcoholism have so far had limited promise (173 ,174). A tentative association between versions of the D4 receptor and novelty-seeking (175) was reported. However, this finding was not replicated, and studies of other dopamine receptor genes to alcoholism have been negative (133 ,176 ,177 ,178 ,179 ,180 ,181 and 182).

Opiates

Preclinical Studies

Ethanol modulates opioid neuropeptides in regionally specific ways. Acute ethanol administration to animals or ethanol perfusion of cultured pituitary or hypothalamic tissue increased (183) or had no effect on (184) tissue content or release of ß-endorphin. Ethanol also raised enkephalin and dynorphin levels in some brain regions in some studies (183 ,185). Chronic ethanol administration to rodents decreased pituitary ß-endorphin processing and reduced hypothalamic mRNA levels of the peptide precursors proopiomelanocortin (POMC) and prodynorphin (186 ,187).

In vitro, ethanol has a biphasic effect on µ opioid receptor binding. Ethanol concentrations, in the range of 10 to 25 mM, produce a small, but significant, increases in the binding of µ receptor ligands, whereas higher concentrations of ethanol inhibit ligand binding to the µ receptor (188). Under physiologic conditions, the µ receptor may be more sensitive to ethanol-induced inhibition than the δ receptor (189). Chronic ethanol administration reduces the density and function of striatal and accumbens µ opioid receptors (190). In contrast, chronic ethanol administration also modulates the binding and function of δ opioid receptors (191 ,192).

Opiates: Clinical Correlates

Naltrexone appears to reduce the rewarding effects of ethanol and ethanol consumption in social drinkers (41 ,193). Also, naltrexone maintenance appears to reduce the pleasurable aspects of ethanol consumed during treatment for alcoholism (194 ,195). This property appears to contribute to the capacity of opiate antagonists to reduce ethanol consumption in alcoholic patients (196 ,197).

The contributions of endogenous opiate systems to the rewarding effects of ethanol are further supported by evidence of abnormalities in these systems in alcoholic patients. However, the current data do not yield a clear picture of these abnormalities. Postmortem studies have described both increased µ-receptor density (198) and decreased µ-receptor affinity (199). CSF and plasma studies have also suggested the existence of reductions in ß-endorphin levels and ethanol-stimulated increases in plasma ß-endorphin levels (200 ,201). Naltrexone reductions in ethanol effects appear to be particularly evident in individuals at high risk for developing alcoholism (202). These data suggest a genetic component underlying the efficacy of naltrexone treatment for alcoholism. To date, there have been negative studies evaluating the association of the proenkephalin A gene and µ opiate receptor gene with alcoholism (203 ,204).

THE NEURAL CIRCUITRY OF ALCOHOL ABUSE AND DEPENDENCE: INSIGHTS FROM NEUROIMAGING AND NEUROPSYCHOLOGY

Part of "100 - Ethanol Abuse, Dependence, and Withdrawal: Neurobiology and Clinical Implications "

Structure (Computed Tomography, Magnetic Resonance Imaging, Postmortem)

Preclinical studies describe alcoholism-related neurotoxicity. These toxic effects appear to reflect a combination of the neurotoxic effects of ethanol, the interaction of ethanol with nutritional deficiencies, and the neurotoxic consequences of ethanol withdrawal (205). In rats, extended consumption of an ethanol diet did not produce anatomic deficits. However, ethanol withdrawal was associated with reductions in dendritic arborization and neuronal loss (206).

Brain shrinkage and neuronal loss has been documented in the brain tissue from individuals with alcoholism in both cortical and limbic regions (207). With respect to brain volume, postmortem research suggests that white matter loss appears to be more prominent than gray matter loss (208).

Neuronal loss appears to be primarily loss of pyramidal neurons, with relative sparing of interneurons (208). Another study was unable to replicate neuronal loss using uniform sampling and unbiased neuron counting methods (209). However, even when neurons are extant, they may exhibit structural deficits consistent with neurotoxicity (210). Generally, cortical and limbic brain shrinkage and neuronal loss may be more prominent in individuals whose alcohol dependence is complicated by Wernicke's encephalopathy and Korsakoff's psychosis (211). The Wernicke-Korsakoff syndrome has been associated with abnormalities in frontal cortex, as well as in several subcortical structures including the thalamus, hippocampus, mamillary bodies, and amygdala (10).

Structural neuroimaging studies are consistent with the findings in postmortem research. Cortical and limbic volumetric losses in ethanol-dependent patients have been described using computed axial tomography (CAT) and magnetic resonance imaging (MRI) (212 ,213). Gray and white matter volumetric losses are progressive with heavy drinking and are most prominent in frontal and temporal cortex (213 ,214 and 215). Ethanol-dependent individuals also show sulcal and ventricular enlargement (215), as well as hippocampal volume reduction (216). Reductions in the volume of the corpus callosum has also been described (217). Brain atrophy documented in structural neuroimaging studies is most more prominent with advancing age in adults (Fig. 100.4) (218). This age-related effect may reflect an age-related vulnerability

to ethanol in older populations, the interaction of aging processes, and the neurotoxicity of alcoholism, i.e., the “premature aging of the brain” (219), as well as the accumulated impact of long-standing alcoholism on older populations. Thus, atrophy may not be detectable in young healthy ethanol-dependent populations (220). In contrast, ethanol-dependent adolescents show hippocampal volumetric changes not seen in the studies of healthy young adults (221). The study in adolescents raises the possibility that adolescents show disruptive effects on brain development or an increased sensitivity to the neurotoxic effects of ethanol.

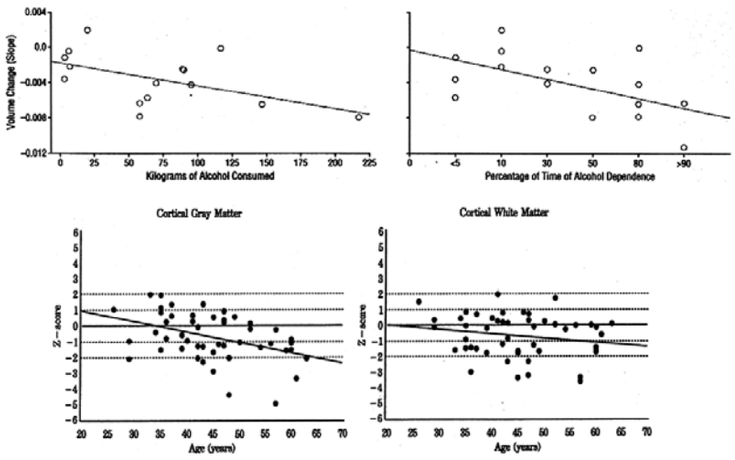


FIGURE 100.4. The **top two panels** illustrate the relationship between drinking severity and duration with gray and white matter volume loss in ethanol-dependent patients. The **bottom two panels** illustrate the interaction between alcoholism and volume loss with age. **Top figures**: The relationship between cortical gray matter rate of change and the amount of alcohol consumed during the follow-up period (*left*) (Spearman rho = -0.52, p = .04) and between the cortical gray matter rate of change and the estimated amount of time during past 5 years that alcoholic patients (group 1) met *Diagnostic and Statistical Manual of Mental Disorders*, third edition revised (DSM-III-R) criteria for alcohol dependence during the follow-up period (*right*) (Spearman rho = -0.53, p =.04). One alcoholic patient who reported 950 kg of alcohol consumption is omitted from the **left panel** so as not to distort scaling. The *darker circle* represents two patients with overlapping values. (From Pfefferbaum A, Sullivan EV, Rosenbloom MJ, et al. A controlled study of cortical gray matter and ventricular changes in alcoholic men over a 5-year interval. *Arch Gen Psychiatry* 1998;55:905–912.) **Bottom figures**: Cortical gray (*left*) and white matter (*right*) volume changes with age in ethanol-dependent patients. Data from individual ethanol-dependent patients are expressed as age-corrected Z-scores plotted as a function of age. (From Pfefferbaum A, Lim KO, Zipursky RB, et al. Brain gray and white matter volume loss accelerates with aging in chronic alcoholics: a quantitative MRI study. *Alcohol Clin Exp Res* 1992;16:1078–1089, with permission.)

Over the initial years of sobriety, there is recovery in the volumes of gray and white matter and reductions in sulcal and ventricular volume (222). There are differences in the rate of particular brain regions and particular tissue types with regard to the rate of recovery (222 ,223). The relatively rapid recovery of white matter volume with sobriety does not appear to reflect the return of tissue water displaced by ethanol, i.e., tissue rehydration (224).

Nutritional status, neurologic complications of ethanol withdrawal, and hepatic function appear to be related to findings in structural neuroimaging studies. Ethanol withdrawal seizures have been linked to neurotoxicity in these patients (225). Supporting this association, temporal cortex white matter loss was particularly associated with ethanol withdrawal seizures (226). Although cortical atrophy has been described in ethanol-dependent patients with good nutritional status (227), the Wernicke-Korsakoff syndrome and hepatic cirrhosis are generally associated with more prominent MRI volumetric deficits in cortical and limbic structures than ethanol-dependent patients who are otherwise healthy (228).

To date, there has been very little study of structural factors that might predispose individuals to develop alcoholism. One risk factor, antisocial personality disorder, appears to be independently associated with reductions in frontal cortex gray matter volume (229). Thus, the observation that frontal gray matter volume loss is present in young ethanol-dependent patients could reflect a combination of the vulnerability to alcoholism and atrophic effects of ethanol dependence (214).

Magnetic Resonance Spectroscopy

Magnetic resonance spectroscopy (MRS) has been applied to the evaluation of structural deficits in alcoholic patients in a limited fashion. Proton-MRS ([^{1}H]MRS) enables the measurement of N-acetyl-aspartate (NAA), a constituent of viable neurons. One study found reductions in the NAA/creatine ratio in the frontal cortex and cerebellum of ethanol-dependent patients (230 ,231). A phosphorus-MRS ([^{31}P]MRS) study found reductions in phosphodiester and phosphocreatine levels in cortical white matter in ethanol-dependent patients (232).

Findings From Functional Neuroimaging

Functional neuroimaging studies described reductions in cortical blood flow and metabolism associated with ethanol withdrawal. There is a limited understanding of the extent to which these studies also reflect genetic or alcohol-related structural abnormalities. Studies of ethanol intoxication suggest that it reduces cortical metabolism in humans (233). Clinical studies of withdrawal, mostly conducted during or following medications for detoxification, predominately describe reductions in regional cerebral perfusion or glucose metabolism in frontal and temporal cortex (234 ,235). More pronounced deficits were observed in patients with evidence of cortical atrophy based on structural neuroimaging (236), years of ethanol use, age (237), and multiple ethanol detoxifications (238). Cerebral perfusion and metabolic deficits may attenuate over the initial months of sobriety (237), and improvement may continue over several years (239). Antisocial personality (240), a risk factor for developing alcoholism, but not family history of alcoholism (241), was associated with more pronounced volumetric deficits.

Across several studies, reductions in resting frontal cortical perfusion and metabolic rate is associated with reduced performance on cognitive tests that engage the frontal cortex (242). Similarly, cerebellar metabolic deficits are associated with behavioral evidence of cerebellar dysfunction (236).

Behavioral Studies

The most profound cognitive deficits associated with alcoholism are the memory impairments arising from nutritional deficiency, as in Korsakoff syndrome (10), or hepatoxicity, as in hepatic encephalopathy (243). Although the most severe consequences of alcoholism for cognition may not reflect the direct toxic effects of ethanol on the brain, many patients exhibit cognitive deficits independent of these factors that reflect the combined impact of age, familial vulnerability for alcoholism, adaptations to ethanol effects on the brain, perhaps degree of liver injury (244), presence of comorbid depression, and ethanol-related neurotoxicity (245). Cognitive function is further compromised in those patients who continue to drink, due to the direct effects of ethanol on cognition (246).

The familial vulnerability to alcoholism and traits associated with that vulnerability are associated cognitive deficits and educational achievement (247). Although reductions in attention, planning, visual-spatial learning, and impulse control have been described in children of alcoholics (248), these findings are not universal (249). These findings may be largely accounted for by comorbid traits such as antisocial personality, similar to both MRI and event-related potential (ERP) findings in alcohol dependent patients (250 ,251). Consistent with this view, the familial risk for cognitive deficits associated with alcoholism is particularly associated with poor prognosis of the parent in alcoholism treatment,

a characteristic of antisocial alcoholism (252). Familial history of alcoholism appears to compound the negative consequences of social drinking on cognitive function (253). However, children of alcoholics exhibit attenuated impairment in attention and memory relative to individuals with a familial alcoholism history (254). Thus it is possible that cognitive responses to ethanol may also contribute to the risk for making the transition from social drinking to alcoholism.

Although traits associated with the vulnerability to alcoholism contribute to cognitive deficits seen in ethanoldependent patients, they do not account for the deficits in patients (255). Ethanol-dependent patients show many impairments in cognitive function. Deficits in the level of performance and efficiency of verbal skills, learning and memory, problem-solving and abstracting, and perceptual-motor skills have been described (256). Cognitive deficits may reflect in part the degree of neurotoxicity related to alcoholism. For example, young ethanol-dependent patients may show normal brain volumes on MRI and normal cognitive function (257 ,258). With repeated episodes of ethanol withdrawal and advancing age, cognitive deficits become more pronounced (259).

Cognitive function improves over the initial year of sobriety; however, the domains of cognitive function do not recover at the same rate and recovery may be partial (260 ,261). Poor cognitive function at the time that treatment is initiated appears to predict improved alcohol-related treatment outcomes (262). However, progress in treatment may be reflected in improved cognitive function (263). Consistent with the view that treatment may modulate cognitive recovery, some cognitive deficits in recovering patients appear to respond to cognitive rehabilitation (264).

In summary, the behavioral studies describe the display of deficits in cognitive functions that may have implications for circuitry dysfunction in alcoholism: executive function deficits associated with the prefrontal cortex, visual-spatial deficits associated with the parietal cortex, and learning/memory deficits that may involve the hippocampus and related temporal cortical structures. Overall, these studies are consistent with the findings related to reduced tissue volume on MRI, reductions in cortical metabolism with fluorodeoxyglucose (FDG)-PET (242), and information processing deficits in ERP studies (265). This overlap implies a connection between alterations in brain structure, function, and behavior related to alcoholism.

THE INTERPLAY OF THE NEURAL CIRCUITRY AND NEUROCHEMISTRY OF ALCOHOLISM: IMPLICATIONS FOR TREATMENT

Part of "100 - Ethanol Abuse, Dependence, and Withdrawal: Neurobiology and Clinical Implications "

Ethanol has multiple specific effects on amino acid, monoamine, and neuropeptide neurotransmitter systems, and these effects contribute to its complex array of behavioral effects in animals and humans. Direct effects of ethanol on excitatory and inhibitory neurotransmission also are modulated by direct and indirect effects of ethanol on ion channels. Further, the cellular consequences of exposure to ethanol are modulated by ethanol-sensitive regulatory enzymes, such as PKA and PKC.

The diversity of ethanol targets in the brain and frequent co-localization of the neural circuitry bearing the sites of ethanol action raise the possibility of convergent mechanisms underlying the rewarding effects of ethanol in the brain. In this regard, there is growing evidence that the interplay of the prefrontal cortex (PFC) and limbic structures including the nucleus accumbens (NAc) and amygdala generally plays an important role in reward (266 ,267). Ethanol has many component actions that directly inhibit the output of the NAc including NMDA receptor antagonism, GABA facilitation, and enhancement in 5-HT and dopamine release (267). Drugs that antagonize the inhibitory effects of ethanol in the NAc, such as naltrexone (268), may play an important role in the treatment of alcoholism even if μ opiate receptors are not a major site of action for many of the behavioral effects of ethanol.

Abnormalities in PFC development may be an important factor influencing the vulnerability to alcoholism, in part by altering the interplay of PFC and NAc that underlies reward. As noted above, the vulnerability to alcoholism, particularly in the case of individuals with antisocial personality disorder or impulsive traits, appears to be associated with behavioral, physiologic, and structural evidence of PFC dysfunction. An important and unanswered question is, How does PFC dysfunction contribute to the vulnerability to alcoholism? Hypotheses have been presented that suggest that alcoholism is just one of several forms of impulsive behavior that these individuals fail to inhibit due to a general deficiency in behavioral inhibition or as a consequence of the failure to anticipate the negative consequences of alcoholism (269 ,270).

These cognitive hypotheses may be complemented by a reward dysfunction hypothesis resting on a consideration of the impact of PFC dysfunction on mechanisms underlying reward. The PFC input into limbic structures responsible for reward is critical to the experience, anticipation, and seeking of reward (271 ,272). The activation of PFC outputs to limbic structure causes a release of glutamate that may serve to activate, for example, the GABAergic output neurons of the NAc, and this action opposes direct effects of ethanol (268). From this perspective, PFC activation serves as a “brake” on reward mechanisms. In fact, it is tempting to think of this PFC-NAc interplay as a pathway contributing to the capacity of human judgment to restrain impulsive reward-related behavior. Yet abnormal PFC input would also be expected to disturb NAc function with respect to both the processing of rewarding stimuli generally and drugs of abuse in particular. Thus, it may not be surprising that

the familial vulnerability to alcoholism is associated with both PFC functional deficits and alterations in the capacity of drugs representing multiple component actions of ethanol (GABA, NMDA, opiate) to generate rewarding or aversive experiences in humans. If so, then genes controlling corticolimbic neurodevelopment and genes encoding the individual targets or intracellular mediators of ethanol action may provide a diversity of potential foci for the study of the genetics of alcoholism.

The cellular adaptations to chronic ethanol underlie tolerance to ethanol effects and withdrawal symptoms that develop upon the initiation of abstinence. As reviewed in this chapter, acute withdrawal symptoms reflect an enhancement of glutamatergic function and a deficit in GABAergic function. When examined beyond these generalizations, the neurobiology of ethanol dependence appears to be complex. For example, dependence-related adaptations may be reflected in absolute numbers of receptors in binding studies or perhaps changes in receptor subunit composition. Also, protracted withdrawal may produce functional changes in the opposite direction of acute withdrawal. Withdrawal is a critical juncture in the treatment of alcoholism because withdrawal symptoms may present an immediate medical risk, motivate relapse to ethanol use, induce withdrawal-related neuroplasticity that can increase risk for subsequent withdrawal-related medical complications, and may promote neurotoxicity. In light of these issues, novel treatments for withdrawal aim to achieve multiple therapeutic objectives (273).

ACKNOWLEDGMENTS

Part of "100 - Ethanol Abuse, Dependence, and Withdrawal: Neurobiology and Clinical Implications "

This work was supported by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) (grants KO2 AA 00261-01 and RO1 AA11321-01A1) and the Department of Veterans Affairs (Alcohol Research Center, Clinical Neurosciences Division, National Center for Posttraumatic Stress Disorder). The authors thanks Ms. Shawndra Poharski for her assistance in preparing this manuscript.

REFERENCES

1. Martin CS, Earleywine M. Ascending and descending rates of change in blood alcohol concentrations and subjective intoxication ratings. *J Subst Abuse* 1990;2:345–352.

2. Peterson JB, Rothfleisch J, Zelazo PD, et al. Acute alcohol intoxication and cognitive functioning. *J Stud Alcohol* 1990;51:114–122.

3. Krull KR, Smith LT, Parsons OA. Simple reaction time event-related potentials: effects of alcohol and diazepam. *Prog Neuropsychopharmacol Biol Psychiatry* 1994;18:1247–1260.

4. Heath AC, Martin NG. Genetic differences in psychomotor performance decrement after alcohol: a multivariate analysis. *J Stud Alcohol* 1992;53:262–271.

5. Moeller FG, Dougherty DM, Lane SD, et al. Antisocial personality disorder and alcohol-induced aggression. *Alcohol Clin Exp Res* 1998;22:1898–1902.

6. Linnoila M. Effect of drugs and alcohol on psychomotor skills related to driving. *Ann Clin Res* 1974;6:7–18.

7. Trevisan L, Boutros N, Petrakis I, et al. Complications of alcohol withdrawal: pathophysiological insights. *Alcohol Health Res World* 1998;22:61–66.

8. De Soto CB, O'Donnell WE, Allred LJ, et al. Symptomatology in alcoholics at various stages of abstinence. *Alcohol Clin Exp Res* 1985;9:505–512.

9. Victor M, Adams RD, Collins GH. *The Wernicke-Korsakoff syndrome and related neurologic disorders due to alcoholism and malnutrition*. Philadelphia: FA Davies, 1989.

10. Kopelman MD. The Korsakoff syndrome. *Br J Psychiatry* 1995;166:154–173.

11. Franks NP, Lieb WR. Inhibitory synapses. Anaesthetics set their sites on ion channels [news; comment]. *Nature* 1997;389:334–335.

12. Mihic SJ, Ye Q, Wick MJ, et al. Sites of alcohol and volatile anaesthetic action on GABA (A) and glycine receptors [see comments]. *Nature* 1997;389:385–389.

13. Alexander SPH, Peters JA, eds. *Receptor and ion channel nomenclature supplement*. Trends in pharmacological sciences (series). 1998.

14. Sucher NJ, Awobuluyi M, Choi YB, et al. NMDA receptors: from genes to channels. *Trends Pharmacol Sci* 1996;17:348–355.

15. Grant KA, Lovinger DM. Cellular and behavioral neurobiology of alcohol: receptor-mediated neuronal processes. *Clin Neurosci* 1995;3:155–164.

16. Popp R, Lickteig R, Lovinger D. Factors that enhance ethanol inhibition of N-methyl-D-aspartate receptors in cerebellar granule cells. *J Pharmacol Exp Ther* 1999;289:1564–1574.

17. Miyakawa T, Yagi T, Kitazawa H, et al. Fyn-kinase as a determinant of ethanol sensitivity: relation to NMDA-receptor function. *Science* 1997;278:698–701.

18. Lovinger DM. Developmental decrease in ethanol inhibition of N-methyl-D-aspartate receptors in rat neocortical neurons: relation to the actions of ifenprodil. *J Pharmacol Exp Ther* 1995;274:164–172.

19. Kuner T, Schoepfer R, Korpi ER. Ethanol inhibits glutamate-induced currents in heteromeric NMDA receptor subtypes. *Neuroreport* 1993;5:297–300.

20. Chandler LJ, Sumners C, Crews FT. Ethanol inhibits NMDA receptor-mediated excitotoxicity in rat primary neuronal cultures. *Alcohol Clin Exp Res* 1993;17:54–60.

21. Imperato A, Di Chiara G. Preferential stimulation of dopamine release in the nucleus accumbens of freely moving rats by ethanol. *J Pharmacol Exp Ther* 1986;239:219–228.

22. Moghaddam B, Bolinao ML. Biphasic effect of ethanol on extracellular accumulation of glutamate in the hippocampus and the nucleus accumbens. *Neurosci Lett* 1994;178:99–102.

23. Grunze HC, Rainnie DG, Hasselmo ME, et al. NMDA-dependent modulation of CA1 local circuit inhibition. *J Neurosci* 1996;16:2034–2043.

24. Grant K. Strategies for understanding the pharmacological effects of ethanol with drug discrimination procedures. *Pharmacol Biochem Behav* 1999;64:261–267.

25. Valverius P, Crabbe J, Hoffman P, et al. NMDA receptors in mice bred to be prone or resistant to ethanol withdrawal seizures. *Eur J Pharmacol* 1990;184:185–189.

26. Balster R, Wiley J, Tokarz M, et al. Effects of ethanol and NMDA antagonists on operant behavior in ethanol withdrawal seizure-prone and -resistant mice. *Behav Pharmacol* 1993;4:107–113.

27. Gulya K, Grant KA, Valverius P, et al. Brain regional specificity and time-course of changes in the NMDA receptor-ionophore complex during ethanol withdrawal. *Brain Res* 1991;547:129–134.

28. Grant KA, Snell LD, Rogawski MA, et al. Comparison of the effects of the uncompetitive N-methyl-D-aspartate antagonist (+-)-5-aminocarbonyl-10,11-dihydro-5H-dibenzo[a, d] cyclohepten-5,10-imine (ADCI) with its structural analogs dizocilpine (MK-801) and carbamazepine on ethanol withdrawal seizures. *J Pharmacol Exp Ther* 1992;260:1017–1022.

29. Snell L, Szabó G, Tabakoff B, et al. Gangliosides reduce the development of ethanol dependence without affecting ethanol tolerance. *J Pharmacol Exp Ther* 1996;279:128–136.

30. Snell LD, Nunley KR, Lickteig RL, et al. Regional and subunit specific changes in NMDA receptor mRNA and immunoreactivity in mouse brain following chronic ethanol ingestion. *Brain Res Mol Brain Res* 1996;40:71–78.

31. Trevisan L, Fitzgerald LW, Brose N, et al. Chronic ingestion of ethanol up-regulates NMDAR1 receptor subunit immunoreactivity in rat hippocampus. *J Neurochem* 1994;62:1635–1638.

32. Iorio K, Reinlib L, Tabakoff B, et al. Chronic exposure of cerebellar granule cells to ethanol results in increased NMDA receptor function. *Mol Brain Res* 1992;41:1142–1148.

33. Rossetti ZL, Carboni S. Ethanol withdrawal is associated with increased extracellular glutamate in the rat striatum. *Eur J Pharmacol* 1995;283:177–183.

34. Hoffman PL. Glutamate receptors in alcohol withdrawalinduced neurotoxicity. *Metab Brain Dis* 1995;10:73–79.

35. Krystal JH, Petrakis IL, Webb E, et al. Dose-related ethanol-like effects of the NMDA antagonist, ketamine, in recently detoxified alcoholics. *Arch Gen Psychiatry* 1998;55:354–360.

36. Schutz CG, Soyka M. Dextromethorphan challenge in alcohol-dependent patients and controls [letter; comment]. *Arch Gen Psychiatry* 2000;57:291–292.

37. Krystal JH, D'Souza DC, Karper LP, et al. Interactive effects of subanesthetic ketamine and haloperidol. *Psychopharmacology* 1999;145:193–204.

38. Krystal JH, Petrakis IL. Dextromethorphan challenge in alcohol dependent patients and controls—reply. *Arch Gen Psychiatry* 2000;57:292.

39. Perez-Reyes M, White WR, McDonald SA, et al. Interaction between ethanol and dextroamphetamine: effects on psychomotor performance. *Alcohol Clin Exp Res* 1992;16:75–81.

40. Madonick SM, D'Souza DC, Brush L, et al. Naltrexone blockade of ketamine intoxication in healthy humans. *Alcohol Clin Exp Res* 1999;23:103A(no. 586).

41. Swift RM, Whelihan W, Kuznetsov O, et al. Naltrexoneinduced alterations in human ethanol intoxication. *Am J Psychiatry* 1994;151:1463–1467.

42. Krystal JH, Karper LP, Bennett A, et al. Interactive effects of subanesthetic ketamine and subhypnotic lorazepam in humans. *Psychopharmacology* 1998;135:213–229.

43. Anand A, Charney DS, Cappiello A, et al. Lamotrigine attenuates ketamine effects in humans: support for hyperglutamatergic effects of NMDA antagonists. *Arch Gen Psychiatry* 2000;57:270–276.

44. Krupitsky E, Burakov A, Romanova T, et al. Interactive ethanol-like cognitive and behavioral effects of nimodipine and ketamine in abstinent alcoholics. *Alcohol Clin Exp Res* 2000;24:12A.

45. Michaelis EK, Freed WJ, Galton N, et al. Glutamate receptor changes in brain synaptic membranes from human alcoholics. *Neurochem Res* 1990;15:1055–1063.

46. Freund G, Anderson KJ. Glutamate receptors in the frontal cortex of alcoholics. *Alcohol Clin Exp Res* 1996;20:1165–1172.

47. Tsai G, Gastfriend DR, Coyle JT. The glutamatergic basis of human alcoholism. *Am J Psychiatry* 1995;152:332–340.

48. Brown ME, Anton RF, Malcolm R, et al. Alcohol detoxification and withdrawal seizures: clinical support for a kindling hypothesis. *Biol Psychiatry* 1988;23:507–514.

49. Krystal JH, Webb E, Grillon C, et al. Evidence of acoustic startle hyperreflexia in recently detoxified early onset male alcoholics: modulation by yohimbine and m-chlorophenylpiperazine (mCPP). *Psychopharmacology* 1997;131:207–215.

50. Krystal J, Karper L, D'Souza DC, et al. Differentiating NMDA dysregulation in schizophrenia and alcoholism using ketamine. *Schizophr Res* 1995;15:156.

51. Krystal J, Petrakis I, Krasnicki S, et al. Altered responses to agonists of the strychnine-insensitive glycine NMDA coagonist (SIGLY) site in recently detoxified alcoholics. *Alcohol Clin Exp Res* 1999;22:94A.

52. Petrakis IL, Boutros NN, O'Malley SM, et al. NMDA dysregulation in individuals with a family vulnerability to alcoholism. *Alcohol Clin Exp Res* 2000;24:12A.

53. Spanagel R, Zieglgansberger W. Anti-craving compounds for ethanol: new pharmacological tools to study addictive processes. *Trends Pharmacol Sci* 1997;18:54–59.

54. Sass H, Soyka M, Mann K, et al. Relapse prevention by acamprosate. Results from a placebo-controlled study on alcohol dependence [published erratum appears in *Arch Gen Psychiatry* 1996;53(12):1097]. *Arch Gen Psychiatry* 1996;53:673–680.

55. Mason BJ. Acamprosate: new preclinical and clinical findings. 2000 Scientific Meeting of the Research Society on Alcoholism. Denver, CO, 2000;3A.

56. Zeise ML, Kasparov S, Capogna M, et al. Acamprosate (calcium acetylhomotaurinate) decreases postsynaptic potentials in the rat neocortex: possible involvement of excitatory amino acid receptors. *Eur J Pharmacol* 1993;231:47–52.

57. Madamba SG, Schweitzer P, Zieglgansberger W, et al. Acamprosate (calcium acetylhomotaurinate) enhances the N-methyl-D-aspartate component of excitatory neurotransmission in rat hippocampal CA1 neurons in vitro. *Alcohol Clin Exp Res* 1996;20:651–658.

58. Allan AM, Harris RA. Acute and chronic ethanol treatments alter GABA receptor-operated chloride channels. *Pharmacol Biochem Behav* 1987;27:665–670.

59. Suzdak P, Glowa J, Crawley J, et al. A selective imidazobenzodiazepine antagonist of ethanol in the rat. *Science* 1986;234:1243–1247.

60. Simson PE, Criswell HE, Breese GR. Ethanol potentiates gamma-aminobutyric acid-mediated inhibition in the inferior colliculus: evidence for local ethanol/gamma-aminobutyric acid interactions. *J Pharmacol Exp Ther* 1991;259:1288–1293.

61. Hevers W, Lüddens H. The diversity of GABAA receptors. *Mol Neurobiol* 1998;18:35–86.

62. Wafford K, Burnett D, Leidenheimer N, et al. Ethanol sensitivity of the GABAA receptor expressed in Xenopus oocytes requires 8 amino acids contained in the gamma-2L subunit. *Neuron* 1991;7:27–33.

63. Harris RA, McQuilkin SJ, Paylor R, et al. Mutant mice lacking the gamma isoform of protein kinase C show decreased behavioral actions of ethanol and altered function of gamma-aminobutyrate type A receptors [see comments]. *Proc Natl Acad Sci USA* 1995;92:3658–3662.

64. Bowers B, Owen E, Collins A, et al. Decreased ethanol sensitivity and tolerance development in gamma-protein kinase C null mutant mice is dependent on genetic background. *Alcohol Clin Exp Res* 1999;23:387–397.

65. Lin A, Freund R, Palmer M. Sensitization of gamma-aminobutyric acid-induced depressions of cerebellar Purkinje neurons to the potentiative effects of ethanol by beta adrenergic mechanisms in rat brain. *J Pharmacol Exp Ther* 1993;265:426–432.

66. Freund RK, Palmer MR. 8-Bromo-cAMP mimics beta-adrenergic sensitization of GABA responses to ethanol in cerebellar Purkinje neurons in vivo. *Alcohol Clin Exp Res* 1996;20:408–412.

67. Montpied P, Morrow AL, Karanian JW, et al. Prolonged ethanol inhalation decreases gamma-aminobutyric acid A receptor alpha subunit mRNAs in the rat cerebral cortex. *Mol Pharmacol* 1991;39:157–163.

68. Hemmingsen R, Braestrup C, Nielsen M, et al. The benzodiazepine/GABA receptor complex during severe ethanol intoxication and withdrawal in the rat. *Acta Psychiatr Scand* 1982;65:120–126.

69. Buck KJ. Molecular genetic analysis of the role of GABAergic systems in the behavioral and cellular actions of alcohol. *Behav Genet* 1996;26:313–323.

70. Parsian A, Cloninger CR. Human GABAA receptor alpha 1 and alpha 3 subunits genes and alcoholism. *Alcohol Clin Exp Res* 1997;21:430–433.

71. Cowley DS, Roy-Byrne PP, Radant A, et al. Eye movement effects of diazepam in sons of alcoholic fathers and male control subjects. *Alcohol Clin Exp Res* 1994;18:324–332.

72. Schuckit MA, Tsuang JW, Anthenelli RM, et al. Alcohol challenges in young men from alcoholic pedigrees and control families: a report from the COGA project. *J Stud Alcohol* 1996;57:368–377.

73. de Wit H, McCracken SG. Ethanol self-administration in males with and without an alcoholic first-degree relative. *Alcohol Clin Exp Res* 1990;14:63–70.

74. Volkow ND, Wang GJ, Begleiter H, et al. Regional brain metabolic response to lorazepam in subjects at risk for alcoholism. *Alcohol Clin Exp Res* 1995;19:510–516.

75. de Wit H. Diazepam preference in males with and without an alcoholic first-degree relative. *Alcohol Clin Exp Res* 1991;15:593–600.

76. Chutuape MA, de Wit H. Preferences for ethanol and diazepam in anxious individuals: an evaluation of the self-medication hypothesis. *Psychopharmacology* 1995;121:91–103.

77. Adinoff B, Kramer GL, Petty F. Levels of gamma-aminobutyric acid in cerebrospinal fluid and plasma during alcohol withdrawal. *Psychiatry Res* 1995;59:137–144.

78. Behar K, Rothman D, Petersen K, et al. Preliminary evidence of reduced cortical GABA levels in localized 1H NMR spectra of alcohol dependent and hepatic encephalopathy patients. *Am J Psychiatry* 1999;156:952–954.

79. Petty F, Kramer GL, Davis LL, et al. Plasma gamma-aminobutyric acid (GABA) predicts outcome in patients with alcohol dependence. *Prog Neuropsychopharmacol Biol Psychiatry* 1997;21:809–816.

80. Sherif FM, Tawati AM, Ahmed SS, et al. Basic aspects of GABA-transmission in alcoholism, with particular reference to GABA-transaminase. *Eur Neuropsychopharmacol* 1997;7:1–7.

81. Tran VT, Snyder SH, Major LF, et al. GABA receptors are increased in brains of alcoholics. *Ann Neurol* 1981;9:289–292.

82. Korpi ER. Role of GABAA receptors in the actions of alcohol and in alcoholism: recent advances. *Alcohol Alcohol* 1994;29:115–129.

83. Freund G, Ballinger WE Jr. Decrease of benzodiazepine receptors in frontal cortex of alcoholics. *Alcohol* 1988;5:275–282.

84. Dodd PR, Kril JJ, Thomas GJ, et al. Receptor binding sites and uptake activities mediating GABA neurotransmission in chronic alcoholics with Wernicke encephalopathy. *Brain Res* 1996;710:215–228.

85. Gilman S, Koeppe RA, Adams K, et al. Positron emission tomographic studies of cerebral benzodiazepine-receptor binding in chronic alcoholics. *Ann Neurol* 1996;40:163–171.

86. Abi-Dargham A, Krystal JH, Anjivel S, et al. Alterations of benzodiazepine receptors in type II alcoholics measured with SPECT and [123I]iomazenil. *Am J Psychiatry* 1998;155:1550–1555.

87. Lingford-Hughes AR, Acton PD, Gacinovic S, et al. Reduced levels of GABA-benzodiazepine receptor in alcohol dependency in the absence of grey matter atrophy. *Br J Psychiatry* 1998;173:116–122.

88. Volkow ND, Wang GJ, Hitzemann R, et al. Decreased cerebral response to inhibitory neurotransmission in alcoholics. *Am J Psychiatry* 1993;150:417–422.

89. Volkow ND, Wang GJ, Overall JE, et al. Regional brain metabolic response to lorazepam in alcoholics during early and late alcohol detoxification. *Alcohol Clin Exp Res* 1997;21:1278–1284.

90. Nutt D, Adinoff B, Linnoila M. Benzodiazepines in the treatment of alcoholism. *Recent Dev Alcoholism* 1989;7:283–313.

91. Lheureux P, Askenasi R. Efficacy of flumazenil in acute alcohol intoxication: double blind placebo-controlled evaluation. *Hum Exp Toxicol* 1991;10:235–239.

92. Potokar J, Coupland N, Glue P, et al. Flumazenil in alcohol withdrawal: a double-blind placebo-controlled study. *Alcohol Alcohol* 1997;32:605–611.

93. Melchior CL, Allen PM. Interaction of pregnanolone and pregnenolone sulfate with ethanol and pentobarbital. *Pharmacol Biochem Behav* 1992;42:605–611.

94. Devaud LL, Purdy RH, Morrow AL. The neurosteroid, 3 alpha-hydroxy-5 alpha-pregnan-20-one, protects against bicuculline-induced seizures during ethanol withdrawal in rats. *Alcohol Clin Exp Res* 1995;19:350–355.

95. Grant KA, Azarov A, Bowen CA, et al. Ethanol-like discriminative stimulus effects of the neurosteroid 3 alpha-hydroxy-5 alpha-pregnan-20-one in female Macaca fascicularis monkeys. *Psychopharmacology* 1996;124:340–346.

96. Brot MD, Koob GF, Britton KT. Anxiolytic effects of steroid hormones during the estrous cycle. Interactions with ethanol. *Recent Dev Alcoholism* 1995;12:43–59.

97. Rapkin AJ, Morgan M, Goldman L, et al. Progesterone metabolite allopregnanolone in women with premenstrual syndrome. *Obstet Gynecol* 1997;90:709–714.

98. Grant KA, Azarov A, Shively CA, et al. Discriminative stimulus effects of ethanol and 3 alpha-hydroxy-5 alpha-pregnan-20-one in relation to menstrual cycle phase in cynomolgus monkeys (Macaca fascicularis). *Psychopharmacology* 1997;130:59–68.

99. Charette L, Tate DL, Wilson A. Alcohol consumption and menstrual distress in women at higher and lower risk for alcoholism. *Alcohol Clin Exp Res* 1990;14:152–157.

100. de Wit H, Rukstalis M. Acute effects of triazolam in women: relationships with progesterone, estradiol and allopregnanolone. *Psychopharmacology* 1997;130:69–78.

101. Bollmann JH, Helmchen F, Borst JG, et al. Postsynaptic Ca2+ influx mediated by three different pathways during synaptic transmission at a calyx-type synapse. *J Neurosci* 1998;18:10409–10419.

102. Schiller J, Schiller Y, Clapham DE. NMDA receptors amplify calcium influx into dendritic spines during associative pre- and postsynaptic activation [see comments]. *Nature Neurosci* 1998;1:114–118.

103. Catterall WA. Structure and function of neuronal Ca2+ channels and their role in neurotransmitter release. *Cell Calcium* 1998;24:307–323.

104. Jones SW. Overview of voltage-dependent calcium channels. *J Bioenerget Biomembr* 1998;30:299–312.

105. Green KL, Grant KA. Effects of L-type voltage-sensitive calcium channel modulators on the discriminative stimulus effects of ethanol in rats. *Alcohol Clin Exp Res* 1999;23:806–814.

106. Leslie S. Sedative-hypnotic drugs: interaction with calcium channels. *Alcohol Drug Res* 1986;6:371–377.

107. Mullikin-Kilpatrick D, Treistman S. Inhibition of dihydropyridine-sensitive Ca2+ channels by ethanol in undifferentiated and nerve growth factor-treated PC12 cells: interaction with the inactivated state. *J Pharmacol Exp Ther* 1995;272:489–497.

108. Messing R, Sneade A, Savidge B. Protein kinase C participates in up-regulation of dihydropyridine-sensitive calcium channels by ethanol. *J Neurochem* 1990;55:1383–1389.

109. Little H, Dolin S, Whittington M. Possible role of calcium channels in ethanol tolerance and dependence. *Ann NY Acad Sci* 1988;560:465–466.

110. Brennan C, Crabbe J, Littleton J. Genetic regulation of dihydropyridine-sensitive calcium channels in brain may determine susceptibility to physical dependence on alcohol. *Neuropharmacology* 1990;29:429–432.

111. Wang X, Dayanithi G, Lemos J, et al. Calcium currents and peptide release from neurohypophyseal terminals are inhibited by ethanol. *J Pharmacol Exp Ther* 1991;259:705–711.

112. Twombly D, Herman M, Kye C, et al. Ethanol effects on two types of voltage-activated calcium channels. *J Pharmacol Exp Ther* 1990;254:1029–1037.

113. McMahon T, Anderson R, Metten P, et al. Protein kinase Ce mediates up-regulation of N-type calcium channels by ethanol. *Mol Pharmacol* 2000;57:53–58.

114. Krystal JH, Petrakis IL, Webb E, et al. Dose-related ethanol-like effects of the NMDA antagonist, ketamine, in recently detoxified alcoholics. *Arch Gen Psychiatry* 1998;55:354–360.

115. Banger M, Benkert O, Roschke J, et al. Nimodipine in acute alcohol withdrawal state. *J Psychiatr Res* 1992;26:117–123.

116. Grant KA, Colombo G, Gatto GJ. Characterization of the ethanol-like discriminative stimulus effects of 5-HT receptor agonists as a function of ethanol training dose. *Psychopharmacology* 1997;133:133–141.

117. Krystal JH, Webb E, Cooney N, et al. Specificity of ethanollike effects elicited by serotonergic and noradrenergic mechanisms. *Arch Gen Psychiatry* 1994;51:898–911.

118. Buydens-Branchey L, Branchey M, Fergeson P, et al. Hormonal, psychological, and alcohol craving changes after m-chlorophenylpiperazine administration in alcoholics. *Alcohol Clin Exp Res* 1997;21:220–226.

119. George DT, Benkelfat C, Rawlings RR, et al. Behavioral and neuroendocrine responses to m-chlorophenylpiperazine in subtypes of alcoholics and in healthy comparison subjects. *Am J Psychiatry* 1997;154:81–87.

120. Buydens-Branchey L, Branchey M, Fergeson P, et al. The meta-chlorophenylpiperazine challenge test in cocaine addicts: hormonal and psychological responses. *Biol Psychiatry* 1997;41:1071–1086.

121. Krystal JH, Webb E, Cooney NL, et al. Serotonergic and noradrenergic dysregulation in alcoholism: m-chlorophenylpiperazine and yohimbine effects in recently detoxified alcoholics and healthy comparison subjects. *Am J Psychiatry* 1996;153:83–92.

122. Hommer D, Andreasen P, Rio D, et al. Effects of m-chlorophenylpiperazine on regional brain glucose utilization: a positron emission tomographic comparison of alcoholic and control subjects. *J Neurosci* 1997;17:2796–2806.

123. Callahan PM, Cunningham KA. Involvement of 5-HT2C receptors in mediating the discriminative stimulus properties of m-chlorophenylpiperazine (mCPP). *Eur J Pharmacol* 1994;257:27–38.

124. Robertson DW, Bloomquist W, Wong DT, et al. mCPP but not TFMPP is an antagonist at cardiac 5HT3 receptors. *Life Sci* 1992;50:599–605.

125. To ZP, Bonhaus DW, Eglen RM, et al. Characterization and distribution of putative 5-HT7 receptors in guinea-pig brain. *Br J Pharmacol* 1995;115:107–116.

126. Pauwels PJ, Palmier C, Wurch T, et al. Pharmacology of cloned human 5-HT1D receptor-mediated functional responses in stably transfected rat C6-glial cell lines: further evidence differentiating human 5-HT1D and 5-HT1B receptors. *Naunyn Schmiedebergs Arch Pharmacol* 1996;353:144–156.

127. Shi WX, Nathaniel P, Bunney BS. Ritanserin, a 5-HT2A/2C antagonist, reverses direct dopamine agonist-induced inhibition of midbrain dopamine neurons. *J Pharmacol Exp Ther* 1995;274:735–740.

128. Boess FG, Monsma FJ Jr, Carolo C, et al. Functional and radioligand binding characterization of rat 5-HT6 receptors stably expressed in HEK293 cells. *Neuropharmacology* 1997;36:713–720.

129. Seibyl JP, Krystal JH, Price LH, et al. Effects of ritanserin on the behavioral, neuroendocrine, and cardiovascular responses to meta-chlorophenylpiperazine in healthy human subjects. *Psychiatry Res* 1991;38:227–236.

130. Johnson BA, Jasinski DR, Galloway GP, et al. Ritanserin in the treatment of alcohol dependence—a multi-center clinical trial. Ritanserin Study Group. *Psychopharmacology (Berl)* 1996;128:206–215.

131. Gelernter J, Kranzler H, Cubells JF. Serotonin transporter protein (SLC6A4) allele and haplotype frequencies and linkage disequilibria in African- and European-American and Japanese populations and in alcohol-dependent subjects. *Hum Genet* 1997;101:243–246.

132. Nielsen DA, Goldman D, Virkkunen M, et al. TPH replication study: not [letter; comment]. *Arch Gen Psychiatry* 1996;53:964–965.

133. Sander T, Harms H, Podschus J, et al. Dopamine D1, D2 and D3 receptor genes in alcohol dependence. *Psychiatr Genet* 1995;5:171–176.

134. Nielsen DA, Goldman D, Virkkunen M, et al. Suicidality and 5-hydroxyindoleacetic acid concentration associated with a tryptophan hydroxylase polymorphism. *Arch Gen Psychiatry* 1994;51:34–38.

135. Daoust M, Lhuintre JP, Ernouf D, et al. Ethanol intake and 3H-serotonin uptake. II: A study in alcoholic patients using platelets 3H-paroxetine binding. *Life Sci* 1991;48:1977–1983.

136. Rausch JL, Monteiro MG, Schuckit MA. Platelet serotonin uptake in men with family histories of alcoholism. *Neuropsychopharmacology* 1991;4:83–86.

137. Tiihonen J, Kuikka JT, Bergstrom KA, et al. Single-photon emission tomography imaging of monoamine transporters in impulsive violent behaviour. *Eur J Nucl Med* 1997;24:1253–1260.

138. Chen HT, Casanova MF, Kleinman JE, et al. 3H-paroxetine binding in brains of alcoholics. *Psychiatry Res* 1991;38:293–299.

139. Kranzler HR, Burleson JA, Brown J, et al. Fluoxetine treatment seems to reduce the beneficial effects of cognitive-behavioral therapy in type B alcoholics. *Alcohol Clin Exp Res* 1996;20:1534–1541.

140. Naranjo CA, Poulos CX, Bremner KE, et al. Fluoxetine attenuates alcohol intake and desire to drink. *Int Clin Psychopharmacol* 1994;9:163–172.

141. Cornelius JR, Salloum IM, Ehler JG, et al. Fluoxetine in depressed alcoholics. A double-blind, placebo-controlled trial. *Arch Gen Psychiatry* 1997;54:700–705.

142. Malcolm R, Anton RF, Randall CL, et al. A placebo-controlled trial of buspirone in anxious inpatient alcoholics. *Alcohol Clin Exp Res* 1992;16:1007–1013.

143. Grant KA, Barrett JE. Blockade of the discriminative stimulus effects of ethanol with 5-HT3 receptor antagonists. *Psychopharmacology* 1991;104:451–456.

144. Swift RM, Davidson D, Whelihan W, et al. Ondansetron alters human alcohol intoxication. *Biol Psychiatry* 1996;40:514–521.

145. Johnson BA, Roache JD, Javors MA, et al. Ondansetron for reduction of drinking among biologically predisposed alcoholic patients: A randomized controlled trial. *JAMA* 2000;284:963–971.

146. Koob G, Bloom F. Cellular and molecular mechanisms of drug dependence. *Science* 1988;242:715–723.

147. Mereu G, Fadda F, Gessa G. Ethanol stimulates the firing rate of nigral dopaminergic neurons in unanesthetized rats. *Brain Res* 1984;292:63–69.

148. Brodie M, Shefner S, Dunwiddie T. Ethanol increases the firing rate of dopamine neurons of the rat ventral tegmental area in vitro. *Brain Res* 1990;508:65–69.

149. Weiss F, Lorang M, Bloom F, et al. Oral alcohol self-administration stimulates dopamine release in the rat nucleus accumbens: genetic and motivational determinants. *J Pharmacol Exp Ther* 1993;267:250–258.

150. Hodge C, Chappelle A, Samson H. Dopamine receptors in the medial prefrontal cortex influence ethanol and sucrose-reinforced responding. *Alcohol Clin Exp Res* 1996;20:1631–1638.

151. Gonzales R, Weiss F. Suppression of ethanol-reinforced behavior by naltrexone is associated with attenuation of the ethanol-induced increase in dialysate dopamine levels in the nucleus accumbens. *J Neurosci* 1998;18:10663–10671.

152. Blomqvist O, Ericson M, Johnson D, et al. Voluntary ethanol intake in the rat: effects of nicotinic acetylcholine receptor blockade or subchronic nicotine treatment. *Eur J Pharmacol* 1996;314:257–267.

153. Weiss F, Parsons L, Schulteis G, et al. Ethanol self-administration restores withdrawal-associated deficiencies in accumbal dopamine and 5-hydroxytryptamine release in dependent rats. *J Neurosci* 1996;16:3474–3485.

154. Rossetti Z, Melis F, Carboni S, et al. Marked decrease of extraneuronal dopamine after alcohol withdrawal in rats: reversal by MK-801. *Eur J Pharmacol* 1991;200:371–372.

155. Rossetti Z, Isola D, De Vry J, et al. Effects of nimodipine on extracellular dopamine levels in the rat nucleus accumbens in ethanol withdrawal. *Neuropharmacology* 1999;38:1361–1365.

156. Grzanna R, Molliver ME. The locus coeruleus in the rat: an immunohistochemical delineation. *Neuroscience* 1980;5:21–40.

157. Aston-Jones G, Foote S, Bloom F. Low doses of ethanol disrupt sensory responses to brain noradrenergic neurons. *Nature* 1982;296:857–860.

158. Shefner S, Tabakoff B. Basal firing of rat locus coeruleus neurons affects sensitivity to ethanol. *Alcohol* 1985;2:239–243.

159. Tabakoff B, Hoffman P. Effect of alcohol on neurotransmitters and their receptors and enzymes. In: Begleiter H, Kissin B, eds. *The pharmacology of alcohol and alcohol dependence*. New York: Oxford University Press, 1996.

160. Ahlenius S, Carlsson A, Engel J, et al. Antagonism by alpha methyltyrosine of the ethanol-induced stimulation and euphoria in man. *Clin Pharmacol Ther* 1973;14:586–591.

161. McDougle CJ, Krystal JH, Price LH, et al. Noradrenergic response to acute ethanol administration in healthy subjects: comparison with intravenous yohimbine. *Psychopharmacology* 1995;118:127–135.

162. Alkana RL, Parker ES, Cohen HB, et al. Reversal of ethanol intoxications in humans: an assessment of the efficacy of propranolol. *Psychopharmacology* 1976;51:29–37.

163. Alkana RL, Parker ES, Cohen HB, et al. Reversal of ethanol intoxication in humans: an assessment of the efficacy of L-dopa, aminophylline, and ephedrine. *Psychopharmacology* 1977;55:203–212.

164. Balldin J, Alling C, Gottfries CG, et al. Changes in dopamine receptor sensitivity in humans after heavy alcohol intake. *Psychopharmacology* 1985;86:142–146.

165. Volkow ND, Wang GJ, Fowler JS, et al. Decreases in dopamine receptors but not in dopamine transporters in alcoholics. *Alcohol Clin Exp Res* 1996;20:1594–1598.

166. Glue P, Sellman JD, Nicholls MG, et al. Studies of alpha-2-adrenoceptor function in abstinent alcoholics. *Br J Addict* 1989;84:97–102.

167. Borg S, Czarnecka A, Kvande H, et al. Clinical conditions and concentrations of MOPEG in the cerebrospinal fluid and urine of male alcoholic patients during withdrawal. *Alcohol Clin Exp Res* 1983;7:411–415.

168. Blum K, Noble EP, Sheridan PJ, et al. Allelic association of human dopamine D2 receptor gene in alcoholism [see comments]. *JAMA* 1990;263:2055–2060.

169. Parsian A, Todd RD, Devor EJ, et al. Alcoholism and alleles of the human D2 dopamine receptor locus. Studies of association and linkage. *Arch Gen Psychiatry* 1991;48:655–663.

170. Gejman PV, Ram A, Gelernter J, et al. No structural mutation in the dopamine D2 receptor gene in alcoholism or schizophrenia. Analysis using denaturing gradient gel electrophoresis. *JAMA* 1994;271:204–208.

171. Gelernter J, Goldman D, Risch N. The A1 allele at the D2 dopamine receptor gene and alcoholism. A reappraisal [see comments]. *JAMA* 1993;269:1673–1677.

172. Lawford BR, Young RM, Rowell JA, et al. Bromocriptine in the treatment of alcoholics with the D2 dopamine receptor A1 allele [see comments]. *Nature Med* 1995;1:337–341.

173. Penick EC, Powell BJ, Campbell J, et al. Pharmacological treatment for antisocial personality disorder alcoholics: a preliminary study. *Alcohol Clin Exp Res* 1996;20:477–484.

174. Naranjo CA, Dongier M, Bremner KE. Long-acting injectable bromocriptine does not reduce relapse in alcoholics. *Addiction* 1997;92:969–978.

175. Sander T, Harms H, Dufeu P, et al. Dopamine D4 receptor exon III alleles and variation of novelty seeking in alcoholics. *Am J Med Genet* 1997;74:483–487.

176. Muramatsu T, Higuchi S, Murayama M, et al. Association between alcoholism and the dopamine D4 receptor gene. *J Med Genet* 1996;33:113–115.

177. Adamson MD, Kennedy J, Petronis A, et al. DRD4 dopamine receptor genotype and CSF monoamine metabolites in Finnish alcoholics and controls. *Am J Med Genet* 1995;60:199–205.

178. Chang FM, Ko HC, Lu RB, et al. The dopamine D4 receptor gene (DRD4) is not associated with alcoholism in three Taiwanese populations: six polymorphisms tested separately and as haplotypes. *Biol Psychiatry* 1997;41:394–405.

179. Gorwood P, Martres MP, Ades J, et al. Lack of association between alcohol-dependence and D3 dopamine receptor gene in three independent samples. *Am J Med Genet* 1995;60:529–531.

180. Malhotra AK, Virkkunen M, Rooney W, et al. The association between the dopamine D4 receptor (D4DR) 16 amino acid repeat polymorphism and novelty seeking. *Mol Psychiatry* 1996;1:388–391.

181. Parsian A, Chakraverty S, Fisher L, et al. No association between polymorphisms in the human dopamine D3 and D4 receptors genes and alcoholism. *Am J Med Genet* 1997;74:281–285.

182. Sullivan PF, Fifield WJ, Kennedy MA, et al. No association between novelty seeking and the type 4 dopamine receptor gene (DRD4) in two New Zealand samples. *Am J Psychiatry* 1998;155:98–101.

183. Schulz R, Wuster M, Duka T, et al. Acute and chronic ethanol treatment changes endorphin levels in brain and pituitary. *Psychopharmacol (Berl)* 1980;68:221–227.

184. Gianoulakis C, Barcomb A. Effect of acute ethanol in vivo and in vitro on the b-endorphin system in the rat. *Life Sci* 1987;40:19–28.

185. Gianoulakis C. The effect of ethanol on the biosynthesis and regulation of opioid peptides. *Experientia* 1989;45:428–435.

186. Dave J, Eiden L, Karanian J, et al. Ethanol exposure decreases pituitary corticotropin-releasing factor binding, adenylate cyclase activity, proopiomelanocortin biosynthesis, and plasma b-endorphin levels in the rat. *Endocrinology* 1986;118:280–286.

187. Gulya K, Orpana A, Sikela J, et al. Prodynorphin and vasopressin mRNA levels are differentially affected by chronic ethanol ingestion in the mouse. *Mol Brain Res* 1993;20:1–8.

188. Tabakoff B, Hoffman P. Alcohol interactions with brain opiate receptors. *Life Sci* 1983;32:197–204.

189. Hoffman P, Chung C, Tabakoff B. Effects of ethanol, temperature, and endogenous regulatory factors on the characteristics of striatal opiate receptors. *J Neurochem* 1984;43:1003–1010.

190. Tabakoff B, Urwyler S, Hoffman P. Ethanol alters kinetic characteristics and function of striatal morphine receptors. *J Neurochem* 1981;37:518–521.

191. Pfeiffer A, Seizinger B, Herz A. Chronic ethanol imbibition interferes with d, but not with m-opiate receptors. *Neuropharmacology* 1981;20:1229–1232.

192. Charness M. Ethanol and opioid receptor signalling. *Experientia* 1989;45:418–427.

193. Davidson D, Swift R, Fitz E. Naltrexone increases the latency to drink alcohol in social drinkers. *Alcohol Clin Exp Res* 1996;20:732–739.

194. O'Malley SS, Jaffe AJ, Rode S, et al. Experience of a "slip" among alcoholics treated with naltrexone or placebo. *Am J Psychiatry* 1996;153:281–283.

195. Volpicelli JR, Watson NT, King AC, et al. Effect of naltrexone on alcohol "high" in alcoholics. *Am J Psychiatry* 1995;152:613–615.

196. Volpicelli JR, Alterman AI, Hayashida M, et al. Naltrexone in the treatment of alcohol dependence [see comments]. *Arch Gen Psychiatry* 1992;49:876–880.

197. O'Malley SS, Jaffe AJ, Chang G, et al. Naltrexone and coping skills therapy for alcohol dependence. A controlled study. *Arch Gen Psychiatry* 1992;49:881–887.

198. Ritchie T, Noble EP. [3H]naloxone binding in the human brain: alcoholism and the TaqI A D2 dopamine receptor polymorphism. *Brain Res* 1996;718:193–197.

199. Tabakoff B, Hoffman PL, Valverius P, et al. Characteristics of receptors and enzymes in brains of human alcoholics. *Alcohol* 1985;2:419–423.

200. Genazzani AR, Nappi G, Facchinetti F, et al. Central deficiency of beta-endorphin in alcohol addicts. *J Clin Endocrinol Metab* 1982;55:583–586.

201. Gianoulakis C, de Waele JP, Thavundayil J. Implication of the endogenous opioid system in excessive ethanol consumption. *Alcohol* 1996;13:19–23.

202. King AC, Volpicelli JR, Frazer A, et al. Effect of naltrexone on subjective alcohol response in subjects at high and low risk for future alcohol dependence. *Psychopharmacology* 1997;129:15–22.

203. Bergen AW, Kokoszka J, Peterson R, et al. Mu opioid receptor gene variants: lack of association with alcohol dependence. *Mol Psychiatry* 1997;2:490–494.

204. Chan RJ, McBride AW, Thomasson HR, et al. Allele frequencies of the preproenkephalin A (PENK) gene CA repeat in Asians, African-Americans, and Caucasians: lack of evidence for different allele frequencies in alcoholics. *Alcohol Clin Exp Res* 1994;18:533–535.

205. Harper C, Kril J. An introduction to alcohol-induced brain damage and its causes. *Alcohol Alcohol Suppl* 1994;2:237–243.

206. Cadete-Leite A, Alves MC, Paula-Barbosa MM, et al. Quantitative analysis of basal dendrites of prefrontal pyramidal cells after chronic alcohol consumption and withdrawal in the adult rat. *Alcohol Alcohol* 1990;25:467–475.

207. Harper C. The neuropathology of alcohol-specific brain damage, or does alcohol damage the brain? *J Neuropathol Exp Neurol* 1998;57:101–110.

208. Kril JJ, Halliday GM, Svoboda MD, et al. The cerebral cortex is damaged in chronic alcoholics. *Neuroscience* 1997;79:983–998.

209. Jensen GB, Pakkenberg B. Do alcoholics drink their neurons away? *Lancet* 1993;342:1201–1204.

210. McMullen PA, Saint-Cyr JA, Carlen PL. Morphological alterations in rat CA1 hippocampal cell dendrites resulting from chronic ethanol consumption and withdrawal. *J Comp Neurol* 1984;225:111–118.

211. Kril JJ. The contribution of alcohol, thiamine deficiency and cirrhosis of the liver to cerebral cortical damage in alcoholics. *Metab Brain Dis* 1995;10:9–16.

212. Pfefferbaum A, Rosenbloom M, Crusan K, et al. Brain CT changes in alcoholics: effects of age and alcohol consumption. *Alcohol Clin Exp Res* 1988;12:81–87.

213. Jernigan TL, Butters N, DiTraglia G, et al. Reduced cerebral grey matter observed in alcoholics using magnetic resonance imaging. *Alcohol Clin Exp Res* 1991;15:418–427.

214. Pfefferbaum A, Sullivan EV, Mathalon DH, et al. Frontal lobe volume loss observed with magnetic resonance imaging in older chronic alcoholics. *Alcohol Clin Exp Res* 1997;21:521–529.

215. Pfefferbaum A, Sullivan EV, Rosenbloom MJ, et al. A controlled study of cortical gray matter and ventricular changes in alcoholic men over a 5-year interval. *Arch Gen Psychiatry* 1998;55:905–912.

216. Sullivan EV, Marsh L, Mathalon DH, et al. Anterior hippocampal volume deficits in nonamnesic, aging chronic alcoholics. *Alcohol Clin Exp Res* 1995;19:110–122.

217. Hommer D, Momenan R, Rawlings R, et al. Decreased corpus callosum size among alcoholic women. *Arch Neurol* 1996;53:359–363.

218. Pfefferbaum A, Sullivan EV, Rosenbloom MJ, et al. Increase in brain cerebrospinal fluid volume is greater in older than in younger alcoholic patients: a replication study and CT/MRI comparison. *Psychiatry Res* 1993;50:257–274.

219. Leber WR, Parsons OA. Premature aging and alcoholism. *Int J Addict* 1982;17:61–88.

220. Krabbendam L, Visser PJ, Derix MM, et al. Normal cognitive performance in patients with chronic alcoholism in contrast to patients with Korsakoff's syndrome. *J Neuropsychiatry Clin Neurosci* 2000;12:44–50.

221. De Bellis MD, Clark DB, Beers SR, et al. Hippocampal volume in adolescent-onset alcohol use disorders. *Am J Psychiatry* 2000;157:737–744.

222. Pfefferbaum A, Sullivan EV, Mathalon DH, et al. Longitudinal changes in magnetic resonance imaging brain volumes in abstinent and relapsed alcoholics. *Alcohol Clin Exp Res* 1995;19:1177–1191.

223. Shear PK, Jernigan TL, Butters N. Volumetric magnetic resonance imaging quantification of longitudinal brain changes in abstinent alcoholics [published erratum appears in *Alcohol Clin Exp Res* 1994;18(3):766]. *Alcohol Clin Exp Res* 1994;18:172–176.

224. Mann K, Mundle G, Langle G, et al. The reversibility of alcoholic brain damage is not due to rehydration: a CT study. *Addiction* 1993;88:649–653.

225. Tsai G, Coyle JT. The role of glutamatergic neurotransmission in the pathophysiology of alcoholism. *Annu Rev Med* 1998;49:173–184.

226. Sullivan EV, Marsh L, Mathalon DH, et al. Relationship between alcohol withdrawal seizures and temporal lobe white matter volume deficits. *Alcohol Clin Exp Res* 1996;20:348–354.

227. Nicolas JM, Estruch R, Salamero M, et al. Brain impairment in well-nourished chronic alcoholics is related to ethanol intake. *Ann Neurol* 1997;41:590–598.

228. Barthauer L, Tarter R, Hirsch W, et al. Brain morphologic characteristics of cirrhotic alcoholics and cirrhotic nonalcoholics: an MRI study. *Alcohol Clin Exp Res* 1992;16:982–985.

229. Raine A, Lencz T, Bihrle S, et al. Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Arch Gen Psychiatry* 2000;57:119–27; discussion 128–129.

230. Jagannathan NR, Desai NG, Raghunathan P. Brain metabolite changes in alcoholism: an in vivo proton magnetic resonance spectroscopy (MRS) study. *Magn Reson Imaging* 1996;14:553–557.

231. Seitz D, Widmann U, Seeger U, et al. Localized proton magnetic resonance spectroscopy of the cerebellum in detoxifying alcoholics. *Alcohol Clin Exp Res* 1999;23:158–163.

232. Meyerhoff DJ, MacKay S, Sappey-Marinier D, et al. Effects of chronic alcohol abuse and HIV infection on brain phosphorus metabolites. *Alcohol Clin Exp Res* 1995;19:685–692.

233. Volkow ND, Hitzemann R, Wolf AP, et al. Acute effects of ethanol on regional brain glucose metabolism and transport. *Psychiatry Res* 1990;35:39–48.

234. Volkow ND, Hitzemann R, Wang GJ, et al. Decreased brain metabolism in neurologically intact healthy alcoholics. *Am J Psychiatry* 1992;149:1016–1022.

235. Wang GJ, Volkow ND, Roque CT, et al. Functional importance of ventricular enlargement and cortical atrophy in healthy subjects and alcoholics as assessed with PET, MR imaging, and neuropsychologic testing [see comments]. *Radiology* 1993;186:59–65.

236. Gilman S, Adams K, Koeppe RA, et al. Cerebellar and frontal hypometabolism in alcoholic cerebellar degeneration studied with positron emission tomography. *Ann Neurol* 1990;28:775–785.

237. Volkow ND, Wang GJ, Hitzemann R, et al. Recovery of brain glucose metabolism in detoxified alcoholics. *Am J Psychiatry* 1994;151:178–183.

238. George MS, Teneback CC, Malcolm RJ, et al. Multiple previous alcohol detoxifications are associated with decreased medial temporal and paralimbic function in the postwithdrawal period. *Alcohol Clin Exp Res* 1999;23:1077–1084.

239. Johnson-Greene D, Adams KM, Gilman S, et al. Effects of abstinence and relapse upon neuropsychological function and cerebral glucose metabolism in severe chronic alcoholism. *J Clin Exp Neuropsychol* 1997;19:378–385.

240. Kuruoglu AC, Arikan Z, Vural G, et al. Single photon emission computerised tomography in chronic alcoholism. Antisocial personality disorder may be associated with decreased frontal perfusion. *Br J Psychiatry* 1996;169:348–354.

241. Adams KM, Gilman S, Johnson-Greene D, et al. The significance of family history status in relation to neuropsychological test performance and cerebral glucose metabolism studied with positron emission tomography in older alcoholic patients. *Alcohol Clin Exp Res* 1998;22:105–110.

242. Adams KM, Gilman S, Koeppe RA, et al. Neuropsychological deficits are correlated with frontal hypometabolism in positron emission tomography studies of older alcoholic patients. *Alcohol Clin Exp Res* 1993;17:205–210.

243. Tarter RE, Hegedus AM, Van Thiel DH. Neuropsychiatric sequelae of portal-systemic encephalopathy: a review. International *J Neurosci* 1984;24:203–216.

244. Tarter RE, Sandford SL, Hays AL, et al. Hepatic injury correlates with neuropsychologic impairment. *Int J Neurosci* 1989;44:75–82.

245. Schafer K, Butters N, Smith T, et al. Cognitive performance of alcoholics: a longitudinal evaluation of the role of drinking history, depression, liver function, nutrition, and family history. *Alcohol Clin Exp Res* 1991;15:653–660.

246. Shelton MD, Parsons OA. Alcoholics' self-assessment of their neuropsychological functioning in everyday life. *J Clin Psychol* 1987;43:395–403.

247. Hegedus AM, Alterman AI, Tarter RE. Learning achievement in sons of alcoholics. *Alcohol Clin Exp Res* 1984;8:330–333.

248. Tarter RE, Jacob T, Bremer DA. Cognitive status of sons of alcoholic men. *Alcohol Clin Exp Res* 1989;13:232–235.

249. Schuckit MA, Butters N, Lyn L, et al. Neuropsychologic deficits and the risk for alcoholism. *Neuropsychopharmacology* 1987;1:45–53.

250. Gillen R, Hesselbrock V. Cognitive functioning, ASP, and family history of alcoholism in young men at risk for alcoholism. *Alcohol Clin Exp Res* 1992;16:206–214.

251. Parsons OA, Nixon SJ. Neurobehavioral sequelae of alcoholism. *Neurol Clin* 1993;11:205–218.

252. Ozkaragoz T, Satz P, Noble EP. Neuropsychological functioning in sons of active alcoholic, recovering alcoholic, and social drinking fathers. *Alcohol* 1997;14:31–37.

253. Alterman AI, Bridges KR, Tarter RE. The influence of both drinking and familial risk statuses on cognitive functioning of social drinkers. *Alcohol Clin Exp Res* 1986;10:448–451.

254. Erblich J, Earleywine M. Children of alcoholics exhibit attenuated cognitive impairment during an ethanol challenge. *Alcohol Clin Exp Res* 1999;23:476–482.

255. Nixon SJ, Tivis R, Parsons OA. Behavioral dysfunction and cognitive efficiency in male and female alcoholics. *Alcohol Clin Exp Res* 1995;19:577–581.

256. Glenn SW, Parsons OA. Neuropsychological efficiency measures in male and female alcoholics. *J Stud Alcohol* 1992;53:546–552.

257. Tarter RE, Buonpane N, Wynant C. Intellectual competence of alcoholics. *J Stud Alcohol* 1975;36:381–386.

258. Eckardt MJ, Stapleton JM, Rawlings RR, et al. Neuropsychological functioning in detoxified alcoholics between 18 and 35 years of age. *Am J Psychiatry* 1995;152:53–59.

259. Glenn SW, Parsons OA, Sinha R, et al. The effects of repeated withdrawals from alcohol on the memory of male and female alcoholics. *Alcohol Alcohol* 1988;23:337–342.

260. Fabian MS, Parsons OA. Differential improvement of cognitive functions in recovering alcoholic women. *J Abnorm Psychol* 1983;92:87–95.

261. Jenkins RL, Parsons OA. Recovery of cognitive abilities in male alcoholics. *Curr Alcohol* 1979;7:229–237.

262. Alterman AI, Kushner H, Holahan JM. Cognitive functioning and treatment outcome in alcoholics. *J Nerv Ment Dis* 1990;178:494–499.

263. Leber WR, Parsons OA, Nichols N. Neuropsychological test results are related to ratings of men alcoholics' therapeutic progress: a replicated study. *J Stud Alcohol* 1985;46:116–121.

264. Goldman MS, Klisz DK, Williams DL. Experience-dependent recovery of cognitive functioning in young alcoholics. *Addict Behav* 1985;10:169–176.

265. Parsons OA, Sinha R, Williams HL. Relationships between neuropsychological test performance and event-related potentials in alcoholic and nonalcoholic samples. *Alcohol Clin Exp Res* 1990;14:746–755.

266. Wise RA. Neurobiology of addiction. *Curr Opin Neurobiol* 1996;6:243–251.

267. Koob GF, Roberts AJ, Schulteis G, et al. Neurocircuitry targets in ethanol reward and dependence. *Alcohol Clin Exp Res* 1998;22:3–9.

268. Nie Z, Madamba SG, Siggins GR. Ethanol inhibits glutamatergic neurotransmission in nucleus accumbens neurons by multiple mechanisms. *J Pharmacol Exp Ther* 1994;271:1566–1573.

269. Giancola PR, Moss HB, Martin CS, et al. Executive cognitive functioning predicts reactive aggression in boys at high risk for substance abuse: a prospective study. *Alcohol Clin Exp Res* 1996;20:740–744.

270. Peterson JB, Pihl RO. Information processing, neuropsychological function, and the inherited predisposition to alcoholism. *Neuropsychol Rev* 1990;1:343–369.

271. Carlezon WA Jr, Wise RA. Rewarding actions of phencyclidine and related drugs in nucleus accumbens shell and frontal cortex. *J Neurosci* 1996;16:3112–3122.

272. Schultz W, Tremblay L, Hollerman JR. Reward prediction in primate basal ganglia and frontal cortex. Neuropharmacology 1998;37:421–429.

273. Malcolm R, Myrick H, Roberts JS, et al. The effects of lorazepam and carbamazepine on single vs. multiple previous alcohol withdrawals in an outpatient randomized trial. *Alcohol Clin Exp Res* 2000;24:38A.

274. Pfefferbaum A, Lim KO, Zipursky RB, et al. Brain gray and white matter volume loss accelerates with aging in chronic alcoholics: a quantitative MRI study. *Alcohol Clin Exp Res* 1992;16:1078–1089.

101
Alcoholism Pharmacotherapy

Joseph R. Volpicelli

Suchitra Krishnan-Sarin

Stephanie S. O'Malley

Joesph R. Volpicelli: Department of Psychiatry, University of Pennsylvania, Veterans Affairs Medical Center, Philadelphia, Pennsylvania.

Suchitra Krishnan-Sarin, Stephanie S. O'Malley: Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.

Alcoholism remains one of the most common and significant medical problems in the United States and internationally. For example, in the United States, over 4% of the general population is alcohol dependent and another 5 to 10 million people drink hazardously at least several times per month (1). The economic and medical costs of alcoholism and alcohol abuse continue to escalate. Most recent figures put the economic costs of alcohol-related expenses at $176 billion annually in the United States (2). This includes the economic costs of increased health care expenses, lost productivity at work, and legal expenses. Similarly, although there have been some reductions in the number of motor vehicle deaths attributed to excessive alcohol drinking, the overall number of alcohol-related annual deaths is 105,000 in the United States (3).

Current psychosocial approaches to alcohol addiction are moderately effective, with perhaps as many as half the patients receiving treatment becoming abstinent or significantly reducing episodes of binge drinking (4). In the past two decades significant progress has been made in understanding the pharmacology of alcohol and why some people become dependent. This has led to the development of several medications that have been shown in research studies to improve treatment outcomes. This chapter reviews some of the possible neurobiological mechanisms involved in alcohol reward and dependence, and how medications can affect these systems to facilitate treatment. We introduce future directions for research such as the use of combinations of medications that may have additive or synergistic effects on improving treatment, and discuss the role of psychosocial support to facilitate the effectiveness of pharmacotherapy.

PHARMACOLOGIC TREATMENTS FOR ALCOHOL DETOXIFICATION

Part of "101 - Alcoholism Pharmacotherapy "

The first step in the pharmacologic treatment of alcoholism is to help patients safely detoxify from alcohol. Although historically, alcohol detoxification has occurred in inpatient setting, increasingly alcohol detoxification is being conducted in ambulatory settings. Except in the case of medical or psychiatric emergencies, outcome studies generally show that successful detoxification can safely and effectively be carried out in ambulatory setting using medications such as benzodiazepines (5 ,6). In addition, the use of anticonvulsants has received recent interest.

Benzodiazepines

Benzodiazepines are γ-aminobutyric acid (GABA) agonists that metaanalysis of placebo-controlled double-blind studies have consistently shown to be safe and effective (7). Benzodiazepines differ widely in their pharmacologic half-life, and this has been a factor in the choice of which benzodiazepines to use for detoxification. For example, one popular approach is to use a benzodiazepine with a long half-life such as chlordiazepoxide as a loading dose and let the benzodiazepine self-taper (8). The advantage of this technique is that the dose can be administered in the physician's office, which precludes problems with patience noncompliance. A second approach is to use shorter acting benzodiazepines and titrate the dose depending on symptoms. In a recent study, oxazepam was used as needed depending on the severity of withdrawal symptoms as assessed by the Clinical Institute Withdrawal Assessment for Alcohol-revised (CIWAA-R). As needed oxazepam resulted in effective alcohol withdrawal management with a lower total amount of oxazepam over a shorter duration compared to routine dosing (9).

Anticonvulsants

Several anticonvulsants have been used instead of benzodiazepines for alcohol withdrawal. Anticonvulsants have the

advantage of no abuse potential and a theoretical advantage of reducing kindling, a sensitization of withdrawal symptoms that occurs after multiepisodes of alcohol withdrawal. In one randomized study comparing valproate, a GABAergic agent, with phenobarbital, both medications were effective in reducing withdrawal symptoms, and there were no reliable differences between mediations with the exception of less hostility in the phenobarbital group (10). Carbamazepine has also been used as an alternative to benzodiazepines to attenuate alcohol withdrawal symptoms (11). Although its mechanism of action remains unknown, research generally shows that carbamazepine is as effective as benzodiazepines. Disadvantages of carbamazepine include a rather narrow therapeutic window, the need to monitor serum levels, and hepatotoxic effects. For patients with a history of alcohol withdrawal seizures, however, anticonvulsants such as carbamazepine may be a useful alternative to benzodiazepines (12).

PHARMACOLOGIC TREATMENTS TO REDUCE ALCOHOL RELAPSE

Part of "101 - Alcoholism Pharmacotherapy "

Disulfiram

The aversive agent disulfiram has been available for the treatment of alcoholism since 1949. Disulfiram works by inhibiting the liver enzyme that catalyzes the oxidation of acetaldehyde, a toxic by-product of alcohol, resulting in an aversive reaction to alcohol consumption. In this way, disulfiram is thought to deter drinking by making the negative consequences of drinking more certain, immediate, and aversive than they would be otherwise. Provided that the patient takes the disulfiram, the decision about whether or not to drink is probably shifted toward abstinence when faced with opportunities to drink based on the knowledge of the disulfiram-ethanol interaction. In randomized controlled clinical trials, however, disulfiram has not been shown to be effective in the absence of supervision of ingestion, probably due to poor compliance (13). With supervision and positive contingencies for taking disulfiram, however, the effectiveness of disulfiram appears to be enhanced (14). As an alternative to behavioral methods for enhancing compliance, pharmacologic methods such as implants have been developed. However, these efforts have been unsuccessful perhaps because these implants have not yielded adequate disulfiram blood concentration required to produce a reaction to alcohol (15 ,16).

Opioid Antagonists

Background

The role of the alcohol-induced activation of the endogenous opioid system in the reinforcing effects of alcohol has been well established in dozens of animal models of alcohol drinking (17 ,18 ,19 ,20 ,21 ,22 ,23 ,24 ,25 ,26 ,27 ,28 ,29 ,30 ,31 ,32 and 33). These studies have consistently demonstrated that alcohol enhances the release of endogenous opioids, and alcohol preference is reduced when opioid receptors are blocked.

Alcohol increases the release of opioid peptides *in vivo*, particularly in rats and humans with a genetic predisposition for excessive alcohol drinking (34 ,35). For example, Gianoulakis and colleagues (34) have found that in humans peripheral levels of ß-endorphin increase in family history–positive subjects following a moderate dose of alcohol, whereas there is no increase in ß-endorphin for social drinkers without a family history of alcoholism. Moreover, Froehlich and colleagues (36) have also demonstrated that alcohol-induced ß-endorphin responses both prior to and following alcohol administration are significantly heritable.

Genetic preference for alcohol drinking has been shown to be associated with differences in opioid receptors and opioid peptides (37 ,38). Nonpreferring (NP) rats exhibit differences in the densities of µ opioid receptors in certain brain reward regions compared to alcohol-preferring rats. Transgenic mice lacking ß-endorphin have been shown to exhibit decreased preference for alcohol compared with wild-type mice (39).

Nonspecific and specific opioid antagonists have been found to reduce alcohol self-administration in rodents and monkeys (19 ,22 ,25 ,31 ,40 ,41 ,42 and 43). Preclinical studies have also evaluated the efficacy of antagonists specific for the µ and δ opioid receptors in reducing alcohol drinking. The µ opioid receptor antagonist ß-funaltrexamine (ß-FNA) and the δ opioid receptor antagonists naltrindole (NTI) and naltriben (NTB) have all been shown to reduce alcohol drinking (17 ,18 ,41). Recent evidence also suggests a role for the δ opioid receptors in mediating the aversive effects of alcohol as indicated by an increase in conditioned taste aversion in alcohol preferring (P) rats in the presence of the δ opioid receptor antagonist NTI (44).

Taken together, these preclinical studies in animals and humans support the model that alcohol drinking is reinforcing at least in part because of its effects on enhancing the release of endogenous opioids. The use of opioid antagonists as an effective agent in the treatment of alcoholism is strongly predicted by these preclinical studies.

Pharmacokinetics, Pharmacodynamics, and Safety

Naltrexone, an opioid antagonist, was originally developed for use in the prevention of relapse in detoxified opiate addicts. Naltrexone has a half-life of approximately 4 hours, and 6-ß-naltrexol, its major metabolite, has a half-life of 12 hours. Rapidly absorbed, naltrexone reaches peak plasma levels between 60 and 90 minutes. Naltrexone undergoes first-pass hepatic metabolism, and there is some evidence of dose-related hepatotoxicity at doses four to five times higher than the currently recommended 50-mg daily dosage.

In alcohol-dependent patients, adverse events reported by at least 2% of those participating in an open-label safety study were nausea (10%), headache (8%), dizziness (4%), nervousness (4%), fatigue (4%), insomnia (3%), vomiting (3%), anxiety (2%), and somnolence (2%) (45). In addition to these new-onset adverse events, naltrexone is contraindicated for patients who are currently opioid dependent, are in acute opioid withdrawal, or require opioid analgesics for management of pain, and those with acute hepatitis or liver failure. Special considerations are involved in the management of medical emergencies requiring pain management because naltrexone is an opioid antagonist. Although there has been little formal research on drug–drug interactions, with the exception of opiate-containing medications, subjects on naltrexone who were on concurrent treatment with antidepressant therapy did not experience any increase in adverse events relative to those not on antidepressant therapy in the aforementioned safety trial.

Efficacy

Naltrexone is currently approved for use in the treatment of alcoholism in the United States, Canada, and many European and Asian countries. The efficacy of naltrexone has been tested in several double-blind placebo controlled trials (Table 101.1).

Published Study	No. of Subjects	Therapy	Medication Dose	Duration (Weeks)	Results			
					Craving	TTFD[a]	Relapse[b]	PDD[c]
Volpicelli et al., 1992 (48)	70	Intensive multimodal	Naltrexone 50 mg/day	12	+	0	+	+
Volpicelli et al., 1997 (47)	97	Relapse prevention	Naltrexone 50 mg/day	12	0	0	+	+
O'Malley et al., 1992 (48)	97	Coping skills or supportive	Naltrexone 50 mg/day	12	+/0	0	+	+
Mason et al., 1994 (49)	21	CBT	Nalmefene 40 mg/day	12	0	0	+	+
Mason et al., 1994 (49)	105	CBT	Nalmefene 20 or 80 mg/day	12	0	0	+	0
Oslin et al., 1997 (50)	44		Naltrexone 50 mg/day	12	NR	0	+/0	0
Anton et al., 1999 (51)	131	CBT	Naltrexone 50 mg/day	12	+/0	0	+	+

Plus sign means a significant difference in favor of the medication group.
Minus sign means a significant difference in favor of the placebo group.
A plus/minus sign is a trend in favor of the medication group or a significant difference in a subsample.
[a]Time to first drink or total abstinence.
[b]Relapse refers to time to first episode of hazardous drinking (survival analysis)
[c]Percent drinking days: cumulative days abstinent or percent days abstinent.
CBT, cognitive behavioral treatment; NR, result not reported.
Adapted from Garbutt et al., JAMA Vol. 281, 14:1318–1325.

TABLE 101.1. DOUBLE-BLIND, PLACEBO-CONTROLLED TRIALS OF OPIOID ANTAGONISTS FOR THE TREATMENT OF ALCOHOL DEPENDENCE

In general, these studies have been 12 weeks in duration, with one study (52) reporting on a 6-month follow-up period. Samples have been composed primarily of male subjects (ranging from 71% to 100%) without other complicating psychiatric or substance abuse problems, although there have been smaller studies in specialized populations, including those who use cocaine and alcohol (53) and older alcoholics (50). The behavioral interventions provided in conjunction with naltrexone include day-hospital treatment, cognitive behavioral therapy, and supportive therapy. These studies have tested the efficacy of a 50-mg daily dose against placebo; although several studies in progress are evaluating the utility of higher doses (e.g., up to 100 mg daily). The majority of studies, which have found naltrexone to be superior to placebo in treatment outcomes, have initiated treatment in subjects following a period of abstinence ranging from 5 to 7 days (46 ,47 and 48 ,51). Other ongoing studies are testing whether an opioid antagonist can be effectively used in a treatment sample to help subjects reduce and possibly initiate a period of abstinence or effectively control binge drinking.

The most consistent finding in the studies of alcohol-dependent subjects is that naltrexone decreases the risk of drinking at hazardous levels and the percentage of drinking days. In three studies, this finding was observed in the overall sample (46 ,48 ,51), and an additional investigation found that naltrexone significantly reduced hazardous drinking in a secondary analysis of subjects who were good treatment compliers (47). In contrast, no evidence of efficacy was found in a recent randomized study (54) comparing placebo,

naltrexone 50 mg daily, and nefazodone. In this investigation, naltrexone therapy was associated with a higher incidence of adverse effects, poorer medication compliance, and greater attrition than placebo, leading the authors to suggest that adverse events may limit the effectiveness of naltrexone.

Although the optimal duration of therapy with naltrexone is unknown, efficacy data are available for 12 weeks. A 6-month follow-up study (55) found that subjects who had originally been treated with naltrexone for 12 weeks were less likely to experience a day of heavy drinking during the follow-up period or to meet criteria for a diagnosis of alcohol dependence than subjects treated with placebo during that time period. However, there was evidence that the effects of naltrexone appeared to decline over time, raising the question of whether longer-term therapy may be needed. In this regard, initial evidence supporting the potential value of longer-term naltrexone therapy for some patients has been reported at scientific conferences.

Nalmefene, a newer opioid antagonist that is structurally similar to naltrexone, has also been reported to reduce the risk of relapse to heavy drinking. In a 3-month double-blind pilot study, there was initial evidence of reduced risk of heavy drinking among subjects treated with 40-mg doses of nalmefene compared to 10-mg or 0-mg doses (49). In a larger double-blind study (56) in which patients were randomized to placebo, 20 mg daily or 80 mg daily, lower relapse rates were observed for patients treated with nalmefene (combined across the 20-mg and 80-mg doses).

Since the initial published reports of naltrexone for use in alcoholism, several smaller studies have been conducted evaluating its potential in special populations of alcoholics. For example, the use of naltrexone in treating individuals with comorbid alcohol and cocaine use disorders have included one open-label study using 150 mg of naltrexone per day (57) that showed a positive effect for naltrexone, and one double-blind, placebo-controlled study using 50 mg of naltrexone per day in 64 subjects, which was negative (53). Pending additional research with larger samples at higher doses, naltrexone treatment does not appear indicated for the management of individuals with concurrent cocaine and alcohol use disorders. A small study in older alcohol-dependent men suggests that it may be efficacious (50), and in an open-label trial it was found to be helpful for adolescents (58).

The finding of reduced risk of relapse following a lapse has led to considerable interest and research into possible mechanisms underlying this effect. In the clinical trials, alcohol-dependent subjects retrospectively reported feeling less “high” (59) and lower levels of craving and incentive to continue drinking (60). Fixed alcohol dose administration studies in non–alcohol-dependent subjects suggest that naltrexone may attenuate some of the positive mood altering effects (e.g., stimulation), but not the aversive effects of alcohol (e.g., cognitive impairment, sedation) (61 ,62). High rates of nausea have been noted in alcohol administration studies in non–alcohol-dependent subjects maintained on naltrexone, suggesting that naltrexone may make alcohol more aversive in some subjects, particularly at higher doses of alcohol and naltrexone (63). These fixed alcohol dose administration studies involve rapid consumption of large amounts of alcohol. Interestingly, nausea in interaction with alcohol has not been a common complaint in the clinical trials. This suggests that these alcohol-dependent individuals either may be less vulnerable to nausea or may limit their alcohol intake (both the rate of consumption and the amount consumed) to levels that do not cause nausea in interaction with naltrexone. Direct evidence that naltrexone treatment is associated with reduced speed of drinking and the number of drinks consumed has been obtained using ad libitum drinking paradigms (64). Evidence that naltrexone reduces craving or urge to drink is also accruing from this body of research (63 ,64).

Summary

The evidence suggests that naltrexone 50 mg daily is efficacious in reducing the risk of heavy drinking and in increasing the percentage of days abstinent. Although the hypothesized effect of naltrexone on reduction of craving has been somewhat elusive in the clinical trials, laboratory studies provide support for this hypothesis. Additional studies are under way to test the optimal duration of therapy and the efficacy of alternative doses. Although the side-effect profile of naltrexone is acceptable, efforts to minimize adverse events should be investigated given that these events are associated with reduced compliance with therapy, and compliance has been linked to treatment outcome (65).

Acamprosate

Background

Ethanol has also been shown to alter levels of, and have high affinity for receptors of, two other neurotransmitters, glutamate and GABA. *In vitro* studies indicate that ethanol inhibits function of the glutamatergic *N*-methyl-D-aspartate (NMDA) receptor by inhibiting ion flux through this ionotropic receptor. Both *in vitro* and preclinical *in vivo* studies have also demonstrated that ethanol modulates NMDA-mediated release of other neurotransmitters such as acetylcholine, dopamine, and norepinephrine. Microinjections of glutamate antagonists into the nucleus accumbens of rats not dependent on alcohol has been shown to significantly decrease self-administration of alcohol. In preclinical studies, chronic alcohol administration results in an up-regulation of NMDA receptors, and NMDA antagonists given during withdrawal from alcohol have been shown to suppress withdrawal seizures. Clinical studies indicate that increased cerebrospinal fluid (CSF) levels of glutamate during

ethanol withdrawal may be associated with the development of seizures, and that repeated withdrawals increases the risk of seizures. Similarly, ethanol has also been shown to modulate the GABA system particularly $GABA_A$receptor function. Chronic administration of ethanol results in decreases in the messenger RNA (mRNA) and protein for the α subunit of the $GABA_A$receptor. GABA levels are also found to be reduced in the brain and CSF of recently detoxified alcoholics. Moreover, drugs that modulate $GABA_A$ receptor function such as benzodiazepines, barbiturates, and anticonvulsants have been shown to suppress the symptoms of ethanol withdrawal. Given the above evidence, there has been increased interest in examining the effects of agents that alter gluatamate and GABA function on alcohol drinking.

Acamprosate (calcium acetyl-homotaurine), a homotaurine derivative is a structural analogue of GABA and an upper homologue of taurine. It displays high binding capacity with GABA receptors, and it also shows functional activity in direct and indirect tests of GABA activity (66). Studies suggest it also inhibits NMDA receptors and reduces glutamate concentrations, particularly in the nucleus accumbens (67 ,68). A number of preclinical studies have shown that acamprosate produces dose-dependent decreases in alcohol consumption, with complete suppression of drinking seen at a dose of 400 mg/kg. It has also been shown that acamprosate diminishes reinstatement of alcohol drinking in the alcohol-dependent rat.

Pharmacodynamics, Pharmacokinetics, and Safety

Acamprosate has low bioavailability (10%), is not metabolized by the liver, and is primarily excreted through the kidney, with an excretion half-life of 18 hours (69 ,70). Acamprosate has an excellent safety profile. The most common adverse effect distinguishing acamprosate from placebo is diarrhea; other reported side effects that may be associated with acamprosate are rash and changes in libido. Drug interaction studies indicate that acamprosate does not interact with a variety of medications prescribed to individuals with alcohol dependence (e.g., antidepressants, anxiolytics, disulfiram, hypnotics, or neuroleptics) (69).

Efficacy

Acamprosate is approved for use as a treatment for alcohol dependence in most European countries and in many Latin American countries as well as in Australia, South Africa, and Hong Kong. The efficacy of acamprosate has been evaluated in over ten published placebo-controlled trials ranging from 3 to 12 months, with follow-up periods ranging from 0 to 12 months following the discontinuation of therapy (Table 101.2).

TABLE 101.2. DOUBLE-BLIND, PLACEBO-CONTROLLED TRIALS OF ACAMPROSATE FOR THE TREATMENT OF ALCOHOL DEPENDENCE

Published Study	No. of Subjects	Duration (Weeks)	Results			
			Craving	TTFD[a]	Relapse[b]	PDD[c]
Lhuintre et al., 1990 (71)	569	12	NR[d]	NR[d]	NR[d]	NR[d]
Pelc et al., 1992 (72)	102	24	0	+	NR	+
Ladewig et al., 1993 (73)	61	24	NR	0/+	NR	+
Paille et al., 1995 (74)	538	52	0/+	+	NR	+
Sass et al., 1996 (75)	272	48	NR	+	NR	+
Whitworth et al., 1996 (76)	448	52	NR	+	NR	+
Roussaux et al., 1996 (77)	127	52	0	0	NR	NR
Geerlings et al., 1997 (78)	262	24	NR	+	NR	+
Barrias et al., 1997 (79)	302	52	NR	+	NR	+
Pelc et al., 1997 (80)	188	12	+	+	NR	+
Poldrugo, 1997 (81)	246	26	0	+	NR	+

Plus sign means a significant difference in favor of the medication group.
Minus sign means a significant difference in favor of the placebo group.
A plus/minus sign is a trend in favor of the medication group or a significant difference in a subsample.
[a]Time to first drink or total abstinence.
[b]Relapse refers to time to first episode of hazardous drinking (survival analysis).
[c]Percent drinking days: cumulative days abstinent or percent days abstinent.
[d]The primary outcome for this study was gamma-glutamyl transpeptidase (GGT) and was significantly lower in acamprosate compared to placebo.
NR, result not reported.
Adapted from Garbutt et al., JAMA Vol. 281, 14:1318–1325.

The treatment period generally began following completion of inpatient detoxification. With respect to dose, earlier studies typically adjusted the dose of acamprosate for body weight, whereas more recent studies have used a fixed dose of 1,998 mg/day, with two 333-mg tablets given three times per day (six tablets per day). The nature of the concurrent behavioral interventions was not specified and typically was that used by a particular site. The primary outcome measures included retention in treatment and measures of abstinence, such as rate of abstinence preceding study visits, continuous abstinence (i.e., completing the study without having a drink), or a measure of cumulative abstinence duration (e.g., total number of days abstinent or percentage of days abstinent during treatment the study). Information about the actual quantity of alcohol consumed on a nonabstinent day was rarely reported.

Summarizing across the studies in Table 101.2 , the majority of studies find an advantage of acamprosate over placebo on measures of total abstinence, time to first drink, and/or cumulative abstinence duration (71 ,72 ,73 and 74 ,77 ,78 ,79 ,80 and 81). For example, in an early 12-week trial, Lhuintre et al. (71) found that abstinence rates for patients treated with acamprosate were nearly double (61%) that of patients treated with placebo (32%). Paille et al. (74) demonstrated that the effects of acamprosate on measures of abstinence were dose dependent. Specifically, point prevalence measures of abstinence at 6 months were 18.6% in the placebo group, 27.7% in the 1.3-g/day condition, and 34.7% in the 2-g/day condition. Similar dose effects were found on retention in treatment. In a study of 272 severely dependent alcoholics who had been abstinent 14 to 28 days prior to acamprosate treatment, 43% of the acamprosate-treated patients were continuously abstinent compared to 21% of those who received placebo over the course of 48 weeks (75). Although overall abstinence rates were lower in a different sample of severely dependent alcoholics with only 5 days of abstinence pretreatment (76), differences in abstinence rates were found favoring acamprosate over placebo during the 360-day treatment period. The advantage of acamprosate over placebo continued once acamprosate was discontinued after 6 and 12 months of active treatment.

At this time, there is no information available about the optimal duration of treatment with acamprosate. Although the duration of treatment has varied across studies (e.g., 3 to 12 months), there are no studies that have examined the effect of treatment periods of different lengths or the value of continued acamprosate in treatment responders. Given that differences between acamprosate and placebo appear to emerge after 2 to 3 months of treatment and generally persist after treatment is discontinued, studies addressing the potential value of short- versus long-term treatment are warranted to guide clinical practice.

Whether results comparable to those obtained in European

studies will be obtained in the United States is of great interest. A 21-site, 6-month, double-blind, placebo-controlled trial has recently been conducted to determine safety and efficacy of acamprosate in 601 U.S. alcoholdependent patients (82). This study tested the efficacy of a 2-g daily dose against placebo and includes an exploratory 3-g dose, given the absence of rate-limiting side effects. In contrast to European studies in which patients were randomized into study treatments following inpatient detoxification, the U.S. study allowed for randomization at as early as 2 days of abstinence. The results of this study are not published as of this time.

Summary

The evidence suggests that acamprosate can have a positive effect on measures of abstinence from alcohol following inpatient detoxification. These effects appear to be dose dependent, favoring the higher doses of acamprosate that have been tested. Although it is hypothesized that the efficacy of the acamprosate is due to its effects on conditioned withdrawal and withdrawal-related craving (83 ,84 and 85), ratings on analogue scales of craving have not distinguished acamprosate and placebo-treated patients in the clinical trials to date. In addition, the potential effect of acamprosate on alcohol reward and drinking following a lapse in abstinence is not understood at this time, because the majority of studies collected information on abstinence only, and there are no laboratory studies examining the interactions of acamprosate and alcohol. The results of the U.S. trial, however, may provide additional information because daily reports of drinking quantity were obtained.

Serotoninergic Medications

Background

The use of medications that affect the serotonin (5-HT) system was initially anticipated by clinical observations regarding similarities between alcoholism and mood, anxiety, impulse control, and antisocial personality disorders. Given

the presumed relationship between these various disorders and a dysfunction in the serotonin system, this clinical observation led to speculation that alcohol dependence was also related to some serotonin dysfunction. Several lines of preclinical research in animals and social drinkers support the notion that alcohol drinking compensates for some deficiency in serotoninergic activity. Most of these have consistently shown certain precursors to reducing alcohol drinking. More specifically, studies conducted in animals selectively bred for high alcohol drinking (HAD) or low alcohol drinking (LAD) behavior indicate that tissue content of serotonin and its metabolite 5-hydroxyindoleacetic acid are substantially lower in certain brain regions of the alcohol-preferring (P) animals compared with NP and in the HAD compared with the LAD rats (86). Smith and Weiss (87) have recently shown that ethanol-naive P rats have higher basal levels 5-HT release compared with NP rats, whereas chronic alcohol treated P rats had decreased extracellular levels of 5-HT in comparison to NP rats. However, although acute administration of alcohol results in increased levels of serotonin in the brain and periphery of alcohol-naive animals, this release is not altered by a genetic predisposition toward high alcohol drinking (88).

The evidence on densities of serotonin receptors in rats with a genetic predisposition to alcohol drinking is controversial. Alcohol-preferring (P) rats have higher 5-HT_{1A} binding and lower 5-HT_{1B} and 5-HT_2 binding in several brain regions when compared with NP rats (89). In contrast, the replicate HAD and LAD lines do not display the same differences in receptor densities, and in the alcohol-drinking fawn-hooded rats the densities of 5-HT_{1A} receptors were lower and those of the 5-HT_2 receptors higher compared to that of LAD Wistar rats (see ref. 90 for review). Preclinical studies indicate that 5-HT_{1A} agonists and serotonin reuptake inhibitors reduce ethanol intake in P and HAD rats as well as in unselected rat lines (86 ,91). In contrast, the role of the 5-HT_2 and 5-HT_3 receptor systems in alcohol drinking behavior is controversial (see ref. 90 for review). Although some studies indicate that ethanol drinking is reduced by both 5-HT_2 receptor agonists and antagonists, other investigators report no effects with antagonists of 5-HT_2 receptors. Similarly, the role of the 5-HT_3 receptor system in mediating ethanol drinking is also controversial, with reductions in drinking seen in paradigms using continuous access to alcohol, but little efficacy being observed in paradigms using limited access to alcohol. In contrast, studies using serotonin uptake inhibitors such as fluoxetine reported robust decreases in alcohol drinking in the P rats (86 ,92).

Pharmacodynamics, Pharmacokinetics, and Safety

There are currently five Food and Drug Administration (FDA) approved selective serotonin reuptake inhibitors (SSRIs) available today: fluoxetine (Prozac), fluvoxamine (Luvox), paroxetine (Paxil), sertraline (Zoloft), and citalopram (Celexa). SSRIs have in common the ability to block the reuptake of serotonin, and this functionally enhances serotoninergic activity. Fluoxetine is characterized by a long plasma half-life with a range of 1 to 4 days and its active metabolite norfluoxetine has a half-life of up to 2 weeks. In contrast, the half-life of the other SSRIs varies between 21 hours for paroxetine and 36 hours for citalopram (93). The long half-life for fluoxetine offers some pharmacokinetic advantage for patients who are less compliant with taking their medications, and fluoxetine has less of a discontinuation syndrome compared to the shorter duration SSRIs paroxetine and sertraline (94). SSRIs are inhibitors of cytochrome P-450 isoenzymes, with paroxetine an especially strong inhibitor of the P-450-2D6 isoenzyme, whereas fluvoxamine is an especially potent inhibitor of P-450-1A2. Thus there are important drug–drug interactions when SSRIs are combined with medications that are metabolized by the P-450 system. Despite their common mechanism of action, there are important pharmacokinetic and pharmacodynamic differences. Despite their name, SSRIs are not completely selective in affecting just serotonin reuptake. For example, sertraline and to a lesser extent fluoxetine are relatively potent dopamine reuptake inhibitors, and the various SSRIs can also block the reuptake of norepinephine (95). In addition, the SSRIs also antagonize muscarinic and histaminic receptors leading to anticholinergic and sedative side effects. Of the most disturbing side effects to SSRIs, initial nausea and sexual dysfunction are the most common.

Efficacy

None of the SSRIs is currently approved for the treatment of alcoholism. The results of several placebo-controlled double-blind studies using SSRIs for the treatment of alcohol dependence have led to conflicting results. In an Italian study with 81 subjects randomized to placebo, fluvoxamine, or citalopram, both of the SSRI groups showed a higher incidence of continuous abstinence compared to the placebo group (96). Similarly, in a Finnish study of 62 randomized subjects, citalopram was more effective then placebo in alcohol drinking outcomes (97). These studies are not consistent with two American trials. For example, in a 12-week trial using fluoxetine in a general sample of alcohol-dependent subjects, there were no overall differences between the medication and placebo groups (98). At doses of up to 60 mg per day in a group of 101 subjects who also received weekly sessions of relapse prevention therapy, fluoxetine did not reduce any measure of alcohol drinking.

Although the overall results of SSRIs for alcoholism treatment are generally negative, there may be subtypes of patients who benefit from treatment with SSRIs and other serotoninergic medications (Table 101.3). For example, in a study of 51 alcoholics with severe comorbid major depression,

subjects randomized to fluoxetine experienced less depression and less alcohol drinking than placebo-treated subjects (99). At 1-year follow-up the results for both depression and alcohol continued to favor the fluoxetine group (100).

Study	No. of Subjects	Alcohol/Subtype	Medication	Duration (Weeks)	Results			
					Craving	TTFD[a]	Relapse[b]	PDD[c]
Janiri et al., 1996 (101)	50	AD	Fluoxetine	8	NR	NR	NR	+
Tiihonen et al., 1996 (97)	62	AD	Citralopram	12	NR	NR	NR	NR
Cornelius et al., 1997 (99)	51	MD/AD	Fluoxetine	12	NR	NR	NR	+
Kranzler et al., 1996 (102)	60	AD/type A	Fluoxetine	14	NR	NR	0	0
	35	AD/type B			NR	NR	0	–
Pettinatti et al., 2000 (103)	55	AD/type A	Sertraline	14	NR	NR	0	+
	45	AD/type B			NR	NR	0	0
Malec et al., 1996 (104)	57	AD	Buspirone	12	0	NR	NR	0
Fawcett et al., 2000 (105)	156	AD	Buspirone	24	NR	0	NR	0
Malcolm et al., 1992 (106)	67	GAD/AD	Buspirone	24	NR	0	0	0
Tollefson et al., 1992 (107)	51	GAD/AD	Buspirone		+/0	NR	NR	NR
Kranzler, 1994 (108)	61	GAD/AD	Buspirone	12	NR	NR	+	+
Wiesbeck, 1999 (109)	493	AD	Ritanserin	24	0	0	0	0
Sellers et al., 1994 (110)	71	AD (mild)	Ondansetron	6	NR	NR	NR	+/0
Johnson et al., 2000 (111)	161	AD/early onset	Ondansetron	11	NR	NR	NR	+
	160	AD/late onset			NR	NR	NR	0

Plus sign means a significant difference in favor of the medication group.
Minus sign means a significant difference in favor of the placebo group.
A plus/minus sign is a trend in favor of the medication group or a significant difference in a subsample.
[a]Time to first drink or total abstinence.
[b]Relapse refers to time to first episode of hazardous drinking (survival analysis).
[c]Percent drinking days: cumulative days abstinent or percent days abstinent.
AD, alcohol dependence; GAD, generalized anxiety disorder; MD, mood disorder; NR, result not reported.
Adapted from Garbutt et al., JAMA Vol. 281, 14:1318–1325.

TABLE 101.3. DOUBLE-BLIND, PLACEBO-CONTROLLED TRIALS OF SEROTONINERGIC AGENTS FOR THE TREATMENT OF ALCOHOL DEPENDENCE

Given that there may be important subgroups of alcoholics who self-medicate with alcohol, Kranzler and colleagues (102) subsequently reanalyzed their data after using a k-cluster technique to identify type A and type B alcoholics. Type B alcoholics are thought to reflect some underlying serotoninergic dysfunction because they tend to be more impulsive, have more emotional distress, and have increased severity of alcohol dependence. Contrary to the prior predictions, type B alcoholics drank more when given fluoxetine compared to placebo subjects (102). There were no medication differences in type A alcoholics. Similarly, Pettinati and colleagues (103) found that in a 14-week placebo-controlled trial of sertraline (200 mg per day), there was no main effect of sertraline on any alcohol drinking measure, but a significant alcoholic subtype by medication interaction. Subjects with presumed serotoninergic dysfunctions (type B) treated with sertraline tended to drink more than placebo-treated subjects, whereas the less severe type A subgroup of alcoholics showed a favorable response to sertraline in several drinking measures.

Other Serotoninergic Medications (Buspirone, Ritanserin, Ondansetron)

There are a variety of other medications that affect the serotonin system but work through different mechanisms than the reuptake inhibitors. The results are mixed and suggest that these medications may be effective only for certain subtypes of alcoholics.

Buspirone

The results of buspirone, a serotonin 1A partial agonist, on alcohol drinking are mixed and may depend on the subgroup of alcoholics studied (104 ,105 ,106 ,107 and 108). As Table 101.3 shows, in a general sample of alcoholics buspirone shows little evidence of clinical efficacy. In anxious alcoholics,

however, studies consistently show that buspirone reduces anxiety in subjects and may reduce alcohol drinking and, in one study, days in which alcohol was craved (107).

Ritanserin

In a large multicentered placebo-controlled trial, 493 subjects were randomized to receive placebo or one of three doses of ritanserin, a 5-HT_2 receptor antagonist, with a treatment duration of 6 months. The results of the study showed no differences between the placebo group and any of the three medication groups (109).

Ondansetron

More encouraging are the results of the 5-HT_3 antagonist ondansetron. A 6-week placebo-controlled study of 71 patients, found that 0.5 mg/d but not 4 mg/d of ondansetron reduced alcohol intake, although this effect was at the level of a trend (p = .06) (110). Johnson and colleagues (111) found that among early-onset alcoholics (onset prior to age 25), 4 µg/kg reduced the intensity of drinking compared to placebo. Among subjects receiving ondansetron, the percent of days abstinent was about 70% compared to 50% for those subjects treated with placebo (111). There were no differences between the medication and placebo groups for late-onset alcoholics.

Summary

Although there are suggestions that serotoninergic mechanisms are involved in excessive drinking, the results of using serotoninergic medications for alcoholism treatment are inconsistent. An understanding of which patients may be helped by which serotoninergic medications is complicated by the heterogeneous nature of alcohol-dependent patients and the subtle pharmacokinetic and pharmacodynamic differences among serotoninergic medications. The finding that some subtypes of alcoholics may do worse while taking serotoninergic medications is of considerable clinical interest, because many alcoholic-dependent patients may be taking a serotoninergic medication to treat a comorbid psychopathology. Given the widespread use of SSRIs and other serotoninergic medications, it is likely that there are more alcoholics patients taking serotoninergic medications than those taking all the medications specifically approved for the treatment of alcohol dependence combined. Rather than improve their drinking status, use of these medications may be interfering with alcohol recovery in some patients.

Tricyclic Antidepressants (TCAs)

The tricyclic antidepressants (e.g., impramine, desipramine, amitriptyline) represent a rather large class of medications that have been used for several decades to successfully treat mood and anxiety disorders. This class of medications, like the SSRIs, blocks the reuptake of serotonin but are far less specific in their actions. They also block the reuptake of norepinephine and dopamine, and antagonize muscarinic and histaminic receptors to varying degrees. The effect of TCAs on antagonizing muscarinic receptors and histaminic receptors give TCAs substantial anticholinergic and sedative effects. These anticholinergic effects include dry mouth, constipation, and tachycardia. The antihistamine effects include drowsiness and sedation. The TCAs are metabolized in the liver by the cytochrome P-450-2D6. Thus TCAs can interact with medications that are also metabolized by the P-450 system. Of note, alcohol can induce liver enzyme activity and reduce plasma TCA levels.

Efficacy

Studies using TCAs for the treatment of comorbid depression and alcohol dependence have generally shown that TCAs effectively reduce symptoms of depression but have little effect of alcohol drinking. For example, Mason and colleagues (112) tested the effectiveness of desipramine in a double-blind, placebo-controlled trial of 71 alcohol-dependent subjects with (28 subjects) or without (41 subjects) concurrent symptoms of depression (112). Overall, desipramine did not reduce alcohol drinking, but it was effective in reducing depression scores in those subjects with coexisting depression. Similarly, McGrath et al. (113) studied a group of 69 subjects with a history of depression that either predated or occurred independently of alcohol abuse. In this double-blind, placebo-controlled study, imipramine combined with relapse prevention therapy was effective in improving depression but had little effect on alcohol drinking. Among subjects who showed a good clinical response on depressive symptoms, there was evidence that imipramine was associated with greater reductions in alcohol drinking compared to placebo. In summary, the results of these small-scale studies provide suggestive evidence that there is a subgroup of patients with coexisting depression who may benefit from TCAs.

Lithium

The use of lithium for the treatment of alcoholism was suggested on the basis of clinical observations that many patients with mood disorders, particularly bipolar disorder, report alcohol use as a way to control mood instability. Early small-scale trials of lithium in the treatment of alcoholics were encouraging (114). For example, there were some data that among patients who received therapeutic levels of lithium, there were improved treatment outcomes (115). However, in a large multicenter placebo-controlled trial with 457 male alcoholics involving both depressed and nondepressed alcoholics, there were no significant improvements in alcohol drinking outcomes overall or in the depressed subgroup (116).

Similarly, in a recent double-blind, placebo-controlled study there were no significant reductions in alcohol drinking for a general population of alcoholics (105). Based on these larger, well-designed studies, the use of lithium to treat alcoholism does not receive empirical support. Its use in controlling bipolar symptoms may still be important for those with coexisting bipolar disorder and alcoholism.

Combination Therapy

Research on rational combinations of medications to treat alcoholism is an area that is rapidly developing. Given that the acute and chronic effects of alcohol involve a number of neurotransmitter systems, a therapeutic approach targeting more than one system may be more effective than monotherapy. In addition, medications may be combined to target distinct aspects of the process of relapse (craving, abstinence, and/or relapse following an initial lapse in abstinence) in order to help a larger number of individuals with alcohol dependence. Finally, combination therapy with efficacious agents may permit the use of lower doses of one or both medications, thereby potentially improving tolerability and compliance with treatment and maximizing treatment outcome.

A number of preclinical studies using rodent models have examined the effect of combining naltrexone and other agents thought to alter alcohol intake, including fluoxetine (117 ,118 and 119), a thyrotropin-releasing hormone analogue TA-0910 (120), the calcium channel blocker isradipine (121), the 5-HT_3 antagonist ondansetron (122), and the 5-HT_{1A} antagonist WA-100635 (123). The majority, but not all (121), of these studies have found at least an additive effect of combining naltrexone with these agents. Whether or not similar effects will be obtained in human subjects is under investigation for the combination of naltrexone and ondansetron (124) and the SSRI sertraline, with very small preliminary reports suggesting some optimism for continuing to investigate these approaches to combination therapy (124 ,125).

The possibility that disulfiram can be used to augment the efficacy of acamprosate has been evaluated in secondary analyses of a double-blind, placebo-controlled study (126). In this study, 118 Swedish subjects were randomized to either acamprosate or placebo, and disulfiram use was permitted on a voluntary basis. Comparisons of subjects who took disulfiram in combination with either acamprosate (*n* = 24) or placebo (*n* = 22) and those who received acamprosate or placebo only indicated that combined use of acamprosate and disulfiram was associated with the highest number of continuous abstinent days compared to the other three groups. These findings are of interest; however, they must be interpreted cautiously because subjects taking disulfiram were self-selected, differed from those who did not use disulfiram on a number of measures, and had much more frequent contact with the treatment program due to the fact that disulfiram administration was supervised.

There is considerable interest in the potential effect of combining acamprosate and naltrexone for the treatment of alcohol dependence. These agents target different neurobiological systems altered by alcohol drinking and dependence, and have been found to influence different aspects of the relapse process. Acamprosate has been shown to have its primary effect on measures of abstinence, whereas naltrexone is most noteworthy for its effect of reducing the risk of relapse following a lapse in abstinence. Finally, these two medications are eliminated through different pathways (hepatic metabolism for naltrexone and excretion for acamprosate). Preliminary data supporting the safety of this combination derived from laboratory studies of normal volunteers (56) and alcohol-dependent subjects (124). A large-scale multisite evaluation of the efficacy of these two medications alone and in combination when provided with behavioral interventions of different intensities is planned (127).

PHARMACOLOGY AND INTERACTION WITH BEHAVIORAL INTERVENTIONS

Part of "101 - Alcoholism Pharmacotherapy "

Psychosocial Treatment Approaches

Medications for the treatment of alcoholism are generally given in the context of psychosocial treatment. There are a variety of psychosocial approaches to alcoholism treatment and little evidence that one type of treatment is superior to others. Project MATCH (Matching Alcoholism Treatments to Client Heterogeneity) provides the clearest presentation of our state-of-the-art psychosocial treatments (4). In this large multicentered study, over 1,700 subjects were randomly assigned to motivational enhancement treatment (MET), cognitive behavioral treatment (CBT), or twelve-step facilitation (TSF). The results clearly demonstrate that subjects presenting for treatment and receiving some type of psychosocial intervention general reduce their alcohol consumption. For example, alcohol was consumed on about 75% of the days prior to starting treatment and then with treatment was consumed on less than 25% of days. These were few differences between the three psychosocial conditions and limited evidence that one type of treatment was better for a particular type of patient (4).

Inspection of the various double-blind studies on the effectiveness of medications to reduce alcohol drinking generally shows that the psychosocial interventions alone have a dramatic effect in reducing alcohol drinking. For example, as shown in the project MATCH data, it is common for baseline drinking to occur on the average of 60% to 75% of days prior to starting the study and for placebo subjects to reduce drinking to less the 20% of the days (46 ,47 and 48 ,51). The additional benefit of pharmacotherapy is thus an

adjunct to the considerable benefit of participating in a clinical trial that includes a psychosocial intervention.

Special Issues in the Use of Psychopharmacology in the Treatment of Alcoholism

Despite the convincing clinical treatment research results, the clinical use of medications for the treatment of alcoholism lags behind the pharmacotherapy of other psychiatric disorders. With the exception of medications that treat alcohol withdrawal symptoms, the medications presented here do not provide any immediate relief of symptoms. The long-term beneficial effects have been shown in terms of reductions in slips from abstinence or relapses to heavy drinking, but to the individual patient these outcomes are difficult to attribute to the use of a medication. A successful outcome for the medication is the absence of some adverse clinical event that may or may not happen in the future. In contrast, medications used to treat other psychiatric disorders do provide relief of emotional distress even if the relief is delayed by several weeks, as is the case for antidepressants in the treatment of mood disorders. In general, there are no obvious rewarding properties of taking a medication to treat alcoholism. Coupled with the fact that the immediate effect of drinking alcohol is to feel good or reduce some unpleasant feeling, it makes it a challenge to help patients become and remain compliant with taking their prescribed medication.

Given the lack of easily experienced positive effects from taking medications for alcoholism, it is not surprising that medication compliance is an important factor in the efficacy of medications. For example, in a 12-week, double-blind, placebo-controlled trial using naltrexone in conjunction with addiction counseling, with a total sample of 98 randomized subjects, naltrexone had a modest effect in reducing alcohol relapse rates. However, among subjects who took at least 80% of their prescribed medication, the relative effectiveness of naltrexone was much improved, as 52% of placebo subjects relapsed compared to 14% of the naltrexone subjects (47).

Compliance-Enhancement Techniques

To enhance motivation to remain in treatment and comply with taking medication, several behavioral interventions have been implemented. For example, the BRENDA approach, developed at the University of Pennsylvania (128), incorporates various behavioral strategies such as giving people feedback, developing an empathic therapeutic relation, working collaboratively with the patient to develop treatment goals, and continuing to assess treatment adherence.

A comparison of compliance rates of patients treated with the BRENDA approach to historical compliance rates at the Treatment Research Center at the University of Pennsylvania (103) suggests that BRENDA can enhance treatment and medication compliance. A randomized controlled study is currently under way to directly compare the BRENDA approach to cognitive behavioral therapy and simple physician medication management.

Integration of Behavioral and Pharmacotherapies

In a sense all pharmacotherapy studies are combined behavioral and pharmacotherapy studies. The effectiveness of medication is superimposed in a context of behavioral treatment. Thus, all the studies reported here reflect the additional benefits of an active medication group superimposed on the benefit of a behavioral treatment. The behavioral treatment intervention of subjects presenting for treatment has rather large effects as reflected on the improvement seen in the placebo groups in pharmacotherapy trials.

It remains to be determined how the behavioral interventions interact with pharmacotherapy, but there is a potential for additive or even synergistic effects of combining behavioral and pharmacologic treatments. Just as different medications may address different biochemical mechanism to improve treatment outcome, the integration of medications with psychosocial interventions can address different aspects of recovery. As discussed above, behavioral strategies can enhance medication compliance and treatment retention, thus giving the medication a better chance to be effective. Similarly, pharmacotherapy can reduce the chance of relapse to clinically significant drinking and increase the chance the patient will stay in treatment sufficiently to learn new behavioral coping skills. For example, a medication such as naltrexone can act immediately to reduce the severity of a slip and a return to hazardous drinking. When combined with cognitive and behavioral strategies to cope with triggers for relapse, the synergistic effects of the combined approach can be seen when the naltrexone is stopped, as the patient can now rely on learned skills to avoid and cope with a lapse.

CONCLUSION

Part of "101 - Alcoholism Pharmacotherapy "

The past two decades have shown dramatic changes in the understanding of the pharmacology of alcohol. From understanding alcohol’s nonspecific effects on membranes to alcohol’s specific effects on neurotransmitter systems and second messengers, dramatic advances in the field have led to newer more effective treatments. Recent understanding of the pharmacology of alcohol has led to the development of new medications that improve treatment outcome and help show why some people are vulnerable to becoming addicted to alcohol. Of the new medications, the opiate antagonist naltrexone and acamprosate offer the most immediate promise. For specific populations, serotoninergic medications, tricyclic antidepressants, and mood stabilizers offer

hope for treatment. The use of these medications alone or in combinations remains fertile avenues for research. Finally, special challenges are involved with the clinical use of medications for alcoholism treatment. Psychosocial treatments designed to improve motivation to remain in treatment and adhere to the medication regimen are important adjuncts to pharmacologic treatment. The use of the compliance-enhancing techniques can be safely and effectively integrated into primary care models, thus bring addiction treatment to a wide range of health care providers. Ultimately, the intensity and/or nature of the behavioral intervention may interact with the effects of medication to determine the ultimate outcome of treatment. Given dramatic reductions in the availability of intensive treatment, such as inpatient rehabilitation, and the fact that few individuals seek specialized alcoholism treatment, the availability of effective pharmacotherapies should extend the range of patients who can be successfully managed with less intensive behavioral interventions and increase the probability that individuals with alcohol dependence are identified in primary care settings and offered treatment.

REFERENCES

1. Grant BF. Prevalence and correlates of alcohol use and DSM-IV alcohol dependence in the United States: results of the National Longitudinal Alcohol Epidemiologic Survey. *J Stud Alcohol* 1997;58(5):464–473.

2. Rice DP. Economic costs of substance abuse, 1995. *Proc Assoc Am Physicians* 1999;111(2):119–125.

3. McGinnis JM, Foege WH. Mortality and morbidity attributable to use of addictive substances in the United States. *Proc Assoc Am Physicians* 1999;111(2):109–118.

4. Group PMR. Matching Alcoholism Treatments to Client Heterogeneity: Project MATCH posttreatment drinking outcomes [see comments]. *J Stud Alcohol* 1997;58(1):7–29.

5. Hayashida M, et al. Comparative effectiveness and costs of inpatient and outpatient detoxification of patients with mild-to-moderate alcohol withdrawal syndrome [see comments]. *N Engl J Med* 1989;320(6):358–365.

6. Sausser GJ, Fishburne SB Jr, Everett VD. Outpatient detoxification of the alcoholic. *J Fam Pract* 1982;14(5):863–867.

7. Holbrook AM, et al. Meta-analysis of benzodiazepine use in the treatment of acute alcohol withdrawal. *CMAJ* 1999;160(5):649–655.

8. Schaffer A, Naranjo CA. Recommended drug treatment strategies for the alcoholic patient. *Drugs* 1998;56(4):571–585.

9. Reoux JP, Miller K. Routine hospital alcohol detoxification practice compared to symptom triggered management with an Objective Withdrawal Scale (CIWA-Ar) [In Process Citation]. *Am J Addict* 2000;9(2):135–144.

10. Rosenthal RN, et al. A pilot open randomized trial of valproate and phenobarbital in the treatment of acute alcohol withdrawal. *Am J Addict* 1998;7(3):189–197.

11. Thome J, Wiesbeck GA, Vince GH. [Carbamazepine in treatment of alcohol withdrawal syndrome—an overview of current research]. *Fortschr Neurol Psychiatr* 1994;62(4):125–133.

12. Wolanczyk T, Wojnar M, Cedro A. [Carbamazepine in the treatment of alcohol withdrawal]. *Psychiatr Pol* 1997;31(2):189–196.

13. Fuller RK, et al. Disulfiram treatment of alcoholism. A Veterans Administration cooperative study. *JAMA* 1986;256(11):1449–1455.

14. Chick J, et al. Disulfiram treatment of alcoholism. *Br J Psychiatry* 1992;161:84–89.

15. Wilson A, et al. Disulfiram implantation: a dose response trial. *J Clin Psychiatry* 1984;45(6):242–247.

16. Johnsen J, Morland J. Disulfiram implant: a double-blind placebo controlled follow-up on treatment outcome. *Alcohol Clin Exp Res* 1991;15(3):532–536.

17. Krishnan-Sarin S, et al. The delta opioid receptor antagonist naltrindole attenuates both alcohol and saccharin intake in rats selectively bred for alcohol preference. *Psychopharmacology (Berl)* 1995;120(2):177–185.

18. Krishnan-Sarin S, et al. The delta 2-opioid receptor antagonist naltriben selectively attenuates alcohol intake in rats bred for alcohol preference. *Pharmacol Biochem Behav* 1995;52(1):153–159.

19. Krishnan-Sarin S, et al. Effect of mu opioid receptor blockade on alcohol intake in rats bred for high alcohol drinking. *Pharmacol Biochem Behav* 1998;59(3):627–635.

20. June HL, et al. The delta 2-opioid receptor antagonist naltriben reduces motivated responding for ethanol. Psychopharmacology (Berl) 1999;147(1):81–89.

21. June HL, et al. The opioid receptor antagonist nalmefene reduces responding maintained by ethanol presentation: preclinical studies in ethanol-preferring and outbred Wistar rats. *Alcohol Clin Exp Res* 1998;22(9):2174–2185.

22. Froehlich JC, et al. Naloxone attenuates voluntary ethanol intake in rats selectively bred for high ethanol preference. *Pharmacol Biochem Behav* 1990;35(2):385–390.

23. Reid LD, et al. Periodic naltrexone and propensity to take alcoholic beverage. *Alcohol Clin Exp Res* 1996;20(8):1329–1334.

24. Reid LD, et al. Opioids and intake of alcoholic beverages. *NIDA Res Monogr* 1986;75:359–362.

25. Reid LD, Hunter GA. Morphine and naloxone modulate intake of ethanol. *Alcohol* 1984;1(1):33–37.

26. Ulm RR, Volpicelli JR, Volpicelli LA. Opiates and alcohol self-administration in animals. *J Clin Psychiatry* 1995;56(suppl 7):5–14.

27. Volpicelli JR, Davis MA, Olgin JE. Naltrexone blocks the post-shock increase of ethanol consumption. *Life Sci* 1986;38(9):841–847.

28. Stromberg MF, et al. A comparison of the effects of the opioid antagonists naltrexone, naltrindole, and beta-funaltrexamine on ethanol consumption in the rat. *Alcohol* 1998;15(4):281–289.

29. Stromberg MF, Volpicelli JR, O'Brien CP. Effects of naltrexone administered repeatedly across 30 or 60 days on ethanol consumption using a limited access procedure in the rat. *Alcohol Clin Exp Res* 1998;22(9):2186–2191.

30. Altshuler HL, Phillips PE, Feinhandler DA. Alteration of ethanol self-administration by naltrexone. *Life Sci* 1980;26(9):679–688.

31. Myers RD, Borg S, Mossberg R. Antagonism by naltrexone of voluntary alcohol selection in the chronically drinking macaque monkey. *Alcohol* 1986;3(6):383–388.

32. Myers RD, Lankford MF. Suppression of alcohol preference in high alcohol drinking rats: efficacy of amperozide versus naltrexone. *Neuropsychopharmacology* 1996;14(2):139–149.

33. Hubbell CL, et al. Opioidergic, serotonergic, and dopaminergic manipulations and rats' intake of a sweetened alcoholic beverage. *Alcohol* 1991;8(5):355–367.

34. Gianoulakis C, et al. Different pituitary beta-endorphin and adrenal cortisol response to ethanol in individuals with high and low risk for future development of alcoholism. *Life Sci* 1989;45(12):1097–1109.

35. Thiagarajan AB, Mefford IN, Eskay RL. Single-dose ethanol administration activates the hypothalamic-pituitary- adrenal axis: exploration of the mechanism of action. *Neuroendocrinology* 1989;50(4):427–432.

36. Froehlich JC, et al. Analysis of heritability of hormonal responses to alcohol in twins: beta-endorphin as a potential biomarker of genetic risk for alcoholism. *Alcohol Clin Exp Res* 2000;24(3):265–277.

37. de Waele JP, Kiianmaa K, Gianoulakis C. Distribution of the mu and delta opioid binding sites in the brain of the alcohol-preferring AA and alcohol-avoiding ANA lines of rats. *J Pharmacol Exp Ther* 1995;275(1):518–527.

38. de Waele JP, Gianoulakis C. Characterization of the mu and delta opioid receptors in the brain of the C57BL/6 and DBA/2 mice, selected for their differences in voluntary ethanol consumption. *Alcohol Clin Exp Res* 1997;21(4):754–762.

39. Grisel JE, et al. Ethanol oral self-administration is increased in mutant mice with decreased beta-endorphin expression. *Brain Res* 1999;835(1):62–67.

40. Badia-Elder NE, et al. Naloxone retards the expression of a genetic predisposition toward alcohol drinking. *Psychopharmacology (Berl)* 1999;144(3):205–212.

41. Froehlich JC, et al. Importance of delta opioid receptors in maintaining high alcohol drinking. *Psychopharmacology* 1991;103(4):467–472.

42. Hubbell CL, et al. Consumption of ethanol solution is potentiated by morphine and attenuated by naloxone persistently across repeated daily administrations. *Alcohol* 1986;3(1):39–54.

43. Marfaing-Jallat P, Miceli D, Le Magnen J. Decrease in ethanol consumption by naloxone in naive and dependent rats. *Pharmacol Biochem Behav* 1983;18(suppl 1):537–539.

44. Froehlich JC, et al. Contribution of the opioid system to alcohol aversion and alcohol drinking behavior. *J Pharmacol Exp Ther* 1998;287(1):284–292.

45. Croop RS, Faulkner EB, Labriola DF. The safety profile of naltrexone in the treatment of alcoholism. Results from a multicenter usage study. The Naltrexone Usage Study Group. *Arch Gen Psychiatry* 1997;54(12):1130–1135.

46. Volpicelli JR, et al. Naltrexone in the treatment of alcohol dependence. *Arch Gen Psychiatry* 1992;49(11):876–880.

47. Volpicelli JR, et al. Naltrexone and alcohol dependence: role of subject compliance. *Arch Gen Psychiatry* 1997;54:737–742.

48. O'Malley SS, et al. Naltrexone and coping skills therapy for alcohol dependence. A controlled study. *Arch Gen Psychiatry* 1992;49(11):881–887.

49. Mason BJ, et al. A double-blind, placebo-controlled pilot study to evaluate the efficacy and safety of oral nalmefene HCl for alcohol dependence. *Alcohol Clin Exp Res* 1994;18(5):1162–1167.

50. Oslin D, et al. Naltrexone as an adjunctive treatment for older patients with alcohol dependence. *Am J Geriatr Psychiatry* 1997;5(4):324–332.

51. Anton RF, et al. Naltrexone and cognitive behavioral therapy for the treatment of outpatient alcoholics: results of a placebo-controlled trial. *Am J Psychiatry* 1999;156(11):1758–1764.

52. O'Malley S, et al. Six-month follow-up of naltrexone and psychotherapy for alcohol dependence. *Arch Gen Psychiatry* 1996;53:217–224.

53. Hersh D, Van Kirk JR, Kranzler HR. Naltrexone treatment of comorbid alcohol and cocaine use disorders. *Psychopharmacology (Berl)* 1998;139(1 and 2):44–52.

54. Kranzler HR, et al. Naltrexone vs. Nefazodone for treatment of alcohol dependence. A placebo-controlled trial. *Neuropsychopharmacology* 2000;22(5):493–503.

55. O'Malley S, et al. Naltrexone in the treatment of alcohol dependence: a combined analysis of two trials. *Psychiatr Ann* 1995;25:681–688.

56. Mason BJ, et al. A double-blind, placebo-controlled study of oral nalmefene for alcohol dependence. *Arch Gen Psychiatry* 1999;56(8):719–724.

57. Oslin DW, et al. The effects of naltrexone on alcohol and cocaine use in dually addicted patients. *J Subst Abuse Treat* 1999;16(2):163–167.

58. Lifrak PD, et al. Naltrexone for alcoholic adolescents [letter]. *Am J Psychiatry* 1997;154(3):439–441.

59. Volpicelli J, et al. Effect of naltrexone on alcohol "high" in alcoholics. *Am J Psychiatry* 1995;152(4):613–615.

60. O'Malley SS, et al. Experience of a "slip" among alcoholics treated with naltrexone or placebo. *Am J Psychiatry* 1996;153(2):281–283.

61. Swift RM, et al. Naltrexone-induced alterations in human ethanol intoxication. *Am J Psychiatry* 1994;151(10):1463–1467.

62. King AC, et al. Naltrexone biotransformation and incidence of subjective side effects: a preliminary study. *Alcohol Clin Exp Res* 1997;21(5):906–909.

63. McCaul ME, et al. Naltrexone alters subjective and psychomotor responses to alcohol in heavy drinking subjects. *Neuropsychopharmacology* 2000;22(5):480–492.

64. Davidson D, et al. Effects of naltrexone on alcohol self-administration in heavy drinkers. *Alcohol Clin Exp Res* 1999;23(2):195–203.

65. O'Malley S, Krishnan-Sarin S, Farren C, et al. Naltrexone-induced nausea in patients treated for alcohol dependence: clinical predictors and evidence for opioid-mediated effects. *J Clin Psychopharmacol* 2000;20(1):69–76.

66. Daoust M, et al. Acamprosate modulates synaptosomal GABA transmission in chronically alcoholised rats. *Pharmacol Biochem Behav* 1992;41(4):669–674.

67. Berton F, et al. Acamprosate enhances N-methyl-D-apartate receptor-mediated neurotransmission but inhibits presynaptic GABA(B) receptors in nucleus accumbens neurons. *Alcohol Clin Exp Res* 1998;22(1):183–191.

68. Dahchour A, et al. Central effects of acamprosate: part 1. Acamprosate blocks the glutamate increase in the nucleus accumbens microdialysate in ethanol withdrawn rats. *Psychiatry Res* 1998;82(2):107–114.

69. Saivin S, et al. Clinical pharmacokinetics of acamprosate. *Clin Pharmacokinet* 1998;35(5):331–345.

70. Wilde MI, Wagstaff AJ. Acamprosate. A review of its pharmacology and clinical potential in the management of alcohol dependence after detoxification. *Drugs* 1997;53(6):1038–53.

71. Lhuintre JP, et al. Acamprosate appears to decrease alcohol intake in weaned alcoholics. *Alcohol Alcohol* 1990;25(6):613–622.

72. Pelc I, et al. Calcium acetyl homotaurinate for maintaining abstinence in weaned alcoholic patients; a placebo controlled double-blind multi-centre study. In: Naranjo C, Sellers E, eds. *Novel pharmacological interventions for alcoholism*. New York: Springer-Verlag, 1992:348–352.

73. Ladewig D, et al. [Acamprosate—a stabilizing factor in long-term withdrawal of alcoholic patients]. *Ther Umsch* 1993;50(3):182–188.

74. Paille FM, et al. Double-blind randomized multicentre trial of acamprosate in maintaining abstinence from alcohol. *Alcohol Alcohol* 1995;30(2):239–247.

75. Sass H, et al. Relapse prevention by acamprosate. Results from a placebo-controlled study on alcohol dependence [published erratum appears in *Arch Gen Psychiatry* 1996;53(12):1097]. *Arch Gen Psychiatry* 1996;53(8):673–680.

76. Whitworth AB, et al. Comparison of acamprosate and placebo in long-term treatment of alcohol dependence [see comments]. *Lancet* 1996;347(9013):1438–1442.

77. Roussaux JP, Hers D, Ferauge M. [Does acamprosate diminish the appetite for alcohol in weaned alcoholics?]. *J Pharm Belg* 1996;51(2):65–68.

78. Geerlings PJ, Ansoms C, van den Brink W. Acamprosate and prevention of relapse in alcoholics. Eur Addict Res 1997;3:129–137.

79. Barrias JA, et al. Acamprosate: multicenter Portugese efficacy and tolerance evaluation study. Psiquiatr Clin 1997;18:149–160.

80. Pelc I, et al. Efficacy and safety of acamprosate in the treatment of detoxified alcohol-dependent patients. A 90-day placebo-controlled dose-finding study. *Br J Psychiatry* 1997;171:73–77.

81. Poldrugo F. Acamprosate treatment in a long-term community-based alcohol rehabilitation programme. *Addiction* 1997;92(11):1537–1546.

82. Mason BJ, Ownby RL. Acamprosate for the treatment of alcolhol dependence: a review of double-blind, placebo-controlled trials. *CNS Spectrums* 2000;5(2):58–69.

83. Spanagel R, Zieglgansberger W. Anti-craving compounds for ethanol: new pharmacological tools to study addictive processes. *Trends Pharmacol Sci* 1997;18(2):54–59.

84. Kratzer U, Schmidt WJ. The anti-craving drug acamprosate inhibits the conditioned place aversion induced by naloxone-precipitated morphine withdrawal in rats. *Neurosci Lett* 1998;252(1):53–56.

85. Littleton J, Acamprosate in alcohol dependence: how does it work? *Addiction* 1995;90(9):1179–1188.

86. McBride WJ, Chernet E, McKinzie DL, et al. Quantitative autoradiography of mu-opioid receptors in the CNS of alcohol-naive alcohol-preferring P and -nonpreferring NP rats. *Alcohol* 1998;16(4):317–323.

87. Smith AD, Weiss F. Ethanol exposure differentially alters central monamine neurotransmission in alcohol-preferring versus -nonpreferring rats. *J Pharmacol Exp Ther* 1999;288(3):1223–1228.

88. Yoshimoto K, McBride WJ, Lumeng L, et al. Ethanol enhances the release of dopamine and serotonin in the nucleus accumbens of HAD and LAD lines of rats. *Alcohol Clin Exp Res* 1992;16(4):781–785.

89. McBride WJ, Chernet E, Russell RN, et al. Regional CNS densities of monamine receptors in alcohol-naive alcohol-preferring P and -nonpreferring NP rats. *Alcohol* 1997;14(2):141–148.

90. McBride WJ, Li TK. Animal models of alcoholism: neurobiology of high alcohol-drinking behavior in rodents. *Crit Rev Neurobiol* 1998;12(4):339–369.

91. LeMarquand D, Pihl RO, Benkelfat C. Serotonin and alcohol intake, abuse, and dependence: findings of animal studies [review]. *Biol Psychiatry* 1994;36(6):395–421.

92. Murphy JM, Wealler MB, Gatto GJ, et al. Monoamine uptake inhibitors attenuate ethanol intake in alcohol-preferring (P) rats. *Alcohol* 1985;2(2):349–352.

93. Hiemke C, Hartter S. Pharmacokinetics of selective serotonin reuptake inhibitors. *Pharmacol Ther* 2000;85(1):11–28.

94. Rosenbaum JF, et al. Selective serotonin reuptake inhibitor discontinuation syndrome: a randomized clinical trial [see comments]. *Biol Psychiatry* 1998;44(2):77–87.

95. Goodnick PJ, Goldstein BJ.Selective serotonin reuptake inhibitors in affective disorders—I. Basic pharmacology. *J Psychopharmacol* 1998;12(3):S5–20.

96. Angelone SM, et al. Effects of fluvoxamine and citalopram in maintaining abstinence in a sample of Italian detoxified alcoholics. *Alcohol Alcohol* 1998;33(2):151–156.

97. Tiihonen J, et al. Citalopram in the treatment of alcoholism: a double-blind placebo-controlled study. *Pharmacopsychiatry* 1996;29(1):27–29.

98. Kranzler HR, et al. Placebo-controlled trial of fluoxetine as an adjunct to relapse prevention in alcoholics. *Am J Psychiatry* 1995;152(3):391–397.

99. Cornelius JR, et al. Fluoxetine in depressed alcoholics. A double-blind, placebo-controlled trial [see comments]. *Arch Gen Psychiatry* 1997;54(8):700–705.

100. Cornelius JR, et al. Fluoxetine versus placebo in depressed alcoholics: a 1-year follow-up study. *Addict Behav* 2000;25(2):307–310.

101. Janiri L, Gobbi G, Mannelli P, et al. Effect of fluoxetine at antidepressant doses on short-term outcome of detoxified alcoholics. *Int Clin Psychopharmacol* 1996;11:109–117.

102. Kranzler HR, et al. Fluoxetine treatment seems to reduce the beneficial effects of cognitive-behavioral therapy in type B alcoholics. *Alcohol Clin Exp Res* 1996;20(9):1534–1541.

103. Pettinati HM, et al. Improving naltrexone response: an intervention for medical practitioners to enhance medication in alcohol dependent patients. *J Addict Dis* 2000;19:71–83.

104. Malec E, et al. Buspirone in the treatment of alcohol dependence: a placebo-controlled trial. *Alcohol Clin Exp Res* 1996;20(2):307–312.

105. Fawcett J, et al. Pharmacological treatments for alcoholism: revisiting lithium and considering buspirone [In Process Citation]. *Alcohol Clin Exp Res* 2000;24(5):666–674.

106. Malcolm R, et al. A placebo-controlled trial of buspirone in anxious inpatient alcoholics. *Alcohol Clin Exp Res* 1992;16(6):1007–1013.

107. Tollefson GD, Montague-Clouse J, Tollefson SL. Treatment of comorbid generalized anxiety in a recently detoxified alcoholic population with a selective serotonergic drug (buspirone). *J Clin Psychopharmacol* 1992;12(1):19–26.

108. Kranzler HR, et al. Buspirone treatment of anxious alcoholics. A placebo-controlled trial. *Arch Gen Psychiatry* 1994;51(9):720–731.

109. Wiesbeck GA, et al. Ritanserin in relapse prevention in abstinent alcoholics: results from a placebo-controlled double-blind international multicenter trial. Ritanserin in Alcoholism Work Group. *Alcohol Clin Exp Res* 1999;23(2):230–235.

110. Sellers EM, et al. Clinical efficacy of the 5-HT3 antagonist ondansetron in alcohol abuse and dependence. *Alcohol Clin Exp Res* 1994;18(4):879–885.

111. Johnson BA, et al. Ondansetron for reduction of drinking among biologically predisposed alcoholic patients: a randomized controlled trial [In Process Citation]. *JAMA* 2000;284(8):963–971.

112. Mason BJ, et al. A double-blind, placebo-controlled trial of desipramine for primary alcohol dependence stratified on the presence or absence of major depression [see comments]. *JAMA* 1996;275(10):761–767.

113. McGrath PJ, et al. Imipramine treatment of alcoholics with primary depression: a placebo-controlled clinical trial. *Arch Gen Psychiatry* 1996;53(3):232–240.

114. Lejoyeux M, Ades J. [Prescription of lithium in alcoholic patients]. *Presse Med* 1991;20(25):1176–1180.

115. Fawcett J, et al. A double-blind, placebo-controlled trial of lithium carbonate therapy for alcoholism. *Arch Gen Psychiatry* 1987;44(3):248–256.

116. Dorus W, et al. Lithium treatment of depressed and nondepressed alcoholics [see comments]. *JAMA* 1989;262(12):1646–1652.

117. Zink RW, Rohrbach K, Froehlich JC. Naltrexone and fluoxtine act synergistically to decrease alcohol intake. *Alcohol Clin Exp Res* 1997;21(suppl:104A.

118. Gardell LR, et al. Combination of naltrexone and fluoxetine on rats' propensity to take alcoholic beverage. *Alcohol Clin Exp Res* 1997;21(8):1435–1439.

119. Le AD, et al. Effects of naltrexone and fluoxetine on alcohol self-administration and reinstatement of alcohol seeking induced by priming injections of alcohol and exposure to stress. *Neuropsychopharmacology* 1999;21(3):435–444.

120. Rezvani AH, et al. Combination pharmacotherapy: a mixture of small doses of naltrexone, fluoxetine, and a thyrotropin-releasing hormone analogue reduces alcohol intake in three strains of alcohol-preferring rats [In Process Citation]. *Alcohol Alcohol* 2000;35(1):76–83.

121. Gardell LR, et al. Isradipine and naltrexone in combination with isradipine interact with a period of abstinence to reduce rats' intakes of an alcoholic beverage. *Alcohol Clin Exp Res* 1997;21(9):1592–1598.

122. Le AD, Sellers EM. Interaction between opiate and 5-HT3 receptor antagonists in the regulation of alcohol intake. *Alcohol Alcohol Suppl* 1994;2:545–549.

123. Zhou FC, et al. Additive reduction of alcohol drinking by 5-HT lA antagonist WAY 100635 and serotonin uptake blocker fluoxetine in alcohol-preferring P rats. *Alcohol Clin Exp Res* 1998;22:266–269.

124. Johnson BA, Ait-Daoud N, Prihoda TJ. Combining ondansetron and naltrexone effectively treats biologically predisposed alcoholics: from hypotheses to preliminary clinical evidence [In Process Citation]. *Alcohol Clin Exp Res* 2000;24(5):737–742.

125. Farren CK, Catapano D, O'Malley S. Sertraline with naltrexone vs naltrexone alone in the treatment of alcohol dependence. *Alcohol Clin Exp Res* 1997;21(suppl):64A.

126. Besson J, et al. Combined efficacy of acamprosate and disulfiram in the treatment of alcoholism: a controlled study. *Alcohol Clin Exp Res* 1998;22(3):573–579.

127. O'Malley S, Mattson M. Design and rationale for COMBINE, a multi-site study on combining medications and behavioral intervention for alcohol dependence. Presented at the Research Society on Alcoholism 23rd Annual Scientific Meeting, Denver, CO, June 24–29, 2000.

128. Volpicelli JR, et al. Combining medication and psychosocial treatments for addictions: the BRENDA approach. New York: Guilford, 2001.

102

Pathophysiology and Treatment of Cocaine Dependence

Thomas R. Kosten

Thomas R. Kosten: Department of Psychiatry, Yale University School of Medicine, VA Connecticut Healthcare System, West Haven, Connecticut.

PHARMACOTHERAPY TARGETS IN STIMULANT DEPENDENCE

Part of "102 - Pathophysiology and Treatment of Cocaine Dependence "

Pharmacotherapy can help to initiate abstinence and prevent relapse among the estimated 2 million stimulant-dependent users (1). These 2 million dependent users include some of the residual of long-term users from the peak of this epidemic 15 years ago, but a steady stream of new users and casualties is also accumulating. Between 1991 and 1998 the 30-day prevalence of cocaine abuse among eighth, tenth, and twelfth graders had increased more than twofold (1). Casualties from stimulant use also continue to accumulate, cocaine involvement in emergency room accident and violence cases remains prominent, and recent National Institute of Justice figures show that male and female arrestees in major cities display 40% to 80% cocaine-positive urines (2). These emergency room episodes have remained stable after a 78% increase from 1991 to 1994. Now localized epidemics of amphetamine abuse are developing, particularly in the western United States. The dangers associated with stimulant use are enormous and include increased risk of HIV infection, possible detrimental effects on the unborn and newborn, increased crime and violence, as well as medical, financial, and psychological problems. Because of these consequences, the task of identifying, characterizing, and developing treatments is more important than ever.

Models of Treatment: Neurobiological

To initiate abstinence among stimulant abusers, pharmacotherapy can be directed toward the abnormalities that stimulant dependence appears to create in normal neurobiology. This medication role in renormalizing these alterations parallels the role of medications for initiating abstinence from alcohol or opioids, where pharmacotherapy can reduce the harm from seizures or widespread physiologic withdrawal symptoms. Although discontinuation of stimulant dependence is not associated with severe medical complications, abstinence initiation does produce symptoms of dysphoria (3). These symptoms can be pharmacotherapy targets, and more broadly the disrupted cognition of stimulant abusers can be targeted to facilitate behavioral and cognitive psychotherapies, which have demonstrated efficacy for these disorders. Thus, in addition to abstinence initiation and relapse prevention to stimulant use, surrogate targets include withdrawal symptoms such as craving and dysphoria as well as cognitive impairment, which can result from disrupted neurobiology. Because depressive symptoms are relatively common among stimulant abusers in the early phases of abstinence, antidepressants for stimulant abusers were one of the first interventions studied in controlled trials. Although these medications have a checkered history of failures and successes, some recent data suggest that the depressed stimulant abuser may benefit from antidepressants (4 ,5). This benefit includes reductions in stimulant abuse as well as the depressive symptoms, and is consistent with the recent findings among depressed alcoholics and older studies in depressed methadone maintained patients (6). However, stimulants may induce a depressive syndrome, and these secondary or drug-induced depressions are less clear targets for pharmacotherapeutic intervention (5 ,6).

A useful concept in treatment of these patients is renormalization of disrupted neurobiology. Abnormalities in neurotransmitter receptors and transporters that have been noted in animal models have been confirmed in human neuroimaging studies of the dopamine neurotransmitter systems (7 ,8). Neuroendocrine challenge studies show functional defects consistent with these neuroimaging findings, and norepinephrine systems that stimulants might also disrupt show parallel pharmacologic-challenge abnormalities such as lowered thresholds for yohimbine induction of panic attacks (9 ,10 and 11). These three neurotransmitter systems show

the direct actions of chronic stimulants, but other neurotransmitter systems are indirectly affected including glutamate, γ-aminobutyric acid (GABA), and κ opioid systems (12). Abnormalities in any of these systems are appropriate targets for pharmacotherapy and have been targeted by clinical trials using a range of available agents that are reviewed below.

Although reversible abnormalities in neurotransmitter systems offer the potential for renormalization with subchronic treatment, maintenance treatment may be essential for irreversible changes or genetic predispositions. Examples of such abnormalities among stimulant abusers have been suggested for the postsynaptic dopamine receptor, which may be irreversibly down-regulated compared to normal individuals and even symptomatically resemble Parkinson's disease. Among candidates for genetic predispositions are polymorphisms in the dopamine transporter (DAT) associated with paranoia such as tandem repeats at the SLC6A3 site (13). Recently, the homozygous 10 tandem repeat, which along with the 9 repeat are the most common variants, has also shown a functional correlate of reduced DAT binding in humans (14). A critical association relevant to cocaine abusers is the up-regulation of the DAT after chronic cocaine in many abusers (7). Those who possess this 10 repeat polymorphism are not likely to have their DAT become up-regulated and therefore be successful candidates for medications that target postsynpatic dopamine targets rather than the dopamine transporter. Thus, patients with identified genetic polymorphisms in the DAT could be given more effectively targeted maintenance pharmacotherapies to prevent relapse.

Abnormalities in cerebral blood flow also appear to be common among stimulant abusers and may contribute to cognitive dysfunction (15 ,16). The basis for these perfusion defects appears to be a combination of platelet abnormalities leading to "sticky" platelets and vasospasm from repeated vasoconstriction induced by repeated stimulant use (18). The pharmacotherapies developing for stroke including antiplatelet agents such as clopidogrel and vasodilators such as the calcium channel blockers hold promise for this condition.

Finally, a way to sustain abstinence might be pharmacologic blockade with antibodies, enzymes, or receptor antagonists. The dopamine receptor antagonists such as haloperidol for the D2 receptor or Schering 39166 for the D1 receptor have not been successful, although some argument has been made that a partial agonist with its antagonism only expressed at higher doses might be effective (19). The other more peripheral approach is to prevent or at least slow the entry of stimulants into the brain using antibodies to the stimulant or augmenting the enzymes responsible for metabolic disposition of the stimulant. Because rapid entry of stimulants into the brain appears essential for their reinforcing properties, a delay in this entry might be as effective as fully preventing entry by retarding the stimulant in the bloodstream (3). Although augmenting cholinesterase activity (the enzyme that metabolizes cocaine to an inactive metabolite) remains to be clinically tested, active immunization has been studied in humans. Consistent with the animal studies, humans produced substantial quantities of antibody to an active immunization, but a reduction in cocaine use among outpatients has yet to be tested, to correlate with the reduced self-administration of cocaine observed in the animal studies (20).

CLINICAL ASPECTS OF STIMULANT USE

Part of "102 - Pathophysiology and Treatment of Cocaine Dependence "

The rewarding effects of cocaine and amphetamine are influenced by the route of administration because some routes (e.g., intravenous administration) produce more immediate onset of euphoria. The euphoria appears to depend on occupancy of the DAT (21). The preferred method of selfadministering cocaine has been snorting and, more recently, smoking. Amphetamines come in a variety of forms (e.g., pill, liquid, or powder form), but are usually taken orally or intravenously. The effects of route of administration and pharmacokinetics was extensively covered in the previous edition of this chapter (2).

Stimulant use may range from low dose to high dose and from infrequent to chronic or binge patterns. Depending on the dosage, pattern, and duration of use, stimulants can produce several drug-induced states that differ in clinical characteristics. Moderate to high doses of stimulants can produce stimulant intoxication that may or may not be pleasant. The intoxicated person may show signs of hyperawareness, hypersexuality, hypervigilance, and psychomotor agitation. Often the symptoms of stimulant-induced intoxication resemble mania. The intoxicated person should be monitored by the medical staff until the symptoms of intoxication diminish. If the intoxication does not return to baseline level within 24 hours, mania may be present and treatment for manic disorder may be required (3).

With increased dosage and duration of administration, stimulants can also produce a state of mental confusion and excitement, known as stimulant delirium. Delirium is associated with becoming disoriented and confused, as well as anxious and fearful. Extreme medical caution is needed when treating delirium because such symptoms may indicate stimulant overdose. For instance, crack cocaine addicts who overdose need careful monitoring for seizures, cardiac arrhythmias, stroke, and pulmonary complications. Overdose management has been reviewed in detail (22), but a syndrome of hyperthermia and agitation might be most safely managed with high doses of benzodiazepines (23).

During high-dose stimulant use, often seen during binge episodes, individuals can experience stimulant-induced psychosis characterized by delusions, paranoid thinking, and stereotyped compulsive behavior. When they are delusional, close clinical monitoring is essential and it may be necessary

to employ short-term treatment with neuroleptics to ameliorate the psychosis. It is more common for amphetamine than cocaine to induce psychosis, perhaps due to the difficulty in maintaining high chronic levels of cocaine in the body. Also, stimulant-induced psychosis in humans may be related to the dose and duration of administration of amphetamine, although cocaine psychosis and paranoia may be related to psychiatric predisposition (24).

Stimulant withdrawal, which occurs following cessation of cocaine or amphetamine use, can produce a wide range of dysphoric symptoms. Following binge use, individuals may initially experience a “crash” period, which is characterized by symptoms of depression, anxiety, agitation, and intense drug craving, although controlled studies have shown minimal withdrawal symptoms (3 ,8).

Treatment of stimulant abuse requires a comprehensive assessment of the patient’s psychological, medical, forensic, and drug use history. Moreover, because information obtained from chemically dependent persons may be incomplete or unreliable, it is important that the patients receive a thorough physical including blood and supervised urine samples for analysis. The clinician needs to be aware that polydrug abuse is common. Patients may ingest large amounts of one or more drugs at potentially lethal doses, and therefore it is important that the physician be aware of the dangers of possible drug combinations, such as cocaine and alcohol or heroin.

Pharmacologic intervention may be necessary during stimulant-induced drug states. For instance, neuroleptics may be useful in controlling stimulant-induced psychosis or delirium, and during withdrawal when depression may set in, antidepressants may be an appropriate choice for treatment medication. Treatment medications can be given on an inpatient or outpatient basis. However, if medications are used for outpatient treatment, it is critical to warn the patient of the potential adverse interactions between cocaine and the prescribed treatment medication. For instance, high blood pressure could result from the release of epinephrine by cocaine combined with the reuptake blockade by the tricyclic (25), although later in the course of treatment tricyclics decrease the sensitivity of the postsynaptic adrenergic receptors. Finally monitoring of treatment is essential using urine toxicologies as well as self-reports. The frequency of monitoring may be as infrequent as weekly, but three times weekly is optimal. There appears to be no advantage to quantitative over simple qualitative results based on typical cutoffs such as 300 ng/mL of benzoylecgonine for cocaine use in routine clinical practice.

HUMAN TESTING PARADIGMS FOR NEW MEDICATIONS

Part of "102 - Pathophysiology and Treatment of Cocaine Dependence "

Human Laboratory

The human laboratory setting in which cocaine or amphetamine is administered to volunteer subjects has been a critical paradigm for testing potential pharmacotherapies for stimulant dependence (12 ,25 ,26). Variations on this paradigm have used visual, tactile, aural, or cognitive cues to induce craving for these abused substances. In both experimental settings, the outcome measures have been subjective responses such as euphoria, unpleasant feelings or craving itself, as well as estimates of how much the drug is worth to the participant (e.g., dollar value). The induction of craving for more cocaine after a small to modest dose of cocaine is called the priming effect, and modulation of this priming effect can be an important role for a treatment medication in reducing relapse (27). This reduction in relapse would occur by preventing a “slip,” that is a single use of cocaine in a patient who wants to remain abstinent, from leading into a full relapse to binge cocaine usage.

To more precisely operationalize this human laboratory model, self-administration has been introduced. In that paradigm the subject can self-administer cocaine or amphetamine repeatedly within a range dictated by medical safety considerations. The subject is offered the alternative of getting cocaine or various other rewards that have a range of monetary values. The subject is thereby asked not only to estimate a worth of the cocaine, but also to actually choose to get that amount or to get the cocaine. In these paradigms the behavior of drug taking can be more clearly approximated and a medication that might block the effects of cocaine could be detected. This blocking effect would presumably lead the subject to prefer the alternative reward over the cocaine after the first test dose, when the medication is present, because the reinforcing effects of cocaine would be reduced (12). Although this model has theortical appeal and has shown the expected subject behaviors with various doses of cocaine, it has yet to be tested with an appropriate medication to judge its sensitivity for blocking agents.

In all of these paradigms, the key outcome of cocaine or amphetamine interactions with potential pharmacotherapies yields not only surrogate efficacy data, but also medical safety data. Cardiovascular measures in particular can be carefully monitored after both acute and subchronic dosing with potential medications. The baseline effects of these treatment medications can be assessed in escalating dose regimes, and then dose–response evaluations using escalating doses of cocaine or amphetamine can be examined. Subjective responses may also be important to assess dysphoric interactions between the medication and the abused drugs. These reactions might help in reducing stimulant abuse, although they might also discourage compliance with the medication. Overall, this is a powerful paradigm for medication development because of its potential to inform the clinical trials process with information about how the outpatients in a trial are likely to respond to the new medication, when a study participant uses a stimulant. Its utility as a rapid screening procedure for eliminating medications from further outpatient testing has yet to be demonstrated, but

this may be a future use of these highly controlled paradigms as we obtain gold standard agents with demonstrated efficacy in outpatient trials.

Neuroimaging

A newer technology for human laboratory assessment of potential medications is neuroimaging of either functional activity or receptor and transporter occupancy (28). Functional activity can involve either cerebral blood flow (CBF) or metabolic activity using fluorodeoxyglucose (FDG). The use of FDG as a medication development strategy has been examined in a recent study of selegiline combined with cocaine administration. In this study, selegiline reduced the euphoria from acute cocaine administration, and positron emission tomography (PET) imaging using FDG showed that the cocaine-induced changes in metabolic activity were blocked by the administration of selegiline (29). This surrogate marker provided an interesting correlate of the attenuation of cocaine's subjective effects, because other outpatient work has suggested that selegiline might reduce cocaine abuse in outpatients. Because similar studies of subjective effects alone have not had corresponding predictive validity for outpatient efficacy, these neuroimaging measures may have promise as a more rapid screening tool for medications.

Another medications development approach using neuroimaging focuses on the CBF defects that have been observed in cocaine abusers and on the neuropsychological deficits that persist even during sustained abstinence (6 ,15 ,17). As reviewed below, these CBF defects may be responsive to pharmacotherapy. The therapeutic implication is that by resolving these CBF defects, cognitive functioning might improve, and the response to cognitive behavioral therapies also might be enhanced.

Outpatient Randomized Clinical Trials

Outpatient clinical trials remain the standard approach to assessing efficacy of a medication. Although many principles of conducting randomized placebo-controlled clinical trials in psychopharmacology apply to these studies, some specific considerations are relevant to outcome measures that are not found in other areas. Urine toxicology is a most informative outcome that can be analyzed with both quantitative and qualitative approaches. The urines are typically obtained three times per week for maximum sensitivity to repeated stimulant use based on the duration that detectable metabolite levels remain after use. Analyses are most frequently done with cutoff scores of 300 ng/mL, for example, with the cocaine metabolite benzoylecgonine, with any level above this being considered an indication of cocaine use within the last 3 days. More complex analyses have been proposed using quantitative levels either directly with gas chromotography–mass spectroscopy for quantitation or immunoassays for semiquantitation. This semiquantitation can be combined with self-reported use and compared to urine levels obtained prior to the urine in question to estimate new use of cocaine, because with three times weekly toxicologies a heavy daily cocaine abuser can stop using for 2 or 3 days and yet still have a positive urine (30) (e.g., positive on Monday and Wednesday when the last use was on Sunday). Although the goal of treatment is often complete abstinence, the sensitivity of these urine tests can be enhanced by these data manipulations. Thus, self-reported decreases in stimulant use may be important as a treatment outcome even when the goal may be abstinence initiation. Treatment retention is also critical in getting toward this goal of abstinence initiation, in order to keep the patient available for intervention.

In these outpatient studies, relapse prevention is a conceptual outcome that follows abstinence initiation. Relapse as defined by recurrent use or dependence after "sustained abstinence" first requires a definition of sustained abstinence, particularly among the binge users of stimulants. These patients may use weekly or even less often in binges that last up to several days. Although meeting the criteria of the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV) for stimulant dependence, the periods of nonuse can manifest reasonable psychosocial functioning, yet each time the patient returns to a binge this return is not a relapse. Simple definitions of sustained abstinence can just be defined by a period of cocaine- or amphetamine-free urines that lasts three, four, or perhaps ten times longer than the typical inter-binge interval. For current investigations, an important prognostic stratification is evolving based on sustained abstinence; patients who are abstinent during the 2 to 5 weeks before entering a medication trial have better treatment outcomes than those who continue to use up to their entry into treatment (31). Longer-term relapse prevention has also been an area where psychotherapy may synergize with pharmacotherapy (32). For example, sustained abstinence with desipramine treatment for cocaine dependence was enhanced by relapse-prevention cognitive behavioral therapy when examined at 6 and 12 month follow-up. Relapse was significantly higher after attaining abstinence with the medication alone than with both medication and the behavioral therapy.

SPECIFIC MEDICATIONS

Part of "102 - Pathophysiology and Treatment of Cocaine Dependence "

A large number of medications have been used for a variety of cocaine-related effects, including treatment of cocaine withdrawal or cocaine craving, and initiation and maintenance of abstinence. Although many of these medications have appeared to be promising in open trials, randomized, placebo-controlled clinical trials have not shown any medications to have substantial efficacy for cocaine dependence. Many studies have included small sample sizes and have been hampered by large dropout rates. Diagnostic criteria

have varied across clinical trials (some studies enroll patients meeting diagnostic criteria for cocaine dependence or for cocaine abuse, and others do not specify patient diagnosis). Many larger studies have examined patients with primary opiate dependence on methadone maintenance. Although these patients tend to be more available for follow-up because of their need to report to a clinic daily for methadone treatment, it is likely that they are different from patients with primary cocaine use disorders. Therefore, results obtained in studies enrolling methadone-maintained cocaine abusers may not apply to other patient groups. Outcome variables differ among clinical trials, making it difficult to determine a medication's effectiveness. Studies that utilized self-reports without confirmation by urine toxicology screen may not be reflective of cocaine use by study participants.

Antidepressants

Desipramine, a tricyclic antidepressant agent, was one of the first medications to be studied as a treatment for cocaine dependence. It is one of the most extensively studied pharmacotherapies for cocaine dependence to date (4). The initial study of desipramine suggested its efficacy based on self-report primarily, and two subsequent studies in methadone-maintained samples based on urine toxicology found no difference from placebo (33 ,34 and 35). A large clinical trial examined the efficacy of desipramine and psychotherapy, alone and in combination, as a treatment for ambulatory cocaine abusers (32). In this 12-week, double-blind, placebo-controlled trial, 139 subjects were assigned to one of four conditions: relapse prevention therapy plus desipramine, clinical management plus desipramine, relapse prevention plus placebo, and clinical management plus placebo. The mean dose of desipramine was 200 mg daily and was adjusted by a nonblinded psychiatrist in response to plasma concentration (target ranges 300 to 750 ng/mL) and side effects. All groups showed significant improvement in treatment retention and a reduction in cocaine use at 12 weeks, but there were no significant main effects for psychotherapy, pharmacotherapy, or the combination. Lower severity patients (cocaine use 1 to 2.5 g/week) had improved abstinence initiation when treated with desipramine. Desipramine was significantly more effective than placebo in reducing cocaine use during the first 6 weeks of treatment. Depressed subjects had a greater reduction in cocaine use than nondepressed subjects and had a better response to relapse prevention therapy. This finding of a desipramine response among depressed patients was confirmed among the depressed patients on methadone. A subsequent study by Nunes et al. (5) also found that depressed cocaine abusers on methadone had a significant reduction in cocaine use on imipramine, but not placebo. They did not find a significant effect in the nondepressed patients. Finally, a recent study with desipramine in methadone- and buprenorphine-maintained cocaine- and opioid-dependent patients found a reduction in both opioid and cocaine abuse with desipramine (36). A recent report of desipramine in depressed cocaine abusers found no difference from placebo; however, those patients whose depression remitted also showed a substantial reduction in cocaine use (37). Thus, these tricyclic antidepressants do not have well-demonstrated utility even in the depressed cocaine abusers, who can be a substantial subgroup comprising up to 40% of those presenting for treatment (3 ,6).

Several well-controlled human laboratory and outpatient clinical trials with fluoxetine have been conducted in patients with cocaine use disorders. A double-blind, placebo-controlled, cocaine administration study examined the interaction of cocaine with fluoxetine at 0, 20, 40, or 60 mg daily on an ascending schedule (38), and found that the 40- and 60-mg doses of fluoxetine decreased subjective effects of cocaine. Fluoxetine has been utilized in outpatient clinical trials in both methadone-maintained, cocaine-dependent patients and in patients with primary cocaine use disorders. An open study in methadone-maintained, cocaine-dependent patients found that fluoxetine at 45 mg daily significantly reduced self-reported use and quantitative urine benzoylecgonine concentrations during 9 weeks of treatment (39). More recently, fluoxetine has not reduced cocaine positive urines more than placebo in either methadone-maintained or primary cocaine abusers (40). The consensus of these studies is that fluoxetine may not have a clinical role among unselected cocaine abusers, and side effects have limited its use in several studies.

Bupropion is a second-generation antidepressant that enhances dopaminergic and noradrenergic transmission, but has little effect on serotoninergic neurotransmission. Although a pilot study suggested efficacy, a large multicenter study in methadone-maintained patients showed little benefit in cocaine dependence (41).

Dopaminergic Agents (DA)

The most widely accepted explanation of cocaine-induced euphoria is that dopamine reuptake inhibition results in increased extracellular dopamine concentration in the mesolimbic and mesocortical reward pathways in the brain (42). This basis for euphoria has suggested that dopamine antagonists might reduce cocaine use, but few human laboratory studies have supported their use, and controlled outpatient trials with both D1 and D2 antagonists have not been supportive. Although a laboratory study suggested attenuation of cocaine effects by the D1 antagonist Schering 39166, a multisite outpatient trial found no dose response and no difference from placebo in cocaine use (43; Ko, personal communication, 1999). The D2 antagonists such as haloperidol and flupenthixol have had minimal effects on euphoria in human cocaine administration studies (44), and flupenthixol has not been superior to placebo in an outpatient trial (45).

Because relative DA hypofunction induced by cocaine abuse may underlie craving and withdrawal symptoms that are often observed in recently abstinent cocaine-dependent patients (3), DA agonists may be of use. Bromocriptine is an agonist with high affinity for the D2 receptor. In human studies, pretreatment with either bromocriptine 2.5 or 5 mg 2 hours prior to cocaine administration had no effect on cocaine euphoria (46). Although early work supported its use, even in an early double-blind clinical trial, bromocriptine at 5 to 7.5 mg daily was poorly tolerated, with high dropout rates (47). In another small double-blind, placebo-controlled trial Moscovitz et al. (48) gave bromocriptine 1.25 mg three times daily or placebo to cocaine-abusing patients presenting to an emergency room for minor medical complaints. They found no difference in retention (bromocriptine group 43%, placebo group 31%), but those randomized to bromocriptine had more urine toxicology screens negative for cocaine (67%) than those randomized to placebo (31%). Cocaine administration studies have found a lack of effects with pergolide (49). A placebo-controlled outpatient study of pergolide found no difference in cocaine use and significant side effects, in spite of early pilot work in 21 patients suggesting good responses in 16 of 21 patients (50). Most recently, the D1 agonist ABT431 has been examined in human cocaine administration studies and found to reduce cocaine-induced craving (51). This compound is only available intravenously, but this offers promise for related compounds such as the D3 partial agonist recently reported in the animal laboratory (19).

Amantadine increases dopaminergic transmission, but whether the mechanism is DA release, direct effects on DA receptors, or DA reuptake blockade is unclear. One study examined the effects of acute amantadine (200 or 400 mg) and chronic amantadine (100 mg twice daily for 4 days) followed by insufflation of cocaine 0.9 mg/kg (52). Although the acute 200-mg dose of amantadine was associated with a decrease in cocaine “high,” chronic administration of amantadine 100 mg twice daily was associated with increased “high” in male subjects after cocaine administration. The effectiveness of amantadine was evaluated in a double-blind, placebo-controlled trial in which 42 patients in a day treatment program were randomized to amantadine 100 mg twice daily (n = 21) to be taken over 10 days or placebo (n = 21). Urine toxicology screens showed that those who had received amantadine were significantly more likely to be free of cocaine (p <.05) at the 2-week and 1-month follow-up visits (53).

L-deprenyl is a monoamine oxidase type B inhibitor that specifically inhibits the metabolism of DA. A study in five human volunteers examined the effects of 10 mg L-deprenyl alone and in combination with cocaine, but found no attenuation of cocaine effects (54). More recently, it was found to attenuate some subjective effects of cocaine, and an outpatient trial showed reduced cocaine use reported in comparison to placebo (29; Vocci, personal communication, 2000).

Methylphenidate (MP) is a stimulant and DA agonist primarily used in the treatment of childhood attention-deficit/hyperactivity disorder. MP is a DA agonist with pharmacologic properties that include DA release, and it has similar levels of binding to the DAT as cocaine. Grabowski et al. (55) have reported that it does not increase cocaine use and retains patients better than placebo, but have not shown a reduction in cocaine use compared to placebo.

Mazindol is a DA reuptake inhibitor that is without abuse liability and it has been suggested that it might antagonize the effects of cocaine as a treatment. A report on the effects of cocaine alone and in combination with mazindol at 1 or 2 mg orally in cocaine abusing volunteers found that the combination significantly increased heart rate and blood pressure (56). Mazindol did not alter the subjective effects of cocaine. One 12-week, double-blind, placebo-controlled clinical trial of mazindol 2 mg daily in cocainedependent subjects reported no difference from placebo (57). Mazindol was also not well tolerated, with 16 of 33 patients dropping out, and the average length of treatment was 5 weeks. A similar trial in methadone maintained patients found limited efficacy for those patients who had been cocaine abstinent for at least 2 weeks before starting mazindol (58).

Nonspecific Anticraving Agents

A number of other agents have been tested to reduce the desire or craving for cocaine. The rationales have broadly involved mechanisms such as sensitization and kindling as well as neurotransmitter systems that are indirectly affected by cocaine such as the opioid, excitatory amino acid/glutamate, and GABAergic systems. For most of these approaches, outpatient clinical trials have been quite limited. Medications include GABA agents such as baclofen, opioid antagonists such as naltrexone, calcium channel blockers such as nifedipine, antikindling agents such as carbamazepine, and disulfiram. Finally, stress responses and the associated elevation of cortisol have been considered as potentially important in cocaine craving induction and as a therapeutic agent. However, a cocaine administration study showed no reduction in cocaine effects or self-administration with the cortisol synthesis inhibitor ketoconazole in spite of significant reductions in cortisol levels (59).

Carbamazepine (CBZ) is an anticonvulsant medication hypothesized to have potential as a treatment for cocaine craving and abuse because of its ability to block cocaine-induced “kindling” in rodents. A double-blind, placebo-controlled, crossover study of the interaction of 400 mg of CBZ daily for 5 days with cocaine found no effects on subjective response to cocaine (60). A double-blind, placebo-controlled study in outpatients included a 20-day, controlled, fixed-dose (CBZ 200 mg or 400 mg or placebo)

trial in 30 volunteers and found that cocaine use was unchanged (61). Another study in 183 cocaine abusers randomized to CBZ 400 or 800 mg daily or placebo showed that CBZ at 400 mg was associated with a significant decrease in cocaine-positive urines and a reduction in cocaine craving (62). However, three other double-blind, placebo-controlled studies with CBZ treatment in over 150 subjects found no significant difference in cocaine use, cocaine-positive urine samples, or depressive symptoms measured by the Beck Depression Inventory (63 ,64 and 65). Plasma CBZ levels of 5.6 ± 0.8 µg/mL were achieved by week 4 in these studies. Thus, confidence in this medication has waned.

Naltrexone is an opioid antagonist that has been examined as a treatment agent for cocaine abuse. One study examined the effects of cocaine after 10 days of treatment with naltrexone 50 mg or placebo in a double-blind, randomized, within-subjects design (66). Some cocaine-induced subjective effects were less during naltrexone than placebo administration. A placebo-controlled outpatient study of naltrexone found no efficacy (67).

The calcium channel blockers and antagonists of glutaminergic function have also been examined as anticraving agents and protective agents to minimize cardiovascular cerebral damage from cocaine. The calcium channel antagonist nifedipine has been studied and shows some promise (68). Nimodipine showed a reduction in the effects of intravenous cocaine as well as reductions in acute cocaine-related cardiovascular toxicity, but lamotrigine did not reduce cocaine effects in a similar placebo-controlled crossover study (69 ,70). Memantine, a glutamate inhibitor, showed no efficacy in reducing cocaine effects acutely (71). Outpatient placebo-controlled studies have not been done with these agents, however.

Much enthusiasm has developed for the use of agents targeting the GABA system, particularly for vigabatrin, which antagonizes the breakdown of GABA (72). Unfortunately, this agent is not available in the United States, and its side effects of bitemporal hemianopsia may preclude its use in cocaine abusers (73). However, baclofen, which is a direct agonist for the $GABA_B$ receptor, has shown some reduction in cocaine self-administration in animals and some utility in reducing cocaine abuse among humans (74). No other controlled trials have been published with this or related GABA agents, but several have gotten preliminary screenings in the National Institute on Drug Abuse (NIDA) medications development program including tiagabine, which also enhances GABA levels (Vocci, personal communication, 2000).

Disulfiram is an aldehyde dehydrogenase inhibitor used in treating alcoholism, a common coexisting problem among cocaine abusers. One study in six cocaine-dependent volunteers examined the effect of disulfiram 250 mg on responses to intranasal cocaine (2 mg/kg) using a randomized double-blind, placebo-controlled design (75). Although disulfiram induced no significant differences in cocaine "high," it decreased craving for cocaine. Plasma cocaine concentration following cocaine administration was significantly greater while on disulfiram, and this may have contributed to the decreased craving and increased dysphoria observed in some subjects. Carroll et al. (76) found that cocaine use was significantly reduced in the disulfiram group compared to psychotherapy alone, with patients who abused both alcohol and cocaine. The patients reported a significantly lower percentage of cocaine use days and fewer days of cocaine use, and fewer positive urine screens for cocaine were observed.

In surveys of cocaine abusers, 65% have reported significant problems in concentration and 57% reported memory problems, and formal testing suggests some sustained abnormalities in memory and concentration among abusers (3 ,16). Initial studies of recovering cocaine-dependent patients have revealed impairments of short-term memory, attention, and complex psychomotor and simple motor abilities, but the data are limited (16 ,77). Reaction time, motoric signs of central nervous system (CNS) dysfunction, and EEG evidence of residual CNS hyperexcitability may also persist (78).

These problems may be associated with structural or functional brain damage caused by cocaine including strokes (16). Structural imaging using computed tomographic scanning and magnetic resonance imaging (MRI) have shown enlarged ventricles and sulci in cocaine abusers (79). Functional neuroimaging studies have shown focal reductions in regional cerebral blood flow (rCBF) among chronic cocaine abusers (15 ,16 and 17). These defects also appear to be persistent for several weeks of abstinence at least, and can be associated with neuropsychological deficits (15 ,16 and 17 ,80). The ischemic damage from cocaine can lead to neuronal degeneration, as suggested by phosphorus magnetic resonance spectroscopy (^{31}P-MRS), in which abstinent cocaine abusers showed abnormally high levels of phosphomonoesters and low levels of nucleotide triphosphates compared to normals (81).

The etiology of decreases in rCBF following cocaine may involve vasoconstriction (82) and platelet abnormalities. The vasoconstriction may respond to calcium channel blockers (83). Abnormal platelets may produce thrombosis in cerebral vessels and produce blood flow alterations (18). In autopsy studies platelet-rich coronary thrombi (sometimes in otherwise normal vessels) and accelerated atheromatous lesions are found and could be ascribed to platelet activation and platelet α-granule release (16). Because platelet granule release appears to be completely inhibited by aspirin under *shear* conditions (analogous to flowing blood), and aspirin prevents thrombotic complications, a preliminary test of 4 weeks of aspirin therapy led to a 50% improvement in cerebral perfusion (16). In a placebo-controlled study that has just been completed, aspirin significantly reduced perfusion defects on single photon emission computed tomography (SPECT) imaging (84 ,85).

Peripheral Blocking Agents Targeting Cocaine Itself

Although the simplest peripheral blocking approach of passively injecting polyclonal antibodies to cocaine into a human might be useful for cocaine overdoses, these antibodies would not last very long and might be of limited use as a sustained treatment. For any type of relapse prevention, the immune response elements must remain at relatively high levels for periods of several weeks or months, which is best done by active immunization (86). However, three other approaches using catalytic antibodies, monoclonal passive antibodies, or injections of butrylcholinesterase have some promise (87). With all these peripheral cocaine-blocking agents, the amount of cocaine entering the brain is partially blocked or its rate of entry is reduced. Either of these effects can cause a very significant reduction in the high or rush from cocaine. All four of these approaches can also be combined and used together with the pharmacotherapies described above. The only approach that has been tested in humans is active immunization (86). The initial animal studies showed excellent production of a highly specific antibody to cocaine. With active immunization the amount of inhibition of cocaine entering the brain ranged from 30% to 63% at 30 seconds after cocaine injection in rats. This amount of inhibition was sufficient to extinguish cocaine self-administration in the rat model.

In the initial human study of this vaccine, it was well tolerated with virtually no side effects using a dose of 1,000 µg given with two booster injections over a 3-month period (88). The vaccine produced substantive quantities of antibody that was related to both the dose of vaccine and the number of booster injections. Thus, further studies of its potential efficacy in relapse prevention for abstinent cocaine abusers appear warranted.

PSYCHOTHERAPIES

Part of "102 - Pathophysiology and Treatment of Cocaine Dependence "

Professional Psychotherapy vs. Drug Counseling

Because of the limited efficacy of pharmacotherapy, the success of behavioral and psychotherapies is important to consider. Two major approaches have been evaluated. First, the use of professional therapies such as cognitive behavioral therapy and supportive expressive therapies has been examined. Second, contingency management as a form of behavioral therapy has gotten much attention and reasonable success. These therapies have now been extensively studied and are increasingly being examined as treatments that might be complemented by emerging pharmacotherapies. However, nonprofessional drug counseling also holds much promise and many be more readily available to community programs.

The most extensive examination of psychotherapy for cocaine dependence has been the NIDA Collaborative Cocaine Treatment Study. It was a large, multisite psychotherapy clinical trial for outpatients who met the DSM-IV criteria for cocaine dependence. For 480 randomized patients, four treatments were compared over an 18-month period. All treatments included group drug counseling. One treatment also added cognitive therapy, one added supportive-expressive psychodynamic therapy, and one added individual drug counseling. The final group had drug counseling alone. Two specific interaction hypotheses, one involving psychiatric severity and the other involving degree of antisocial personality characteristics, were examined, but no major findings related to these hypotheses have been found (88 ,89).

All of the therapies were manual guided and treatment was intensive, including 36 possible individual sessions and 24 group sessions for 6 months. Patients were assessed monthly during active treatment and at 9 and 12 months after baseline. Primary outcome measures were the Addiction Severity Index–Drug Use Composite score and the number of days of cocaine use in the past month. Compared with the two psychotherapies and with group drug counseling (GDC) alone, individual drug counseling plus GDC showed the greatest improvement on the Addiction Severity Index–Drug Use Composite score. Individual group counseling plus GDC was also superior to the two psychotherapies on the number of days of cocaine use in the past month. Hypotheses regarding the superiority of psychotherapy to GDC for patients with greater psychiatric severity and the superiority of cognitive therapy plus GDC compared with supportive-expressive therapy plus GDC for patients with antisocial personality traits or external coping style were not confirmed. Thus, compared with professional psychotherapy, a manual-guided combination of intensive individual drug counseling and GDC showed promise for the treatment of cocaine dependence (90).

Cognitive Behavioral Therapy (CBT)

In spite of these overall discouraging results, cognitive behavioral treatments have been among the most frequently evaluated psychosocial approaches for the treatment of substance use disorders and have a comparatively strong level of empirical support (91 ,92). To date, more than 24 randomized controlled trials have evaluated the effectiveness of cognitive behavioral relapse prevention treatment on substance use outcomes among adult tobacco smokers and alcohol, cocaine, marijuana, opiate, and other types of substance abusers (93). Overall, these studies suggest that the average effect size for CBT compared with control or comparison conditions is 0.36 (Feingold, unpublished data/APA presentation), which is consistent with a moderate effect. Review of this group of studies suggests that, across substances of abuse but most strongly for smoking, there is good evidence of the effectiveness of CBT compared with no-treatment

controls (93). This body of literature also suggests that outcomes in which CBT may hold particular promise include reduced severity of relapses when they occur, enhanced durability of effects, and patient-treatment matching, particularly for patients at higher levels of impairment along dimensions such as psychopathology or dependence severity. A review of this series of studies can be found in Carroll (93).

To help cocaine-dependent individuals meet the treatment goal of abstinence and relapse prevention, CBT treatment has two critical components. The first is a thorough *functional analysis* of the role cocaine and other substances play in the individual's life. A functional analysis is simply an exploration of cocaine use with respect to its antecedents and consequences. The second critical component of CBT is skills training. In CBT, a substantial portion of every session is devoted to the teaching and practice of coping skills; in fact, CBT can be thought of as a highly individualized training program that helps cocaine abusers unlearn old habits associated with cocaine abuse and learn or relearn more healthy skills and habits. Other important features of CBT are fostering the motivation for abstinence, teaching coping skills, changing reinforcement contingencies, fostering management of painful affects, and improving interpersonal functioning.

In a study comparing supportive therapy to CBT for pharmacotherapy of cocaine dependence, 121 individuals meeting DSM-III-R criteria for cocaine dependence were randomly assigned to one of the four treatment conditions: (a) CBT in combination with desipramine, (b) CBT plus placebo, (c) clinical management (ClM) plus desipramine, and (d) ClM plus placebo (33). Cocaine outcomes were comparable whether the patient received CBT or ClM, or whether the patient received desipramine or placebo, but patients with more severe cocaine use were retained longer in treatment, attained longer periods of abstinence, and had fewer urine screens positive for cocaine when treated with CBT compared with ClM. CBT also was more effective than supportive ClM in retaining depressed subjects in treatment and in reducing cocaine use (94). Thus, CBT has been useful for medication development as a platform for clinical trials because it meets the guidelines for an effective platform. Specifically, it is strong enough to hold patients in treatment, but not so strong as to eliminate the possibility for any medication effects. As counterexamples, treatments such as clinical management tend to be too weak to hold patients, although day treatments tend to produce very high rates of abstinence without any medications, but can serve as excellent means to inducing initial abstinence.

Contingency Management Procedures

Contingency management (CM) procedures are based on a behavioral perspective of drug abuse, which views drugs as powerful reinforcers maintaining high rates of behavior aimed at administering the drugs, even in the absence of physical dependence (95). In substance abusers, drugs can therefore be seen as being the predominant reinforcers exerting control over a large portion of these individuals' behavioral repertoire, whereas in nonsubstance abusers more socially acceptable reinforcers influence behavior. Thus, the goal of drug abuse treatment is to decrease behavior maintained by drug reinforcers and increase behavior maintained by nondrug reinforcers. CM procedures are one method of accomplishing this goal, by presenting rewards or incentives contingent upon documented drug abstinence (positive contingencies), withdrawing privileges contingent upon documented drug use (negative contingencies), or a combination of the two.

Higgins and colleagues (95 ,96 and 97) have demonstrated that CM procedures in combination with a community reinforcement approach (CRA) facilitate initial abstinence in primarily cocaine-dependent persons. In the first, 12-week study (95), the CM procedure consisted of vouchers with a monetary value, which were presented upon evidence of drug abstinence (i.e., cocaine-free urine) during weeks 1 to 12. The vouchers increased in value for every consecutively drug-free urine and were exchangeable for client-therapist agreed-upon retail items and activities related to treatment goals. Treatment retention was significantly higher in the behavioral treatment than standard drug counseling group. In addition, 85% of clients receiving the behavioral treatment achieved at least 3 weeks of abstinence as compared to 33% of clients receiving standard drug abuse counseling. In the second study (96), the CM procedure was modified, in that vouchers exchangeable for goods and services in weeks 1 to 12 and lottery tickets in weeks 13 to 24 were presented contingent upon documented drug abstinence. As before, treatment retention was significantly higher in the behavioral treatment than standard drug counseling group. Similarly, 68% and 42% of the clients in the behavioral treatment group achieved at least 8 and 16 weeks, respectively, of continuous cocaine abstinence as opposed to 11% and 5% in the standard drug abuse counseling group. In the third, 24-week study (97), cocaine-dependent individuals were randomized to receive either behavioral treatment without incentives or behavioral treatment with incentives (i.e., vouchers exchangeable for goods and services). The group that received the incentives showed significantly greater treatment retention (75% vs. 40%) and longer duration of continuous abstinence (11.7 vs. 6.0 weeks) than the group not receiving the incentives. Overall, the findings of these studies suggest that incentives contingent on drug abstinence can be a powerful intervention for facilitating cocaine abstinence in primary cocaine abusers, although separating the CRA effects has not been done.

This voucher system also has been examined in a 12-week clinical trial for its ability to facilitate cocaine abstinence in methadone-maintained cocaine abusers (98 ,99). The contingency group subjects achieved significantly longer durations

of sustained abstinence than yoked-controls (mean of 5.0 vs. 0.8 weeks, respectively), with 47% of contingency subjects achieving at least 7 weeks vs. 6% of controls achieving at least 2 weeks of sustained cocaine abstinence. These findings suggest that vouchers also can be used as incentives for drug abstinence in opioid-dependent cocaine abusers using a CM procedure similar to that employed by Higgins et al. (97).

There are also problems with CM. One issue with CM procedures is that the therapeutic effects tend to be impermanent following withdrawal of the intervention. This issue of continued efficacy after stopping medications has been addressed in a very limited way, mostly due to the lack of medications showing equivalent efficacy to these contingency approaches (32). Also, there are no mechanisms available to support CM in standard clinical programs, although some new approaches are being developed (100 ,101). Because vouchers are used to support treatment goals, therapists must work with patients to evaluate appropriate use of the vouchers, and treatment staff generally must assist in making voucher purchases. These restrictions impose considerable program costs over and above the costs of the vouchers. The delay between the time the reinforcement (purchase of goods or services) is provided and the time that the behavior being reinforced (abstinence, as evidenced by a drug-free urine) occurs may decrease the value to the patient (but not the actual program cost) of the reinforcement. The efficacy of CM in studies with cocaine-dependent patients also appears to be considerably more modest at best than in the earlier studies. Iguchi and his colleagues (102) compared voucher-based CM used to reinforce either drug-free urine samples (UA group) or treatment plan tasks (TP tasks) and a no-voucher standard treatment group (STD) during methadone maintenance treatment. The value of the vouchers was set considerably lower than in other studies of CM and did not increase in value for successive drug-free urine samples or completion of therapeutic tasks. The authors also did not use the CRA that Higgins has used, although their TP intervention included many of the CRA elements. There were no significant main effects of treatment group on rates of drug-free urine samples. Rates of drug-free urine samples remained relatively unchanged in either the UA or STD groups, whereas they increased over time in the TP group. Finally, CM is not effective for all patients—for example, 10 of 19 (53%) CM-treated methadone-maintained patients failed to achieve more than 3 consecutive weeks of cocaine abstinence in the study reported by Silverman and his colleagues (98 ,99), and resumption of drug use following discontinuation of CM is also a problem. Although increasing the value, schedule, or duration of the vouchers may lead to higher and more sustained rates of abstinence (103), alternatives to CM should also be explored. Considering that drug-dependent patients continue illicit drug use despite extremely high immediate and longer-term costs, increasing patient internal motivation may be more cost-effective than increasing the value of the vouchers or monetary rewards for abstinence. Of additional concern is the possibility that failure to earn vouchers may contribute to demoralization and a lack of perceived self-efficacy for succeeding in stopping drug use and thus contribute to a cycle of drug use and failure.

In summary, despite its promise, there are a number of limitations of CM for the treatment of patients with cocaine dependence: (a) CM is of limited efficacy in this population. (b) CM is labor intensive and difficult to implement. (c) CM is costly and not supported by current funding mechanisms. (d) The failure to obtain vouchers in CM may contribute to demoralization. (e) There is a possible rebound in drug use or dissipation of effects after discontinuation of CM.

CM’s potential utility as a platform for pharmacotherapy has yet to be fully explored, but recent reviews suggest it may have a modest effect size of 0.25 and that various approaches can be used to apply it in community settings (101 ,104).

SUMMARY

Part of "102 - Pathophysiology and Treatment of Cocaine Dependence "

No medications are currently approved by the Food and Drug Administration (FDA) for cocaine dependence, but we have developed several leads for medications based on our understanding of the neurobiology and clinical phenomenology of stimulants. Based on neurobiological abnormalities in dopamine receptors and transporters after chronic stimulant use, studies have examined both dopamine agonists and antagonists, but not shown clinical efficacy. Based on clinical phenomenology, antidepressants have been tried in depressed cocaine abusers who may reduce their cocaine use with desipramine, other tricyclics, serotonin reuptake inhibitors, and bupropion. Among unselected stimulant abusers these antidepressants may be quite limited, but when depressive symptoms are reduced, cocaine abstinence also appears to follow. Cerebral blood flow (CBF) defects also appear to be relatively common among stimulant abusers and to correlate with neuropsychological deficits. These CBF defects in cocaine abusers may respond to antistroke medications, and this potential for remediation builds on a rapidly evolving field of stroke pharmacotherapy. Finally, vaccines are under development that may reduce cocaine’s rewarding effects and prevent relapse among abstinent formerly dependent patients.

Methods for screening medications as potential pharmacotherapies have used human laboratory studies employing cocaine administration as a surrogate efficacy assessment. Although this method needs validation with a gold standard of medications that have demonstrated efficacy in outpatient randomized clinical trials, these laboratory settings have been helpful in assessing medical safety during cocaine interactions. Neuroimaging of cerebral blood flow and of

“receptor” binding also holds promise for medication development.

With all of these pharmacotherapies the behavioral platform for their delivery is critical in retaining the patient in treatment and maintaining compliance with the medications. As a behavioral disorder, stimulant dependence is quite responsive to contingency management using a variety of reinforcers and schedules of reinforcement. Vouchers to purchase prosocial goods and services are the most common reinforcer used to initiate and maintain stimulant-free urines (95 ,97). Reinforcement schedules are typically on a one-to-one fixed ratio initially, with a progressive increase in the ratio of reinforcement and escalation in reinforcers as longer periods of abstinence are attained. The major problem with this approach has been maintaining abstinence after the reinforcers are withdrawn completely and developing a mechanism to obtain these types of reinforcers outside of a research setting. A more typical time limited therapy for clinical programs is cognitive behavioral therapy. Cognitive behavioral therapies have been examined in conjunction with pharmacotherapy, particularly using antidepressants, and have shown interesting additive effects (32). For example, at 1-year follow-up after a 3-month treatment period, those patients who got both the pharmacotherapy and the cognitive therapy showed more sustained abstinence than those who got either therapy alone. The behavioral treatments may also be most useful for abstinence initiation, particularly the contingency management and cognitive behavioral therapy approaches (32). Overall, the long-term outcome at 1 year is substantially enhanced by the use of psychotherapy in combination with medications.

ACKNOWLEDGMENTS

Part of "102 - Pathophysiology and Treatment of Cocaine Dependence "

This work was supported by the National Institute on Drug Abuse grants P50-DA04060 and P50-DA12762.

REFERENCES

1. National Institute on Drug Abuse. *National Survey Results on Drug Use from Monitoring the Future Survey*. Rockville, MD: U.S. Government Printing Office, 1998.

2. Johanson C-E, Schuster CR. Cocaine. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:1685–1697.

3. Gawin FH, Ellinwood EH. Cocaine and other stimulants: actions, abuse and treatment. *N Engl J Med* 1988;318(18):1173–1181.

4. Levin FR, Lehman AF. Meta-analysis of desipramine an adjunct in the treatment of cocaine addiction. *J Clin Pharmacol* 1991;11:374–378.

5. Nunes EV, Quitkin FM, Brady R, et al. Imipramine treatment of methadone maintenance patients with affective disorder and illicit drug use. *Am J Psychiatry* 1991;148:667–669.

6. Kosten TR, Markou A, Koob GF. Depression and stimulant dependence: neurobiology and pharmacotherapy. *J Nerv Ment Dis* 1998;186:737–745.

7. Malison RT, Best SE, van Dyck CH, et al. Elevated striatal dopamine transporters during acute cocaine abstinence as measured by [^{123}I]beta-CIT SPECT. *Am J Psychiatry* 1998;155(6):832–834.

8. Volkow ND, Fowler JS, Wolf AP, et al. Effect of chronic cocaine abuse on postsynaptic dopamine receptors. *Am J Psychiatry* 1990;147(6):719–724.

9. Aronson SC, Black JE, McDougle CJ, et al. Serotonergic mechanisms of cocaine effects in humans. *Psychopharmacology* 1995;119:179–185.

10. McDougle CJ, Price LH, Palumbo JM, et al. Dopaminergic responsivity during cocaine abstinence. *Psychiatry Res* 1992;43:77–85.

11. McDougle CJ, Black JE, Malison RT, et al. Noradrenergic dysregulation during discontinuation of cocaine use in addicts. *Arch Gen Psychiatry* 1994;51:713–719.

12. Johanson C-E, Fischman MM. The pharmacology of cocaine related to its abuse. *Pharmacol Rev* 1989;41:3–52.

13. Gelernter J, Kranzler HR, Satel SL et al. Genetic association between dopamine transporter protein alleles and cocaineinduced paranoia. *Neuropsychopharmacology* 1994;11(3):195–200.

14. Jacobsen LK, Staley JK, Malison RT, et al. Elevated central serotonin transporter binding availability in acutely abstinent cocaine dependent patients. *Am J Psychiatry* 2000;157:1134–1140.

15. Holman BL, Mendelson J, Garada B, et al. Regional cerebral blood flow improves with treatment in chronic cocaine polydrug users. *J Nucl Med* 1993;34:723–727.

16. Kosten TR. Pharmacotherapy of cerebral ischemia in cocaine dependence. *Drug Alcohol Depend* 1998;49:133–144.

17. Volkow ND, Mullani N, Gould L, et al. Cerebral blood flow in chronic cocaine users: a study with positron emission tomography. *Br J Psychiatry* 1988;152:641–648.

18. Rinder HM, Ault KA, Jatlow PI, et al. Platelet alpha-granule release in cocaine users. *Circulation* 1994;90:1162–1167.

19. Pilla M, Perachon S, Sautel F, et al. Selective inhibition of cocaine seeking behavior by a partial dopamine D3 receptor agonist. *Nature* 1999;400:371–375.

20. Kosten TR, Bond J, Sparenborg S, et al. Human cocaine vaccine: safety and immungenicity. *Proc CoII Prob Drug Depend* 2000;60(Suppl. 1):S250.

21. Volkow ND, Wang GJ, Fischman MW, et al. Relationship between subjective effects of cocaine and dopamine transporter occupancy. *Nature* 1997;386(6627):827–830.

22. Gay GR. Clinical management of acute and chronic cocaine poisoning. *Ann Emerg Med* 1982;11:562–572.

23. Kosten TR, Kleber TR. Rapid death during cocaine abuse: variant of the neuroleptic malignant syndrome? *Am J Drug Alcohol Abuse* 1988;14:335–346.

24. King GR, Ellinwood EH. Amphetamines and other stimulants. In: Lowinson JH, Ruiz P, Millman RB, et al., eds. *Substance abuse: a comprehensive textbook*, third ed. Baltimore: Williams & Wilkins, 1997:207–223.

25. Fischman MW, Schuster CR, Resnekov I, et al. Cardiovascular and subject effects of intravenous cocaine administration in humans. *Arch Gen Psychiatry* 1976;10:535–546.

26. Fischman MW, Johanson CE. Ethical and practical issues involved in behavioral pharmacology research that administers drugs of abuse to human volunteers. *Behav Pharmacol* 1998;9(7):479–498.

27. Jaffe JH, Cascella NG, Kumor KM, et al. Cocaine-induced cocaine craving. *Psychopharmacology* 1989;97:59–64.

28. Fowler JS, Volkow ND, Ding YS, et al. Positron emission tomography studies of dopamine-enhancing drugs [review]. *J Clin Pharm* 1999;Suppl:13S–16S.

29. Bartzokis G, Beckson M, Newton T, et al. Selegiline effects on cocaine-induced changes in medial temporal lobe metabolism and subjective ratings of euphoria. *Neuropsychopharmacology* 1999;20(6):582–590.

30. Preston KL, Silverman K, Schuster CR, et al. Assessment of cocaine use with quantitative urinalysis and estimation of new uses. *Addiction* 1997;92(6):717–727.

31. Preston KL, Silverman K, Higgins ST, et al. Cocaine use early in treatment predicts outcome in a behavioral treatment program. *J Consult Clin Psychol* 1998;66(4):691–696.

32. Carroll KM, Rounsaville BJ, Gordon LT, et al. Psychotherapy and pharmacotherapy for ambulatory cocaine abusers. *Arch Gen Psychiatry* 1994;51:177–187.

33. Arndt IO, Dorozynsky L, Woody GE, et al. Desipramine treatment of cocaine dependence in methadone-maintained patients. *Arch Gen Psychiatry* 1992;49(11):888–893.

34. Gawin FH, Kleber HD, Byck R, et al: Desipramine facilitation of initial cocaine abstinence. *Arch Gen Psychiatry* 1989;46:117–121.

35. Kosten TR, Morgan CM, Falcioni J, et al. Pharmacotherapy for cocaine-abusing methadone-maintained patients using amantadine or desipramine. *Arch Gen Psychiatry* 1992;49:894–899.

36. Oliveto AH, Feingold A, Schottenfeld RS, et al. Desipramine in opioid-dependent cocaine abusers maintained on buprenorphine vs. methadone. *Arch Gen Psychiatry* 1999;56:812–820.

37. Nunes EV, McDowell D, Rothenberg J. Desipramine treatment for cocaine dependent patients with depression. *Proc Coll Prob Drug Depend* 2000;60(Suppl. 1):S160 (abst).

38. Walsh SL, Preston KL, Sullivan JT, et al. Fluoxetine alters the effects of intravenous cocaine in humans. *J Clin Psychopharmacol* 1994;14(6):396–407.

39. Batki SL, Manfredi LB, Jacob P, et al. Fluoxetine for cocaine dependence in methadone maintenance: quantitative plasma and urine cocaine/benzoylecgonine concentrations. *J Clin Psychopharmacol* 1993;13:243–250.

40. Grabowski J, Rhoades H, Elk R, et al. Fluoxetine is ineffective for treatment of cocaine dependence or concurrent opiate and cocaine dependence: two placebo-controlled double-blind trials. *J Clin Psychopharmacol* 1995;15(3):163–174.

41. Margolin A, Kosten TR, Avants SK, et al. A multicenter trial of bupropion for cocaine dependence in methadone-maintained patients. *Drug Alcohol Depend* 1995;49:125–131.

42. Koob GF. Drugs of abuse: anatomy, pharmacology and function of reward pathways. *Trends Pharmacol Sci* 1992;13:177–184.

43. Romach MK, Glue P, Kampman K, et al. Attenuation of the euphoric effects of cocaine by the dopamine D1/D5 antagonist ecopipam (SCH 39166). *Arch Gen Psychiatry* 1999;56(12):1101–1106.

44. Sherer MA, Kumor KM, Jaffe JH. Effects of intravenous cocaine are partially attenuated by haloperidol. *Psychiatry Res* 1989;27:117–125.

45. Khalsa E, Jatlow P, Gawin F. Flupenthixol and desipramine treatment of crack users: double blind results. *NIDA Res Monogr* 1994;141:439.

46. Kumor K, Sherer M, Jaffe J. Effects of bromocriptine pretreatment on subjective and physiological responses to IV cocaine. *Pharmacol Biochem Behav* 1989;33:829–837.

47. Tennant FS, Sagherian A. Double-blind comparison of amantadine and bromocriptine for ambulatory withdrawal from cocaine dependence. *Arch Intern Med* 1987;147:109–112.

48. Moscovitz H, Brookoff D, Nelson L. A randomized trial of bromocriptine for cocaine users presenting to the emergency department. *J Gen Intern Med* 1993;8:1–4.

49. Haney M, Foltin RW, Fischman MW. Effects of pergolide on intravenous cocaine self-administration in men and women. *Psychopharmacology* 1998;137(1):15–24.

50. Malcolm R, Moore JW, Kajdasz DK. Pergolide mesylate: adverse events occurring in the treatment of cocaine dependence. *Am J Addictions* 1997;6(2):117–123.

51. Haney M, Collins ED, Ward AS, et al. Effect of a selective dopamine D1 agonist (ABT-431) on smoked cocaine selfadministration in humans. *Psychopharmacology* 1999;143(1):102–110.

52. Sholar MB, Lukas SE, Kouri E, et al. Acute and chronic amantadine pre-treatment attenuates some of cocaine's effect in human subjects. *NIDA Res Monogr* 1994;141:431.

53. Alterman AI, Droba M, Antelo RE, et al. Amantadine may facilitate detoxification of cocaine addicts. *Drug Alcohol Depend* 1992;31:1929.

54. Haberny KA, Walsh SL, Ginn DH, et al. Absence of acute cocaine interactions with the MAO-B inhibitor selegiline. *Drug Alcohol Depend* 1995;39(1):55–62.

55. Grabowski J, Roache J, Schmitz J, et al. Replacement medication for cocaine dependence: methylphenidate. *J Clin Psychopharmacol* 1997;17:485–488.

56. Preston KL, Sullivan JT, Berger P, et al. Effects of cocaine alone and in combination with mazindol in human cocaine abusers. *J Pharmacol Exp Ther* 1993;267:296–307.

57. Stine SM, Krystal JH, Kosten TR, et al. Mazindol treatment for cocaine dependence. *Drug Alcohol Depend* 1995;39:245–252.

58. Margolin A, Avants SK, Kosten TR. Mazindol for relapse prevention to cocaine abuse in methadone-maintained patients. *Am J Drug Alcohol Abuse* 1995;21(4):469–481.

59. Ward AS, Collins ED, Haney M, et al. Ketoconazole attenuates the cortisol response but not the subjective effects of smoked cocaine in humans. *Behav Pharmacology* 1998;9(7):577–586.

60. Hatsukami D, Keenan R, Halikas J, et al. Effects of carbamazepine on acute responses to smoked cocaine-base in human cocaine users. *Psychopharmacology* 1991;104:120–124.

61. Halikas J, Kuhn K, Carlson G, et al. The effect of carbamazepine on cocaine use. *Am J Addict* 1992;1:129–139.

62. Halikas JA, Crosby RD, Pearson VL, et al. Carbamazepine improves early cocaine treatment retention and treatment effectiveness. *ACNP Annual Meeting Abstracts* 1993;175.

63. Cornish JW, Maany I, Fudala PJ, et al. Carbamazepine treatment for cocaine dependence. *Drug Alcohol Depend* 1995;38(3):221–227.

64. Kranzler HR, Bauer LO, Hersh D, et al. Carbamazepine treatment of cocaine dependence: a placebo-controlled trial. *Drug Alcohol Depend* 1995;38:203–211.

65. Montoya ID, Levin FR, Fudala PJ, et al. Double-blind comparison of carbamazepine and placebo for treatment of cocaine dependence. *Drug Alcohol Depend* 1995;38(3):213–219.

66. Kosten TR, Silverman DG, Fleming J, et al. Intravenous cocaine challenges during naltrexone maintenance: a preliminary study. *Biol Psychiatry* 1992;32:543–548.

67. Hersh D, Van Kirk JR, Kranzler HR. Naltrexone treatment of comorbid alcohol and cocaine use disorders. *Psychopharmacology* 1998;139:44–52.

68. Muntaner C, Kumor KM, Nagoshi C, et al. Effects of nifedipine pretreatment on subjective and cardiovascular responses to intravenous cocaine in humans. *Psychopharmacology (Berl)* 1991;105(1):37–41.

69. Kosten TR, Woods SW, Rosen MI, et al. Interactions of cocaine with nimodipine: a brief report. *Am J Addictions* 1999;8(1):77–81.

70. Winther LC, Saleem R, McCance-Katz EF, et al. Effects of lamotrigine on behavioral and cardiovascular responses to cocaine in human subjects. *Am J Drug Alcohol Abuse* 2000;26(1):47–59.

71. Collins ED, Ward AS, McDowell DM, et al. The effects of memantine on the subjective, reinforcing and cardiovascular effects of cocaine in humans. *Behav Pharmacol* 1998;9(7):587–598.

72. Gerasimov MR, Dewey SL. Gamma-vinyl gamma-aminobutyric acid attenuates the synergistic elevations of nucleus accumbens dopamine produced by a cocaine/heroin (speedball) challenge. *Eur J Pharmacol* 1999;380(1):1–4.

73. Kalviainen R, Nousiainen I, Mantyjarvi M, et al. Vigabatrin, a gabaergic antiepileptic drug, causes concentric visual field defects. *Neurology* 1999;53(5):922–926.

74. Ling W, Shoptaw S, Majewska D. Baclofen as a cocaine anti-craving medication: a preliminary clinical study [letter]. *Neuropsychopharmacology* 1998;18(5):403–404.

75. McCance EF, Kosten TR, Jatlow P. Disulfiram effects on acute cocaine administration. *Drug Alcohol Depend* 1998;52(1):27–39.

76. Carroll KM, Nich C, Ball SA, et al. Treatment of cocaine and alcohol dependence with psychotherapy and disulfiram [research report]. *Addiction* 1998;93(5):713–728.

77. Berry J, van Gorp WG, Herzberg DS, et al. Neuropsychological deficits in abstinent cocaine abusers: preliminary findings after two weeks of abstinence. *Drug Alcohol Depend* 1993;32(3):231–237.

78. Bauer LO. Motoric signs of CNS dysfunction associated with alcohol and cocaine withdrawal. *Psychiatry Res* 1993;47:69–77.

79. Pascual-Leone A, Dhuna A, Anderson DC. Cerebral atrophy in habitual cocaine abusers: a planimetric CT study. *Neurology* 1991;41(1):34–38.

80. Woods SW, O'Malley SS, Martini BL, et al. SPECT regional cerebral blood flow and neuropsychological testing in non-demented HIV-positive drug abusers. *Prog Neuropsychopharmacol Biol Psychiatry* 1991;15:649–662.

81. Christensen JD, Kaufman MJ, Levin JM, et al. Abnormal cerebral metabolism in polydrug abusers during early withdrawal: a ^{31}P MR spectroscopy study. *MRM* 1996;35(5):658–663.

82. Silverman DG, Kosten TR, Jatlow PI, et al. Decreased digital flow persists after the abatement of cocaine-induced hemodynamic stimulation. *Anesth Anal* 1997;84:46–50.

83. Johnson B, Barron B, Fang B, et al. Isradipine prevents global and regional cocaine-induced changes in brain blood flow: a preliminary study. *Psychopharmacology* 1998;136(4):335–341.

84. Gottschalk PCH, Seibyl J, Rinder H, Kosten TR. Treatment of cocaine-related cerebral perfusion deficits with anti-platelet agents: SPM analyses of early abstinence. Yale University School of Medicine and VA-Connecticut Healthcare System. *Proc Coll Prob Drug Depend* 2000;60(Suppl. 1):S117; (abst).

85. Volkow ND, Hitzemann R, Fowler JS, et al. Regional brain metabolic activity during different stages of cocaine withdrawal. *J Nucl Med* 1991;32:960.

86. Fox BS. Development of a therapeutic vaccine for the treatment of cocaine addiction. *Drug Alcohol Depend* 1997;48:153–158.

87. Sparenborg S, Vocci F, Zukin S. Peripherial cocaine-blocking agents: new medications for cocaine dependence. *Drug Alcohol Depend* 1997;48:149–151.

88. Crits-Christoph P, Siqueland L, Blaine J, et al. The National Institute on Drug Abuse collaborative cocaine treatment study: rationale and methods. *Arch Gen Psychiatry* 1997;54(8):721–726.

89. Crits-Christoph P, Siqueland L, Chittams J, et al. Training in cognitive, supportive-expressive, and drug counseling therapies for cocaine dependence. *J Consult Clin Psychol* 1998;66(3):484–492.

90. Crits-Christoph P, Siqueland L, Blaine J, et al. Psychosocial treatments for cocaine dependence: National Institute on Drug Abuse collaborative cocaine treatment study. *Arch Gen Psychiatry* 1999;56(6):493–502.

91. Maude-Griffin PM, Hohenstein JM, Humfleet GL, et al. Superior efficacy of cognitive-behavioral therapy for urban crack cocaine abusers: main and matching effects. *J Consult Clin Psychol* 1998;66:832–837.

92. General Accounting Office. *Cocaine treatment: early results from various approaches*. Washington, DC: GAO, 1996.

93. Carroll KM. Relapse prevention as a psychosocial treatment approach: a review of controlled clinical trials. *Exp Clin Psychopharmacol* 1996;4:46–54.

94. Carroll KM, Nich C, Rounsaville BJ. Differential symptom reduction in depressed cocaine abusers treated with psychotherapy and pharmacotherapy. *J Nerv Ment Dis* 1995;183:251–259.

95. Higgins ST, Delaney DD, Budney AJ, et al. A behavioral approach to achieving initial cocaine abstinence. *Am J Psychiatry* 1991;148:1218–1224.

96. Higgins ST, Budney AJ, Bickel WK, et al. Achieving cocaine abstinence with a behavioral approach. *Am J Psychiatry* 1993;150:763–769.

97. Higgins ST, Budney AJ, Bickel WK, et al. Incentives improve outcome in outpatient behavioral treatment of cocaine dependence. *Arch Gen Psychiatry* 1994;51:568–576.

98. Silverman K, Higgins ST, Brooner RK, et al. Sustained cocaine abstinence in methadone maintenance patients through voucher-based reinforcement therapy. *Arch Gen Psychiatry* 1996;53:409–415.

99. Silverman K, Wong CJ, Umbricht-Schneiter A, et al. Broad beneficial effects of reinforcement of cocaine abstinence in methadone patients. *J Consult Clin Psychol* 1998;66:803–810.

100. Petry NM, Martin B, Cooney JL, et al. Give them prizes, and they will come: contingency management for treatment of alcohol dependence. *J Consult Clin Psychol* 2000;68:250–257.

101. Petry NM. A comprehensive guide to the application of contingency management procedures in clinical settings. *Drug Alcohol Depend* 2000;58:9–25.

102. Iguchi MY, Belding MA, Morral AR, et al. Reinforcing operants other than abstinence in drug abuse treatment: an effective alternative for reducing drug use. *J Consult Clin Psychol* 1997;65:421–428.

103. Kirby KC, Marlowe DB, Festingjer DS, et al. Schedule of voucher delivery influences initiation of cocaine abstinence. *J Consult Clin Psychol* 1998;66:761–767.

104. Griffith JD, Rowan-Szal GA, Roark RR, et al. Contingency management in outpatient methadone treatment: a meta-analysis. *Drug Alcohol Depend* 2000;58:55–66.

103

Application of Imaging Technologies in the Investigation of Drug Addiction

Nora D. Volkow

Joanna S. Fowler

Nora D. Volkow: Medical Department, Brookhaven National Laboratory, Upton, New York.

Joanna S. Fowler: Chemistry Department, Brookhaven National Laboratory, Upton, New York.

Brain imaging can be used to assess the following in the human brain: (a) morphology [computed tomography (CT) and magnetic resonance imaging (MRI)]; (b) electrical and magnetic signals [electroencephalography (EEG) and magnetoencephalography (MEG)]; (c) neurotransmission [positron emission tomography (PET) and single photon emission computed tomography (SPECT)]; (d) tissue composition [magnetic resonance spectroscopy (MRS)]; and (e) blood flow and metabolism [functional MRI (fMRI), PET, SPECT, and dynamic CT]. Table 103.1 summarizes the spatial and temporal resolution and the sensitivity for the various imaging modalities.

	Parameter Measured	Temporal Resolution	Spatial Resolution	Sensitivity
MEG	Function	1 ms	5 mm	
EEG		1 ms	10–15 mm	
CT	Structure	ms		
MRI	Structure	ms	1.0–1.5 mm	10^{-3} molar
	Function	3–5 s	mm	
	Biochemistry	10–20 min	cm	
[b]PET	Function	45 s	4 mm	10^{-2} molar
[b]SPECT	Biochemistry	15 min	4 mm	
	Pharmacokinetics	60 s	4 mm	

[a]The spatial and temporal resolution and the sensitivity cited for PET and MRI correspond to those of currently available commercial instruments. Research instruments have been developed that have better performance.
[b]Require the use of radiotracers and hence repeated studies with these modalities are limited by radiation dosimetry to the subjects.
CT, computed axial tomography; EEG, electoencephalography; MEG, magnetoencephalography; MRI, magnetic resonance imaging; PET, positron emission tomography; SPECT, single photon emission computed tomography.

TABLE 103.1. IMAGING MODALITIES USED TO INVESTIGATE THE LIVING HUMAN BRAIN a

This chapter focuses mainly on the application of PET, SPECT, and MRI for the investigation of the effects of drugs of abuse in the human brain and their relationship with their reinforcing, addictive, and toxic effects. A brief description of these imaging techniques follows.

PET and SPECT are nuclear medicine instruments that detect and measure the spatial distribution and movement of radioisotopes in tissues of living subjects. PET measures compounds labeled with positron emitting radioisotopes and SPECT with single photon emitting radioisotopes. An advantage of the positron emitters is that some of these are isotopes for the natural elements of life (^{11}C, ^{15}O, ^{13}N), and this feature enables labeling of compounds without affecting their pharmacologic properties. Although labeling an organic compound with a single photon emitter such as ^{123}I results in a compound that is different from the parent compound, many iodine-substituted radiotracers with high biological selectivity and affinity for specific molecular targets have been developed. The positron emitters used for imaging have shorter half-lives than the single photon emitters. Both types of isotopes can be used to label ligands for specific receptor, transporter, or enzymatic systems to be used with PET or SPECT to quantify these parameters in living human brains. In addition, PET tracers such as [^{18}F] or [^{11}C]-labeled deoxyglucose (FDG, CDG) and [^{15}O]-labeled water can be used to measure regional brain glucose metabolism and cerebral blood flow (CBF), and SPECT tracers such as ^{99m}Tc hexamethylpropyleneamineoxime (HMPAO) can be used to measure CBF.

MRI is an imaging instrument that can distinguish elements in tissue on the basis of their magnetic properties. This information can be used to obtain images that reflect brain structure, brain function, or chemical composition. Information on structure in the brain can be obtained on the basis of differences in chemical composition between gray and white matter. For structural brain imaging, this is mostly accomplished by proton analysis, which enables the assessment of the water content of tissues. Information on brain function is derived from the differences in magnetic properties of oxygenated versus deoxygenated hemoglobin (blood oxygenation-dependent or BOLD contrast). During activation of a brain region, an excess of arterial blood is delivered into the area, with concomitant changes in the ratio of deoxyhemoglobin to oxyhemoglobin. Concentration on a wide variety of compounds that reflect metabolic state of the tissue and cell integrity can be obtained with MRS. MRS can also be used to measure the concentration and metabolism of compounds such as ^{13}C-glucose.

The most widespread application of imaging in the study of drugs of abuse has been its use to assess brain function, which can be done using imaging modalities that measure electrical activity, CBF, or brain metabolism. Of the modalities used for functional imaging, fMRI has the highest spatial resolution. Conversely, MEG and EEG are the imaging technologies with the highest temporal resolution, which enables the examiner to assess the temporal displacement of activation signals as they propagate in brain on the order of a few milliseconds (1).

The effect of drugs of abuse on neurotransmission has also been investigated. This effect depends on biochemical processes that occur at very low concentrations (nanomolar-picomolar range). PET and SPECT have the highest sensitivity of all currently available imaging techniques, and they can measures concentrations in the nanomolar-picomolar range, which are the physiologic concentrations at which neurotransmitter processes occur (2 ,3 and 4).

PHARMACOLOGIC PROPERTIES OF DRUGS OF ABUSE IN THE HUMAN BRAIN

Part of "103 - Application of Imaging Technologies in the Investigation of Drug Addiction "

The investigation of the pharmacologic properties of drugs entails studies of their pharmacokinetics (primarily using PET and the [^{11}C]-labeled drug) as well as their pharmacodynamics (using PET or SPECT and a radiotracer with specificity for a particular molecular or biochemical target or using PET, SPECT, and fMRI to assess brain function). Because these studies are done in awake human subjects, one can investigate the relationship between the behavioral effects of drugs and their effects on brain function and neurochemistry.

Pharmacokinetics

PET can be used to measure the absolute uptake, their regional distribution, and the kinetics of [^{11}C]-labeled drugs in the human brain. Moreover, the labeled drug can also be used to determine the target organs for the drug and thus can provide information on potential organ toxicity. Table 103.2 shows the various addictive drugs that have been labeled with a positron emitter and whose distribution has been evaluated with PET.

Drug Class	Specific Drug	Reference or Review
Psychostimulants	Cocaine	5,6
	Methylphenidate	113
	Metamphetamine	114
Opiates	Morphine	115
	Heroin	115
	Codeine	115
	Buprenorphine	116
	Methadone	117
Cannabinoids	THC	118
Nicotine	Nicotine	119–120
Caffeine	Caffeine	112
LSD	LSD	121

TABLE 103.2. DRUGS WITH ABUSE LIABILITY THAT HAVE BEEN LABELED WITH A POSITRON EMITTER (CARBON-11)

An example of the value of this strategy is its use in the investigation of the pharmacokinetics of cocaine in the human brain, as assessed with [^{11}C]cocaine (5), and a comparison with methylphenidate (MP), a drug used in the treatment of attention-deficit disorder that, like cocaine, blocks the dopamine (DA) transporter (DAT) but is much less abused than cocaine, as assessed with [^{11}C]MP (6). Cocaine and MP were found to have a large brain uptake (7% to 10% injected dose) and to have an almost identical pattern of distribution in the human brain, where they bound

predominantly to the striatum and where the specific binding for both drugs was to the DATs. Both drugs had a very fast rate of uptake, with peak concentrations in striatum achieved for cocaine between 4 and 6 minutes and for MP between 6 and 10 minutes after injection (6). However, their clearances differed; MP's clearance from striatum (half-life longer than 90 minutes from peak uptake) was significantly slower than that of cocaine's (half-life of 20 minutes from peak uptake) (Fig. 103.1). For both drugs, their fast uptake in striatum paralleled the temporal course for the experience of "high" reported by subjects given pharmacologic doses of intravenous cocaine or of MP. However, whereas for cocaine the rate of clearance paralleled the decline in the "high," for MP the "high" declined while there was still significant binding of the drug in brain (Fig. 103.1). Because it was the "rate of uptake" that was associated with the "high" for both drugs and not the presence of the drug in brain, investigators postulated from this observation that the rate of clearance may affect the propensity of a drug to promote frequent repeated administration. Although the rate at which psychostimulants enter the brain had been recognized as an important variable in their reinforcing effects (7), the relevance of their rate of clearance had not. These pharmacokinetic studies provided evidence that the rate of drug clearance is relevant in their reinforcing effects. In the case of cocaine, the fast rate of clearance enables repeated, frequent administration that is characteristic of cocaine bingeing (cocaine is taken every 15 to 30 minutes), whereas for MP, its relatively slow clearance from brain is likely to produce accumulation and toxicity that thus prevents frequent repeated administration.

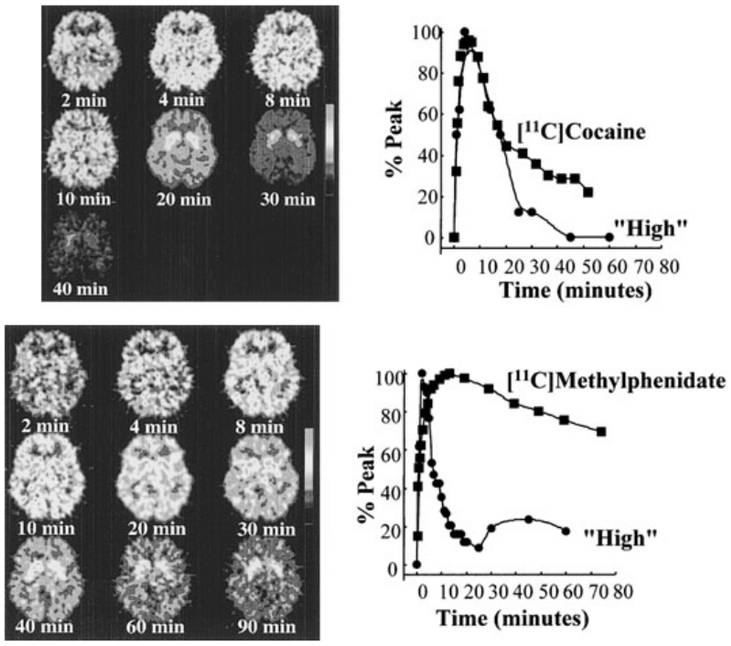


FIGURE 103.1. Left: Images at the level of the striatum obtained with [^{11}C]cocaine and with [^{11}C]methylphenidate at different times after radiotracer injection. **Right:** Time activity curves for radiotracer concentration in striatum and temporal course for the "high" expressed as a percentage from peak after pharmacologic doses of intravenous cocaine **(upper panel)** and of intravenous methylphenidate **(lower panel)**. (Modified from Volkow ND, Ding U, Fowler JS, et al. Is methylphenidate like cocaine? Studies on their pharmacokinetics and distribution in human brain. *Arch Gen Psychiatry* 1995;52:456–463, with permission.) See color version of figure.

Pharmacodynamics

Multiple parameters pertaining to the mechanisms of action of the drug of abuse can be investigated with imaging. These include measurement of the efficacy of the drug of abuse at the molecular target that is associated with the reinforcing effects of the drug of abuse (i.e., DAT for cocaine) and assessment on the effects of the drug of abuse on DA concentration and on brain function. These parameters can be assessed both in nonaddicted control subjects and in addicted patients to determine whether there are differences in the responses between them.

Drug Efficacy

The efficacy of the drug of abuse at the molecular target is illustrated with studies done to investigate the levels of DAT blockade achieved by reinforcing doses of cocaine. With PET and appropriate radiotracers, it is possible to measure the levels of DAT occupancy achieved by drugs that block DAT in human subjects reproducibly (8). The levels of DAT occupancy by different doses of intravenous cocaine were assessed with PET and [^{11}C]cocaine in active cocaine abusers (9). This study showed that cocaine is very effective in blocking DAT; at the doses commonly used by cocaine abusers (0.3 and 0.6 mg/kg), cocaine blocked more than 60% of the DAT. This study also showed that the higher the levels of DAT blockage, the higher the intensity of the "high," and that for cocaine to induce a "high" it had to block more than 60% of DAT function. A similar study done with intravenous MP showed that the ED_{50} (the dose required to block 50% of the DAT) was half that of cocaine (MP, 0.075 mg/kg; cocaine, 0.13 mg/kg) (10). As for cocaine, the magnitude of the DAT occupancy was significantly associated with the intensity of the "high," a finding corroborating the importance of DAT blockade in the "high." The differences in the ED_{50} between cocaine and MP are compatible with differences in their affinities for DAT (the inhibition constant or K_i of DA uptake corresponds to 640 and 390 nM, respectively) (11). In analyzing the implications of the similar *in vivo* efficacy for DAT blockade by cocaine and MP, regarding the low abuse potential of MP, it is important to emphasize that the similarities were observed after intravenous administration, which

is not the route of administration used in the treatment of attention-deficit/hyperactivity disorder. Because the rapidity of drug effects is an important variable in the reinforcing effects of drugs of abuse (12) and routes of administration affect drug pharmacokinetics, the results with intravenous MP cannot be extrapolated to oral MP.

The high levels of DAT occupancy achieved by cocaine and MP contrast with the results obtained for other drugs of abuse such as benzodiazepines. SPECT studies measuring the levels of receptor occupancy by the benzodiazepine drug lorazepam showed that only a few receptors are occupied at pharmacologic doses (13), findings that support the notion that in humans there is a “reserve” of benzodiazepine receptors.

Effects on Dopamine Concentration

Because the ability of drugs of abuse to increase extracellular DA concentration is considered crucial for their reinforcing effects, the estimation of DA changes becomes particularly relevant. PET and SPECT enable one to carry such measures in the human brain using radioligands that bind with relatively low affinity to DA D2 receptors (i.e., [^{11}C]raclopride, [^{123}I]IBZM) and compete with DA for binding to DA receptors. For this purpose, subjects are scanned twice, at baseline and after administration of the drug of abuse, and the difference in the binding of the radioligand between both conditions is mostly a reflection of drug induced changes in extracellular DA. Studies to measure changes in DA concentration induced by drugs of abuse in the human brain have been carried out for amphetamine, cocaine, and MP (14 ,15 and 16). These studies showed that these three psychostimulant drugs significantly increase extracellular DA, and, in the case of intravenous MP, the magnitude of druginduced DA changes was closely correlated with the intensity of the self-reports of “high.” In fact, for intravenous MP, the changes in DA were a better predictor of the “high” than the levels of DAT blockade, a finding that indicates that the effects of DAT blocker drugs are not only a function of the levels of DAT blockade but also of the amount of DA being released by the terminal (17).

Effects on Regional Brain Function

The most widely used imaging approach for the investigation of drugs of abuse has been to assess the effects of acute drug administration on brain glucose metabolism or CBF. This allows analysis of the brain regions that are most sensitive to the effects of the drug, and because the studies are done in awake human study subjects, it allows an analysis of the relation between regional changes in metabolism or flow and the behavioral effects of the drug. Although most drugs of abuse decrease regional brain glucose metabolism, their effects on CBF are increased by some drugs and decreased by others. This discrepancy between metabolism and CBF is probably an indication of the vasoactive properties of many of these pharmacologic agents, a property that is relevant for understanding their toxicity as it relates to cerebrovascular disease. The discrepancy could also reflect the finding that changes in metabolism reflect an average of the changes that occur over the uptake period of FDG (30 to 35 minutes), whereas those from blood flow reflect activity that occurs between 2 and 5 seconds for fMRI and 60 seconds for PET and [^{15}O]water.

CHRONIC EFFECTS OF DRUGS OF ABUSE IN THE HUMAN BRAIN

Part of "103 - Application of Imaging Technologies in the Investigation of Drug Addiction "

Imaging studies have been done to assess neurochemical and functional changes in the brain of addicted subjects that are associated with the process of addiction as well as changes associated with drug toxicity. Functional imaging strategies have also been used to assess the brain region involved in drug-related states such as drug craving. (See Chapter 110 .)

Drug Toxicity

Drug toxicity can be assessed with imaging techniques for brain as well as for other organs. Toxicity from drugs has been documented in abusers of cocaine, methamphetamine, and ecstasy, and the findings from these studies are covered under the subsection of the drug class. In addition, the ability to label the drug with a positron emitter and to follow its distribution in the human body and the availability of radiotracers that allow one to monitor organ function provide a mechanism for evaluating potential toxicity of drugs to organs other than brain. For example, PET studies done with [^{11}C]cocaine have shown significant accumulation in human heart (18), a finding leading to the question whether this could be associated with cocaine’s cardiotoxic properties. Cocaine was shown to induce a long-lasting inhibition of the norepinephrine transporter in heart using 6-[^{18}F]fluoronorepinephrine or [^{11}C]hydroxyephedrine (19 ,20). Cocaine is a local anesthetic, and its accumulation in heart could result in direct myocardial toxicity. At the same time, inhibition of the norepinephrine transporter by cocaine interferes with a protective mechanism of the heart to remove circulating catecholamines. Thus, these two separate mechanisms operating in parallel are likely to contribute to the highly cardiotoxic properties of cocaine.

Cocaine

Toxicity

Studies using PET and [^{15}]-labeled water provided the first documentation of abnormalities in CBF in cocaine abusers

(Fig. 103.2) (21). The patchy distribution of these CBF defects in brain suggested that they were secondary to cocaine's vasoactive effects (e.g., vasoconstriction, and platelet aggregation), rather than its regional neuroactive properties. These imaging findings, which appeared as defects seen with small strokes and hemorrhages, corroborated clinical reports of strokelike symptoms associated with cocaine use. These PET findings were subsequently replicated in several SPECT studies of CBF in chronic cocaine abusers (reviewed in ref. 22). More recent studies with MRI documented hyperintense lesions in white matter suggestive of subclinical anoxic vascular events in cocaine abusers that were also ascribed to the vasoactive effects of cocaine (23 ,24). The vasoconstricting effects of cocaine in human brain were corroborated by MRI studies showing significant reductions in cerebral blood volume (23%) (25) and CBF after acute cocaine administration (26).

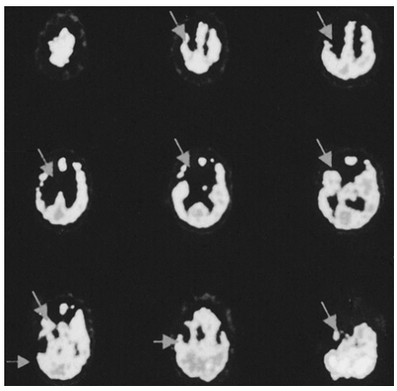


FIGURE 103.2. Brain images obtained with PET and [^{15}O]water to measure cerebral blood flow in a cocaine abuser. Note the regional decreases in blood flow *(arrows)*. (Modified from Volkow ND, Mullani N, Gould KL, et al. Cerebral blood flow in chronic cocaine users: a study with positron emission tomography. *Br J Psychiatry* 1988;152:641–648, with permission.) See color version of figure.

Using MRS, it is possible to assess tissue composition differentially for neurons and glial cells in brain. This information can be used, in turn, to determine whether there is neuronal damage or glial proliferation. In cocaine abusers, MRS studies reported increases in total creatine (+7%) and myoinositol (+18%) in white matter but no changes in *N*-acetyl aspartate, which is a marker for neuronal content (27). This finding was interpreted as reflecting alterations of nonneuronal but not of neuronal cells in cocaine abusers.

Brain Glucose Metabolism and Function

In contrast to the marked defects in CBF reported in cocaine abusers, the functional changes as assessed with brain glucose metabolism are not as pronounced and vary significantly as a function of detoxification. Also different from the patchy distribution of the CBF defect are the decrements in metabolism that tend to localize to cortical projections of the DA system. In recently detoxified cocaine abusers (less than 1 week), brain glucose metabolism was reported to be significantly higher in orbitofrontal cortex and in striatum than in control subjects (28). Metabolic activity was highest in subjects tested during the initial 72 hours after withdrawal, and cocaine abusers who had the highest brain metabolic values had also the highest subjective ratings for craving. In contrast, cocaine abusers tested between 1 and 4 months of detoxification showed significant reductions in metabolic activity in prefrontal cortex, orbitofrontal cortex, and anterior cingulate gyrus (Fig. 103.3) (29). Thus, the orbitofrontal cortex, which is hypermetabolic during early cocaine discontinuation, becomes hypometabolic with protracted cocaine withdrawal.

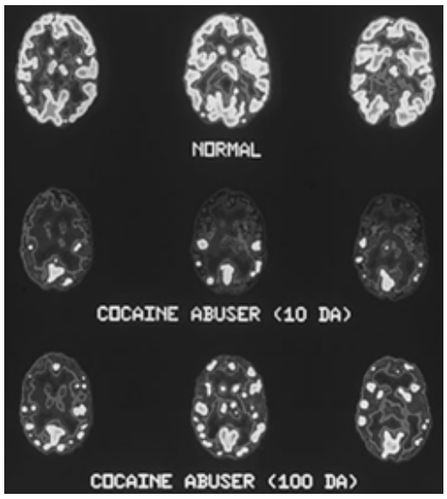


FIGURE 103.3. Images at the level of the striatum obtained with PET and FDG to measure regional brain glucose metabolism in a control subject and in a cocaine abuser tested at two different time points of the detoxification. Notice the marked reduction in metabolism in the frontal cortex. See color version of figure.

In addition to the studies measuring resting metabolism or CBF, the effects of pharmacologic challenges in cocaine abusers have been compared with controls. Intravenous cocaine was found to reduce brain glucose metabolism in cortical and subcortical brain regions as measured by FDG PET (30). In contrast, an fMRI study of acute cocaine administration revealed widespread activation in various cortical and subcortical brain regions, and the temporal course paralleled that of cocaine-induced "rush" (31). Because other acute pharmacologic interventions used to study cocaine abusers were for the most part chosen to target a specific neurotransmitter system, these aspects are discussed under the appropriate neurotransmitter headings.

Dopamine System

Ii has been hypothesized that decreased DA activity could underlie cocaine addiction (32). PET studies done to assess whether there are changes in DA brain activity in cocaine

abusers have used a multitracer approach to assess the relationship between levels of DA D2 receptors and regional brain metabolism in cocaine abusers during early cocaine withdrawal and after cocaine detoxification. Studies in cocaine abusers tested during early cocaine withdrawal (less than 1 week) revealed significant decreases in DA D2 receptor availability when compared with controls (Fig. 103.4) (33).

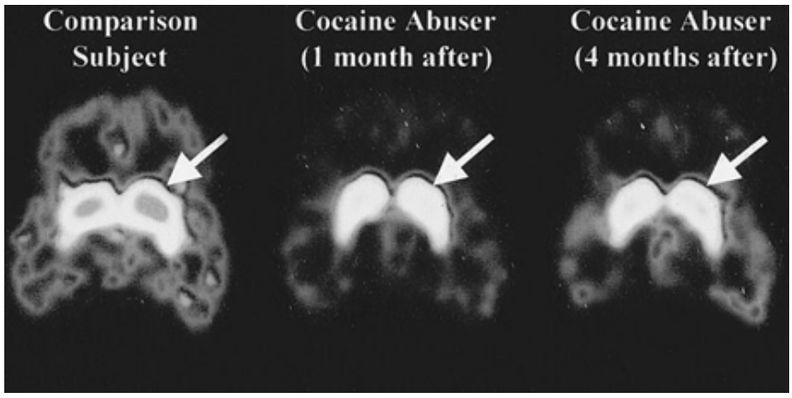


FIGURE 103.4. Images at the level of the striatum obtained with PET and [^{18}F]NMSP to measure dopamine D2 receptors in a control subject and in a cocaine abuser tested at two different time points of the detoxification. Notice the marked reduction in dopamine D2 receptors during both early and protracted detoxification. See color version of figure.

Studies in cocaine abusers tested between 1 and 4 months of detoxification also showed significant reductions in DA D2 receptor availability as assessed with [^{11}C]raclopride (34). The reductions in DA D2 receptor availability persisted on repeated testing 3 months after completing the inpatient detoxification program. In these patients, the reductions in DA D2 receptors were significantly correlated with metabolic activity in prefrontal cortex, orbitofrontal cortex, and cingulate gyrus (34). Lower values for D2 receptors were associated with lower metabolism in orbitofrontal cortex, cingulate gyrus, and prefrontal cortex, a finding suggesting an association between DA activity and the function of these frontal brain regions. The persistence of the decreased D2 function raise the question of long-term cocaine-induced changed versus preexisting DA system deficits that could increase vulnerability to cocaine dependence.

Abnormalities in orbitofrontal cortex and cingulate gyrus have also been reported for patients with obsessive-compulsive disorders (35), with whom cocaine abusers share the compulsive quality of their behaviors. In patients with obsessive-compulsive disorders, this feature manifests itself in specific behavioral rituals, and in cocaine abusers, it manifests as an obsession to procure the drug and in the repetitive pattern of cocaine self-administration. In laboratory animals, destruction of the orbitofrontal cortex leads to the emergence of repetitive behaviors that cannot be easily terminated (36), and a similar syndrome can be generated by the destruction of the mesocortical DA pathway (37). Thus, it has been postulated that DA-mediated dysregulation of the orbitofrontal cortex and the anterior cingulate gyrus may be one of the mechanisms responsible for the compulsive administration of cocaine during a “binge” and for the loss of control experienced by the drug abusers when exposed to cocaine or cocaine-related cues (38).

Studies in cocaine abusers to assess the DA terminal have mostly used ligands for the DAT. The results from studies are not consistent; PET studies done with [^{11}C]cocaine as a DAT ligand in actively abusing as well as in detoxified cocaine abusers have failed to show changes in DAT availability (reviewed in ref. 39), whereas SPECT studies done with [^{123}I]β-CIT showed significant DAT increases (17% to 25%) during states of acute (up to 96 hours) drug abstinence (40). Studies with [^{18}F]6-fluorodopa (6-FDOPA), which is an index of DA synthesis that also serves as a marker of the DA terminal, revealed significant reductions in recently detoxified cocaine abusers (11 to 30 days) when compared with control subjects or early abstinent cocaine addicts (1 to 10 days) (41). Reasons for the disparity between these studies are likely to reflect not only differences in the subjects studied but also differences in the effects of cocaine on the targets studied (i.e., it is possible that cocaine increases the expression of DAT while decreasing the synthesis of DA in the terminal). The discrepancies between the two types of DAT studies could reflect differences between the radiotracers and the models used. Because of the disparities, the effects of cocaine on the DA terminal are still not clear. However, because no study has documented reductions in DAT in cocaine abusers, this provides with evidence that cocaine does not induce degeneration of the DA terminal in humans.

Studies in cocaine abusers to assess DA release by the DA terminal have been done using PET and the DA D2 radioligand [^{11}C]raclopride. Studies to assess DA release were performed with and without administration of MP, which is a drug that, like cocaine, blocks DAT. In humans, the measures of MP-induced DA changes are reproducible (42), and they are similar in magnitude to those induced by equivalent doses of cocaine (43). Studies comparing the changes in [^{11}C]raclopride binding between cocaine abusers and control subjects showed that the response of cocaine abusers was 50% lower than that of controls. The “high” induced by intravenous MP was also more intense in controls than in cocaine abusers, whereas in cocaine abusers but not in controls, MP induced intense cocaine craving. This finding indicates that cocaine-dependent patients release less DA in the striatum and have a blunted “high” relative to controls when they are given MP. These results provide evidence that cocaine addiction does not imply an enhanced pleasurable response nor is there a sensitized DA response to the drug. Rather, the reduced DA release and blunted “high” are compatible with cross-tolerance between cocaine and intravenous MP.

The marked decrease in DA brain function in the cocaine abusers (reduction in DA D2 receptors, DA synthesis, and release) may lead to a decrease in activation of DA-modulated

reward circuits that are important in drive and motivation. Thus, one could postulate that the decreased in DA activity in cocaine abusers may make normal reinforcers less effective, and these patients may be taking the drug to compensate for the decreased stimulation of DA reward pathways. The decrease in DA function may also contribute to the dysphoria and the anhedonia experienced by these patients during cocaine withdrawal. Thus, strategies to enhance DA brain function in cocaine abusers may help these individuals to engage in activities that may help them to avoid a relapse.

GABA System

Cocaine enhances DA brain activity, and DA signals are transferred by γ-aminobutyric acid (GABA)ergic pathways (44). These make the GABA pathways a particularly susceptible target for cocaine's effects. PET studies have shown significant reductions in striatal DA D2 receptors in cocaine abusers (33 and 34). Because D2 receptors are predominantly located on GABA cells (45), reductions of these receptors suggest involvement of GABA pathways in cocaine abusers. The GABA system has been evaluated in cocaine abusers with functional imaging techniques. These studies assessed the brain regional responsivity to GABA stimulation in cocaine abusers and controls (46). Brain responsivity to GABA stimulation was assessed by measuring the brain metabolic responses to lorazepam, a drug that facilitates GABA neurotransmission. Although plasma lorazepam concentration was significantly higher in controls that in drug abusers, lorazepam-induced sleepiness in cocaine abusers was significantly more intense than in controls. Lorazepam reduced whole-brain metabolism, the decrements were greater in drug abusers (21±3 %) than in controls (13±7 %), and the differences were largest in striatum, thalamus, and parietal cortex. Because lorazepam-induced sleepiness was correlated with changes in thalamic metabolism, this finding suggests that the increased sedation in cocaine abusers results from the enhanced sensitivity of the thalamus to lorazepam. These results support the notion of disruption of GABA activity in the brain of cocaine abusers. The extreme sedative effects observed for some of the cocaine abusers after lorazepam administration should alert clinicians to potential untoward reactions in the use of these drugs in active cocaine abusers.

MRS was used to assess the concentration of GABA levels in brain comparing cocaine abusers and controls (47). GABA measurements were localized to the occipital cortex (volume, 9 cc). The cocaine abusers showed a significant decrease (23%) in GABA in comparison with controls. In contrast, macromolecule levels were not significantly different between controls and cocaine abusers. These data corroborate an involvement of cerebral GABA levels in cocaine abusers.

Opioid System

The endogenous opioid system has been implicated in the reinforcing actions of cocaine and other addictive drugs. μ-Opioid receptor binding was measured in cocaine-dependent subjects using PET and [^{11}C]carfentanil (48). μ-Opioid binding was increased in several brain regions of the cocaine addicts in proportion to the severity of cocaine craving experienced at the time. The up-regulation of μ-opioid receptor binding persisted after 4 weeks of detoxification. These findings provide evidence for the involvement of the opioid system in cocaine addiction.

Alcohol

Imaging studies in patients with alcoholism have been done to measure CBF, brain glucose metabolism (baseline and with pharmacologic challenges), benzodiazepine receptors, DA D2 receptors, and DATs and serotonin transporters in brain.

Brain Metabolism and Cerebral Blood Flow

Most of the nonstructural imaging studies have been done to investigate brain metabolic and CBF changes in patients with chronic alcoholism with and without neurologic impairment (reviewed in refs. 49 and 50). Patients with alcoholism and Korsakoff encephalopathy showed decreased metabolism in prefrontal, parietal, and temporal cortices, and patients with alcoholism and neurologic symptoms other than Korsakoff encephalopathy showed decreased metabolism in frontal and parietal cortices. Studies in patients with alcoholism who have no evidence of neurologic impairment have also consistently shown evidence of frontal abnormalities (reviewed in ref. 51). Decrements in metabolism were most accentuated in the older patients with alcoholism with longer histories of alcohol consumption. The degree of brain metabolic recovery with detoxification was evaluated with PET in patients with alcoholism who were evaluated at different times of the detoxification (52 ,53). These studies showed that brain metabolism increased significantly during detoxification, predominantly during the first 16 to 30 days of detoxification. However, decreased metabolic activity in orbitofrontal cortex persisted (Fig. 103.5) (9). Most PET studies in patients with alcoholism have been done in male patients, and little is known about changes in female patients with alcoholism. One PET study measured regional brain metabolism in recently detoxified female patients with alcoholism and compared it with that in age-matched female controls (54). This study showed no differences between patients with alcoholism and female control subjects. These results did not support a higher toxicity for the effects of alcohol in the female than in the male brain, in which most studies have consistently reported lower metabolism in the frontal region. However, this study was confounded

by the finding that the severity of alcohol use in these female patients with alcoholism was less than that of the male patients with alcoholism previously investigated in PET studies. The female subjects in this study were mostly premenopausal, and thus the lack of metabolic abnormalities may have reflected not only the lower alcohol severity but also the protective effects of estrogens.

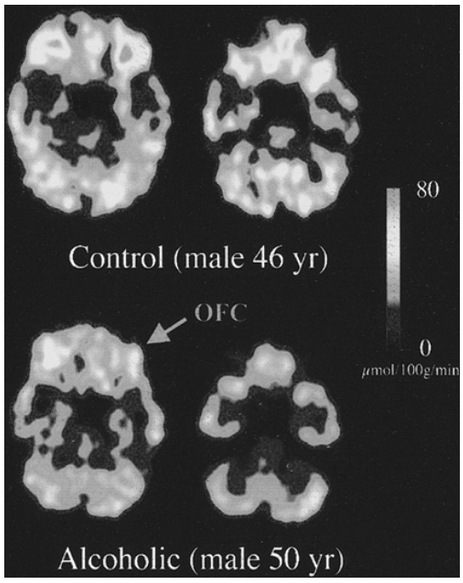


FIGURE 103.5. Images at the level of the orbitofrontal cortex obtained with PET and FDG to measure regional brain glucose metabolism in a control subject and in detoxified patient with alcoholism. Notice the reduction in metabolism in the orbitofrontal cortex. See color version of figure.

In addition to the studies measuring resting metabolism or CBF in patients with alcoholism, multiple studies have been done comparing regional brain metabolic and CBF responses to various pharmacologic challenges between control subjects and patients with alcoholism. Because most of the pharmacologic interventions were chosen to target a specific neurotransmitter system, we will discuss the findings from these studies under the neurotransmitter heading. In the case of alcohol, which is believed to exert its effects through multiple neurotransmitter systems, its effects on brain glucose metabolism (55 ,56) and CBF (57) were evaluated with PET. Such studies showed that acute alcohol administration decreased brain glucose metabolism (Fig. 103.6). When compared with controls, patients with alcoholism showed a significantly larger reduction in metabolism despite showing less subjective response to the intoxicating properties of ethanol (55). In control subjects but not in patients with alcoholism, the subjective response to the intoxicating effects of ethanol was significantly correlated with the brain metabolic decrements (55). This seemingly paradoxic response in patients with alcoholism was interpreted as reflecting their tolerance to ethanol-induced decrements in metabolism.

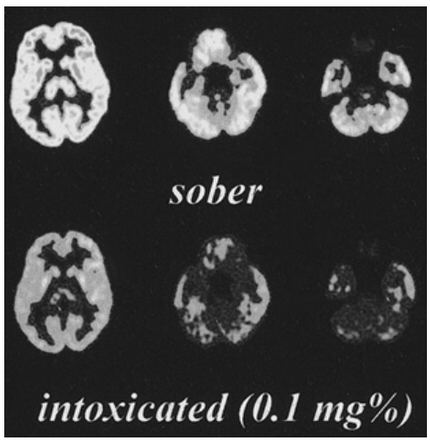


FIGURE 103.6. Images at the level of the striatum and of the orbitofrontal cortex and cerebellum obtained with PET and FDG to measure the effects of acute alcohol administration (1 g/kg orally) on brain glucose metabolism in a control subject. Notice the marked reduction in metabolism during alcohol intoxication. See color version of figure.

GABA System

The effects of benzodiazepines, which facilitate GABA neurotransmission, on brain glucose metabolism (58 ,59 and 60) and CBF (61) have also been evaluated with PET. Such studies have shown that, similar to ethanol, benzodiazepines decrease brain glucose metabolism, and the effects are more pronounced in the occipital cortex, the area of the brain with the highest density of benzodiazepine receptors (62). Benzodiazepines also decrease CBF, and as for metabolism, the largest changes are observed in the occipital cortex. Studies comparing the response to benzodiazepines between control subjects and patients with alcoholism showed a significantly lower response in the orbitofrontal cortex in patients with alcoholism than in controls (46 ,53). The blunted response to lorazepam in orbitofrontal cortex persisted after detoxification, a finding suggesting that it was not the result of withdrawal and that hypoactivity in this brain region may in part reflect abnormal GABAergic function (benzodiazepines facilitate GABAergic neurotransmission).

Imaging studies have also been done to measure changes in benzodiazepine receptors using PET and [^{11}C]flumazenil in patients with alcoholism. One study showed that although there were no changes in the levels of receptor between controls and patients with alcoholism, the latter had a significantly larger variability for B_{max} (receptor concentration) measures than controls (63). However, more recent studies have consistently reported significant decreases in

benzodiazepine receptors in patients with alcoholism, predominantly in frontal brain regions including cingulate gyrus and orbitofrontal cortex (64 ,65), but also in cerebellum (66). The reductions in benzodiazepine receptors in patients with alcoholism reported by these imaging studies are consistent with postmortem studies and may indicate either a toxic effect of alcoholism on benzodiazepine receptors or a vulnerability factor for developing alcoholism. The reductions in benzodiazepine receptors in the orbitofrontal cortex of patients with alcoholism could account for the blunted response to benzodiazepines reported in this brain region (53 ,59).

Dopamine System

DA D2 receptors were evaluated in patients with alcoholism with PET and [^{11}C]raclopride and showed significant reductions in DA D2 receptor availability when compared with controls (67 ,68). No significant correlations were found between DA D2 receptor availability and days of detoxification (1 to 68 weeks). One of these studies also measured DATs in a subgroup of the alcoholics in whom reductions in DA D2 receptors were detected and reported no changes in DAT availability (68). This finding was interpreted as evidence of GABAergic involvement in patients with alcoholism because DA D2 receptors in striatum are mainly localized in GABA cells.

DAT availability in patients with alcoholism has been measured by various PET and SPECT studies. The results have not been consistent. SPECT studies with [^{123}I]β-CIT reported that a group of violent patients with alcoholism had increases (8%) and nonviolent patients with alcoholism had decreases in DATs (25%) when compared with controls (69). SPECT studies in nonviolent patients with late-onset alcoholism have also reported a reduction in DATs (70). However, a PET study done with [^{11}C]D-threo-MP and a SPECT study done with [^{123}I]β-CIT showed no changes in DATs in patients with alcoholism (71 ,68). These discrepancies are likely to reflect in differences in the time since detoxification. One SPECT study showed that whereas DAT levels were significantly reduced in patients with alcoholism within the first few days of last alcohol use, the levels returned to normal 4 weeks after detoxification (72). PET studies with 6-[^{18}F]-FDOPA (a marker for the DA synthesis in the DA terminal) in patients late-onset (type 1) alcoholism showed higher striatal 6-[^{18}F]-FDOPA uptake in the patients with alcoholism than in the controls, a finding that was interpreted as a compensatory mechanism to low postsynaptic DA function (73).

Serotonin System

The effects of m-chlorophenylpiperazine (mCPP), a mixed serotonin agonist-antagonist drug, on brain glucose metabolism was compared in patients with alcoholism and in controls. This study showed that mCPP-induced activation in thalamus, orbitofrontal cortex, caudate, and middle frontal gyrus was significantly blunted in patients with alcoholism when compared with controls (74). This finding was interpreted as reflecting a hyporesponsive striatothalamoorbitofrontal circuit in patients with alcoholism. The abnormal response to mCPP suggests an involvement of the serotonin system in the abnormalities seen in this circuit in patients with alcoholism.

The availability of serotonin transporters, which serve as markers for the serotonin terminals, was measured with SPECT and [^{123}I]β-CIT in patients with alcoholism. This study showed a significant reduction in the availability of brainstem serotonin transporters in patients with alcoholism that was significantly correlated with lifetime alcohol consumption and with ratings of depression and anxiety during withdrawal (75). As for the prior study, this finding provides evidence of a role for serotonin in alcoholism and in its involvement with depressive symptoms during withdrawal.

Opioid System

The effects of an oral naltrexone challenge on CBF in patients with alcoholism during detoxification was studied with SPECT and HMPAO. At baseline, patients with alcoholism showed lower CBF in left orbitofrontal cortex and prefrontal cortex than controls. After naltrexone, a significant regional CBF decrease was found in basal ganglia and the left mesial temporal region, which are structures rich in opioid receptors. These results were interpreted as supporting the involvement of the opioid system in alcohol dependence (76).

Spectroscopic Studies

Patients with alcoholism had a significant reduction of the cerebellar *N*-acetylaspartate–to-creatine ratio, which was interpreted as reflecting neuronal loss and a reduction of the choline-to-creatine ratio, which was interpreted as reflecting cell membrane modification or myelin alterations (77).

Subjects at Risk of Alcoholism

The regional brain metabolic response to lorazepam was evaluated in subjects with a positive family history of alcoholism (FHP) and was compared with that of subjects without a family history of alcoholism (FHN) (78). At baseline, FHP subjects showed lower cerebellar metabolism than FHN, and when challenged with lorazepam, they also showed a blunted response in cerebellum and in anterior cingulate gyrus. Lorazepam-induced changes in cerebellar metabolism were significantly correlated with motor impairment. The blunted cerebellar sensitivity to benzodiazepines in FHP could account for the decreased sensitivity to the motor effects of alcohol and benzodiazepines in FHP subjects.

The decreased cerebellar baseline metabolism in FHP subjects as well as the blunted cerebellar response to lorazepam challenge may reflect disrupted activity of benzodiazepine-GABA receptors in cerebellum of FHP subjects.

Opiates

The effects of morphine on brain glucose metabolism were evaluated in polydrug abusers (79). This study showed that morphine reduced glucose metabolism by 10% in whole brain and by about 5% to 15% in telencephalic areas and the cerebellar cortex. Morphine-induced metabolic decrements were not significantly related to subjective measures of euphoria. The effects of acute fentanyl, a synthetic opiate, on CBF were measured with PET and [^{15}O]water. Fentanyl administration was associated with significant increases in regional CBF in cingulate, orbitofrontal, and medial prefrontal cortices, as well as caudate nuclei, areas known to be involved in reward and addiction (80).

Resting CBF was measured with SPECT and ^{99m}Tc-HMPAO in heroin abusers tested 1 week after opiate discontinuation and then retested 2 weeks later (81). The initial scans demonstrated significant CBF defects in the frontal, parietal, and temporal cortices. Two weeks later, the SPECT scans showed improvement. The results from this study provide evidence that long-term use of opiates results in perfusion abnormalities that are partially reversible with short-term abstinence.

Dopamine System

Using PET and [^{11}C]raclopride, opioid-dependent subjects were found to have lower DA D2 receptor availability than controls. Naloxone-induced withdrawal in opioid-dependent subjects did not change [^{11}C]raclopride binding, a finding indicating that withdrawal does not alter synaptic DA in the striatum as measured by this method (82).

Opioid System

To date, there have been no published studies of opioid abusers using these opiate receptor radioligands to study heroin abusers (see earlier).

Spectroscopic Studies

Phosphorus magnetic resonance spectroscopy (^{31}P MRS) at 1.5 T was performed on polysubstance abusing men (cocaine and heroin dependence) (83). The phosphorus metabolite signal expressed as a percentage of total phosphorus signal was 15% higher for phosphomonoesters, 10% lower for nucleotide triphosphates (ß-NTP), and 7% lower for total nucleotide phosphates in polydrug abusers compared with controls. These findings were interpreted as suggesting that long-term drug abuse or withdrawal results in changes in cerebral high-energy phosphates and in phospholipid metabolites.

Marijuana

Marijuana is the most widely used illicit drug of abuse in the United States. Despite its widespread use, the mechanisms by which ð^{9}-tetrahydrocannabinol (THC) (the main psychoactive substance of marijuana) exerts its psychoactive effects are still not known. Relatively few imaging studies have been done to assess the effects of acute and chronic marijuana use in the human brain.

Brain Metabolism and Cerebral Blood Flow

SPECT studies assessed the effect of THC intoxication on CBF in chronic marijuana users (84 ,85). Acute marijuana administration led to decreases in CBF in subjects who were not experienced marijuana smokers, whereas it increased CBF in subjects who were experienced smokers. In a more recent study, these investigators extended these findings to a larger groups of subjects and documented increases in CBF in anterior cingulate gyrus and in the insula in marijuana users (86). Interpretation of the effects of THC on CBF is confounded by the vasoactive properties of THC (87). Thus, it is difficult to separate the effects of THC that are related to its action on nervous tissue from those that are related to its vasoactive effects. This problem is obviated when using deoxyglucose to measure brain glucose metabolism because this agent is insensitive to fluctuations in CBF (88).

The effects of THC on regional brain glucose metabolism have been evaluated in nonabusing controls (89) as well as in marijuana abusers (90). The whole-brain metabolic response to the effects of THC was variable among individuals; some subjects had increases, some had decreases, and some did not show change. Despite these variable responses in whole-brain metabolism, there was a very consistent pattern of metabolic activation by THC. That is, under THC intoxication, most of the subjects showed activation of the cerebellum. The cerebellar activation by THC was significant both for the absolute and for the relative measures. In marijuana abusers, THC also increased metabolism in the anterior cingulate gyrus and in the orbitofrontal cortex. A more recent study assessing the effects of marijuana on CBF also reported an increase in cerebellar flow during intoxication (91). The highly localized concentration of cannabinoid receptors in the cerebellum (92) supports involvement of the cannabinoid receptors in the metabolic and CBF response during THC intoxication. Activation of the cerebellum by THC could explain the disruption in motor coordination and proprioception during THC intoxication. Cannabinoid receptors are also localized in other discrete areas, namely, hippocampus, substantia nigra, pars reticulata,

and globus pallidus. These are too small to be measured with the spatial resolution of the PET instrument used.

SPECT studies compared CBF in subjects with attention-deficit disorder who had a history of marijuana abuse with those who did not (93). Decreased perfusion in the prefrontal cortex was seen in both marijuana users and nonusers. However, the marijuana users also demonstrated marked decreased activity in the temporal lobes, which was ascribed to chronic marijuana use.

Cannabinoid Receptors

Attempts to investigate THC in the living brain with PET by using the labeled drug with a positron emitter have been unsuccessful because of the highly lipophilic nature of THC. This was also a limitation for δ^8-tetrahydrocannabinol, an analog of δ^9-THC, which was labeled with ^{18}F. Its uptake and distribution showed widespread uptake in the baboon brain with no particular pattern of localization (94). This pattern of distribution mainly reflected nonspecific binding because pretreatment with δ^8-THC did not affect the uptake of [^{18}F]δ^8-THC in brain. A promising alternative may be the use of THC antagonists with high receptor affinities (95).

Cigarettes and Nicotine

Even though there are 45 million cigarette smokers in the United States and there are 400,000 deaths per year associated with smoking, surprisingly little is known about the neurochemical actions of tobacco smoker exposure on the human brain, and very few imaging studies have been performed. The acute administration of intravenous nicotine has been reported to reduce brain glucose metabolism (96). In addition, PET studies with [^{11}C]nicotine have shown that cigarette smokers have increased binding in brain that could reflect the up-regulation in nicotinic receptors sites reported in smokers (97).

Monoamine oxidase A and B (MAO A and B) have been examined in the human brain (98 ,99). MAO breaks down neurotransmitter amines like DA, serotonin, and norepinephrine, as well as amines from exogenous sources. It occurs in two subtypes, MAO A and MAO B, which can be imaged *in vivo* using [^{11}C]clorgyline and [^{11}C]L-deprenyl-D2 and PET. Using these ligands, it was shown that cigarette smokers have a reduction in brain monoamine oxidase B (MAO B) of about 40% relative to nonsmokers and former smokers, and smokers have a 28% reduction in brain MAO A relative to nonsmokers. It is known that nicotine does not inhibit MAO B at physiologically relevant levels. MAO A and B inhibition is associated with enhanced activity of DA, a neurotransmitter involved in reinforcing and motivating behaviors and in movement as well as decreased production of hydrogen peroxide, a source of reactive oxygen species. Inhibition of MAO by cigarette smoke could be one of the mechanisms accounting for the lower incidence of Parkinson disease in cigarette smokers. MAO A and B inhibition by smoke may also account for some of the epidemiologic features of smoking that include a higher rate of smoking in persons with depression and addiction to other substances. In this regard, it is noted that MAO A inhibitors are effective in the treatment of depression.

Ecstasy

The toxicology of the popular illicit drug ecstasy, 3,4-methylenedioxymethamphetamine (MDMA), is covered in Chapter 108 . Studies in laboratory animals have shown that ecstasy induces neurotoxicity to serotonergic neurons. Ecstasy users imaged with the PET ligand [^{11}C]McN-5652 (for 5-hydroxytryptamine transporters), showed decreased global and regional brain 5-hydroxytryptamine transporter binding compared with controls (100). SPECT studies with [^{123}I]β-CIT (a radioligand for DAT and serotonin transporters) confirmed the reductions in serotonin transporters in ecstasy users (101) and provided preliminary evidence of serotonergic neurotoxicity by ecstasy in humans.

Because the cerebrovasculature is regulated partly by the serotonergic system, it was questioned whether ecstasy would affect CBF in humans. For this purpose, CBF was measured with SPECT in ecstasy users tested at baseline and after receiving MDMA (102). Abstinent ecstasy users showed no baseline CBF changes when compared with controls. However, within 3 weeks after MDMA administration, CBF remained decreased in the visual cortex, the caudate, and the superior parietal and dorsolateral frontal regions compared with baseline values. These reductions were interpreted as reflecting transient effects of MDMA on the serotonergic system or the indirect effects of its metabolites on the DA system.

PET and FDG were also used to measure regional brain glucose metabolism in ecstasy users (103). Ecstasy users showed altered activity in amygdala, hippocampus, and Brodmann's area II, findings interpreted as suggesting a long-term effect of ecstasy on brain function.

Spectroscopic studies with proton MRS done in ecstasy users showed normal *N*-acetyl compounds in all brain regions but showed increases in myoinositol concentration (+16.3%) and the myoinositol-to-creatine ratio (+14.1%) in parietal white matter (104). The finding of a normal *N*-acetyl concentration was interpreted as a lack of neuronal injury in recreational ecstasy users, and the increase in myoinositol was seen as an increase in glial content.

Methamphetamine

Methamphetamine is a particularly problematic drug not only in that it is highly addictive but also because animal

studies have shown that it is neurotoxic to DA cells (105). Because the pattern and doses of methamphetamine administered to laboratory animals differ from those used by drug abusers, imaging studies have been performed to determine whether similar pathologic features occur in human methamphetamine abusers. The data in humans are very limited: a postmortem study of 12 methamphetamine abusers (106) and two PET studies, one of six (107) and the other of 15 methamphetamine abusers (108). Both reported reductions in brain DATs. Moreover, for one of the studies, the reductions in DAT were associated with motor slowing and memory impairment and were present even in patients who had been detoxified for more than 11 months (108). These results provide evidence that methamphetamine, at the doses administered by humans, damages the DA terminals, that these effects are long lasting, and that the damage from methamphetamine is functionally significant.

Spectroscopic studies were done with proton MRS in abstinent methamphetamine abusers and showed that the concentration of *N*-acetylaspartate, a neuronal marker, was reduced significantly (-5 to -6%) in the basal ganglia and frontal white matter of methamphetamine users compared with controls (109). The frontal white matter (*N*-acetylaspartate) correlated inversely with the logarithm of the lifetime methamphetamine use. The methamphetamine users also showed significantly reduced total creatine in the basal ganglia (-8%) and increased choline-containing compounds (+13%) and myoinositol (+11%) in the frontal gray matter. These findings were interpreted as providing evidence of long-term neuronal damage in abstinent methamphetamine users.

Caffeine

Caffeine is an antagonist of the G-protein–linked adenosine receptors. Caffeine is a mild stimulant when taken orally, but it has marked effects when administered intravenously, including olfactory hallucinations.

Despite the widespread consumption of caffeine and the potential for confounding effects in studies of other drugs of abuse, caffeine appears to be the subject of only a single PET CBF study. Caffeine (250 mg), given either intravenously or orally, produced an approximately 30% decrease in global CBF measured using [^{15}O]water, without significant regional differences (110). The effect of caffeine lasted for at least 45 minutes. The original purpose of this study was to examine whether patients with panic disorder exhibited a difference response from control subjects, but no differences were found.

The effects of caffeine on regional brain metabolism, as assessed by the changes in brain lactate, were measured with MRS (111). The response of heavy caffeine users was compared with that of caffeine-intolerant persons. Subjects were studied at baseline and 1 hour after caffeine (10 mg/kg). The caffeine-intolerant group but not the heavy caffeine users showed significant increases in brain lactate. Reexposure of the regular caffeine users to caffeine after a caffeine holiday resulted in rises in brain lactate similar in magnitude to those seen in the caffeine-intolerant group. These results provide evidence of caffeine tolerance in the human brain and do not support the role of lactate as a mediator of caffeine intolerance.

Studies using [^{11}C]caffeine showed that its binding in brain was mostly nonspecific, as was expected because of caffeine's low affinity and lack of selectivity for adenosine receptor subtypes (112). Intravenously administered [^{11}C]caffeine resulted in very fast uptake and clearance from brain, contrasted with the slow brain uptake when [^{11}C]caffeine was given orally through a nasogastric tube (brain uptake was increasing even after 2 hours at the end of the study).

CONCLUSION

Part of "103 - Application of Imaging Technologies in the Investigation of Drug Addiction "

Brain imaging using virtually all available methods has proved useful in evaluating the effects of abused drugs. Much has been learned about the mechanisms of action in human subjects as well as the potential for toxic effects. Among the consistent findings across the various drugs of abuse are the following:

1. The pharmacokinetic properties of drugs of abuse affect their reinforcing effects.
2. Many of the drugs of abuse have vasoactive effects, which, in the case of cocaine, can result in cerebrovascular disease.
3. The orbitofrontal cortex and the anterior cingulate gyrus have consistently been shown to be abnormal in addicted subjects, a finding implicating a role in the process of drug addiction.
4. The availability of DA D2 receptors is reduced in most drug abusers who have been investigated. Because DA D2 receptors modulate reward circuits, this could be one of the mechanisms that contributes to drug self-administration.

As new ligands and new methods are developed, an improved understanding of the mechanisms of addiction can be expected.

ACKNOWLEDGMENTS

Part of "103 - Application of Imaging Technologies in the Investigation of Drug Addiction "

This research was supported in part by the US Department of Energy (office of Health and Environmental Research) under Contract DE-ACO2-76CH00016, the Institute of Drug Abuse under grant nos. DA 06891, DA 09490, DA

062278, and the Institute of Alcohol Abuse and Alcoholism, under grant no. AA 09481.

REFERENCES

1. Lounasmaa OV, Hämäläinen M, Hari R, et al. Information processing in the human brain: magnetoencephalographic approach. *Proc Natl Acad Sci USA* 1996;93:8809–8815.

2. Phelps ME, Mazziotta JC, Schelbert H. *Positron emission tomography and autoradiography: principles and applications forthe brain and heart*. New York: Raven, 1986.

3. Jones T. The role of positron emission tomography within the spectrum of medical imaging. *Eur J Nucl Med* 1996;23:207–211.

4. Volkow ND, Fowler JS, Gatley JS, et al. Evaluation of the human brain dopamine system with PET. *J Nucl Med* 1996;37:1242–1256.

5. Fowler JS, Volkow ND, Wolf AP, et al. Mapping cocaine binding in human and baboon brain in vivo. *Synapse* 1989;4:371–377.

6. Volkow ND, Ding U, Fowler JS, et al. Is methylphenidate like cocaine? Studies on their pharmacokinetics and distribution in human brain. *Arch Gen Psychiatry* 1995;52:456–463.

7. Oldendorf WH. Some relationships between addiction and drug delivery to the brain. In: Frankenheim J, Brown RM, eds. *Bioavailability of drugs to the brain and the blood brain barrier*. Research monograph 120. Rockville, MD: National Institute on Drug Abuse, 1992.

8. Fowler JS, Volkow ND, Logan J, et al. Measuring dopamine transporter occupancy by cocaine: radiotracer considerations. *Synapse*, 1998;28:111–116.

9. Volkow ND, Wang GJ, Fischman MW, et al. Relationship between subjective effects of cocaine and dopamine transporter occupancy. *Nature* 1997;386:827–830.

10. Volkow ND, Wang GJ, Fowler JS, et al. Blockade of striatal dopamine transporters by intravenous methylphenidate is not sufficient to induce self reports of “high.” *J Pharmacol Exp Ther* 1999;288:14–20.

11. Ritz MC, Lamb RJ, Goldberg SR, et al. Cocaine receptors on dopamine transporters are related to self-administration of cocaine. *Science* 1987;237:1219–1223.

12. Balster RL, Schuster CR. Fixed-interval schedule of cocaine reinforcement: effects of dose and infusion duration. *J Exp Anal Beh* 1973;20:119–129.

13. Sybirska E, Seibyl JP, Bremner JD, et al. [^{123}I]iomazenil SPECT imaging demonstrates significant benzodiazepine receptor reserve in human and nonhuman primate brain. *Neuropharmacology* 1993;32:671–680.

14. Volkow ND, Wang G-J, Fowler JS, et al. Imaging endogenous dopamine competition with [^{11}C]raclopride in the human brain. *Synapse* 1994;16:255–262.

15. Laruelle M, D’Souza CD, Baldwin RM, et al. Imaging D2 receptor occupancy by endogenous dopamine in humans. *Neuropsychopharmacology* 1997;17:162–174.

16. Schlaepfer TE, Pearlson GD, Wong DF, et al. PET study of competition between intravenous cocaine and [^{11}C]raclopride at dopamine receptors in human subjects. *Am J Psychiatry* 1997;154:1209–1213.

17. Volkow ND, Wang G-J, Fowler JS, et al. Prediction of reinforcing responses to psychostimulants in humans by brain dopamine D2 receptor levels. *Am J Psychiatry* 1999;56:1440–1443.

18. Volkow ND, Fowler JS, Wolf AP, et al. Distribution of ^{11}C-cocaine in human heart, lungs, liver and adrenals: a dynamic PET study. *J Nucl Med* 1992;33:521–525.

19. Fowler JS, Ding YS, Volkow N, et al. PET studies of cocaine inhibition of the myocardial norepinephrine uptake. *Synapse* 1994;16:312–317.

20. Melon PG, DeGrado TR, Nguyen N, et al. Non-invasive assessment of regional myocardial sympathetic neuronal function following intravenous cocaine injection in dogs. *J Nucl Med* 1992;33:994.

21. Volkow ND, Mullani N, Gould KL, et al. Cerebral blood flow in chronic cocaine users: a study with positron emission tomography. *Br J Psychiatry* 1988;152:641–648.

22. Holman BL, Mendelson J, Garada B, et al. Regional cerebral blood flow improves with treatment in chronic cocaine polydrug users. *J Nucl Med* 1993;34:723–727.

23. Bartzokis G, Beckson M, Hance DB, et al. Magnetic resonance imaging evidence of “silent” cerebrovascular toxicity in cocaine dependence. *Biol Psychiatry* 1999;45:1203–1211.

24. Holman BL, Mendelson J, Garada B, et al. Regional cerebral blood flow improves with treatment in chronic cocaine polydrug users. *J Nucl Med* 1993;34:723-727.

25. Kaufman MJ, Levin JM, Maas LC, et al. Cocaine decreases relative cerebral blood volume in humans: a dynamic susceptibility contrast magnetic resonance imaging study. *Psychopharmacology (Berl)* 1998;138:76-81.

26. Gollub RL, Breiter HC, Kantor H, et al. Cocaine decreases cortical cerebral blood flow but does not obscure regional activation in functional magnetic resonance imaging in human subjects. *J Cereb Blood Flow Metab* 1998;18:724-734.

27. Chang L, Mehringer CM, Ernst T, et al. Neurochemical alterations in asymptomatic abstinent cocaine users: a proton magnetic resonance spectroscopy study. *Biol Psychiatry* 1997;42:1105–1114.

28. Volkow ND, Fowler JS, Wolf AP, et al. Changes in brain glucose metabolism in cocaine dependence and withdrawal. *Am J Psychiatry* 1991;148:621–626.

29. Volkow ND, Hitzemann R, Wang G-J, et al. Long-term frontal brain metabolic changes in cocaine abusers. *Synapse* 1992;11:184–190.

30. London ED, Cascella NG, Wong DF, et al. Cocaine induced reduction of glucose utilization in human brain: a study using positron emission tomography and [fluorine 18]fluorodeoxyglucose. *Arch Gen Psychiatry* 1990;47:567–574.

31. Breiter HC, Gollub RL, Weisskoff RM, et al. Acute effects of cocaine on human brain activity and emotion. *Neuron* 1997;19:591–611.

32. Dackis CA, Gold MS. New concepts in cocaine addiction: the dopamine depletion hypothesis. *Neurosci Biobehav Rev* 1985;9:469–477.

33. Volkow ND, Fowler JS, Wolf AP, et al. Effects of chronic cocaine abuse on postsynaptic dopamine receptors. *Am J Psychiatry* 1990;147:7l9-724.

34. Volkow ND, Fowler JS, Wang G-J, et al. Decreased dopamine D2 receptor availability is associated with reduced frontal metabolism in cocaine abusers. *Synapse* 1993;14:169–177.

35. Baxter LR, Phelps ME, Mazziotta J. Local cerebral glucose metabolic rates in obsessive compulsive disorder: a comparison with rates in unipolar depression and normal controls. *Arch Gen Psychiatry* 1987;44:211–218.

36. Kolb B. Studies on the caudate putamen and the dorsomedial thalamic nucleus: implications for mammalian frontal lobe function. *Physiol Behav* 1977;18:237–244.

37. Le Moal M, Simon H. Mesocorticolimbic dopaminergic network: functional and regulatory notes. *Physiol Rev* 1991;71:155–234.

38. Volkow ND, Fowler JS. Addiction, a disease of compulsion and drive: involvement of the orbitofrontal cortex. *Cereb Cortex* 2000;10:318–325.

39. Wang G-J, Volkow ND, Fowler JS, et al. Cocaine abusers do not show age-related losses of dopamine transporters. *Life Sci* 1997;61:1059–1065.

40. Malison RT, Best SE, van Dyck CH, et al. Elevated striatal dopamine transporters during acute cocaine abstinence as measured by [^{123}I] beta-CIT SPECT. *Am J Psychiatry* 1998;155:832–834.

41. Wu JC, Bell K, Najafi A, et al. Decreasing striatal 6-FDOPA uptake with increasing duration of cocaine withdrawal. *Neuropsychopharmacology* 1997;7:402–409.

42. Wang G-J, Volkow ND, Fowler JS, et al. Reproducibility of repeated measures of endogenous dopamine competition with [C-11]raclopride in the human brain. *J Nucl Med* 1999;40:1285–1291.

43. Volkow ND, Fowler JS, Gatley SJ, et al. Comparable changes in synaptic dopamine induced by methylphenidate and by cocaine in the baboon brain. *Synapse* 1998;31:59–66.

44. Scheel-Krüger J. Dopamine-GABA interactions: evidence that GABA transmits, modulates and mediates dopaminergic functions in the basal ganglia and the limbic system. *Acta Neurol Scand* 1986;73:9–49.

45. Aubert I, Ghorayeb I, Normand E, et al. Phenotypical characterization of the neurons expressing the D1 and D2 dopamine receptors in the monkey striatum. *J Comp Neurol* 2000;418:22–32.

46. Volkow ND, Wang G-J, Fowler JS, et al. Enhanced sensitivity to benzodiazepines in active cocaine abusing subjects: a PET study. *Am J Psychiatry* 1998;155:200–206.

47. Hetherington HP, Pan JW, Telang F, et al. Reduced brain GABA levels in cocaine abusers. *Int Soc Magn Reson Med* 2000: 523.

48. Zubieta JK, Gorelick DA, Stauffer R, et al. Increased mu opioid receptor binding detected by PET in cocaine-dependent men is associated with cocaine craving. *Nat Med* 1996;2:1225–1229.

49. Volkow ND, Wang G-J, Doria JJ. Monitoring the brain's response to alcohol with positron emission tomography. *Alcohol Health Res World* 1996;19:196–299.

50. Lyons D, Whitlow CT, Smith HR, et al. Brain imaging: functional consequences of ethanol in the central nervous system. *Recent Dev Alcohol* 1998;14:253–284.

51. Mathew RJ, Wilson WH. Substance abuse and cerebral blood flow. *Am J Psychiatry* 1991;148:292–305.

52. Volkow ND, Wang G-J, Hitzemann R, et al. Recovery of brain glucose metabolism in detoxified alcoholics. *Am J Psychiatry* 1994;151:178–183.

53. Volkow ND, Wang GJ, Overall JE, et al. Regional brain metabolic response to lorazepam in alcoholics during early and late alcohol detoxification. *Alcohol Clin Exp Res* 1997;21:1278–1284.

54. Wang GJ, Volkow ND, Fowler JS, et al. Regional cerebral metabolism in female alcoholics of moderate severity does not differ from that of controls. *Alcohol Clin Exp Res* 1998;22:1850–1854.

55. DeWit H, Metz J, Wagner N, et al. Behavioral and subjective effects of ethanol: relationship to cerebral metabolism using PET. *Alcohol Clin Exp Res* 1990;14:482–489.

56. Volkow ND, Hitzemann R, Wolf AP, et al. Acute effects of ethanol on regional brain glucose metabolism and transport. *Psychiatry Res* 1990;35:39–48.

57. Volkow ND, Guynn R, Marani S, et al. Effects of alcohol on cerebral blood flow. *Psychiatry Res* 1988;24:201–209.

58. DeWit H, Metz J, Wagner N, et al. Effects of diazepam on cerebral metabolism and mood in normal volunteers. *Neurospyschopharmacology* 1991;5:33–41.

59. Volkow ND, Wang G-J, Hitzemann R, et al. Decreased cerebral response to inhibitory neurotransmission in alcoholics. *Am J Psychiatry* 1993;150: 417–422.

60. Volkow ND, Wang G-J, Hitzemann R, et al. Depression of thalamic metabolism by lorazepam is associated with sleepiness. *J Neuropsychopharmacol* 1995;12:123–132.

61. Matthew E, Andreason P, Pettigrew K, et al. Benzodiazepine receptors mediate regional blood flow changes in the living human brain. *Proc Natl Acad Sci USA* 1995;92:2775–2779.

62. Braestrup C, Albrechtsen R, Squires RF. High densities of benzodiazepine receptors in human cortical areas. *Nature* 1977;269:702–704.

63. Litton JE, Neiman J, Pauli S, et al. PET anlysis of [^{11}C]flumazenil binding to benzodiazepine receptors in chronic alcohol-dependent men and healthy controls. *Psychiatry Res* 1993;50:1–13.

64. Lingford-Hughes AR, Acton PD, Gacinovic S, et al. Reduced levels of GABA-benzodiazepine receptor in alcohol dependency in the absence of grey matter atrophy. *Br J Psychiatry* 1998;173:116–122.

65. Gilman S, Koeppe RA, Adams K, et al. Positron emission tomographic studies of cerebral benzodiazepine-receptor binding in chronic alcoholics. *Ann Neurol* 1996;40:163–171.

66. Abi-Dargham A, Krystal JH, Anjilvel S, et al. Alterations of benzodiazepine receptors in type II alcoholic subjects measured with SPECT and [^{123}I]iomazenil. *Am J Psychiatry* 1998;155:1550–1555.

67. Hietala J, West C, Syvälahti E, et al. Striatal D2 dopamine receptor binding characteristics in vivo in patients with alcohol dependence. *Psychopharmacology* 1994;116:285–290.

68. Volkow ND, Wang GJ, Fowler JS, et al. Decreases in dopamine receptors but not in dopamine transporters in alcoholics. *Alcohol Clin Exp Res* 1996;20:1594–1598.

69. Tiihonen J, Kuikka J, Bergström K, et al. Altered striatal dopamine re-uptake site densities in habitual violent and non-violent alcoholics. *Nat Med* 1995;1:654–657.

70. Repo E, Kuikka JT, Bergström KA, et al. Dopamine transporter and D2-receptor density in late-onset alcoholism. *Psychopharmacology (Berl)* 1999;147:314–318.

71. Heinz A, Goldman D, Jones DW, et al. Genotype influences *in vivo* dopamine transporter availability in human striatum. *Neuropsychopharmacology* 2000;22:133–139.

72. Laine TP, Ahonen A, Torniainen P, et al. Dopamine transporters increase in human brain after alcohol withdrawal. *Mol Psychiatry* 1999;4:189–191.

73. Tiihonen J, Vilkman H, Rasanen P, et al. Striatal presynaptic dopamine function in type 1 alcoholics measured with positron emission tomography. *Mol Psychiatry* 1998;3:156–161.

74. Hommer D, Andreasen P, Rio D, et al. Effects of m-chlorophenylpiperazine on regional brain glucose utilization: a positron emission tomographic comparison of alcoholic and control subjects. *J Neurosci* 1997;17:2796–2806.

75. Heinz A, Ragan P, Jones DW, et al. Reduced central serotonin transporters in alcoholism. *Am J Psychiatry* 1998;155:1544–1549.

76. Catafau AM, Etcheberrigaray A, Perez de los Cobos J, et al. Regional cerebral blood flow changes in chronic alcoholic patients induced by naltrexone challenge during detoxification. *J Nucl Med* 1999;40:19–24.

77. Seitz D, Widmann U, Seeger U, et al. Localized proton magnetic resonance spectroscopy of the cerebellum in detoxifying alcoholics. *Alcohol Clin Exp Res* 1999;23:158–163.

78. Volkow ND, Wang GJ, Begleiter H, et al. Regional brain metabolic response to lorazepam in subjects at risk for alcoholism. *Alcohol Clin Exp Res* 1995;19:510–516.

79. London ED, Broussolle EP, Links JM, et al. Morphine-induced metabolic changes in human brain. Studies with positron emission tomography and [fluorine 18]fluorodeoxyglucose. *Arch Gen Psychiatry* 1990;47:73–81.

80. Firestone LL, Gyulai F, Mintun M, et al. Human brain activity response to fentanyl imaged by positron emission tomography. *Anesth Analg* 1996;82:1247–1251.

81. Rose JS, Branchey M, Buydens-Branchey L, et al. Cerebral perfusion in early and late opiate withdrawal: a technetium-99m-HMPAO SPECT study. *Psychiatry Res* 1996;67:39–47.

82. Wang GJ, Volkow ND, Fowler JS, et al. Dopamine D2 receptor availability in opiate-dependent subjects before and after naloxone-precipitated withdrawal. *Neuropsychopharmacology* 1997;16:174–182.

83. Christensen JD, Kaufman MJ, Levin JM, et al. Abnormal cerebral metabolism in polydrug abusers during early withdrawal: a ^{31}P MR spectroscopy study. *Magn Reson Med* 1996;35:658–663.

84. Mathew RJ, Wilson WH, Tant SR. Acute changes in cerebral blood flow associated with marijuana smoking. *Acta Psychiatr Scand* 1989;79:118–128.

85. Tunving K, Thulin SD, Risberg J. Regional cerebral blood flow in long term heavy cannabis use. *Psychiatry Res* 1986;17:15–21.

86. Mathew RJ, Wilson WH, Coleman RE, et al. Marijuana intoxication and brain activation in marijuana smokers. *Life Sci* 1997;60:2075–2089.

87. Nahas GG. Toxicolology and Pharmacology. In: Nahas GG, ed. *Marijuana in science and medicine*. New York: Raven, 1986:109.

88. Sokoloff L, Reivich M, Kennedy C, et al. The ^{11}C deoxyglucose method for the measurement of local cerebral glucose utilization: the conscious and anesthetized albino rat. *J Neurochem* 1977;28:897–916.

89. Volkow ND, Gillespie H, Mullani, N, et al. Cerebellar metabolic activation by delta-9-tetrahydrocannabinol in human brain: a study with positron emission tomography and F-18-2 fluoro-2-deoxyglucose. *Psychiatry Res* 1991;40:69–78.

90. Volkow ND, Gillespie H, Mullani N, et al. Brain glucose metabolism in chronic marijuana users during baseline and during marijuana intoxication. *Psychiatry Res* 1996;67:29–38.

91. Mathew RJ, Wilson WH, Turkington TG, et al. Cerebellar activity and disturbed time sense after THC. *Brain Res* 1998;797:183–189.

92. Herkenham M, Lynn AB, Little MD, et al. Cannabinoid receptor localization in brain. *Proc Natl Acad Sci USA* 1990;87:1932–1936.

93. Amen DG, Waugh M. High resolution brain SPECT imaging of marijuana smokers with AD/HD. *J Psychoactive Drugs* 1998;30:209–214.

94. Charalambous A, Marciniak G, Shiue CY, et al. PET studies in the primate brain and biodistribution in mice using (-)-5'-^{18}F-delta 8-THC. *Pharmacol Biochem Behav* 1991;40:503–507.

95. Gatley SJ, Lan R, Volkow ND, et al. Imaging the brain marijuana receptor: development of a radioligand that binds to cannabinoid CB1 receptors *in vivo. J Neurochem* 1998;70:417–423.

96. Stapleton JM, Henningfield JE, Wong DF, et al. Nicotine reduces cerebral glucose utilization in humans. *NIDA Res Monogr* 1993;132:106.

97. Nyback H, Halldin C, Ahlin A, et al. PET studies of the uptake of (S)- and (R)-[^{11}C]nicotine in the human brain: difficulties in visualizing specific receptor binding *in vivo. Psychopharmacology (Berl)* 1994;115:31–36.

98. Fowler JS, Volkow ND, Wang G-J, et al. Inhibition of MAO B in the brains of smokers. *Nature* 1996;379:733–736.

99. Fowler JS, Volkow ND, Wang G-J, et al. Brain monoamine oxidase A inhibition in cigarette smokers. *Proc Natl Acad Sci USA* 1996;93:14065–14069.

100. McCann UD, Szabo Z, Scheffel U, et al. Positron emission tomographic evidence of toxic effect of MDMA ("ecstasy") on brain serotonin neurons in human beings. *Lancet* 1998;352:1433–1437.

101. Semple DM, Ebmeier KP, Glabus MF, et al. Reduced *in vivo* binding to the serotonin transporter in the cerebral cortex of MDMA ("ecstasy") users. *Br J Psychiatry* 1999;175:63–69.

102. Chang L, Grob CS, Ernst T, et al. Effect of ecstasy [3,4-methylenedioxymethamphetamine (MDMA)] on cerebral blood flow: a co-registered SPECT and MRI study. *Psychiatry Res* 2000;98:15–28.

103. Obrocki J, Buchert R, Vaterlein O, et al. Ecstasy: long-term effects on the human central nervous system revealed by positron emission tomography. *Br J Psychiatry* 1999;175:186–188.

104. Chang L, Ernst T, Grob CS, et al. Cerebral (1)H MRS alterations in recreational 3, 4-methylenedioxymethamphetamine (MDMA, "ecstasy") users. *Magn Reson Imaging* 1999;10:521–526.

105. Seiden LS, Sabol KE. Methamphetamine and methylenedioxymethamphetamine neurotoxicity: possible mechanisms of cell destruction. *NIDA Res Monogr* 1996;163:251–276.

106. Wilson JM, Kalasinsky KS, Levey AI, et al. Striatal dopamine nerve terminal markers in human, chronic methamphetamine users. *Nat Med* 1996;2:699–703.

107. McCann UD, Wong DF, Yokoi F, et al. Reduced striatal dopamine transporter density in abstinent methamphetamine and methcathinone users: evidence from positron emission tomography studies with [^{11}C]WIN-35,428. *J Neurosci* 1998;18:8417-8422.

108. Volkow ND, Chang L, Wang G-J, et al. Reduction in brain dopamine transporters in methamphetamine abusers are associated with motor and cognitive changes. *Proc Soc Neurosci* 1999;25:1563.

109. Ernst T, Chang L, Leonido-Yee M, et al. Evidence for long-term neurotoxicity associated with methamphetamine abuse: a 1H MRS study. *Neurology* 2000;54:1344–1349.

110. Cameron OG, Modell JG, Hariharan M. Caffeine and human cerebral blood flow: a positron emission tomography study. *Life Sci* 1990;47:1141–1146.

111. Dager SR, Layton ME, Strauss W, et al. Human brain metabolic response to caffeine and the effects of tolerance. *Am J Psychiatry* 1999;156:229–237.

112. Ding Y-S, Fowler JS, Volkow ND. Synthesis and preliminary PET studies of [^{11}C]caffeine. *J Nucl Med* 1998;39:238P.

113. Volkow ND, Ding Y-S, Fowler JS, et al. Is methylphenidate like cocaine? Studies on their pharmacokinetics and distribution in human brain. *Arch Gen Psychiatry* 1995;52:456–463.

114. Nakamura H, Hishinuma T, Tomioka Y, et al. Effects of haloperidol and cocaine pretreatments on brain distribution and kinetics of [^{11}C]methamphetamine in methamphetamine sensitized dog: application of PET to drug pharmacokinetic study. *Nucl Med Biol* 1997;24:165–169.

115. Hartvig P, Bergström K, Lindberg B, et al. Kinetics of ^{11}C-labeled opiates in the brain of rhesus monkeys. *J Pharmacol Exp Ther* 1984;230:250–255.

116. Galynker I, Schlyer DJ, Dewey SL, et al. Opioid receptor imaging and displacement studies with [6-O-[^{11}C] methyl]buprenorphine in baboon brain. *Nucl Med Biol* 1996;23:325–331.

117. Ding Y-S, Fowler JS, Volkow ND, et al. Synthesis and positron emission tomographic (PET) baboon studies of [^{11}C]methadone and R-(-)-[^{11}C]methadone. *J Nucl Med* 1996;37:40P.

118. Charalambous A, Marciniak G, Shiue CY, et al. PET studies in the primate brain and biodistribution in mice using (-)-5′-^{18}F-delta 8-THC. *Pharmacol Biochem Behav* 1991;40:503–507.

119. Lundqvist H, Nordberg A, Hartvig P, et al. (S)-(-)-[^{11}C]nicotine binding assessed by PET: a dual tracer model evaluated in the rhesus monkey brain. *Alzheimer Dis Assoc Disord* 1998;12:238–246.

120. Bergstrom M, Nordberg A, Lunell E, et al. Regional deposition of inhaled ^{11}C-nicotine vapor in the human airway as visualized by positron emission tomography. *Clin Pharmacol Ther* 1995;57:309–317.

121. Lever JR, Dannals RF, Wilson AA, et al. Synthesis and *in vivo* characterization of D-(+)-(N1-[^{11}C]methyl)-2-Br-LSD: a radioligand for positron emission tomographic studies of serotonin 5-HT2 receptors. *Int J Radiat Appl Instr B* 1989;16:697–704.

104

Molecular and Cellular Neurobiology and Pathophysiology of Opiate Addiction

Mary Jeanne Kreek

Mary Jeanne Kreek: Laboratory of the Biology of Addictive Diseases, Rockefeller University, New York, New York.

In the years 1994 to 2001 (since the last edition of this book went to press), innumerable and important advances have been made in our understanding of the molecular and cellular neurobiology, as well as pathophysiology, of opiate addiction. Clearly, the greatest advances have come about ultimately because of the first successful cloning of a specific opiate receptor, the δ receptor, achieved in late 1992 by two groups working independently, using expression cloning in a cell line that is known to express the δ-opioid receptor (1 ,2). The reports of the groups of Evans and colleagues from Los Angeles and Kieffer and colleagues from Strasbourg, France, followed by the cloning of μ- and κ-opioid receptors of rodents and in humans by Yu, Uhl, and others, opened new doors for both animal and basic clinical research studies, as well as human molecular genetics studies (1 ,2 ,3 ,4 and 5). Other notable technologic advances have been made recently and are continuing to be made. Possibly the most dramatic of these, from which we will undoubtedly see novel and unexpected findings over the next few years, is the development of microarray technology, to determine the changes in levels of gene expression of literally thousands of genes simultaneously (although not yet with the sensitivity required to detect changes in mRNA levels reflecting gene expression of many neuropeptides and most neuroreceptors), and also even newer microarray technology for identification and screening for human polymorphisms, including single nucleotide polymorphisms (SNPs) (6 ,7). By using these new findings and technologies, as well as by building on earlier and current best techniques, profound advances have been made in each of three areas, of which only a few may be covered briefly herein.

Many of these varied advances have been collated and placed in perspectives of our earlier knowledge in several thoughtful reviews, such as selected reviews of preclinical research (6 ,8 ,9 ,10 ,11 ,12 ,13 ,14 ,15 ,16 ,17 ,18 ,19 ,20 ,21 ,22 ,23 ,24 ,25 and 26), basic clinical research (27 ,28 ,29 ,30 and 31), and molecular biology and genetics (32 ,33 ,34 and 35). All these advances have and will continue to make further revelations concerning each of the addictions, and in particular, for this discussion, opiate addiction.

PRECLINICAL STUDIES OF CHRONIC ADMINISTRATION AND WITHDRAWAL EFFECTS OF OPIATES IN DIVERSE AND NOVEL ANIMAL MODELS

Part of "104 - Molecular and Cellular Neurobiology and Pathophysiology of Opiate Addiction "

Neuropeptide and Neurotransmitter Systems Primarily Affected

Opioid Peptides and Receptors: Molecular, Cell Biological, and Signal Transduction Alterations, and Possible Implications for Pathophysiology of Opiate Addiction

After the definitive discovery of specific opioid receptors in 1973, research began to address what had been a long-standing hypothesis, later apparently to be disproved. The hypothesis was that tolerance to opioids depended on down-regulation or decreased availability of, and thus access to, μ-opioid receptors after chronic μ-opioid agonist (e.g., heroin or morphine) exposure. Later, this could be considered to result from “desensitization” of μ-opioid receptors while still on the cell surface (i.e., phosphorylation or uncoupling of the receptors from their G-protein–coupled signal transduction mechanisms essential for the effects after binding), or, alternatively, to result from a decrease in numbers of receptors on the cell surface (i.e., actual down-regulation), which could be caused either by a long-hypothesized, but only recently documented phenomenon, that of endocytosis or internalization of receptors, or by a decrease in the reappearance

of receptors at the cell surface once internalized (36 ,37 ,38 ,39 ,40 ,41 ,42 ,43 ,44 ,45 ,46 ,47 ,48 ,49 ,50 ,51 and 52). Moreover, a significant decrease in production of new receptors could contribute to a so-called down-regulation. Although the terms *down-regulation* and *up-regulation* have been used loosely with inadequate definitions, the overall concept that chronic μ-opioid agonist administration may cause reduced capacity to bind, or increased capacity to bind, or to have no effect, but each alternative with reduced capacity of activated receptors to have an effect (or “tolerance”), has persisted, and repeatedly studied, with conflicting results. The earliest studies to address this issue of impact of chronic opioid administration effects on binding were conducted to elucidate the well-documented and accepted phenomenon of tolerance, both in cell systems and in whole animals. Morphine was the most common opiate used to determine whether opioids down-regulated or otherwise altered opioid-receptor–binding sites.

Later, the effects of specific opioid antagonists on opioid-receptor binding or density were also conducted, primarily using naltrexone. The binding of opioid antagonists, of course, does not involve coupling; that is, no $G_{i/o}$-protein–coupled signal transduction mechanisms are involved, because the opioid antagonists (according to most current theories) do not activate receptors, but rather prevent activation by endogenous or exogenous opioids. The effect of chronic administration of primarily μ-opioid-receptor antagonists to up-regulate the binding capacity, or density, of μ-opioid receptors has been well established. There is now essentially a consensus from many and diverse studies that the chronic administration of opioid antagonists, primarily naltrexone, will cause a significant up-regulation or increase in density of μ-opioid receptors (53 ,54). There has, however, been some controversy regarding whether opioid antagonist treatment and the resultant up-regulation of μ-opioid receptors leads to a sensitized state, that is, a state in which an opioid agonist would have a greater than usual effect on any system. This has been addressed both in animal models and in humans, with some conflicting results.

From the very beginning of the documentation of the existence of specific opiate receptors, in 1973, although numerous studies have used several different opioid agonists, primarily morphine, given by different regimens, ranging from intermittent injections to repeated pellet implantation, to a few studies using chronic administration by pump, there have been conflicting study results and no consensus on the effects on μ-opioid-receptor binding or density. The results reported from studies conducted in living adult animals, for the most part, have shown no overall net changes in μ-opioid-receptor–binding capacity, that is, no overall changes in opioid-receptor density, as measured by quantitative autoradiography or by classic homogenate binding assay studies, and, more recently, no overall changes in μ-opioid-receptor mRNA levels. The original studies, conducted in living adult animals by the groups that included those who first defined opioid receptors, showed no alterations of opioid receptors during chronic morphine exposure, and this finding altered their initial hypothesis, that such chronic exposure to an opioid agonist would cause down-regulation of receptors (55 ,56 and 57). Subsequent studies using diverse ligands and dosing regimens continued to give varied results, with up-regulation of μ-opioid receptors, down-regulation of μ-opioid receptors, and no change of μ-opioid-receptor density or binding after chronic μ-opioid-agonist administration all reported. The prevailing concept for receptor-agonist ligands and, in this case, specifically agonists for the μ-opioid-receptor system, has been that persistent activation of receptors would generally lead to down-regulation, and conversely, the persistent deprivation of receptors of specific ligands would generally lead to persistent lack of activation of receptors and thus to up-regulation. However, the results are complex and conflicting.

From 1996 to 2000, several intriguing articles appeared concerning the effects of opioid-agonist administration on receptor internalization (44 ,45 ,46 ,47 ,48 ,49 ,50 ,51 and 52). In addition, further and relevant studies on signal transduction, primarily through G-protein–coupling mechanisms, but also alternative mechanisms, appeared and extended our earlier knowledge and may ultimately explain some of the apparently conflicting results concerning receptor binding and density (12 ,13 ,14 and 15 ,18 ,19 ,20 and 21 ,23 ,24 ,34 ,36 ,37 ,38 ,39 ,40 ,41 ,42 and 43).

Starting with the early seminal work of Aghajanian in the late 1970s, chronic morphine administration to whole animals was shown to alter, at a cellular level, the function of neurons, specifically in the locus ceruleus, and also to lead to tolerance and physical dependence (12 ,18 ,20 ,24). These changes included initially the well-established acute inhibition by morphine of adenylyl cyclase, as well as resultant inhibition of the cyclic adenosine monophosphate (cAMP)–dependent cascade, followed by, during chronic morphine treatment, a compensatory increase of activity of adenylyl cyclase, with an increase in the cAMP-dependent cascade, including increases in protein kinase A and increases in phosphorylated proteins, as well as of the cAMP-dependent response element binding protein, or CREB (18 ,20 ,24). Nestler and Aghajanian hypothesized that this up-regulation of the entire cAMP pathway in the locus ceruleus represents a compensatory change to oppose or offset the initial inhibitory effects of morphine and thus could be considered to be one component of tolerance (18 ,20). They also suggested that these increases in the cAMP pathway components could contribute to opiate dependence and thus withdrawal, because these changes could be involved in a variety of functions once no longer opposed by morphine (18 ,20). This concept is of particular relevance because this up-regulation of the entire cAMP pathway during chronic morphine exposure has been shown to occur predictably in the locus ceruleus of all strains and species of rodents studied to date. Because the locus ceruleus is the major noradrenergic nucleus of the brain, diverse noradrenergic functions that are known to be activated in opiate

withdrawal could be affected. Nestler and other groups showed that although these changes occur uniformly in the locus ceruleus neurons, and in a few other brain regions, particularly in the nucleus accumbens, this type of change in the nucleus accumbens is strain dependent, and also such changes do not occur in many other brain regions in any strain or species (12, 18, 20, 24) . They also do not occur in the gastrointestinal tract. Nestler and others did not find a down-regulation of μ-opioid receptors during chronic morphine treatment in the locus ceruleus (18, 20, 24) . They did, however, report an uncoupling of the μ-opioid receptor from its G-protein–coupled inwardly rectifying potassium channels during chronic morphine exposure, with a resultant reduction in the maximal outward current and documented decreased efficiency along with decreased potency of the opioid. This is intriguing in the context of findings of the laboratories of Yu and Kreek, who reported that after binding of the long, 31-residue, endogenous opioid ß-endorphin to the variant μ-opioid receptor coded by the very common SNP, A118G, there is enhancement of activity of these G-protein–coupled inwardly rectifying potassium channels (58) .

Almost all groups, again starting with the earliest work of Aghajanian, as well as more recent work of Nestler and others, have suggested that the locus ceruleus may be primarily involved in expression of opioid physical dependence and thus in opioid withdrawal (20) . Selley and Childers et al. studied the effects of chronic morphine treatment on opioid-receptor–coupled G-protein activity in membranes from the locus ceruleus and showed that chronic morphine treatment decreased the inhibitory G-protein activity in the locus ceruleus and yet did not produce any detectable desensitization, a finding suggesting a potential adaptation at that level (40) . Chronic morphine treatment decreased both basal and opioid stimulated guanosine triphosphatase (GTPase) activity and yet caused no changes in the percentage of stimulation by an opioid agonist. All these results were extended by binding assays using [^{35}S]GTPγS (40) . In further studies, it was found that long-term heroin self-administration also similarly altered the opioid-receptor–activated G proteins in specific brain regions, primarily in specific brainstem nuclei (42) . Decreased μ-opioid-agonist–stimulated [^{35}S]GTPγS binding was observed in the locus ceruleus and in a few related regions during long-term heroin self-administration. These findings were similar to those previously described in animals treated with morphine on a long-term basis (40) . Moreover, the decreased μ-opioid-stimulated [^{35}S]GTPγS binding was found in two additional regions, the thalamus and the amygdala, which may be of importance for the reinforcing effects of drugs of abuse and thus self-administration (42) .

All these scientists mention that the neuronal and molecular basis of opioid tolerance and dependence remains unclear. The opioid receptors involved have all been cloned and have been documented to be part of the G-protein–coupled family of seven transmembrane receptors; there has been further documentation of receptor phosphorylation, desensitization, and uncoupling from G proteins, as well as new studies documenting internalization (endocytosis) of opioid receptors (36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51 and 52) . However, studies in animals continue to produce very conflicting results concerning the effects of chronic opiate administration on opioid-receptor binding and density or number. Similarly, despite documentation by many groups that cellular adaptations may be directly involved in the development of tolerance and dependence, the mechanisms have yet to be fully elucidated. Moreover, multiple other neurotransmitter systems have been implicated, in particular the *N*-methyl-D-aspartate (NMDA)–receptor complex and its ligands.

Chronic morphine administration, with resultant changes in G-protein–coupled signal transduction mechanisms and changes in downstream effectors, such as increases in CREB and phosphorylated CREB, and also other changes such as increases in and accumulation of chronic FRAs (Fos-related antigens), are all nonspecific. For instance, diverse stimuli such as cocaine, opiates, opiate withdrawal, nicotine, other drugs of abuse, and stress have been shown to cause increases in chronic FRAs (12) . CREB is one component of the enhanced cAMP response and is a transcription factor; chronic FRAs have now been identified as isoforms of ΔFosB, which is a splice variant of the *fosB* gene. Each of these enhanced or altered transcription factors can change the levels of expression of many specific genes and in specific brain regions. These increases in CREB and in chronic FRAs, both results of chronic morphine administration, may yield enhanced or altered gene transcription elements and, in turn, changes in levels of expression of specific genes and in specific brain regions. The increases in chronic FRAs after long-term morphine administration occur exclusively in the striatum, whereas the increases in chronic FRAs seen after stress occur in the prefrontal cortex (20) . However, these components are also transient. For instance, Nestler found that chronic FRAs probably persist for only a few weeks after accumulation during chronic opiate or cocaine administration and thus require repeated exposure to a drug of abuse for reappearance or for persistence. These increased and accumulated amounts of CREB and chronic FRAs are tangible examples of neuroplasticity of the brain and document one type of change that may occur and persist with chronic exposure to a drug of abuse. As Nestler warned, however, they are but two of probably many such changes and, although related to specific addiction related phenomena, are not the sole cause of any of the three distinct and separable phenomena of tolerance, physical dependence, or addiction. Moreover, all the resultant gene expression changes that occur, related to CREB and ΔFosB, and that may contribute to an atypical activator protein 1–type transcription factor, are nonspecific changes, with respect to the causative agent. In addition, these transient, but gene-specific, changes in gene expression can

occur in specific brain regions after specific times of exposure to, or withdrawal from exposure to, a drug of abuse such as morphine (but also cocaine and other drugs of abuse).

Direct effects of these transcription factors have been studied only to a limited extent. Nestler hypothesized that the documented increases in amounts and phosphorylation of CREB in the locus ceruleus caused by chronic morphine administration may be directly involved in the regulation of the entire cAMP pathway through the CREB effects as a transcription factor on gene expression (18). His group showed that application of CREB antisense oligonucleotides applied to the locus ceruleus of opiate-dependent rats decreases the opiate withdrawal-induced increases in neuronal firing that are usually seen (18). Further, the laboratory of Nestler showed that accumulation of chronic FRAs during chronic morphine treatment, related to the transient early gene protein products Fos and Jun, which, in turn, join to form a major gene transcription factor, activator protein 1, may play a role in the effects of morphine and also of stimulants (18). For instance, Nestler's group showed in *fosB* knockout mice, in which chronic FRAs are presumed not to be formed or accumulate, an enhanced locomotor response to cocaine (12). Using an effective construct involving both a tetracycline transactivator gene to allow regulation of gene expression and the gene encoding ΔFosB that then can be delivered to a specific brain region, Nestler's group conducted further studies of the effects of drugs of abuse. This gene construct allows overproduction of ΔFosB by the gene insertion, the overexpression of which can be prevented by administration of a tetracycline congener, but it be started again by stopping treatment with the tetracycline congener. The overexpression can be both brain region specific and time specific (12). To date, enhancement of ΔFosB in the striatum has been shown to alter the behavioral response to cocaine (12). Using a different transgenic approach, a viral vector may be used to deliver a desired gene to a specific brain region to yield overexpression. Carlezon and Nestler and their colleagues used such a delivery system to achieve overexpression of the NMDA-receptor component, $GluR_1$, with overexpression in different specific brain regions; in some brain regions, and with this region-specific overexpression, they were able to document increases in opiate reward and, in other brain regions, increases of aversion to opiates (59).

There is an increasing consensus that the reinforcing effects of drugs of abuse, along with possibly physical dependence, are not directly related to tolerance, and they also may not be directly related to any changes in receptor density, number, desensitization, internalization, G-protein uncoupling, or other effects on signal transduction mechanisms. These findings are further supported by the report of Bohn, Lefkowitz, Caron, and colleagues that, in studies in ß-arrestin knockout mice, one sees enhancement and persistence of the antinociceptive effects of morphine (60). Furthermore, Bohn and colleagues reported that no tolerance develops to the antinociceptive effects of morphine during chronic administration, but they also said that there apparently is no impact on physical dependence in the ß-arrestin knockout mice (61).

Dopamine, Other Neurotransmitters, Neuropeptides, and Their Receptors: Molecular, Cell Biological and Signal Transduction Alterations, and Possible Implications for Pathophysiology of Opiate Addiction

The early work of many groups showed that opiates, like most other drugs of abuse, appear to act to enhance dopaminergic tone and through that enhancement achieve some, most, or all of their reinforcing effects. Moreover, through a variety of studies, primarily conducted in animals using either surgical lesions or specifically directed neurotoxins, and also other specific chemicals to enhance or decrease dopaminergic function, along with ultimately microdialysis techniques, researchers showed that enhancement of dopamine tone in the mesolimbic-mesocortical dopaminergic system in particular is associated with the rewarding or reinforcing effects of most or all drugs of abuse. The seminal work by Johnson and North documented unequivocally that one action of μ-opioid agonists, exerted through μ-opioid receptors localized in the ventral tegmental area, is on inhibitory GABAergic interneurons and is one of inhibition of those neurons (62). Thus, by inhibiting these inhibitory neurons, which normally put a brake on the dopaminergic neurons in the ventral tegmental area, the result is activation of the dopaminergic neurons, with enhanced release of dopamine in the nucleus accumbens, as well as in the amygdala and probably in all other regions of the mesolimbic-mesocortical dopaminergic fields (62).

Although many investigators have attributed the reinforcing effects of all drugs of abuse, including heroin and morphine, to actual or presumed enhanced dopamine levels in the nucleus accumbens, through this indirect action for opiates, and for cocaine, through a direct blockade of the dopamine reuptake transporter, there is increasing evidence that dopamine not only is not essential for the reinforcing effects of heroin and morphine, but also does not play a central role in the reinforcing and rewarding effects of opiates.

Studies have been conducted in animals with deletion of the dopamine transporter gene, which many researchers had hypothesized would eliminate cocaine self-administration because of the very high constant levels of dopamine and the lack of further effects by superimposed cocaine (63). This was found to be not the case (63). The dopamine transporter knockout mice were found unequivocally to self-administer cocaine, although the acquisition of that behavior was slower than in the wild-type mice (64). Thus, even for cocaine, dopamine clearly plays a role in the rewarding effects,

but it is not the sole component, nor are changes in dopamine levels essential for the reinforcing effects of cocaine (63 ,64). In that same animal model, the dopamine reuptake transporter knockout mice, it has been found, however, that morphine is more avidly self-administered than in wild-type mice, a finding suggesting a positive interaction between the persistently elevated levels of dopamine and morphine to enhance reward (65). Hemby and Smith and their colleagues also found a synergistic elevation of extracellular dopamine when cocaine was added to heroin in self-administration studies (66). These findings may explain, in part, the common co-dependency in humans of both heroin and cocaine addictions.

With respect to opiates, two very early studies showed that when animals were lesioned to delete the dopaminergic neurons completely in discrete brain regions by use of a neurotoxin, 6-hydroxydopamine, self-administration of morphine proceeded normally as in unlesioned animals. However, in such animals, cocaine self-administration was eliminated.

There have been conflicting results in other studies. For instance, in one study using the technique of *in vivo* fast cyclic voltammetry, it was found that heroin caused a dose-dependent increase in dopamine in the nucleus accumbens during heroin self-administration, and co-administration of a κ-agonist (U-50,488 H) with the heroin, or alternatively, intracerebroventricular administration of dynorphin A, significantly depressed the heroin-stimulated dopamine release (67). Moreover, installation of the κ-synthetic compound or natural ligand dynorphin A alone decreased basal dopamine release, as had also been shown by Claye and others (68). Studies by Xi, Fuller, and Stein thus suggested that the μ-agonist morphine activates the mesolimbic-mesocortical dopaminergic pathway and that κ-opioid-receptor activation offsets, or counterregulates, that activation (67).

However, another set of studies by Hemby, Smith and colleagues showed that systemic self-administration of heroin alone does not cause any elevation in dopamine as determined by *in vivo* microdialysis with the probes in the nucleus accumbens (69). In related studies, these investigators found, as have numerous others, that cocaine caused a striking increase in extracellular dopamine concentrations in the nucleus accumbens, and, moreover, the combination of cocaine and heroin caused a synergistic elevation (66). Their finding that heroin alone failed to cause an increase in dopamine in the nucleus accumbens complemented several earlier findings that heroin self-administration is not attenuated by administration of dopamine antagonists, as well as even earlier studies showing that integrity of dopamine pathways in the nucleus accumbens is not essential for heroin self-administration. These findings document further the early hypothesis of the Kreek laboratory, and many others, that the reinforcing properties of heroin are mediated primarily by dopamine-independent mechanisms and probably by the μ-opioid receptor itself. This hypothesis has been ultimately supported by the findings that μ-opioid-receptor deletion knockout mice have no self-administration of opiates and no rewarding effects of opiates (reviewed in ref. 33).

In another study, which used morphine pellet implantation to develop opioid tolerance and dependence, a reduction in dopamine D2-receptor mRNA levels, but no change in dopamine D1 mRNA levels, was found at the end of the 6 days of morphine exposure (70). The mRNA levels for both dopamine D1 and D2 receptors was reduced after 1 day of withdrawal, and both returned toward normal by the third day after drug withdrawal. These findings may be related to the reduction in dopamine D2-receptor binding, which has been seen in human heroin addicts, by using positron emission tomography. However, curiously in this study, but not in other studies, reductions of mRNA levels for dynorphin and enkephalin genes were found during morphine exposure. In contrast, enhanced dynorphin mRNA levels have been found at least after acute single and multiple intermittent-dose morphine administrations (71 ,72). Further studies will be needed to determine the time course of dynorphin mRNA level changes during morphine exposure. Trujillo, Akil, and their colleagues showed that chronic injection or infusion of morphine caused increases in levels of dynorphin peptides in the dorsal striatum (caudate putamen) but not in the ventral striatum (nucleus accumbens) (73).

Intriguingly, Lee, Henriksen, and colleagues found that only a few (approximately 20%) nucleus accumbens neurons seem to exhibit an inhibitory response after heroin self-administration, along with about 40% of prefrontal cortex neurons showing such inhibition (74). Thus, the multiple changes in signal transduction observed and discussed earlier, including the effects of chronic morphine administration on μ-opioid-receptor–stimulated [^{35}S]GTPγS binding changes, with reduction of [^{35}S]GTPγS binding specifically in the brainstem nuclei, including the dorsal raphe nucleus, the locus ceruleus, the lateral and medial parabrachial nuclei, and the commissural nucleus tractus solitarius, may result from a direct opiate effect or an indirect effect by alteration of the dopaminergic system (40 ,42). Similar findings were made by the group of Sim-Selley, Selley, Childers, and colleagues after chronic heroin self-administration, with the greatest decrease in μ-opioid-receptor–stimulated [^{35}S]GTPγS binding in the brainstem and the lowest alterations in binding in the striatum and cortex (42). Because the changes of dopamine D1-receptor activation would act in one direction and dopamine D2-receptor activation would act in the opposite direction on adenylyl cyclase activity, the effects on these receptors could also influence the effects of μ-opioid-receptor activation, and the changes that have been observed may result exclusively from the opioid effects acting at the μ-opioid receptors or also secondary indirect effects on dopamine receptors.

These and other findings suggest that opiates may act

directly to alter dopaminergic systems both in the ventromedial striatum, that is, the core and shell of the nucleus accumbens, and in the dorsolateral striatum, that is, in the caudate putamen region. Clearly, there are abundant μ-opioid receptors as well as κ-opioid receptors in those regions (26 ,75 ,76 and 77). Work from the Kreek laboratory showed that another drug of abuse, cocaine, when delivered in a binge pattern, which markedly enhances dopaminergic tone, causes an increase in density of μ-opioid receptors and also κ-opioid receptors in those brain regions, and it also alters basal and opioid-regulated adenylyl cyclase activity in these regions (75 ,76 and 77). There have been no similar findings with respect to increasing μ-opioid-receptor density after chronic opioid administration, however. It is not really known to what extent reinforcement or reward resulting from heroin and morphine occurs because of activation directly in these areas, especially the nucleus accumbens and possibly also the amygdala, as contrasted to indirect effects on the ventral tegmental area. The effects on dopamine in each of these different locations and also the different mechanisms involved have not yet been fully elucidated using a model of chronic, high-dose, intermittent but evenly spaced opiate administration, mimicking the human pattern of heroin or morphine abuse and addiction, and also after withdrawal, as well as during reexposure after such opiate administration. During chronic binge pattern cocaine administration, a pattern mimicking the human condition, there is a progressive lowering of basal, as well as cocaine-stimulated, dopamine levels in the extracellular fluid of the caudate putamen and in the nucleus accumbens (78). Noble and Cox clearly defined a role of the dopaminergic system in opioid-receptor desensitization in these brain regions during chronic morphine administration (39).

After chronic opiate administration, Nestler and colleagues found increases in tyrosine hydroxylase in the ventral tegmental area. This is a rate-limiting enzyme in the biosynthesis of dopamine. They also found a reduction in mean size of the ventral tegmental area dopaminergic neurons and decreased axonal transport to the nucleus accumbens (24 ,79). However, there were no changes in numbers of dopaminergic neurons and no changes in the size of nondopaminergic neurons (79). Within ventral tegmental area, infusion of brain-derived neurotrophic factor prevented this morphine-induced reduction in size of dopaminergic neurons (79). Their group also found that chronic morphine administration resulted in an increase of other components related to signal transduction, including the extracellular signal regulated kinases (ERKs), which are effectors for brain-derived neurotrophic factor in the ventral tegmental area (24). However, the time course of these changes and their persistence after morphine withdrawal have yet to be elucidated, and their relation to both physiology and the behaviors of addiction also have not yet been fully explored, although the findings suggest that neurotrophic factors may act in response to the opiate-induced changes in neural integrity, that is, the neuroplasticity after chronic opiate administration that results in impairment of normal neural integrity. Both the chronic morphine-induced injury and the counterregulatory events may alter neural growth, development, and synapse formation, signal transduction, and overall system integrity (24 ,79).

Similarly, the findings that acute and chronic morphine administration and withdrawal may enhance dynorphin gene expression and dynorphin peptides, undoubtedly events mediated in part through action of dopamine D1 receptors, often co-localized on cells with dynorphin gene expression, as well as more direct effects of enhanced transcription factors on dynorphin gene expression, may be again important counterregulatory events, which also represent examples of profound neuroplasticity of the brain. Such findings have also been made during binge pattern cocaine administration (80 ,81). Enhanced dynorphin peptides, in turn, acting at κ-opioid receptors, may reduce dopaminergic tone in many brain regions, including those involved in reward and also locomotor activity, and they may also attenuate opioid withdrawal in dependent animals or humans (6 ,8 ,9 ,10 and 11 ,16). Again, these events must be considered to be a direct result of neuroplasticity and are counterregulatory, the attempt to attenuate, modulate, or even brake the events caused by the rapid changes in dopaminergic tone brought about especially by stimulants such as cocaine, but also to a lesser extent also by opiates.

The changes in signal transduction mechanisms after chronic heroin or morphine administration are undoubtedly primarily the result of the effects of chronic opioid administration. However, because there are also significant changes in dopaminergic tone with enhanced signaling through the dopaminergic pathways, owing to indirect or direct activation of dopamine release, the changes in signal transduction observed may also result from enhanced activation of the dopaminergic neurons, as stated earlier. D1- and D5-type dopaminergic receptors enhance adenylyl cyclase activity, an effect similar to that occurring in the locus ceruleus after chronic, but not acute, morphine administration, in most strains of rodents studied, and also in the nucleus accumbens in some strains of some species. In contrast, activation of the dopaminergic D2 receptors causes a reduction in adenylyl cyclase activity, such as observed during acute morphine administration in all brain regions of strains and species of rodents studied, as well as in all cell systems studied, and an effect that continues to pertain in some specific regions of the brain and other parts of the body during chronic opioid administration. Thus, the observations of alterations in the downstream events of the adenylyl cyclase changes may be the cumulative response of chronic morphine administration on μ-opioid-receptor activation and also of dopamine on dopaminergic D1- and D2-receptor systems.

More recently, the findings of Crain and Shen showed the ability of very small amounts of specific opioid antagonists,

in fact, to enhance the analgesic effects of the μ-opioid-receptor agonists and to prolong the opioid-agonist effects both in animal models and in humans (82). Crain and Shen hypothesized that, although most μ-opioid receptors are coupled with inhibitory $G_{i/o}$ protein, a small proportion may be coupled at the stimulatory G_s protein, which can be suppressed with small amounts of opioid antagonists. These findings of enhanced morphine analgesia are, in part, very similar to the findings of Bohn, Caron, Lefkowitz, and colleagues, in mice with deletion of ß-arrestin (60). These researchers also showed that ß-arrestin is important in several distinct functions, including events leading to the internalization of an agonist bound μ-opioid receptor, which, after the phosphorylation of the bound form, binds to ß-arrestin, along with binding by G-protein–receptor kinases (60). This event of ß-arrestin binding has been described as potentially part of the process that desensitizes, that is, leads to G-protein uncoupling of the μ-opioid receptors, as well as being actually involved in the internalization of endogenous and some exogenous agonist-bound μ-opioid receptors (44 ,45 ,46 ,47 ,48 ,49 ,50 ,51 and 52 ,60). The role of internalization in the development of tolerance and the independent process of dependence remain unclear because there are many conflicting results, including the finding that most exogenous opioid ligands, including morphine, that do not induce prompt internalization of μ-opioid receptors once bound, clearly lead to the development of both tolerance and physical dependence (44 ,45 ,46 ,47 ,48 ,49 ,50 ,51 and 52). In sharp contrast, methadone and etorphine do lead to prompt internalization of μ-opioid receptors, just as do all the natural endogenous opioid peptides such as Met-enkephalin and ß-endorphin (44 ,45 ,46 ,47 ,48 ,49 ,50 ,51 and 52). Intriguingly, in the mice with deletion at the ß-arrestin-2 gene, enhanced morphine analgesia was seen, and further studies revealed that tolerance does not develop to morphine effects, and yet objective signs reflecting the development of physical dependence are present after chronic morphine administration (60). These studies again dissociated the development of tolerance from the development of physical dependence. The studies of the group of Crain, as well as the studies of the group of Caron and Lefkowitz, suggested that either deletion of ß-arrestin or suppression, by opioid antagonists in very small doses, of opioid receptor coupled to G_s, the stimulatory G-protein pathway, will enhance opioid analgesia and also may attenuate or prevent development of tolerance. It is not known whether blocking of the $G_s$coupling alters the development of physical dependence, however. In possibly related studies, Jeziorski and White showed that the NMDA antagonist, MK-801, prevents development of behavioral sensitization during chronic morphine administration, whereas dopamine-receptor antagonists prevent expression, but not development, of sensitization (83). Sensitization has been suggested to be related to drug reward or craving. Possibly in contrast, Churchill, Roques, and Kalivas found that dopamine depletion, such as may happen during chronic opiate, as well as chronic cocaine, administration, augments opioid-induced locomotion (84).

There have been only limited studies of the time course of all these dopaminergic responses during investigator-administered morphine or heroin on an intermittent basis, mimicking the human pattern of heroin abuse, or during chronic self-administration of opiates. It would be assumed that possibly, as with cocaine, one sees a progressive diminution of the responsivity, with a resultant lowering of basal level and stimulant-induced rise of absolute levels of dopamine (78). Numerous human studies suggest this may indeed happen. It has been repeatedly shown in heroin addicts that the short acting μ-opioid agonist heroin will cause a prompt increase in serum prolactin levels, resulting directly from an abrupt lowering of dopamine levels in the tuberoinfundibular dopaminergic systems (85). In humans, and to a greater extent than in rodents, prolactin release is essentially solely under tonic inhibition by dopaminergic tone in the tuberoinfundibular dopaminergic system. However, it was found that during chronic methadone treatment, there is adaptation or tolerance to this phenomenon, an attenuation, but not a complete removal or ablation of this response caused by dopamine lowering and resulting in elevation of serum prolactin levels (85). Even during long-term methadone maintenance treatment, as reported in 1978, it was found that peak plasma levels of the μ-opiate agonist methadone are related temporally to the peak plasma levels of prolactin (85). These findings suggest that the long-acting opioid methadone administered orally continues to have an impact at least on the tuberoinfundibular dopaminergic system, with a lowering of dopaminergic tone, resulting in a modest increase of prolactin levels, although not exceeding upper levels of normal. However, that attenuation occurs suggests that there may be either a lowering of dopaminergic levels and tone in the turberoinfundibular dopaminergic system of that region or, alternatively, an attenuation of responsivity of the μ-opiate-receptor system. It has been shown that the κ-opiate-receptor system similarly plays a role in modulating prolactin levels in humans (86). In normal healthy volunteers, dynorphin A causes a prompt rise in serum prolactin levels, resulting again presumably from a lowering of dopaminergic tone in the tuberoinfundibular system (86). This is a μ, but also a κ-opioid-receptor effect, as documented by use of two different opioid antagonists with different receptor selectivity (86). In preliminary studies, the Kreek laboratory showed that there is altered responsivity both in former heroin addicts and in former cocaine addicts, as well as those with combined heroin opioid and cocaine dependency (87).

Acute morphine administration has been shown to have a variety of profound effects on many other neurotransmitters; this group comprises fast-acting neurotransmitters including excitatory amino acids such as glutamate and slower-acting neurotransmitters such as norepinephrine, epinephrine, and serotonin, as well as dopamine, and a variety

of neuropeptides. Very few studies have been conducted in models using chronic heroin or morphine administration, or self-administration, using long-term, high-dose, regularly spaced intermittent administration or by long-access, high-dose, self-administration, mimicking the human pattern of heroin abuse. Further work will be central to detail the long-term effects and, also of special interest, the effects of the withdrawal and reexposure to mimic relapse. However, qualitatively and quantitatively different changes have been found during chronic morphine or heroin administration by different patterns, dose, and routes of administration.

Physiologic Systems and Behaviors Primarily Altered

Stress Responsivity: Possible Implications for Opiate Addiction

An atypical responsivity to stress and stressors existing on a drug-induced basis or possibly *a priori*, on a genetic or environmental basis, as one component of the "metabolic basis" of heroin addiction was a concept that was hypothesized by the Kreek group in 1964, and it was therefore addressed directly in our prospective studies started at that time and completed in 1972, as well as in other early basic clinical research studies (6 ,85 ,88 ,89 ,90 ,91 and 92). Several laboratories went on to study, in humans, the impact of drugs of abuse and specifically heroin, but also morphine, (as used in a single dose or on a chronic basis in the pharmacotherapy of pain), on one component of stress response, the hypothalamic-pituitary-adrenal (HPA) axis (6 ,93 ,94 ,95 ,96 ,97 ,98 ,99 ,100 ,101 ,102 ,103 ,104 ,105 ,106 ,107 and 108). Long-term studies in animal models came later, however, and were performed by many different groups (6 ,8 ,109 ,110 ,111 ,112 ,113 ,114 ,115 ,116 ,117 and 118). The initiation of these studies was predicated not only on the clinical research, which clearly documented that opiates suppress the HPA axis in humans and continue to do so during the long-term self-administration of short-acting opiates such as heroin, but also, and very importantly, that normalization of this HPA axis occurs during steady-dose long-term methadone maintenance treatment, findings that were made in rigorous studies and reported as early as 1972 (6 ,89 ,90). Studies reported from the late 1970s onward have documented that the endogenous opioids clearly play a tonic modulatory role of inhibiting the hypothalamic-pituitary part of the HPA axis (reviewed in ref. 9). Further more recent studies in humans have shown that this modulation is effected by both μ- and κ-selective opioid ligands (108).

In the middle to late 1980s, several groups began to study the concept that stress and the response to stress, as well as novelty and risk-seeking, may contribute to self-administration of drugs of abuse, including opiates, and parallel studies showed that drugs of abuse including opiates, cocaine, and alcohol perturb components of the stress-responsive systems in animal models. The initial studies measured primarily specific behaviors after assessment of the relative response to novelty or to risk and used different strains of rats, as well as mice. Similarly, more recent studies looked not simply at the acute effects of drugs of abuse, but also at the subacute and chronic effects of drugs of abuse and the impact of withdrawal from such drugs on components of the stress-responsive axis. Even more recent studies went on to study levels of gene expression and the impact of exposure to drugs of abuse over a defined time course of exposure on gene expression, first on "early gene response" and then, more recently, on changes of expression of many other specific genes, in particular, components of stress-responsive axis (6 ,8 ,9 ,109 ,110 ,111 ,112 and 113).

The interactions of the dopaminergic system on the HPA axis as well as the effects of catecholamines on this axis have been studied in both animal models and in humans. It is clear that opiates, like cocaine but to a much lesser extent, cause an elevation in dopaminergic tone, especially in the mesolimbic-mesocortical dopaminergic system. However, as discussed earlier, several groups have shown that although this is a reproducible phenomenon, the mesolimbic-mesocortical dopaminergic system is not essential for heroin or morphine self-administration, and animals that have received lesions abolishing this mesolimbic-mesocortical dopaminergic system readily self-administer opiates such as morphine. This finding is in sharp contrast to that which pertains for cocaine self-administration in which lesions of the mesolimbic-mesocortical dopaminergic system abolish cocaine self-administration. Thus, the role of dopamine in the well-established acute morphine activation of the HPA axis in rodents is of interest, but it may be a related, but not central, component of the mechanism underlying self-administration. More recent studies performed in transgenic mice have had a deletion or knockout of DARPP-32, an obligatory component of the signal transduction mechanisms after activation of primarily dopaminergic D1 receptors; a profound attenuation of the well-established cocaine effect of enhancing hormones of the HPA axis, including adrenocorticotropic hormone (ACTH) and corticosterone levels, was found (110). Parallel studies using this model to explore the impact of this deletion on the well-established acute morphine activation of this axis have yet to be conducted.

Of great interest for many years, and not always recognized by research groups, has been the finding that rodents have the opposite response to acute opiate administration than do humans; that is, activation of the HPA axis occurs. Studies in drug-naive healthy humans, as well as in formerly opiate-dependent healthy humans, and in active heroin addicts have shown that the first, or initial, acute administration of a short-acting opiate, such as morphine or heroin, as well as the first or initial acute administration of a long-acting opioid, such as methadone, will cause suppression of the stress-response systems. Further, in humans, chronic self-administration of short-acting opiates, such as heroin, leads to a continuing suppression of this HPA axis. In contrast,

many rigorous studies have shown with chronic administration of a long-acting opioid, such as methadone, which allows steady-state profusion of μ-opioid receptors in humans and which is provided during methadone maintenance treatment of heroin addiction, one sees normalization of this axis (6 ,8 ,9 ,85 ,89 ,90 ,91 ,92 ,93 ,94 and 95 ,100 and 101).

Zhou et al. modeled this phenomenon in rodents and found that whereas acute intermittent morphine administration causes activation of the HPA axis, delivery of methadone by pump to achieve a steady state, paralleling the situation in humans receiving chronic methadone treatment for management of opiate addiction or chronic pain (with pump delivery essential because methadone has a half-life of 90 minutes in the rat), yields neither alterations in any component of the HPA axis nor alterations in ACTH or corticosterone levels seen (111 ,112). When administered on a chronic basis in humans or in rodents, short-acting opiates such as heroin and morphine cause suppression of the HPA axis and with no sustained activation in rodents. During either spontaneous or naloxone-precipitated withdrawal, one sees activation of the hormones of the HPA axis in all species studied.

Recently, an animal model was designed to mimic more closely the human pattern of heroin administration, with multiple short-acting opiate (morphine) administrations given at evenly spaced intervals over a single day; activation of the HPA axis with elevation of levels of ACTH and corticosterone was found (111). In addition, as part of this initial study of the effects of acute intermittent morphine, but given in a mode more closely similar to that in the human heroin addict, the impact of a superimposed stress on the effects of morphine was also studied (111). A modest stress of water restriction was applied that, like acute morphine, also significantly increased the ACTH levels. However, when morphine was concomitantly administered to the animals undergoing modest water restriction, morphine attenuated the stress-induced elevation of ACTH and corticosterone levels of this axis (111). These findings may have enormous implications for the human condition, in which morphine or heroin may act immediately to attenuate any activation of the HPA axis caused by any one of numerous types of environmental stressors. Rigorous studies have now been conducted showing that another drug of abuse, cocaine, not only causes elevation of ACTH and corticosterone levels, but also initially enhances corticotropin-releasing factor mRNA levels; however, it was also found that chronic binge pattern cocaine administration led ultimately to an attenuation of the still elevated plasma levels of ACTH and corticosterone by 14 days, and at that time corticotropin-releasing factor mRNA levels were significantly lower than basal levels (109). Recently, Zhou and colleagues made similar findings with respect to acute versus chronic ethanol treatment (113).

Various studies in humans, from the Kreek laboratory, and in animal models, by Stewart, Shaham, and Erb and many other investigators, further documented that stress and stressors, in addition to cues of drug use, and “priming,” or reexposure to a drug, may play an important role in relapse to self-administration of drugs of abuse (99 ,103 ,104 ,114 ,115 ,116 ,117 and 118). Moreover, various studies (99 ,103 ,104 ,114 ,115 and 116), such as the work of Piazza and LeMoal, showed that animals with a greater response to novelty or stress and also animals with higher basal levels of the stress-responsive hormone, corticosterone, may more readily begin to acquire self-administration of a drug of abuse, at least of low-dose psychostimulants such as cocaine (reviewed in ref. 116).

Other studies have documented unequivocally that each of the major drugs of abuse highly significantly not only alter the hormone levels of the HPA axis, but also causes alterations of levels of expression of genes of that axis, as well as of similar stress-responsive genes in other parts of the brain, not directly involved in the HPA axis (109 ,112 ,113). Corticotropin-releasing factor, indirectly and directly measured, for instance in the work of Weiss and Koob, was shown to play a potentially very important role in particular aspects of withdrawal from drugs of abuse and in relapse (6 ,8 ,9 ,89 ,90 ,95 ,99 ,103 ,104 ,117 ,118).

Studies in Novel Animal Models

Since the mid-1990s, investigators have increasingly developed and used animal models that more closely mimic human patterns of drug abuse and emulate the pharmacokinetic situation that pertains during treatment of addictions, such as the pharmacotherapy of heroin addiction, which has been successful primarily by using long-acting, specific μ-opioid-receptor–directed agonists, and also a partial agonist, including methadone, L-α-acelytmethadol (LAAM), and more recently buprenorphine (with its abuse potential minimized by the addition of the non–orally bioavailable antagonist naloxone).

One of the earliest of these animal models that closely parallels a human pattern of addiction was the development of the binge pattern cocaine (investigator) administration model. This model mimics the most common pattern of human abuse, that is, multiple self-administrations of cocaine either by the intravenous route of administration or by inhalation (smoking) of the freebase form, known as crack (75 ,76 ,77 and 78 ,80 ,81 ,119 ,120) This model has uniquely allowed identification of molecular neurobiological changes, including increases in μ-opioid-receptor density that has subsequently been identified in human cocaine addicts (75 ,77 ,121). Animal models mimicking the most common human pattern of heroin addiction have really just begun to be used (111). Heretofore, most of the subacute and chronic models used morphine, the major metabolite of heroin, not heroin itself, and they also used morphine pellet implantation, to develop tolerance and dependence with ease and predictability (with such morphine pellets usually implanted every 1, 2, or 3 days, and most commonly using

the NIH-NIDA developed 75-mg morphine pellets developed by the National Institutes of Health and National Institute of Drug Abuse). Although extremely useful and convenient for many studies, this pellet (prolonged exposure, followed by slow withdrawal) approach does not give the features that have been shown in many studies to be profoundly different from when “steady-state” (pump) or “on-off” (intermittent injections) are used. Thus, increasingly, investigator administration models are being developed in which the human pattern of heroin addiction may be mimicked, that is, with heroin or morphine administered at equally spaced intervals during the animal’s awake period, three to six times every day, and with opiate withdrawal over the sleep period, which is most common for the heroin addict.

Similarly, because methadone, the most widely used and efficacious of the μ-opioid agonist treatments for heroin addiction, has, in fact, a very short-acting pharmacokinetic profile in rodents (90 minute half-life in the rat and 60 minutes half-life in the mouse, as contrasted with a 24 hour half-life in humans, and with an even long-acting half-life of the active l(R) enantiomer in humans), to mimic the human situation for treatment in rodents, one must administer methadone in a steady state, using pump technology (122 ,123 ,124 ,125 and 126). When this has been done, very different findings may have been made than when methadone has been administered intermittently, and thus it behaves in the rodent as a short-acting μ-agonist (112).

Over the past several years, it is has also been recognized that whereas opiates and also other drugs of abuse may cause innumerable acute effects, ranging from enhanced early gene expression (e.g., cfos and related Fos peptide changes) to later changes in other gene expressions and resultant neurochemical and behavioral changes, most of these changes disappear, become attenuated, or are altered by opposing or counterregulatory events after subacute or chronic short-acting opiate administration in an on-off pattern, in which setting, for instance, both dynorphin expression and κ-opioid-receptor gene expression become elevated (71 ,72). Increasing numbers of basic laboratory investigators are therefore focusing on studies of subacute and chronic effects of opiates, as well as other drugs of abuse, and then are proceeding to study those effects that persist during and after withdrawal of opiates (and other drugs of abuse) and into the abstinence period, to determine the point of no return or very slow return to normal status and thus the critical turning point in the development of relentless drug self-administration or addiction. Thus, models also have been developed and studies conducted to attempt to model human craving and relapse (or resumption of drug exposure or self-exposure), including the use of cue-induced, stress-induced, and small amounts of drugs of abuse-induced (priming) challenges, as well as in investigator-administered drug. Relevant molecular and neurobiological effects also are being conducted.

Most of these models mentioned earlier are investigator-administered models. There have also been several parallel studies attempting to modify long-existing, self-administration models to more closely parallel the human pattern of drug abuse and addiction (8 ,127 ,128 ,129 and 130). For various important and valid research reasons, self-administration studies, which use rats, mice, or nonhuman primates, primarily have been conducted using short sessions (usually, 1 to 3 hours in length) and in special cages to which each animal is moved for such studies, to provide the repeated cueing of a novel drug-related environment. Some studies, notably by the groups of Koob and Ahmed, Miczek and Tornatzky, and Mantsch et al., are starting to use much longer sessions of self-administration and also with very different unit doses of drug to be self-administered, with the resultant findings of different patterns of acquisition, extinction, and relapse that are probably more relevant to the human disorders of addictions (127 ,128 ,129 and 130).

Most studies since the mid-1980s years also have used relatively low to very low unit doses of the drug to be self-administered (although much higher unit doses were used effectively in some very early studies). These low doses have been used to allow evaluation of the reinforcing or rewarding properties of the drug by measurement of the number of responses, or work performed, and thus willingness to work to receive a unit dose of drug and also thereby to evaluate perturbations, either pharmacologic or behavioral, that may reduce that level of work. However, in human drug abuse and addiction, much larger unit doses of drugs of abuse (heroin or cocaine) are self-administered, and for opiates especially, with longer intervals between self-administrations. Thus, the bolus effect of a very rapid onset of action of a large amount of a short-acting drug such as heroin (or cocaine), self-administered either intravenously or by inhalation with sublimation of freebase drug, is achieved. It has been shown that the rapid rate of rise of amount of drug at a specific site of action, such as the μ-opioid receptor for heroin, is more closely related to the reinforcing effects, and also the rapid offset of drug action is related to the withdrawal or abstinence effects of a drug of abuse. Thus, higher unit dosages of drugs, such as are self-administered in the human situation, will have greater positive and negative reinforcing effects than small doses (8). Numerous small doses may, in fact, more closely begin to model a maintenance or steady-state mode, although the sessions are often too short to be analogous to desired treatment. A few groups are now using much longer sessions of self-administration and also, in some studies, higher unit doses of drug (primarily cocaine, but also heroin or morphine), with the expectation of longer self-administration dosing intervals and much larger total doses self-administered, thus probably with greater impact on molecular, cellular, and neurobiological features and, importantly, a greater magnitude and also qualitatively different and relevant behavioral changes (127 ,128 ,129 and 130).

BASIC CLINICAL RESEARCH

From the mid-1960s, the Kreek group hypothesized that there is a metabolic basis to addictive diseases, and an atypical responsivity to stress and stressors may contribute to the persistence of and relapse to addiction to heroin and also addictions to other specific drugs of abuse. Furthermore, it was hypothesized that such an atypical responsivity to stress and stressors may exist *a priori* on either a genetic or an early environmentally induced basis and may contribute to the initial acquisition of an addiction (6 ,85 ,88 ,89 ,90 and 91). Therefore, prospective studies, which were started in 1964 at the beginning of research on use of the long-acting opioid methadone in the pharmacotherapy of heroin addiction linked with behavioral treatment, included studies to assess the HPA axis component of the stress-responsive system, because this is one critically important component and one that can be evaluated in living humans; additional special studies were also conducted (6 ,85 ,89 ,90 and 91). Those very early studies documented an atypical responsivity of the stress-responsive HPA axis in heroin addicts, with suppression of all aspects of this axis by chronic self-administration of the short-acting opiate heroin, including reduction of plasma levels of hormones and alterations in the feedback control mechanisms, and also abnormal gonadal function with an impact on reproductive biology (6 ,85 ,89 ,90 ,91 and 92). Further, those early prospective studies on the effects of the long-acting opioid methadone, contrasted to the physiologic and pharmacologic effects of short-acting opiates, such as heroin, showed normalization during chronic methadone treatment of diverse physiologic functions disrupted by chronic heroin abuse, with gradual normalization of the stressresponsive HPA axis function over a 3- to 4-month period during steady moderate to high-dose treatment with the long-acting opioid methadone (6 ,8 ,9 ,10 ,85 ,89 ,90 ,93 ,94 ,95 ,96 ,97 ,98 ,99 ,100 ,101 and 102).

Studies also showed reduced responsivity to a chemically induced stress during cycles of heroin addiction and normalized neuroendocrine function of the HPA axis including a normal response to a chemically induced stressor during methadone maintenance treatment (6 ,85 ,89 ,90 and 91). However, further studies that we initially performed in 1983 and 1984 (95 ,99) showed a hyperresponsivity to a chemically induced stressor in medication-free and illicit drug-free former heroin addicts. A hyperresponsivity to an induced stressor, not only in drug- and medication-free former heroin addicts, but also in active cocaine addicts, who were using cocaine alone, and in methadone-maintained persons who continue to be addicted to cocaine was subsequently documented (95 ,98 ,99 ,103 ,104).

Other very important clinical research studies that have been conducted have shown, for instance, that activation of the stress-responsive HPA axis may precede, rather than follow, the signs and symptoms of opiate withdrawal (105 ,106 and 107). The activation of the HPA axis may drive the onset and may contribute to the severity of withdrawal symptoms, rather than result from the unpleasant or noxious qualities of these signs and symptoms (105 ,106 and 107). In addition, further studies of the continuing disruption of the HPA axis during naltrexone treatment and the lack of normalization of the assumed disruptions by heroin of HPA axis function during short-term buprenorphine treatment have been reported (10 ,28 ,96 ,97 ,102).

The Kreek laboratory hypothesized that natural sequence, but shorter, dynorphin A_{1-13} administered intravenously would result in prompt induced elevation of serum prolactin levels in normal healthy volunteer subjects. This was hypothesized because of two sets of previous findings. First, it is well-known that μ-opioid-receptor agonists will effect a rise in serum prolactin levels. Moreover, it has been shown that even during long-term methadone treatment, tolerance or adaptation is not fully developed to this prolactin-releasing effect of methadone (85). The mechanism for this is also known. In humans, prolactin release is essentially completely under tonic inhibition by dopamine. Therefore, an elevation in prolactin levels indicates a spontaneous or induced reduction in dopaminergic tone in the tuberoinfundubilar dopaminergic system. Other studies by several groups showed that synthetic small compounds that are κ agonists may reduce dopaminergic tone in rodents, and Claye et al. showed that the natural peptide dynorphin A_{1-17} instilled into the nucleus accumbens results in a reduction of dopaminergic tone in rats (68). In a study of healthy human volunteers, it was shown that a dose-dependent elevation of serum prolactin levels occurs in response to intravenous administration of dynorphin A_{1-13} (86). Further studies using two different opioid antagonists documented that this effect was mediated by the κ- as well as μ-opioid receptors. It was also shown that females, who have significantly higher basal prolactin levels, responded to a significantly greater extent to this natural peptide κ-opioid-receptor agonist challenge with respect to elevations in serum prolactin levels (86).

In other studies, Specker and Pentel and colleagues found attenuation of opiate withdrawal symptoms in heroin addicts given dynorphin A_{1-13} (131). These studies build onto much earlier studies, which were not well controlled but which suggested that dynorphin peptides may attenuate some of the signs and symptoms of opiate withdrawal. In studies conducted in patients with chronic pain, dynorphin A_{1-13} was shown to augment the analgesia provided by the usual μ-opioid agonists (morphine or methadone), a finding suggesting a positive interaction between the μ- and κ-opioid-receptor agonists and a possible novel approach for providing pain relief (132). All these findings suggest that one could consider clinical research studies using a κ-opioid agonist along with a μ-opioid agonist in a pharmacotherapy of opiate addiction (11).

In another area of basic clinical research related to the neurobiology of heroin addiction or its treatment, imaging

techniques, using positron emission tomography or the related technique of single photon emission computed tomography, with studies of glucose metabolism or blood flow to assess activation or depression of activity of specific brain regions, as well as some studies using ligands directed toward specific types of receptors, including recently the opiate receptors, have been conducted in humans and reported (133 ,134). In addition, some studies using magnetic resonance imaging and functional magnetic resonance imaging have begun to contribute to our information about withdrawal from heroin addiction (135). We hope that, in the future, such imaging studies will contribute even further to our understanding of the neurobiology of the development of and relapse to opiate addiction and will also potentially be able to be related to the apparent normalization of function that can occur during long-acting μ-opioid receptor agonist treatment with methadone, or alternatively with l-α-acelytylmethadol (LAAM) and possibly also (but yet to be studied) with the buprenorphine-naloxone combination. The implications of all of this basic clinical research for treatment have been considered further in reviews (10 ,11 ,27 ,28 ,29 ,30 and 31).

Finally, the first successful cloning of the genes of the specific opiate receptors, starting in late 1992, led to studies to identify polymorphisms of the human opioid receptor and peptide genes and as well as of other genes that have been shown to be affected by drugs of abuse, and specifically for this discussion, by short-acting opiates used illicitly. Many such polymorphisms, including primarily SNPs of the μ-opioid receptor as well as of related genes, have been identified recently (6 ,7 ,32 ,58 ,136 ,137 and 138). Studies of potentially functional changes resulting from those polymorphisms, especially SNPs in the coding region of the genes resulting in amino acid changes, and thus in resultant peptide differences, have been initiated (32 ,58). In addition, a few groups are now studying human molecular genetics of the specific addictive diseases, including heroin addiction. In fact, an epidemiologic study by Tsuang et al. suggested that heroin addiction may have an even greater relative risk attributable to heritable factors than any other addiction, including alcoholism (139).

MOLECULAR BIOLOGY AND GENETICS

Part of "104 - Molecular and Cellular Neurobiology and Pathophysiology of Opiate Addiction "

The completion of cloning of the genes of the endogenous opioid system, following the first reports of the cloning of the δ-opioid receptor in late 1992, allowed the expansion of many types of studies, as well as the initiation of new studies. All the genes of the endogenous opioid system in rodents, as well as in humans, now may be included in molecular neurobiological studies, such as studies of quantification of levels of gene expression (mRNA levels). It is now also possible to look for polymorphisms, including SNPs, in human genes of the endogenous opioid system, as well as genes of related neurotransmitter, neuropeptide, and receptor systems (as discussed earlier). Since 1994, new technologies for such studies have been developed, and during the next decade, undoubtedly, they will be able to be used for novel discoveries of heretofore unrecognized genes and gene products involved in the acquisition, persistence, and relapse to addiction (6 ,7 ,32 ,58 ,136 ,137 and 138). These include use of microarray technologies for measuring gene expression (although to date, these arrays are relatively insensitive and cannot yet detect, let alone measure with precision and accuracy, the small changes, usually less than 50% to 100% increase or decrease, that may be expected in integrated neurobiology for genes of low-abundance encoding neuropeptides and their receptors, such as those of the endogenous opioid system). Nevertheless, use of microarray technology, and with the developing informatics to analyze the vast amounts of the expected resultant data, will undoubtedly reveal novel gene systems involved in the specific addictive diseases. In addition, microarray technology and other new approaches are beginning to be used for the identification of already recognized polymorphisms, including SNPs, and they may be able to be used in the future for the identification of novel polymorphisms (6 ,7 ,32).

The completion of the cloning of the endogenous opioid system has permitted the development of appropriate gene deletion, so-called knockout mouse models (reviewed in refs. 33 ,34 and 35). The single most important and relevant finding with respect to opiate addiction has been the documentation, first by Kieffer and colleagues, that there is no opiate-induced reward, as measured by conditioned place preference in the μ-opioid-receptor knockout mice; in addition, these mice show essentially no self-administration of ethanol (33 ,140). All investigations have shown that μ-opioid-receptor knockout mice have no analgesic response to conventional μ-opioid-receptor agonists such as morphine (reviewed in refs. 33 ,34 and 35). Cloning of these genes also permitted the further use of knock-down, or antisense modeling, as well as gene enhancement using appropriate constructs for gene delivery. Many laboratories have initiated work for conditional knockout or knock-in enhancement of gene expression, with control of time of onset of the deletion or enhancement, as well as in some models, specific brain region–dependent changes.

SUMMARY

Part of "104 - Molecular and Cellular Neurobiology and Pathophysiology of Opiate Addiction "

Many exciting developments stemmed from the initial cloning of each of the three opioid receptors—δ, μ, and κ—in 1992 and 1993, and the subsequent cloning of each of those genes in humans in 1994. Subsequently, many studies have been and can be conducted, using classic techniques, as well as other new modern techniques, such as microarray technology. Various studies on the impact of opiates on gene expression as well as signal transduction systems and integrated physiologic function have been conducted.

Moreover, novel animal models have been developed. Possibly most excitingly of all, further basic clinical research studies have been performed, including studies identifying many polymorphisms of human genes of the endogenous opioid systems. These studies have already given, and will continue to give, increased insights into the pathophysiology as well as molecular and cellular neurobiology and related behavioral changes of opiate addiction, and all these studies have continued to teach us about the enormous capability of the brain to change through neural plasticity.

ACKNOWLEDGMENTS

Part of "104 - Molecular and Cellular Neurobiology and Pathophysiology of Opiate Addiction "

I thank Dr. Yan Zhou, Ms. Sara Handy, and Ms. Susan K. Russo for help in preparation of the manuscript. Funding support was received from the National Institutes of Health National Institute on Drug Abuse Research Center grant P50-DA05130, the National Institutes of Health National Institute on Drug Abuse research scientific award grant KO5-DA00049, the National Institutes of Health National Institute on Drug Abuse research grant RO1-DA12848, the National Institutes of Health National Institute on Drug Abuse research grant RO1-DA09444, the National Institutes of Health National Center for Research Resources (NCRR) General Clinical Research Center grant MO1-RR00102, and the New York State Office of Alcoholism and Substance Abuse Services.

REFERENCES

1. Evans CJ, Keith DE Jr, Morrison H, et al. Cloning of a delta opioid receptor by functional expression. *Science* 1992;258:1952–1955.

2. Kieffer BL, Befort K, Gaveriaux-Ruff C, et al. The delta-opioid receptor: isolation of a cDNA by expression cloning and pharmacological characterization. *Proc Natl Acad Sci USA* 1992;89:12048–12052.

3. Chen Y, Mestek A, Liu J, et al. Molecular cloning and functional expression of a mu-opioid receptor from rat brain. *Mol Pharmacol* 1993;44:8–12.

4. Chen Y, Mestek A, Liu J, et al. Molecular cloning of a rat kappa opioid receptor reveals sequence similarities to the mu and delta opioid receptors. *Biochem J* 1993;295:625–628.

5. Wang JB, Imai Y, Eppler CM, et al. Mu opiate receptor: cDNA cloning and expression. *Proc Natl Acad Sci USA* 1993;90:10230–10234.

6. Kreek MJ. Opiates, opioids, SNP's and the addictions: Nathan B. Eddy Memorial Award for lifetime excellence in drug abuse research lecture. In: Harris LS, ed. *Problems of drug dependence, 1999: proceedings of the 61st annual scientific meeting of the College on Problems of Drug Dependence. NIDA Res Monogr* 2000;180:3–22.

7. LaForge KS, Shick V, Spangler R, et al. Detection of single nucleotide polymorphisms of the human mu opioid receptor gene by hybridization or single nucleotide extension on custom oligonucleotide gelpad microchips: potential in studies of addiction. *Neuropsychiatr Genet* 2000;96:604–615.

8. Kreek MJ, Koob GF. Drug dependence: stress and dysregulation of brain reward pathways. *Drug Alcohol Depend* 1998;51:23–47.

9. Kreek MJ. Opioid receptors: some perspectives from early studies of their role in normal physiology, stress responsivity and in specific addictive diseases. *J Neurochem Res* 1996;21:1469–1488.

10. Kreek MJ. Opiates, opioids and addiction. *Mol Psychiatry* 1996;1:232–254.

11. Kreek MJ. Opiate and cocaine addictions: challenge for pharmacotherapies. *Pharm Biochem Behav* 1997;57:551–569.

12. Nestler EJ, Kelz MB, Chen J. ΔFosB: a molecular mediator of long-term neural and behavioral plasticity. *Brain Res* 1999;835:10–17.

13. Law PY, Loh HH. Regulation of opioid receptor activities. *J Pharmacol Exp Ther* 1999;289:607–624.

14. Akil H, Owens C, Gutstein H, et al. Endogenous opioids: overview and current issues. *Drug Alcohol Depend* 1998;51:127–140.

15. Massotte D, Kieffer BL. A molecular basis for opiate action. *Essays Biochem* 1998;33:65–77.

16. Shippenberg TS, Elmer GI. The neurobiology of opiate reinforcement. *Crit Rev Neurobiol* 1998;12:267–303.

17. Dykstra LA, Preston KL, Bigelow GE. Discriminative stimulus and subjective effects of opioids with *mu* and *kappa* activity: data from laboratory animals and human subjects. *Psychopharmacology (Berl)* 1997;130:14–27.

18. Nestler EJ. Molecular mechanisms of opiate and cocaine addiction. *Curr Opin Neurobiol* 1997;7:713–719.

19. Standifer KM, Pasternak GW. G proteins and opioid receptor-mediated signalling. *Cell Signal* 1997;9:237–248.

20. Nestler EJ, Aghajanian GK. Molecular and cellular basis of addiction. *Science* 1997;278:58–62.

21. Piros ET, Hales TG, Evans CJ. Functional analysis of cloned opioid receptors in transfected cell lines. *Neurochem Res* 1996;21:1277–1285.

22. Schulteis G, Koob GF. Reinforcement processes in opiate addiction: a homeostatic model. *Neurochem Res* 1996;21:1437–1454.

23. Yu L. Protein kinase modulation of mu opioid receptor signalling. *Cell Signal* 1996;8:371–374.

24. Nestler EJ. Under siege: the brain on opiates. *Neuron* 1996;16:897–900.

25. Zaki PA, Bilsky EJ, Vanderah TW, et al. Opioid receptor types and subtypes: the delta receptor as a model. *Annu Rev Pharmacol Toxicol* 1996;36:379–401.

26. Mansour A, Fox CA, Akil H, et al. Opioid-receptor mRNA expression in the rat CNS: anatomical and functional implications. *Trends Neurosci* 1995;18:22–29.

27. Kreek MJ. Long-term pharmacotherapy for opiate (primarily heroin) addiction: opiate agonists. In: Schuster CR, Kuhar MJ, eds. *Pharmacological aspects of drug dependence: toward an integrated neurobehavioral approach*. Berlin: Springer-Verlag, 1996:487–562.

28. Kreek MJ. Long-term pharmacotherapy for opiate (primarily heroin) addiction: opiate antagonists and partial agonists. In: Schuster CR, Kuhar MJ, eds. *Pharmacological aspects of drug dependence: toward an integrated neurobehavioral approach*. Berlin: Springer-Verlag, 1996:563–592.

29. Kreek MJ. Methadone-related opioid agonist pharmacotherapy for heroin addiction: history, recent molecular and neurochemical research and the future in mainstream medicine. *Ann NY Acad Sci* 2000;909:186–216.

30. O'Brien CP. A range of research-based pharmacotherapies for addiction. *Science* 1997;278:66–70.

31. O'Connor PG, Fiellin DA. Pharmacologic treatment of heroin-dependent patients. *Ann Intern Med* 2000;133:40–54.

32. LaForge KS, Kreek MJ, et al. Symposium XIII: allelic polymorphisms of human opioid receptors. Functional studies: Genetic contributions to protection from, or vulnerability to, addictive diseases. In: Harris, LS, ed. Problems of drug dependence, 1999: proceedings of the 61st annual scientific meeting of the College on Problems of Drug Dependence. *NIDA Res Monogr* 2000;180:47–50.

33. Kieffer BL. Opioids: first lessons from knockout mice. *Trends Pharmacol Sci* 1999;20:19–26.

34. Childers SR. Opioid receptors: pinning down the opiate targets. *Curr Biol* 1997;7:695–697.

35. Donovan DM, Miner LL, Sharpe L, et al. Transgenic mice: preproenkephalin and the mu opiate receptor. *Ann NY Acad Sci* 1996;780:19–28.

36. Zhang L, Yu Y, Mackin S, et al. Differential mu opiate receptor phosphorylation and desensitization induced by agonists and phorbol esters. *J Biol Chem* 1996;271:11449–11454.

37. Noble F, Cox BM. Differential desensitization of mu- and delta-opioid receptors in selected neural pathways following chronic morphine treatment. *Br J Pharmacol* 1996;117:161–169.

38. Yu Y, Zhang L, Yin X, et al. Mu opioid receptor phosphorylation, desensitization, and ligand efficacy. *J Biol Chem* 1997;272:28869–28874.

39. Noble F, Cox BM. The role of dopaminergic systems in opioid receptor desensitization in nucleus accumbens and caudate putamen of rat after chronic morphine treatment. *J Pharmacol Exp Ther* 1997;283:557–565.

40. Selley DE, Nestler EJ, Breivogel CS, et al. Opioid receptor-coupled G-proteins in rat locus coeruleus membranes: decrease in activity after chronic morphine treatment. *Brain Res* 1997;746:10–18.

41. Bernstein MA, Welch SP. Mu-opioid receptor down-regulation and cAMP-dependent protein kinase phosphorylation in a mouse model of chronic morphine tolerance. *Mol Brain Res* 1998;55:237–242.

42. Sim-Selley LJ, Selley DE, Vogt LJ, et al. Chronic heroin self-administration desensitizes mu opioid receptor-activated G-proteins in specific regions of rat brain. *J Neurosci* 2000;20:4555–4562.

43. Smith FL, Lohmann AB, Dewey WL. Involvement of phospholipid signal transduction pathways in morphine tolerance in mice. *Br J Pharmacol* 1999;128:220–266.

44. Keith DE, Murray SR, Zaki PA, et al. Morphine activates opioid receptors without causing their rapid internalization. *J Biol Chem* 1996;271:19021–19024.

45. Murray SR, Evans CJ, von Zastrow M. Phosphorylation is not required for dynamin-dependent endocytosis of a truncated mutant opioid receptor. *J Biol Chem* 1998;273:24987–24991.

46. Keith DE, Anton B, Murray SR, et al. Mu-opioid receptor internalization: opiate drugs have differential effects on a conserved endocytic mechanism in vitro and in the mammalian brain. *Mol Pharmacol* 1998;53:377–384.

47. Burford NT, Tolbert LM, Sadee W. Specific G protein activation and mu-opioid receptor internalization caused by morphine, DAMGO and endomorphin I. *Eur J Pharmacol* 1998;342:123–126.

48. Schulz R, Wehmeyer A, Murphy J, et al. Phosducin, beta-arrestin and opioid receptor migration. *Eur J Pharmacol* 1999;375:349–357.

49. Whistler JL, Chuang H, Chu P, et al. Functional dissociation of mu opioid receptor signaling and endocytoses: implications for the biology of opiate tolerance and addiction. *Neuron* 1999;23:737–746.

50. Zaki PA, Keith DE Jr, Brine GA, et al. Ligand-induced changes in surface mu-opioid receptor number: relationship to G protein activation? *J Pharmacol Exp Ther* 2000;292:1127–1134.

51. Trapaidze N, Gomes I, Cvejic S, et al. Opioid receptor endocytosis and activation of MAP kinase pathway. *Mol Brain Res* 2000;76:220–228.

52. Kramer HK, Simon EJ. Mu and delta-opioid receptor agonists induce mitogen-activated protein kinase (MAPK) activation in the absence of receptor internalization. *Neuropharmacology* 2000;39:1707–1719.

53. Unterwald EM, Rubenfeld JM, Imai Y, et al. Chronic opioid antagonist administration upregulates mu opioid receptor binding without altering mu opioid receptor mRNA levels. *Mol Brain Res* 1995;33:351–355.

54. Unterwald EM, Anton B, To T, et al. Quantitative immunolocalization of mu opioid receptors: regulation by naltrexone. *Neuroscience* 1998;85:897–905.

55. Hitzemann RJ, Hitzemann BA, Loh HH. Binding of 3H-naloxone in the mouse brain: effects of ions and tolerance development. *Life Sci* 1974;14:2393–2404.

56. Klee WA, Streaty RA. Narcotic receptor sites in morphine-dependent rats. *Nature* 1974;248:61–63.

57. Simon EJ, Hiller JM. *In vitro* studies on opiate receptors and their ligands. *Fed Proc* 1978;37:141–146.

58. Bond C, LaForge KS, Tian M, et al. Single nucleotide polymorphism in the human mu opioid receptor gene alters beta-endorphin binding and activity: possible implications for opiate addiction. *Proc Natl Acad Sci USA* 1998; 95:9608–9613.

59. Carlezon WA Jr, Haile CN, Coppersmith R, et al. Distinct sites of opiate reward and aversion within the midbrain identified using herpes simplex virus vector expressing GluR1. *J Neurosci* 2000;20:RC62.

60. Bohn LM, Lefkowitz RJ, Gainetdinov RR, et al. Enhanced morphine analgesia in mice lacking beta-arrestin 2. *Science* 1999;286:2495–2498.

61. Bohn LM, Gainetdinov RR, Lin F-T, et al. Mu-Opioid receptor desensitization by beta-arrestin-2 determines morphine tolerance but not dependence. *Nature* 2000;408:720–723.

62. Johnson SW, North RA.Opioids excite dopamine neurons by hyperpolarization of local interneurons. *J Neurosci* 1992;12:483–488.

63. Giros B, Jaber M, Jones SR, et al. Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. *Nature* 1996;379:606–612.

64. Rocha BA, Fumagalli F, Gainetdinov RR, et al. Cocaine self-administration in dopamine-transporter knockout mice. *Nat Neurosci* 1998;1:132–137.

65. Spielewoy C, Gonon F, Roubert C, et al. Increased rewarding properties of morphine in dopamine-transporter knockout mice. *Eur J Neurosci* 2000;12:1827–1837.

66. Hemby SE, Co C, Dworkin SI, et al. Synergistic elevations in nucleus accumbens extracellular dopamine concentrations during self-administration of cocaine/heroin combinations (Speedball) in rats. *J Pharmacol Exp Ther* 1999;288:274–300.

67. Xi ZX, Fuller SA, Stein EA. Dopamine release in the nucleus accumbens during heroin self-administration is modulated by kappa opioid receptors: an *in vivo* fast-cyclic voltammetry study. *J Pharmacol Exp Ther* 1998;284:151–161.

68. Claye LH, Maisonneuve IM, Yu J, et al. Local perfusion of dynorphin A_{1-17} reduces extracellular dopamine levels in the nucleus accumbens. In: Harris LS, ed. Problems of drug dependence, 1996: proceedings of the 58th annual scientific meeting of the College on Problems of Drug Dependence. *NIDA Res Monogr* 1997;174:113.

69. Hemby SE, Martin TJ, Co C, et al. The effects of intravenous heroin administration on extracellular nucleus accumbens dopamine concentrations as determined by *in vivo* microdialysis. *J Pharmacol Exp Ther* 1995;273:591–598.

70. Georges F, Stinus L, Bloch B, et al. Chronic morphine exposure and spontaneous withdrawal are associated with modifications of dopamine receptor and neuropeptide gene expression in the rat striatum. *Eur J Neurosci* 1999;11:481–500.

71. Przewlocka B, Turchan J, Lason W, et al. The effect of single and repeated morphine administration on the prodynorphin system activity in the nucleus accumbens and striatum of the rat. *Neuroscience* 1996;70:749–754.

72. Wang XM, Zhou Y, Spangler R, et al. Acute intermittent morphine increases preprodynorphin and kappa opioid receptor mRNA levels in the rat brain. *Mol Brain Res* 1999;66:184–187.

73. Trujillo KA, Bronstein DM, Sanchez IO, et al. Effects of chronic opiate and opioid antagonist treatment on striatal opioid peptides. *Brain Res* 1995;698:69–78.

74. Lee RS, Criado JR, Koob GF, et al. Cellular responses of nucleus accumbens neurons to opiate-seeking behavior. I. Sustained responding during heroin self-administration. *Synapse* 1999;33:49–58.

75. Unterwald EM, Horne-King J, Kreek MJ. Chronic cocaine alters brain mu opioid receptors. *Brain Res* 1992;584:314–318.

76. Unterwald EM, Cox BM, Kreek MJ, et al. Chronic repeated cocaine administration alters basal and opioid-regulated adenylyl cyclase activity. *Synapse* 1993;15:33–38.

77. Unterwald EM, Rubenfeld JM, Kreek MJ. Repeated cocaine administration upregulates kappa and mu, but not delta opioid receptors. *Neuroreport* 1994;5:1613–1616.

78. Maisonneuve IM, Ho A, Kreek MJ. Chronic administration of a cocaine "binge" alters basal extracellular levels in male rats: an *in vivo* microdialysis study. *J Pharmacol Exp Ther* 1995;272:652–657.

79. Sklair-Tavron L, Shi WX, Lane SB, et al. Chronic morphine induces visible changes in the morphology of mesolimbic dopamine neurons. *Proc Natl Acad Sci USA* 1996;93:11202–11207.

80. Spangler R, Unterwald EM, Kreek MJ. "Binge" cocaine administration induces a sustained increase of prodynorphin mRNA in rat caudate-putamen. *Mol Brain Res* 1993;19:323–327.

81. Spangler R, Ho A, Zhou Y, et al. Regulation of kappa opioid receptor mRNA in the rat brain by "binge" pattern cocaine administration and correlation with preprodynorphin mRNA. *Mol Brain Res* 1996;38:71–76.

82. Crain SM, Shen KF. Antagonists of excitatory opioid receptor functions enhance morphine's analgesic potency and attenuate opioid tolerance/dependence liability. *Pain* 2000;84:121–131.

83. Jeziorski M, White FJ. Dopamine receptor antagonists prevent expression, but not development, of morphine sensitization. *Eur J Pharmacol* 1995;14:235–244.

84. Churchill L, Roques BP, Kalivas PW. Dopamine depletion augments endogenous opioid-induced locomotion in the nucleus accumbens using both mu 1 and delta opioid receptors. *Psychopharmacology (Berl)* 1995;120:347–355.

85. Kreek MJ. Medical complications in methadone patients. *Ann NY Acad Sci* 1978;311:110–134.

86. Kreek MJ, Schluger J, Borg L, et al. Dynorphin A1-13 causes elevation of serum levels of prolactin through an opioid receptor mechanism in humans: gender differences and implications for modulations of dopaminergic tone in the treatment of addictions. *J Pharmacol Exp Ther* 1999;28:260–269.

87. Kreek MJ, Ho A, Borg L. Dynorphin A1-13 administration causes elevation of serum levels of prolactin in human subjects. In: Harris LS, ed. Problems of drug dependence, 1993:proceedings of the 55th annual scientific meeting of the College on Problems of Drug Dependence. *NIDA Res Monogr* 1994;141:108.

88. Dole VP, Nyswander ME, Kreek MJ. Narcotic blockade. *Arch Intern Med* 1966;118:304–309.

89. Kreek MJ. Medical safety, side effects and toxicity of methadone. *Proceedings of the fourth national conference on methadone treatment*. NAPAN, National Institute of Mental Health, 1972:171–174.

90. Kreek MJ. Medical safety and side effects of methadone in tolerant individuals. *JAMA* 1973;223:665–668.

91. Cushman P, Kreek MJ. Some endocrinologic observations in narcotic addicts. In: Zimmerman E, George R, eds. *Narcotic and the hypothalamus*. New York: Raven, 1974:161–173.

92. Stimmel B, Kreek MJ. Dependence, tolerance and withdrawal. In: Stimmel B, ed. *Heroin dependency: medical, economic and social aspects*. New York: Stratton Intercontinental Medical Book, 1975:88–97.

93. Kreek MJ, Wardlaw SL, Friedman J, et al. Effects of chronic exogenous opioid administration on levels of one endogenous opioid (beta-endorphin) in man. In: Simon E, Takagi H, eds. *Advances in endogenous and exogenous opioids*. Tokyo: Kodansha, 1981:364–366.

94. Kreek MJ, Wardlaw SL, Hartman N, et al. Circadian rhythms and levels of beta-endorphin, ACTH, and cortisol during chronic methadone maintenance treatment in humans. *Life Sci* 1983;33:409–411.

95. Kreek MJ, Raghunath J, Plevy S, et al. ACTH, cortisol and beta-endorphin response to metyrapone testing during chronic methadone maintenance treatment in humans. *Neuropeptides* 1984;5:277–278.

96. Kosten TR, Kreek MJ, Raghunath, et al. Cortisol levels during chronic naltrexone maintenance treatment in ex-opiate addicts. *Biol Psychiatry* 1986;21:217–220.

97. Kosten TR, Kreek MJ, Raghunath, et al. A preliminary study of beta-endorphin during chronic naltrexone maintenance treatment in ex-opiate addicts. *Life Sci* 1986;39:55–59.

98. Kreek MJ. Multiple drug abuse patterns and medical consequences. In: Meltzer HY, eds. *Psychopharmacology: the third generation of progress*. New York: Raven, 1987:1597–1604.

99. Kennedy JA, Hartman N, Sbriglio R, et al. Metyrapone-induced withdrawal symptoms. *Br J Addict* 1990;85:1133–1140.

100. Kreek MJ. Rationale for maintenance pharmacotherapy of opiate dependence. In: O'Brien CP, Jaffe JH, eds. *Addictive states*. New York: Raven, 1992:205–230.

101. Kreek MJ. Methadone-related opioid agonist pharmacotherapy for heroin addiction: history, recent molecular and neurochemical research and the future in mainstream medicine. *Ann NY Acad Sci* 2000;909:186–216.

102. Kosten TR, Morgan C, Kreek MJ. Beta-endorphin levels during heroin, methadone, buprenorphine and naloxone challenges: preliminary findings. *Biol Psychiatry* 1992;32:523–528.

103. Schluger J, Bodner G, Gunduz M, et al. Abnormal metyrapone tests during cocaine abstinence. In: Harris LS, ed. Problems of drug dependence, 1997: proceedings of the 59th annual scientific meeting of the College on Problems of Drug Dependence. *NIDA Res Monogr* 1998;178:105.

104. Schluger JH, Borg L, Ho A, et al. Altered HPA axis responsivity to metyrapone testing in methadone maintained former heroin addicts with ongoing cocaine addiction. *Neuropsychopharmacology* 2001;24:568–575.

105. Culpepper-Morgan JA, Inturrisi CE, Portenoy RK, et al. Treatment of opioid induced constipation with oral naloxone: a pilot study. *Clin Pharmacol Ther* 1992;23:90–95.

106. Rosen MI, McMahon TJ, Margolin A, et al. Reliability of sequential naloxone challenge tests. *Am J Drug Alcohol Abuse* 1995;21:453–467.

107. Culpepper-Morgan JA, Kreek MJ. HPA axis hypersensitivity to naloxone in opioid dependence: a case of naloxone induced withdrawal. *Metabolism* 1997;46:130–134.

108. Schluger JH, Ho A, Borg L, et al. Nalmefene causes greater hypothalamic-pituitary-adrenal axis activation than naloxone in normal volunteers: implications for the treatment of alcoholism. *Alcohol Clin Exp Res* 1998;22:1430–1436.

109. Zhou Y, Spangler R, LaForge KS, et al. Corticotropin-releasing factor and CRF-R1 mRNAs in rat brain and pituitary during "binge" pattern cocaine administration and chronic withdrawal. *J Pharmacol Exp Ther* 1996;279:351–358.

110. Zhou Y, Schlussman SD, Ho A, et al. Effects of chronic "binge" pattern cocaine administration on plasma ACTH and corticosterone levels in mice deficient in DARPP-32. *Neuroendocrinology* 1999;70:196–199.

111. Zhou Y, Spangler R, Maggos CE, et al. Hypothalamic-pituitary-adrenal activity and pro-opiomelanocortin mRNA levels in the hypothalamus and pituitary of the rat are differentially modulated by acute intermittent morphine with or without water restriction stress. *J Endocrinol* 1999;163:261–267.

112. Zhou Y, Spangler R, Maggos CE, et al. Steady-state methadone in rats does not change mRNA levels of corticotropin-releasing factor, its pituitary receptor or proopiomelanocortin. *Eur J Pharmacol* 1996;315:31–35.

113. Zhou Y, Franck J, Spangler R, et al. Reduced hypothalamic POMC and anterior pituitary CRF1 receptor mRNA levels after acute, but not chronic, daily "binge" intragastric alcohol administration. *Alcohol Clin Exp Res* 2000;24:1575–1582.

114. Shaham Y, Rajabi H, Stewart J. Relapse to heroin-seeking in rats under opioid maintenance: the effects of stress, heroin priming, and withdrawal. *J Neurosci* 1996;16:1957–1963.

115. Shaham Y, Funk D, Erb S, et al. Corticotropin-releasing factor, but not corticosterone, is involved in stress-induced relapse to heroin-seeking rats. *J Neurosci* 1997;17:2605–2614.

116. Piazza PV, LeMoal M. The role of stress in drug self-administration. *Trends Pharmacol Sci* 1998;19:67–74.

117. Pich EM, Lorang M, Yeganeh M, et al. Increase of extracellular corticotropin-releasing factor-like immunoreactivity levels in the amygdala of awake rats during restraint stress and ethanol withdrawal as measured by microdialysis. *J Neurosci* 1995;15:5439–5447.

118. Richter RM, Pich EM, Koob GF, et al. Sensitization of cocaine-stimulated increase in extracellular levels of corticotropin-releasing factor from the rat amygdala after repeated administration as determined by intracranial microdialysis. *Neurosci Lett* 1995;187:169–172.

119. Branch AD, Unterwald EM, Lee SE, et al. Quantitation of preproenkephalin mRNA levels in brain regions from male Fischer rats following chronic cocaine treatment using a recently developed solution hybridization procedure. *Mol Brain Res* 1992;14:231–238.

120. Unterwald EM, Ho A, Rubenfield JM, et al. Time course of the development of behavioral sensitization and dopamine receptor up-regulation during binge cocaine administration. *J Pharmacol Exp Ther* 1994;270:1387–1396.

121. Zubieta J-K, Gorelick DA, Stauffer R, et al. Increased mu opioid receptor binding detected by PET in cocaine-dependent men is associated with cocaine craving. *Nat Med* 1996;2:1225–1229.

122. Kreek MJ. Methadone disposition during the perinatal period in humans. *Pharmacol Biochem Behav* 1979;11:7–13.

123. Burstein Y, Grady RW, Kreek MJ, et al. Thrombocytosis in the offspring of female mice receiving dl-methadone. *Proc Soc Exp Biol Med* 1980;164:275–279.

124. Inturrisi CE, Verbely K. A gas-liquid chromatographic method for the quantitative determination of methadone in human plasma and urine. *J Chromatogr* 1972;65:361–369.

125. Kreek MJ. Plasma and urine levels of methadone. *NY State J Med* 1973;73:2773–2777.

126. Kreek MJ, Hachey DL, Klein PD. Stereoselective disposition of methadone in man. *Life Sci* 1979;24:925–932.

127. Ahmed SH, Koob GF. Transition from moderate to excessive drug intake: change in hedonic set point. *Science* 1998;282:298.

128. Tornatzky W, Miczek KA. Cocaine self-administration "binges": transition from behavioral and autonomic regulation toward homeostatic dysregulation in rats. *Psychopharmacology* 2000;148:289–298.

129. Mantsch JR, Schlussman, SD, Ho A, et al. Effects of cocaine self-administration on plasma corticosterone and prolactin in rats. *J Pharmacol Exp Ther* 2000;294:239–247.

130. Ahmed, SH, Walker JR, Koob, GF. Persistent increase in the motivation to take heroin in rats with history of drug escalation. *Neuropsychopharmacology* 2000;22:413–421.

131. Specker S, Wanaukul W, Hatsukami DE, et al. Effects of dynorphin A(1-13) on opiate withdrawal in humans. *Psychopharmacology* 1998;137:326–332.

132. Portenoy RK, Caraceni A, Cherny NI, et al. Dynorphin A(1-13) analgesia in opioid-treated patients with chronic pain: a controlled pilot study. *Clin Drug Invest* 1999;17:33–42.

133. Krystal JH, Woods SW, Kosten TR, et al. Opiate dependence and withdrawal: preliminary assessment using single photon emission computerized tomography (SPECT). *Am J Drug Alcohol Abuse* 1995;21:47–63.

134. Kling M, Carson RE, Borg L, et al. Opioid receptor imaging with PET and [^{18}F]cyclofoxy in long-term methadone-treated former heroin addicts. *J Pharmacol Exp Ther* 2000;295:1070–1076.

135. Sell LA, Morris J, Bearn J, et al. Activation of reward circuitry in human opiate addicts. *Eur J Neurosci* 1999;11:1042–1048.

136. Berrettini WH, Hoehe MR, Ferraro TN, et al. Human mu opioid receptor gene polymorphisms and vulnerability to substance abuse. *Addict Biol* 1997;2:303–308.

137. Bergen AW, Kokoszka J, Peterson R, et al. μ-Opioid receptor gene variants: lack of association with alcohol dependence. *Mol Psychiatry* 1997;2:430–434.

138. LaForge KS, Yuferov V, Kreek MJ. Opioid receptor and peptide gene polymorphisms: potential implications for addiction. *Eur J Pharmacol* 2000;410:249–268.

139. Tsuang MT, Lyons MJ, Meyer JM, et al. Co-occurrence of abuse of different drugs in men. *Arch Gen Psychiatry* 1998;55:967–972.

140. Roberts AJ, McDonald JS, Heyser CJ, et al. Mu-Opioid receptor knockout mice do not self-administer alcohol. *J Pharmacol Exp Ther* 2000;293:1002–1008.

105

Current and Experimental Therapeutics for the Treatment of Opioid Addiction

Paul J. Fudala

George E. Woody

Paul J. Fudala: Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania; Department of Behavioral Health Service, Veterans Affairs Medical Center, Philadelphia, Pennsylvania.

George E. Woody: Substance Abuse Treatment Unit, VA Medical Center, Philadelphia, Pennsylvania.

Currently, numerous effective pharmacologic and behavioral therapies are available for the treatment of opioid addiction, and these two types of therapies often are combined to optimize patient management. Newer therapeutic options may take various forms. For example, methadone maintenance is an established treatment modality, whereas the use of buprenorphine and naloxone in an office-based setting represents a new variation on that theme. Clonidine has been used extensively to ameliorate opioid withdrawal signs, whereas lofexidine is a structural analogue that appears to have less hypotensive and sedating effects. The depot dosage form of naltrexone, currently under development, may increase compliance with a medication that has been an effective opioid antagonist but that has been underused secondary to patient nonacceptance. In almost every treatment episode using pharmacotherapy, it is combined with some type of psychosocial or behavioral treatment. Recent research has documented the value of these additional treatments and has provided insight into the ones that are the most effective. This chapter reviews current and experimental treatments for opioid addiction with an emphasis on some of the newer, more promising, and interesting therapies.

TREATMENT PARADIGMS

Part of "105 - Current and Experimental Therapeutics for the Treatment of Opioid Addiction "

Long-Term, Short-Term, Rapid, and Ultrarapid Opioid Detoxification

Detoxification from opioids, for most patients, is only the first phase of a longer treatment process. Most patients seeking treatment have been addicted to heroin or other opioids for 2 to 3 years, some for 30 years or more. Thus, treatment usually involves changes in patients' lifestyles. Although generally ineffective in producing sustained remission unless combined with long-term pharmacologic, psychosocial, or behavioral therapies, detoxification alone continues to be widely used and studied. It is sometimes the only option available for patients who do not meet United States Food and Drug Administration (FDA) criteria for, do not desire, or do not have access to agonist medications such as methadone or methadyl acetate (L-α-acetylmethadol or LAAM).

The detoxification process may include use of opioid agonists (e.g., methadone), partial agonists (e.g., buprenorphine), antagonists (e.g., naloxone, naltrexone), or nonopioid alternatives such as clonidine, benzodiazepines, or nonsteroidal antiinflammatory agents. In many cases, one or more medications are combined, such as naloxone with clonidine and a benzodiazepine. The choice of detoxification medication and the duration of the process depend on numerous factors including patient preference, clinician expertise and experience, type of treatment facility, licensing, and available resources. Ultimately, however, the goal of detoxification is the achievement (and maintenance) of a drug-free state while minimizing withdrawal. Unfortunately, however, detoxification for some patients appears to be used in a punitive manner or as an expedient means to achieve a drug-free state rapidly with no follow-up pharmacologic or behavioral therapy.

Opioid detoxification paradigms are frequently categorized according to their nominal duration: long-term (typically 180 days), short-term (up to 30 days), rapid (typically 3 to 10 days), and ultrarapid (1 to 2 days). These temporal modifiers provide only a coarse description of the paradigm; they do not provide other important information such as the medications used or whether postdetoxification pharmacologic (e.g., naltrexone maintenance), psychosocial, or behavioral therapy is provided. However, some general guidelines typically apply.

The most common detoxification protocols, and those for which the most data are available, are the long-term (typically 180 days) and short-term (up to 30 days) paradigms involving the use of methadone. Unfortunately, these strategies have not generally been associated with acceptable treatment response using relapse to opioid use as an outcome criterion. For example, one study reported that more than half the patients participating in a 180-day detoxification program were using opioids illicitly during the medication-taper phase of the protocol (1). Six-month follow-up indicated that 38.5% of the urine samples (n = 26) tested negative for illicit opioids, only three of 31 patients reported remaining free of illicit opioids for the entire 6 months before follow-up, and 22 participated in some other form of treatment (2). Results from more rapid detoxification evaluations using short- or even intermediate-term (up to 70 days) medication-tapering protocols are even less encouraging and have an unfortunately low success rate. However, provision of additional services such as counseling, behavioral therapy, treatment of underlying psychopathologies, job skills training, and family therapy to address concomitant treatment needs can improve outcome, although success rates remain low, even with these services (3).

Rapid detoxification involves the use an opioid antagonist, typically naltrexone or naloxone, in combination with other medications (such as clonidine and benzodiazepines) to mitigate the precipitated withdrawal syndrome. The procedure is intended to expedite and compress the withdrawal process to minimize discomfort and to decrease treatment time. Ultrarapid detoxification also uses other medications, along with an opioid antagonist, to moderate withdrawal effects. However, rather than being awake as they are during the rapid detoxification process, patients are placed under general anesthesia or, alternatively, are deeply sedated. A comprehensive review of the rapid and ultrarapid detoxification literature (identifying 12 and 9 of each type study, respectively) has been published (4). Rapid detoxification studies were conducted in inpatient facilities, outpatient substance abuse treatment settings, and outpatient primary care facilities; ultrarapid ones were confined to inpatient settings. Patients included those who were heroin dependent as well as those in methadone maintenance treatment.

Only four of the studies reviewed provided follow-up beyond the initial detoxification. Retention on postdetoxification naltrexone maintenance in one rapid detoxification study was 53% at 1 month and 82% in another at 3 months. Only one of the ultrarapid detoxification studies provided follow-up information indicating that all patients (11 of 11) were taking naltrexone 30 days after detoxification (5). A more recently published study (6), in which ultrarapid detoxification was followed by naltrexone maintenance and supportive psychotherapy, indicated that 49 of 72 patients were opioid abstinent 12 months after detoxification. All these studies involved self-selected patients; thus, it is impossible to know the overall effectiveness of this type of intervention.

A major concern regarding ultrarapid detoxification in particular is the occurrence of potentially serious adverse effects, such as respiratory distress (7), or other pulmonary and renal complications (8), during or immediately after the procedure. A high frequency of vomiting has also been reported (9). The degree to which serious adverse effects occur has not yet been determined; however, there have been press reports of sudden death occurring shortly after the procedure that were not caused by relapse to opioid use and overdose.

In spite of the emerging evidence about serious adverse effects, ultrarapid detoxification may be appropriate for highly selected patients based on considerations of previous treatment history, economic factors, and patient choice. However, patients seeking this treatment must be thoroughly informed that serious adverse effects, including sudden unexpected deaths, have occurred in association with this procedure, and its use should probably be limited to inpatient settings where monitoring by anesthesiologists and other highly trained staff is available.

Buprenorphine, a μ-opioid partial agonist, has also been used as a detoxification agent. Results from inpatient (10 ,11 and 12) and outpatient (13 ,14) studies have shown that it is safe and well tolerated, and it mitigates opioid withdrawal signs and symptoms over a range of doses and detoxification schedules. Clonidine, an α_2-adrenergic agonist, has been shown to suppress many of the autonomic signs and symptoms of opioid withdrawal. It can cause pronounced sedation and hypotension but has been used with few problems when appropriate monitoring is available. It does not suppress the subjective discomfort of withdrawal, and probably for that reason, it is not well accepted by most opioid addicts.

Other α_2-adrenergic agonists have also been evaluated to find agents that are equally or more effective, but produce less sedation and hypotension than clonidine. Lofexidine, a medication that was originally promoted as an antihypertensive, has been the most thoroughly studied. When compared with clonidine, it was found to suppress autonomic signs and symptoms of opioid withdrawal equally, but with less sedation and hypotension (15 ,16 and 17). When compared with methadone dose tapering, lofexidine detoxification was associated with opioid withdrawal effects that peaked sooner, but resolved to negligible levels more rapidly (18). In another study (19), an accelerated 5-day lofexidine treatment regimen attenuated opioid withdrawal symptoms more rapidly than 10 days of either lofexidine or methadone, with similar blood pressure responses observed for the two lofexidine groups. Data regarding the potential effectiveness of guanabenz and guanfacine have also been reported, but further studies are required to assess the potential utility of these medications for detoxification treatment. In summary, recent studies have shown that lofexidine is

likely to be a useful opioid detoxification agent whose efficacy approximates that of clonidine but with fewer side effects.

Opioid Agonist Pharmacotherapy

Methadone maintenance was developed by Dole and Nyswander and has become the most commonly used pharmacotherapy for opioid dependence (20). Methadone acts as the μ-opioid receptor, and its ability to suppress opioid withdrawal for 24 to 36 hours after a single oral dose makes it an ideal medication for this purpose. Another μ-opioid agonist, LAAM, received FDA approval for maintenance treatment in 1993. LAAM is a long-acting congener of methadone that suppresses withdrawal for 48 to 72 hours and thus has the advantage of requiring less frequent clinic visits than methadone, which must be taken daily. A third medication, buprenorphine, is far advanced in the FDA approval process. It was mentioned earlier as a detoxification agent and is discussed later and in more detail because it has unique properties that are likely to result in its being used with fewer regulatory controls than methadone and LAAM.

Both methadone and LAAM are Schedule II controlled substances and can be used only for maintenance and detoxification in programs that are licensed and regulated by the FDA and the Drug Enforcement Administration (DEA). The regulations specify who is eligible for treatment, procedures that are required for its administration, the number of take-home doses permitted, and the type of medication storage security needed. Treatment programs have been inspected approximately every 3 years for the past 30 years, and violations have resulted in sanctions ranging from administrative citations to criminal prosecution.

The combination of FDA and DEA regulations has resulted in a treatment system that is separated from the mainstream of other medical care and that consists almost entirely of specially licensed and inspected clinics. Clinics are often located in old buildings that have been converted to comply with regulations but that were never intended for medical use. At the present time, it is estimated that approximately 179,000 patients are being maintained on methadone or LAAM at 940 or more sites, and this number represents only about 20% of the opioid addicts in the United States (21).

This treatment-program regulatory system has been under increasing criticism since the early 1990s. Criticism has come from both patients and treatment providers who believe that the current regulations impose unnecessary burdens and expenses, have done little to improve the quality of treatment, and impede access to care. The importance of these criticisms has been underlined by the recent increase in heroin addiction (22), by evidence that methadone maintenance reduces the incidence of hepatitis and HIV infection, and by the lack of coverage for agonist maintenance by most health insurance plans. The overall effect has been a widening gap between treatment need and availability, and lost treatment opportunities.

As a result, the Institute of Medicine and the National Institutes of Health each made recommendations for regulatory reform (23 ,24). Many of these recommendations are in the process of being carried out and include an overall reduction in regulations and transfer of oversight to accreditation bodies that are approved by the Center for Substance Abuse Treatment, rather than the FDA. Other recommendations include allowing long-term, stable patients to be treated in settings other than methadone clinics where they can receive up to 30 days of take-home medications *(medical maintenance)*, allowing take-home doses for LAAM, and allowing more clinical judgment in determining dosages and take-home schedules. Procedures to prevent diversion include careful screening of patients who receive medical maintenance, return to directly observed medication administration if illicit drug use or diversion is detected, random urine testing, and call-back procedures in which patients will be required to report to the medical treatment setting and to produce the remaining, unused take-home containers.

The appropriate agonist medication dosage has been a subject of both federal and state regulations, although there has been a gradual shift toward allowing more clinical judgment in its determination. Numerous studies have been conducted since the mid-1970s to determine the optimal dose, and, although it is clear that some patients do well on low doses of methadone or LAAM (about 20 to 50 mg), studies have consistently shown that most patients need higher doses if they are to achieve maximum benefit from agonist treatment (25). The results of these methadone dose comparison studies are generally supportive of the guidelines originally proposed by Dole and Nyswander, who recommended doses in the 80- to 120-mg per day range (20). Clear relationships between methadone blood levels and clinical response have not been observed consistently. One study found significant correlations between oral dose and methadone concentration, but only among patients who complained of low dosing (26). These findings suggest that some patients may be more sensitive to dosage changes and that clinical response, including subjective complaints, is a more important guide to adequate dose levels than specific blood levels. No controlled studies have been done examining doses higher than 120 mg; thus, the upper limits of dosing effectiveness are not well understood.

Perhaps the most important pending regulatory change is to amend the Controlled Substances Act with respect to registration requirements for practitioners using drugs approved for detoxification or maintenance that are in Schedules III, IV, and V (27). Physicians who choose to treat persons with opioid dependence under the new regulations will need to notify the Secretary of Health and Human Services in writing of their intent and to show that they

are qualified to provide addiction treatment by virtue of certification or experience. No physician would be allowed to treat more than 30 patients at one time without special approval, according to the legislation as it is now proposed.

This change in the regulations will be especially important for buprenorphine and the buprenorphine-naloxone combination (discussed later), because it will provide better access to treatment for persons who are unwilling or unable to be treated in the current methadone or LAAM system. The overall intent of the proposed regulatory reform is to better integrate maintenance treatment into the mainstream of medical care, to make it more readily available, and to improve its quality.

As mentioned earlier, these changes are likely to influence the ways in which buprenorphine is used in opioid addiction treatment. Buprenorphine is marketed internationally as an analgesic (both without naloxone and with naloxone to deter abuse) and as a treatment for opioid addiction. The most widespread use of buprenorphine is in France, where it was approved for the latter indication in 1996. In the United States, buprenorphine is currently approved only as an analgesic for parenteral administration; approval for opioid addiction treatment is pending. Buprenorphine has been used almost exclusively sublingually in addiction treatment because of its poor oral bioavailability. Most of the early clinical trials used a sublingual solution of buprenorphine formulated in a hydroethanolic vehicle, although a more commercially suitable sublingual tablet formulation is now used.

The greatest advantage of buprenorphine compared with full agonists such as methadone and LAAM is the plateau effect of μ-agonist activity. Parenteral doses as high as 12 mg intravenously (28) have been given to opioid-intolerant patients with only limited adverse effects (e.g. sedation, irritability, nausea, itching). Numerous large trials have confirmed the utility of buprenorphine for agonist maintenance therapy. These studies have included comparisons of buprenorphine with placebo (29 ,30), a buprenorphine-naloxone combination with placebo (30), and a multiple-dose comparison study (31). In one of the most recent trials (32), buprenorphine (given three times weekly) was compared with LAAM (given three times weekly) and methadone (given daily) in a 17-week study. Mean retention in treatment was higher for buprenorphine, LAAM, and high-dose methadone groups compared with low-dose methadone and for high-dose methadone compared with LAAM. Opioid-positive urine samples decreased most for the LAAM-treated group and least for low-dose methadone. Patient self-reports of opioid use did not differ among the groups, but they showed decreases of about 90% over the course of the study.

Buprenorphine has the potential to be abused and can produce addiction. However, most persons who abuse buprenorphine initiated opioid use with other drugs. Abuse may take the form of using greater than prescribed dosages for analgesia, using buprenorphine in place of a more desired but less available opioid such as heroin, or using buprenorphine for its own positive subjective effects (33 ,34). Only one study published to date has characterized the behavioral and physiologic effects of a wide range of buprenorphine analgesic doses in nonusers of opioids (35). The results indicated that buprenorphine, given intravenously, has a low abuse liability in this population.

Buprenorphine, in combination with naloxone, has less potential for abuse than buprenorphine alone (36 ,37). The therapeutic utility of combining naloxone with buprenorphine derives from the low sublingual bioavailability of naloxone compared with buprenorphine. Parenteral misuse of the combination by persons addicted to opioids would be expected to produce antagonist-like effects; thus, most opioid addicts would not be likely to inject the combination more than once. The use of the buprenorphine-naloxone combination product in an office-based setting represents an innovative alternative to the restrictive methadone or LAAM maintenance paradigm described previously. The use of this new drug combination should expand the availability of agonist maintenance treatment with a relatively low risk for abuse or diversion. In addition, the partial agonist activity of buprenorphine results in a much lower risk of overdose death than is the case with methadone or LAAM.

Antagonist Maintenance

Naltrexone is the prototypical opioid antagonist used in abstinence maintenance therapy; this drug blocks the effects of heroin and other opioids through competitive receptor inhibition. Naltrexone has no opioid agonist effects and is a competitive opioid antagonist. It is orally effective and can block opioid effects for 24 hours when administered as a single daily dose of 50 to 60 mg. Higher doses usually will not block opioid effects for 48 to 72 hours though they will provide more cross tolerance to heroin and other opioids during the 24-hour dosing period (38). Despite a favorable adverse event profile (nausea is typically the most common side effect), naltrexone is generally not favored by opioid addicts because, unlike opioid agonists and partial agonists, it produces no positive, reinforcing effects. Furthermore, it may be associated with the precipitation of an opioid withdrawal syndrome if it is used too soon after opioid use stops, an effect that can be minimized by administering a naloxone challenge test before giving naltrexone.

Although the literature on naltrexone treatment spans more than 25 years, work continues on increasing medication compliance and improving outcomes. Some of these more recent efforts include work to develop a depot form that will block opioid effects for 14 to 28 days. This dosage form is currently in phase II clinical trials. At present, a patient treated with naltrexone has only to stop the medication for 1 to 3 days to experience the full effects of subsequent opioid use. A depot dosage form of naltrexone would provide more time for patients to overcome ambivalence about stopping opioid use and could result in more long-term success than has currently been the case. Another variant

on antagonist treatment is nalmefene, an orally effective but somewhat longer-acting (about 48 hours at dosages of 50 to 100 mg per day) opioid antagonist that has been effective for alcohol treatment that may have advantages over naltrexone due to its longer duration of action. The problem will be that addicts may not take it, as has generally been the case with naltrexone (39 ,40).

Psychosocial and Behavioral Treatment

Research has called attention to the finding that, as in other substance use disorders, most patients with opioid dependence and abuse are ambivalent about stopping drug use (41 ,42). This ambivalence presents a therapeutic challenge because it contributes to varying levels of motivation to enter and remain in treatment, to early dropout, and to partial or (in some cases) nontreatment response. Studies have emphasized that treatment providers must be aware of this “normal” ambivalence and make reasonable efforts to resolve it in favor of treatment participation and cessation of drug use (42). Suggestions have been made regarding initial steps to maximize the chances for engagement in treatment and cessation of drug use. These include avoiding unnecessary delays in entering treatment, expressing a hopeful and nonjudgmental attitude, performing a comprehensive evaluation, and developing a treatment plan that is responsive to patients’ self-identified goals (41).

In addition to challenges related to ambivalence, patients often have serious problems with nonopioid substance abuse or with medical, psychiatric, legal, employment, and family or social issues that preexist or result from the addiction. Research has found that addressing these additional problems can be helpful, but they are complex and require coordination between agonist pharmacotherapy staff and other medical and psychosocial services (43 ,44).

The most common type of psychosocial service in opioid agonist treatment is individual drug counseling. Counselors are typically persons at the masters level or below who deliver a behaviorally focused treatment aimed to identify specific problems, to help the patient access services that may not be provided in the clinic (e.g., medical, psychiatric, legal, family or social), to stop substance use, and to improve overall adjustment. Functions that counselors perform include monitoring methadone and LAAM doses and requesting changes when needed, reviewing urine test results, responding to requests for take-homes doses, assisting with family problems, assessing and responding to crises, writing letters for court or social welfare agencies, recommending inpatient treatment when necessary, and providing support and encouragement for a drug-free lifestyle.

Counseling usually addresses both opioid and nonopioid use. Although nicotine (tobacco) use is not always included, the increased emphasis on adverse health effects of smoking has resulted in more attention to stop smoking at all levels, including drug counseling. Counselors and patients typically have weekly, 30- to 60-minute sessions during the first weeks or months of treatment with reductions in frequency to biweekly or monthly depending on progress. The frequency of counseling can vary widely depending on the severity of the patient’s problems, clinic requirements, and counselor workload.

The importance of regular counseling was clearly demonstrated in a study by McLellan and co-workers (43), in which patients were randomly assigned to minimal counseling (one 5- to 10- minute session per month), standard counseling (one 45-minute session per week), or enhanced counseling (standard plus on-site referral to psychiatric, medical, and family or social services). Results showed a dose–response relationship with the minimal condition doing significantly worse than standard and enhanced counseling doing the best overall; however, about 30% of patients did well in the minimal counseling condition. This study clearly demonstrated the positive benefits achieved by drug counseling and showed that, for most patients, counseling is necessary to bring out the maximum benefits of agonist maintenance.

Most counseling is individual, one on one, but some programs use group therapy exclusively. However, most programs use groups only for selected patients with focal problems such as HIV disease, posttraumatic stress disorder, homelessness, loss of close personal relationships, or not at all. Many programs encourage patients to participate in self-help groups, but ask them to be careful to select a group that accepts persons who are receiving opioid agonist maintenance treatment. Some programs have self-help groups that meet regularly on site. Counselors, like psychotherapists, can vary widely in the results they achieve (45). This variability seems more related to the ability to form a positive, helping relationship with the patient than to specific techniques (46).

Contingency management techniques are always included in drug counseling, if for nothing else than to fulfill regulations about requiring progress in treatment as a condition of providing take-home doses, and studies have shown that they can be very helpful. For example, an opportunity to receive take-home medications in return for drug-free urine tests is a powerful motivator for many patients (47). Such a contingency strategy is an example of research with a clear use in general clinical practice because it is easily applied and costs little or nothing beyond standard program costs. More flexibility in dispensing take-home doses as contingencies for positive behaviors could be a positive effect of the regulatory reforms described earlier.

Another contingency that is easily applicable and that some programs have used with positive results is requiring a negative alcohol breath test before dispensing the daily dose of methadone or LAAM. This contingency can be especially useful for patients with alcohol abuse or dependence. Maintenance, counseling, and contingency management are often combined in complex ways, as seen in the following vignette:

A 42-year-old man presented for his sixth episode of

methadone maintenance. He had a long history of alcoholism and was using cocaine regularly. He had done fairly well on methadone as far as illicit opioid use was concerned, but his clinic attendance and ability to comply with clinic rules, especially regarding take-home doses, were severely compromised by alcohol use. In the past, he would remain in treatment for about a year, then become angry over his inability to obtain take-home doses because of positive breathalyzer tests, drop out, and have a relapse to opioid use. He had frequently been offered inpatient detoxification for alcoholism but always refused because “alcohol’s not my problem, heroin’s the problem,” and he could not take time off from work (as a stockperson in a liquor store). When he presented for treatment most recently, he was unemployed (secondary to alcohol problems) and living with his parents, who were threatening to put him out because of drug use. He agreed that, as part of his treatment plan, he would go into the hospital for alcohol detoxification and stabilization on methadone and then be discharged to maintenance therapy. After inpatient discharge, he attended AA-style counseling, requested daily alcohol breath tests, and turned down an offer to return to his job at the liquor store. He remained stable for 3 years on 65 mg per day of methadone with no urine samples positive for opioids (but occasionally positive for cocaine), and he enrolled (and continued) in school.

It is clear that in such a complex but relatively typical case, a single intervention was not enough. Rather, a series of coordinated steps was necessary to achieve a positive treatment response. Although not demonstrated in this vignette, family therapy is another intervention that can be combined with agonist therapy and other psychosocial interventions, and studies have shown that it can be useful as well (48).

Although counseling and other services are effective enhancements of agonist treatment, adherence is often an issue, and clinics vary in the way they respond to this problem. Some remind patients of appointments, others do not permit patients to be medicated unless they keep appointments, and others suspend patients who miss appointments. For nonadherent patients, a very powerful contingency is requiring certain behaviors for patients to remain on the program, a procedure that is often formalized in a *treatment contract*. Here, the patient is given an option of stopping unprescribed drug use, keeping regular counseling appointments, looking for work, or correcting other behaviors that need improvement as a condition for remaining in treatment. Patients who fail are administratively detoxified, suspended for months to years, and referred to another program, although the referrals are not always successful.

The long-term effects of this form of contingency management have not been well studied. For example, relatively little is known about negative effects on patients who may have improved with methadone and counseling, but not to the degree required by the contingency, and who are subsequently discharged for failing to adhere to a treatment contract. One study done in Philadelphia (49) found that among 110 patients who were administratively discharged or dropped out of a Veterans Affairs (VA) maintenance program, 8.2% (nine of 110) died within the following year as compared with only 1% (four of 397) who remained in treatment. Among the 43 patients (from among the 110) who were discharged for failing to adhere to a treatment contract, five (11.6%) died within a year. None of these five patients were in treatment at the time of death, and all died as a result of overdoses. No overdose deaths occurred among patients remaining in treatment, and, interestingly, there were no deaths in those patients who were suspended for violating program rules (mainly drug dealing or giving a false urine specimen). These results are consistent with data from New South Wales, Australia, where there has been a sharp rise in heroin-related deaths. Although it is estimated that 20% to 30% of the heroin addicts in New South Wales are receiving methadone maintenance, only 3% of the 953 heroin-related fatalities occurred among patients receiving methadone maintenance (50). These data emphasize the fine line between contingencies maintained in programs and the dangers associated with program dismissal.

The foregoing data, when considered along with studies showing a protective effect of maintenance on acquiring HIV infection (51), have made some clinicians increasingly hesitant to suspend patients from maintenance treatment for positive urine test results alone. This caution may be especially relevant in environments where the potency of heroin is high, such as Philadelphia, where the average “bag” of heroin is now 71% pure (22).

Therapeutic communities are another psychosocial approach that is often useful for opioid addicts who have a long history of addiction and a strong motivation to become drug free. These programs are very selective, self-governing, long-term (6 to 18 months) residential settings where patients share responsibilities for maintaining the treatment milieu (cleaning, cooking, and leading group therapy). Confrontation of denial and behaviors such as lying and “conning,” combined with group support for healthy, positive change, is used to restructure character and the addictive lifestyle. Medications such as methadone, LAAM, or naltrexone are rarely used; however, medications for specific psychiatric or medical conditions are usually available after careful screening and evaluation. Patients who enter therapeutic communities are often referred by the criminal justice system. Some patients have tried, but not responded, to agonist maintenance on repeated occasions. Although dropout rates are high, studies have shown that more than 80% of patients who complete a course of treatment in a therapeutic community have a sustained remission and demonstrate significant improvement in psychiatric symptoms, employment, and criminal behavior (52 ,53).

Addressing Comorbidity

Patients seeking treatment for opioid dependence are typically dependent on one or more other substances (cocaine, alcohol, benzodiazepines, amphetamines, marijuana, nicotine), and have additional problems in the psychiatric, medical, family or social, employment, or legal areas. In fact, it is rare to find a person with only opioid dependence and no other substance abuse or without a psychiatric, medical, or family or social problem. The presence of these problems, perhaps with the exception of nicotine dependence, tends to magnify the severity of the opioid dependence and makes the patient more difficult to treat.

Among the psychiatric disorders seen in persons with opioid dependence, antisocial personality disorder is one of the most common (54). Diagnostic studies of persons with opioid dependence have typically found rates of antisocial personality disorder ranging from 20% to 50%, as compared with less than 5% in the general population. Posttraumatic stress disorder is also seen with increased frequency.

Opioid-dependent persons are especially at risk for the development of brief depressive symptoms and for episodes of mild to moderate depression that meet symptomatic and duration criteria for major depressive disorder or dysthymia. These syndromes represent both substance-induced mood disorders as well as independent depressive illnesses. Brief periods of depression are especially common during chronic intoxication or withdrawal or in association with psychosocial stressors that are related to the dependence. Insomnia is common, especially during withdrawal; sexual dysfunction, especially impotence, is common during intoxication. Delirium or brief, psychotic-like symptoms are occasionally seen during opioid intoxication (54).

The data on psychiatric comorbidity among opioid addicts and its negative effect on outcome (55) have stimulated research on the effect of combining psychiatric and substance abuse treatment. Several studies have now shown that tricyclic antidepressants can be useful for chronically depressed opioid addicts who are treated with methadone maintenance (56). Two studies have shown that professional psychotherapy can be useful for psychiatrically impaired, methadone-maintained opioid addicts, although another study found no psychotherapy effect (57 ,58 and 59). The main result in most pharmacotherapy and psychotherapy studies with methadone-maintained addicts has usually been a reduction in psychiatric symptoms such as depression, although some studies have shown reductions in substance use as well (56 ,57).

Fewer than 5% of persons with opioid dependence have psychotic disorders such as bipolar illness or schizophrenia; however, these patients can present special problems because programs typically have few psychiatric staff members. As a result, these patients are sometimes excluded from methadone treatment because of the severity of their psychotic disorders. Others are treated with methadone, counseling, and the same antipsychotic or antimanic medications used for nonaddicted patients with similar disorders. Although studies evaluating the outcome of combining opioid agonist treatment with antipsychotic or antimanic medications have not been done, there is little controversy that these medications are useful for opioid addicts with psychotic disorders, and most programs use them with little hesitation.

Women opioid addicts can present special challenges because many have been sexually abused as children, have other psychiatric disorders, and are involved in difficult family or social situations (60). Abusive relationships with addicted men are common, sometimes characterized by situations in which the man exerts control by providing drugs. These complex psychiatric and relationship issues have emphasized the need for comprehensive psychosocial services that include psychiatric assessment and treatment and access to other medical, family, and social services.

Medical comorbidity is a major problem among opioid addicts; HIV infection, AIDS, and hepatitis B and C have become some of the most common illnesses. Sharing injection equipment, including “cookers” and rinse water, and engaging in high-risk sexual behavior are the main routes of infection. Sexual transmission appears to be a more common route for HIV transmission among women than men because the HIV virus is spread more readily from men to women than from women to men. Females patients who are intravenous drug users and who also engage in prostitution or other forms of high-risk sexual behavior are at extremely high risk of HIV infection (60). Cocaine use has been found to be a significant risk factor as a single drug of abuse or when used in combination with heroin or other opioids (61).

As mentioned earlier, mortality is high, and studies have found annual death rates of approximately 10 per 1,000 or greater, which is substantially higher than demographically matched samples in the general population (62). Common causes of death are overdose, accidents, injuries, and medical complications such as cellulitis, hepatitis, AIDS, tuberculosis, and endocarditis. The cocaine and alcohol dependence that is often seen among opioid-dependent persons contributes to medical morbidity by cirrhosis, cardiomyopathy, myocardial infarction, or serious cardiac arrhythmias.

Tuberculosis has become a particularly serious problem among intravenous drug users, especially heroin addicts. In most cases, infection is asymptomatic and is evident only by the presence of a positive tuberculin skin test. However, many cases of active tuberculosis have also been found, especially among those who are infected with HIV who may have a newly acquired infection or reactivation of a prior infection as a result of impaired immune function.

After rising rapidly in the late 1970s and early 1980s, the incidence of new HIV infections among intravenous drug users, of whom opioid-dependent persons constitute

a large proportion, decreased (63). However, as a result of high levels of needle sharing and other risky behavior in the early phases of the epidemic, the prevalence of HIV infection among heroin addicts reached as high as 50% in some areas of the United States (64). Because of the long incubation period before the development of AIDS, it is expected that future years will continue to see high levels of morbidity and mortality associated with HIV infection, although the advent of new pharmacotherapies for HIV has extended many lives.

Studies done over the last several years have identified several important interactions between methadone and drugs used to treat HIV infection. Information is not complete, however, and more studies are needed to map out the extent of these interactions completely. One important interaction is that methadone increases plasma levels of zidovudine; the associated symptoms resemble methadone withdrawal. There have been instances in which methadone doses have been increased in response to complaints of withdrawal, with increasing doses compounding the problem. Another important interaction involves decreased methadone blood levels secondary to nevirapine administration that may be associated with mild to moderate withdrawal. This interaction can be important if the patient discontinues either of these two drugs while taking methadone, because the result may be a sudden rise in methadone blood levels with signs and symptoms of overmedication (65 ,66).

Other medical complications of heroin dependence are seen in children born to opioid-dependent women. Perhaps the most serious is premature delivery and low birth weight, a problem that can be reduced if the mother is receiving methadone maintenance and prenatal care (67). Another is physiologic dependence on opioids, seen in about half the infants born to women maintained on methadone or dependent on heroin or other opioids. Effective treatments for neonatal withdrawal are available, and long-term adverse effects of opioid withdrawal have not been demonstrated. Adverse neonatal effects associated with LAAM or buprenorphine have not been observed, but few studies have been done because neither medication is approved for use in pregnancy.

The possibility that breast-feeding may cause adverse effects in infants of methadone-maintained mothers was studied. It was found that methadone was present in the breast milk of women maintained on doses as high as 180 mg, but the concentration was very low, and no adverse effects were observed in the infants (68). HIV infection is seen in about one-third of infants born to HIV-positive mothers, but this incidence can be reduced to about 10% if HIV-positive pregnant women are given zidovudine before delivery (69). HIV can also be transmitted by breast-feeding, and thus infant formula feeding is recommended for babies of HIV-positive mothers, except in some developing countries, where formula is unavailable or unaffordable. Thorough washing of infants born to HIV-infected mothers immediately after delivery also appears to reduce the incidence of HIV infection.

An important line of research resulting from the data on comorbidity has been studies on the effects of integrating psychiatric and medical care within agonist and other substance abuse treatment programs (70). Clinical experience and National Institute on Drug Abuse demonstration projects have shown that integration of these services with agonist maintenance can be done, and with very positive results, because patients are seen frequently and treatment retention is high (44). Related to this line of research are studies that have shown improved compliance with directly observed antituberculosis pharmacotherapy (71). These findings have important implications for tuberculosis control policies in methadone programs because intravenous drug users are at very high risk of tuberculosis infection and because maintenance programs provide settings in which directly observed therapy can be easily applied. Similar principles apply to administration of psychotropic medication in noncompliant patients with schizophrenia or other major axis I disorders.

Harm Reduction

Harm reduction is concerned with minimizing various negative consequences of addiction. As such, the focus is shifted away from drug use to the consequences of use and its attendant behaviors (72). Examples of harm reduction include needle exchange programs, efforts directed at reducing drug-use-associated behaviors that may result in the transmission of HIV, and making changes in policies (including increasing treatment availability) that reduce heroin use and the criminal behavior associated with drug procurement. Harm reduction refers to reducing harm not only to the individual addict, but also to family, friends, and society generally. Other terms sometimes used synonymously with harm reduction include harm minimization, risk reduction, and risk minimization (73).

Some authors have identified the limitations of harm reduction when it is used as a sole strategy to combat the adverse effects of addiction. For example, Reuter and Caulkins pointed out the benefit of integrating drug use reduction and harm reduction components into a single framework (74), because total harm may be lowered by reducing either component. Roche and colleagues proposed a model for an integrated addiction treatment strategy that incorporates harm reduction and use reduction with abstinence and nonuse (75), in addition to other critical elements such as factors related to culture and gender. Additionally, MacCoun provided a template for integrating harm reduction with prevalence reduction (discouraging the engagement in drug use) and quantity reduction (encouraging the reduction in frequency or extent of drug use) (76).

With regard to opioids, much of the health-related harm from their improper or illicit use is secondary to elements

other than the substances themselves (77). Sequelae of unhygienic methods of administration and poor injection technique are typically more serious than the constipation or other side effects of the drugs themselves, acute overdoses notwithstanding. With regard to opioid addiction treatment, medications such as methadone, LAAM, and buprenorphine, among others (including supervised heroin substitution) used for maintenance agonist treatment, may be considered harm reduction measures. All have the potential to reduce morbidity, mortality, and crime associated with the addict lifestyle. However, in this sense they are no different from other medical therapies such as those used for the treatment of hypertension, diabetes, or asthma.

Needle or syringe exchange represents one of the most controversial strategies in harm reduction. Research indicates that these types of programs may have beneficial effects in numerous areas, including a reduction in the spread of blood-borne infectious disease such as hepatitis and HIV, and acting as a conduit to more comprehensive drug-abuse treatment services (78). In one study (79), the initiation and continuation of syringe exchange program participation among high-risk injection drug users were independently associated with a cessation of syringe sharing. In another study (80), participation in a needle exchange program was associated with patients' entering detoxification treatment for both HIV-infected and noninfected groups. Not all findings have been positive, however. In a study designed to assess the association between risk behaviors and HIV seroprevalence and incidence among injection drug users, risk elevations for HIV associated with needle exchange programs were substantial and consistent despite adjustment for confounding factors (81). However, an examination of potential bias in nonrandomized comparisons (82) suggested that injection drug users participating in needle exchange programs at a given point may include a high proportion of persons whose pattern of drug use puts them at greater risk for blood-borne viral infections. Further, a prospective cohort study found no evidence of a causal association between needle exchange program participation and transmission of HIV (83).

Harm reduction related to psychoactive substance abuse has gone through numerous stages. The current phase has been described as the development of an integrated public health perspective for all drugs in which a multifaceted, strategic approach is taken (84). The direction of this approach will be guided, in part, by whether biases against a harm reduction philosophy can be overcome by those who see it as synonymous with acceptance of drug abuse or legalization, and how harm reduction objectives relate to an overall strategy to improve public health.

IMPACT OF MANAGED CARE

Part of "105 - Current and Experimental Therapeutics for the Treatment of Opioid Addiction "

Efforts to control costs by managed care have resulted in a marked reduction in use of inpatient or residential treatment programs in many locations. Funds saved from these cost reductions have often not been invested in outpatient treatment. A good example is the VA, which administers the largest network of substance abuse treatment programs in the United States. Since the application of managed care policies, the overall amount spent on substance abuse treatment declined by 41%, from $597 million in 1993 to $351 million in 1999. Measured as a percentage of overall VA health care costs, specialized substance abuse care decreased from 4.2% in 1993 to 2.3% in 1999. In contrast, overall VA health care expenditures increased 10% between 1993 and 1999 (85). Most of these reductions were achieved by reducing inpatient beds, with the funds saved allocated to other areas but not to reinvestment in other substance abuse treatment services. The result has been an overall reduction in the total number of veteran patients treated and in the amount of drug counseling provided. As a result, no new methadone programs were opened in the VA despite the recent increase in heroin addiction, evidence of waiting lists for methadone treatment, and cities (such as Portland, Oregon) with serious heroin problems but no agonist maintenance programs in spite of recent increases in heroin overdose deaths.

A focused review of substance abuse programs by the United States Senate Committee on Veterans Affairs found that changes in resource allocation have caused programs to become vulnerable to service disruptions, poor morale, burnout, and reduced motivation and quality of performance and characterized by failures to maintain service levels in accord with the mandates of law (86).

Managed care strategies have also made it very difficult to integrate medical and psychiatric services into agonist maintenance programs. Thus, both old and new pharmacotherapies for opioid addiction described earlier are underused in the VA, the largest substance abuse treatment system. There is every indication that penetration of these new treatments into the opiate treatment field at large has also been slow.

SUMMARY

Part of "105 - Current and Experimental Therapeutics for the Treatment of Opioid Addiction "

New pharmacotherapies, behavioral therapies, and treatment strategies are being developed for opioid addiction. This continued development is important, because more treatment options will encourage treatments that are more individualized and balanced across important dimensions such as patient response, adverse effects, treatment costs, comorbidity, living situations, and overall adjustment. As described earlier, various treatments can be combined to produce better patient outcomes. However, the overall effect of these developments on addiction treatment and public health is very dependent on funding support, which has become a serious problem. Parity legislation may help to

solve funding problems and result in the expansion of treatment to meet patient needs, but the details of how and when more investment in substance treatment may occur are unclear as of this writing.

REFERENCES

1. Banys P, Tusel DJ, See KL, et al. Low (40mg) versus high (80mg) dose methadone in a 180-day heroin detoxification program. *J Subst Abuse Treat* 1994;11:225–232.

2. Reilly PM, Banys P, Tusel DJ, et al. Methadone transition treatment: a treatment model for 180-day methadone detoxification. *Int J Addict* 1995;30:387–402.

3. Kleber HD. Opioids: detoxification. In: Galanter M, Kleber HD, eds. *Textbook of substance abuse treatment*, second ed. Washington, DC: American Psychiatric Association, 1999:251–269.

4. O'Connor PG, Kosten TR. Rapid and ultrarapid opioid detoxification techniques. *JAMA* 1998;279:229–234.

5. Legarda JJ, Gossop M. A 24-h inpatient detoxification treatment for heroin addicts: a preliminary investigation. *Drug Alcohol Depend* 1994;35:91–93.

6. Hensel M, Kox WJ. Safety, efficacy, and long-term results of a modified version of rapid opiate detoxification under general anesthesia: a prospective study in methadone, heroin, codeine, and morphine addicts. *Acta Anaesthesiol Scand* 2000;44:3226–3233.

7. San L, Puig M, Bulbena A, et al. High risk of ultrashort noninvasive opiate detoxification. *Am J Psychiatry* 1995;152:956.

8. Pfab R, Hirtl C, Zilker T. Opiate detoxification under anesthesia: no apparent benefit but suppression of thyroid hormones and risk of pulmonary and renal failure. *J Toxicol Clin Toxicol* 1999;37:43–50.

9. Cucchia AT, Monnat M, Spagnoli J, et al. Ultra-rapid opiate detoxification using deep sedation with or midazolam: short- and long-term results. *Drug Alcohol Depend* 1998;52:243–250.

10. Cheskin LJ, Fudala PJ, Johnson, RE. A controlled comparison of buprenorphine and clonidine for acute detoxification from opioids. *Drug Alcohol Depend* 1994;36:115–121.

11. Parran TV Jr, Adelman CL, Jasinski DR. A buprenorphine stabilization and rapid-taper protocol for the detoxification of opioid dependent patients. *Am J Addict* 1994;3:306–313.

12. Vignau J. Preliminary assessment of a 10-day rapid detoxification programme using high dosage buprenorphine. *Eur Addict Res* 1998;4:29–31.

13. O'Connor PG, Carroll KM, Shi JM, et al. Three methods of opioid detoxification in a primary care setting: a randomized trial. *Ann Intern Med* 1997;127:526–530.

14. Diamant K, Fischer G, Schneider C, et al. Outpatient opiate detoxification treatment with buprenorphine: preliminary investigation. *Eur Addict Res* 1998;4:198–202.

15. Kahn A, Mumford JP, Rogers GA, et al. Double-blind study of lofexidine and clonidine in the detoxification of opiate addicts in hospital. *Drug Alcohol Depend* 1997;44:57–61.

16. Lin SK, Strang J, Su LW, et al. Double blind randomized controlled trial of lofexidine versus clonidine in the treatment of heroin withdrawal. *Drug Alcohol Depend* 1997;48:127–133.

17. Carnwath T, Hardman J. Randomised double-blind comparison of lofexidine and clonidine in the out-patient treatment of opiate withdrawal. *Drug Alcohol Depend* 1998;50:251–254.

18. Bearn J, Gossop M, Strang J. Randomized double-blind comparison of lofexidine and methadone in the in-patient treatment of opiate withdrawal. *Drug Alcohol Depend* 1996;43:87–89.

19. Bearn J, Gossop M, Strang J. Accelerated lofexidine treatment regimen compared with conventional lofexidine and methadone treatment for in-patient opiate detoxification. *Drug Alcohol Depend* 1998;50:227–232.

20. Dole VP, Nyswander M. Successful treatment of 750 criminal addicts. *JAMA* 1968;26:2708–2710.

21. *Addict Treat Forum* 2000;9:2.

22. Community Epidemiology Work Group. *Epidemiologic trends in drug abuse: advance report*. Washington, DC: NIDA, Dept. of Health and Human Services, 1999.

23. Fulco DR, Liverman CT, Earley LE, eds. *The development of medications for the treatment of opiate and cocaine addictions: issues for the government and private sector*. Washington, DC: Institute of Medicine, National Academy Press, 1995.

24. NIH Consensus Statement. *Effective medical treatment of opiate addiction*. Rockville, MD: National Institutes of Health, 1997.

25. Ball JC, Ross A. *The effectiveness of methadone maintenance treatment*. New York: Springer-Verlag, 1991.

26. Hiltunen AJ, Beck O, Hjemdahl P, et al. Rated well-being in relation to plasma concentrations of l- and d-methadone in satisfied and dissatisfied patients on methadone maintenance treatment. *Psychopharmacology* 1999;143:385–393.

27. House Resolution 2634. United States House of Representatives, 2000.

28. Umbricht-Schneiter A, Huestis MA, Cone EJ, et al. Safety of buprenorphine: ceiling for cardio-respiratory effects at high IV doses. *NIDA Res Monogr* 1998;179:225.

29. Johnson RE, Eissenberg T, Stitzer ML, et al. A placebo controlled trial of buprenorphine as a treatment for opioid dependence. *Drug Alcohol Depend* 1995;40:17–25.

30. Fudala PJ, Bridge TP, Herbert S et al. A multisite efficacy evaluation of a buprenorphine/ naloxone product for opiate dependence treatment. *NIDA Res Monogr* 1998;179:105.

31. Ling W, Charuvastra C, Collins JC, et al. Buprenorphine maintenance treatment for opiate dependence: a multicenter, randomized clinical trial. *Addiction* 1998;93:475–486.

32. Johnson RE, Chutuape MA, Strain EC, et al. A comparison of levomethadyl acetate, buprenorphine, and methadone as treatments for opioid dependence. *N Engl J Med* 2000;343:1290–1297.

33. Baumevieille M, Haramburu F, Bégaud B. Abuse of prescription medicines in southwestern France. *Ann Pharmacother* 1997;31:847–850.

34. Dore GM, Hargreaves G, Niven BE, et al. Dependent opioid users assessed for methadone treatment in Otago: patterns of drug use. *NZ Med J* 1997;110:162–165.

35. Zacny JP, Conley K, Galinkin J. Comparing the subjective, psychomotor and physiological effects of intravenous buprenorphine and morphine in healthy volunteers. *J Pharmacol Exp Ther* 1997;282:1187–1197.

36. Fudala PJ, Yu E, Macfadden W, et al. Effects of buprenorphine and naloxone in morphine-stabilized opioid addicts. *Drug Alcohol Depend* 1998b;50:1–8.

37. Mendelson J, Jones RT, Welm S, et al. Buprenorphine and naloxone combinations: the effects of three dose ratios in morphine-stabilized, opiate-dependent volunteers. *Psychopharmacology* 1999;141:37–46.

38. Lee MC, Wagner HN, Tanada S, et al. Duration of occupancy of opiate receptors by naltrexone. *J Nucl Med* 1988;29:1207–1211.

39. Mason BJ, Ritvo EC, Morgan, RO, et al. A double-blind, placebo-controlled pilot study to evaluate the efficacy and safety of oral nalmefene HCl for alcohol dependence. *Alcohol Clin Exp Res* 1994;18:1162–1167.

40. Jones HE, Johnson RE, Fudala PJ, et al. Nalmefene: blockade of intravenous morphine challenge effects in opioid abusing humans. *Drug Alcohol Depend* 2000;60:29–37.

41. Miller WR, Rollnick S. *Motivational interviewing: preparing people to change addictive behavior*. New York: Guilford, 1991.

42. Rollnick S, Mason P, Butler C. *Health behavior change: a guide for practitioners*. New York: Churchhill Livingstone, 1999.

43. McLellan AT, Arndt IO, Metzger DS, et al. The effects of psychosocial services in substance abuse treatment. *JAMA* 1993;269:1953–1959.

44. Umbricht-Schneiter A, Ginn DH, Pabst KM, et al. Providing medical care to methadone clinic patients: referral vs on-site care. *Am J Public Health* 1994;84:207–210.

45. McLellan AT, Woody GE, Luborsky L, et al. Is the counselor an "active ingredient" in substance abuse treatment? *J Nerv Ment Dis* 1988;176:423–430.

46. Luborsky L, McLellan AT, Woody GE, et al. Therapist success and its determinants. *Arch Gen Psychiatry* 1985;42:602–611.

47. Iguchi, MY, Lamb, RJ, Belding, MA, et al. Contingent reinforcement of group participation versus abstinence in a methadone maintenance program. *Exp Clin Psychopharmacol* 1996;4:315–321.

48. Stanton MD, Todd T. *The family therapy of drug abuse and addiction*. New York: Guilford, 1982.

49. Zanis DA, Woody GE. One-year mortality rates following methadone treatment discharge. *Drug Alcohol Depend* 1998;52:257–260.

50. Darke S, Ross J, Zador D, et al. Heroin-related deaths in New South Wales, Australia, 1992–1996. *Drug Alcohol Depend* 2000;60:141–15.

51. Metzger DS, Navaline H, Woody GE. Drug abuse treatment as AIDS prevention. *Public Health Rep* 1998;113S:97–106.

52. Inciardi JA, Martin SS, Butzin CA, et al. An effective model of prison-based treatment for drug-involved offenders. *J Drug Issues* 1997;27:261–278.

53. DeLeon G. Therapeutic communities. In: Galanter M, Kleber HD, eds. *Textbook of substance abuse treatment*. Washington, DC: American Psychiatric Association, 1999:447–462.

54. First MB, Pincus HA, eds. *Diagnostic and statistical manual of mental disorders*, fourth ed (text revision: DSM-IV-TR). Washington, DC: American Psychiatric Association, 1999.

55. McLellan AT, Luborsky L, Woody GE, et al. Predicting response to alcohol and drug abuse treatments: role of psychiatric severity. *Arch Gen Psychiatry* 1983;40:620–625.

56. Nunes EV, Quitkin FM, Donovan S, et al. Imipramine treatment of opiate dependent patients with depressive disorders: a placebo-controlled trial. *Arch Gen Psychiatry* 1998;55:153–160.

57. Woody GE, McLellan AT, Luborsky L, et al. Psychiatric severity as a predictor of benefits from psychotherapy the Penn-VA study. *Am J Psychiatry* 1984;141:1172–1177.

58. Woody GE, Mercer DE. Individual psychotherapy for substance use disorders. In: Gabbard GO, ed. *Treatments of psychiatric disorders*, third ed. Washington, DC: American Psychiatric Association, 2001:849–861.

59. Rounsaville BJ, Glazer W, Wilber CH, et al. Short-term interpersonal psychotherapy in methadone maintained opiate addicts. *Arch Gen Psychiatry* 1982;39:161–166.

60. Blume SB. Addiction in women. In: Galanter M, Kleber HD, eds. *Textbook of substance abuse treatment*. Washington, DC: American Psychiatric Association, 1999:485–494.

61. Booth RE, Kwiatkowski CF, Chitwood DD. Sex related HIV risk behaviors: differential risks among injection drug users, crack smokers, and injection drug users who smoke crack. *Drug Alcohol Depend* 2000;58:219–226.

62. Gronbladh L, Ohlund LS, Gunne LM. Mortality in heroin addiction: impact of methadone treatment. *Acta Psychiatr Scand* 1990;82:223–227.

63. Seage G, Holte SE, Metzger DS, et al. Are US populations appropriate for HIV vaccine trials? The HIVNET Vaccine Preparedness Study *Am J Epidemiol* 2001;153:619–627.

64. DesJarlais D, Marmor M, Friedmann P, et al. HIV Incidence among injection drug users in New York City, 1992–1997: evidence for a declining epidemic. *Am J Public Health* 2000;90:352–359.

65. Altice FL, Friedland GH, Cooney EL. Nevirapine induced opiate withdrawal among injection drug users with HIV infection receiving methadone. *AIDS* 1999;13:957–962.

66. Otero, MJ, Fuentes, A, Sanchez, R, et al. Nevirapine-induced withdrawal symptoms in HIV patients on methadone maintenance programme; an alert. *AIDS* 1999;13:1004–1005.

67. Finnegan, LP. Treatment issues for opioid-dependent women during the perinatal period. *J Psychoactive Drugs* 1991;23:191–201.

68. McCarthy JJ, Posey BL. Methadone levels in human milk. *J Hum Lactation* 2000;16:115–120.

69. Connor, EM, Sperling, RS, Gelber R, et. al. Reduction of maternal-infant transmission of human immunodeficiency virus type 1 with zidovudine treatment. *N Engl J Med* 1994;331:1173–1180.

70. Kessler RC, Nelson CB, McGonagle KA, et al. The epidemiology of co-occurring addictive and mental disorders: implications for prevention and service utilization. *Am J Orthopsychiatry* 1996;66:17–31.

71. Chaulk CP, Moore-Rice K, Rizzo R, et al. Eleven years of community-based directly observed therapy for tuberculosis. *JAMA* 1995;274:945–951.

72. Marlatt GA. Harm reduction: come as you are. *Addict Behav* 1996;21:779–788.

73. Riley D, Sawka E, Conley P, et al. Harm reduction: concepts and practice. A policy discussion paper. *Subst Use Misuse* 1999;34:9–24.

74. Reuter P, Caulkins JP. Redefining the goals of national drug policy: recommendations from a working group. *Am J Public Health* 1995;85:1059–1063.

75. Roche AM, Evans KR, Stanton WR. Harm reduction: roads less travelled to the Holy Grail. *Addiction* 1997;92:1207–1212.

76. MacCoun RJ. Toward a psychology of harm reduction. *Am Psychologist* 1998;53:1199–1208.

77. Kalant H. Differentiating drugs by harm potential: the rational versus the feasible. *Subst Use Misuse* 1999;34:25–34.

78. Normand J, Vlahov D, Moses L, eds. *Preventing HIV transmission: the role of sterile needles and bleach*. Washington, DC: National Research Council, Institute of Medicine. National Academy Press, 1995.

79. Bluthenthal RN, Kral AH, Gee L, et al. The effect of syringe exchange use on high-risk injection drug users: a cohort study. *AIDS* 2000;14:605–611.

80. Strathdee SA, Celentano DD, Shah N, et al. Needle-exchange attendance and health care utilization promote entry into detoxification. *J Urban Health* 1999;76:448–460.

81. Bruneau J, Lamothe F, Franco E, et al. High rates of HIV infection among injection drug users participating in needle exchange programs in Montreal: results of a cohort study. *Am J Epidemiol* 1997;146:994–1002.

82. Hagan H, McGough JP, Thiede H, et al. Volunteer bias in nonrandomized evaluations of the efficacy of needle-exchange programs. *J Urban Health* 2000;77:103–112.

83. Schechter MT, Strathdee SA, Cornelisse PG, et al. Do needle exchange programmes increase the spread of HIV among injection drug users? An investigation of the Vancouver outbreak. *AIDS* 1999;13:F45–F51.

84. Erickson PG. Introduction: the three phases of harm reduction. An examination of emerging concepts, methodologies, and critiques. *Subst Use Misuse* 1999;34:1–7

85. Chen S, Wagner TH, et al. Expenditures for substance abuse treatment in the Department of Veterans Affairs. *1993–1999: report from the Health Economics Resource Center and the Center for Health Care Evaluation*. Menlo Park, CA: VA Palo Alto Health Care System, 2000.

86. Rockefeller JD. Minority staff review of VA programs for veterans with special needs. United States Senate, 1999.

106
Marijuana

Billy R. Martin

William L. Dewey

Vincenzo Di Marzo

Billy R. Martin and William L. Dewey: Department of Pharmacology and Toxicology, Virginia Commonwealth University, Richmond, Virginia.

Vincenzo DiMarzo: Istituto per la Chimica di Molecole Di Interesse Biologico, Consiglio Nazionale delle Ricerche, Naples, Italy.

Marijuana continues to garner considerable attention and is the subject of intense public debate and scientific scrutiny. It is unquestionably one of the most frequently used illicit drugs throughout the world. In Western countries, the pattern of use among age groups has not deviated significantly since the mid-1970s. The most prevalent use occurs in persons who are in their late teens and early twenties. Despite modest declines from the pinnacle of its use in the mid-1970s, there was an upsurge in use during the 1990s. Marijuana smoking is prevalent regardless of age, ethnicity, and sex. Epidemiologic data reveal that the rates of use during the year 2000 by United States students in grades 8, 10, and 12 were 17%, 32%, and 38%, respectively (1). There was a steady increase in daily marijuana use (defined as use on at least 20 occasions in the past 30 days) in all three age groups. For example, 2% of high school seniors used marijuana daily in 1991, with this figure rising to 6% in 1999. Two factors that undoubtedly contribute to the prevalence of marijuana use are the declining perception that marijuana produces harmful effects and the relative ease of acquiring marijuana.

Although marijuana has long been a subject of folklore medicine, interest as a potential therapeutic agent has intensified in recent times, likely as a result of numerous factors. The marijuana debate has increased, correlated with a period in our history marked by increased interest in nontraditional medications, increased awareness of several disease states not readily treated by current medications, and increased discourse about the public policy concerning recreational use of marijuana. Proponents cite a plethora of self-reports regarding the effectiveness of medical marijuana for a wide range of disorders, whereas opponents question its efficacy and point to potential deleterious effects of smoked marijuana. The scientific rationale for deciding the fate of marijuana as a therapeutic agent is often ignored.

As the debate concerning potential therapeutic benefits and health consequences of acute and chronic exposure to marijuana continues, considerable new scientific evidence is emerging regarding the nature of cannabinoid effects both *in vivo* and *in vitro* and the endogenous system through which marijuana acts. The emphasis of this chapter is to summarize recent discoveries of the endogenous system as they relate to both putative adverse effects and therapeutic uses of marijuana and its psychoactive constituent, Δ^9-tetrahydrocannabinol (THC). During the 1990s, numerous breakthroughs occurred that greatly increased our understanding of the cannabinoids. It is now evident that an endogenous cannabinoid system exists. A receptor has been characterized and cloned, second messenger systems identified, a putative endogenous ligand isolated and synthesized, and biochemical pathways for both synthesis and degradation of endogenous ligands described.

PHARMACOLOGY

Part of "106 - Marijuana "

Marijuana has prominent effects on the central nervous system as well as numerous peripheral effects in humans that are well recognized and were reviewed in the previous edition of this book. Acutely, cannabis produces an altered state of consciousness characterized by mild euphoria, relaxation, perceptual alterations including time distortion, and the enhancement of ordinary sensory experiences. Cognitive effects are also marked, such as impaired short-term memory. Motor skills and reaction time are also altered, so skilled activity of various kinds is frequently disrupted. Peripherally, marijuana produces prominent effects on the cardiovascular system characterized by tachycardia, and at high doses it can produce orthostatic hypotension. There are several other effects, such as antiemesis, analgesia, anticonvulsant action, and intraocular pressure lowering, discussed later. However, THC has provided most of the evidence for cannabinoid effects in laboratory animals. As for chronic marijuana use in humans, concerns arise because of potential long-term consequences on cognition and the development

of tolerance and dependence. There has been considerable interest in cannabinoid effects on performance, cognition, and the development of dependence, discussed in the following sections.

Performance

Cannabinoids affect sensory, psychomotor, and cognitive function and the ability to perform certain tasks. There is little dispute that high doses of marijuana can disrupt performance when the task is difficult. As may be expected, the effects of marijuana on performance become more variable as the complexity of the task is simplified and the dose of marijuana is reduced. In a comprehensive review, Chait and Pierri concluded that marijuana, at doses that produce moderate levels of intoxication, can affect a wide range of learned and unlearned behaviors, including simple motor tasks, and more complex psychomotor and cognitive tasks (2). Cannabinoid-induced impairment of flying and driving has been documented in numerous studies, although interpretation of the results remains controversial. THC is frequently found in the blood of drivers involved in automobile accidents, and marijuana use has been associated with impaired field sobriety test performance.

One study compared the effects of marijuana on equilibrium and simulated driving (3). Marijuana smoking that produced a “high” also increased body sway and increased brake latency to a degree comparable to that found in persons with breath alcohol concentrations near 0.05%. Marijuana smoking also acutely produces decrements in smooth-pursuit eye tracking (4). Although robust acute effects of marijuana were found on subjective and physiologic measures, and on smooth-pursuit eye tracking performance, no effects were evident the day after administration, a finding indicating that the residual effects of smoking a single marijuana cigarette are minimal. Furthermore, perceptual motor speed and accuracy, two very important parameters of driving ability, seem to be impaired immediately after marijuana use (5).

Cognition

There is lack of consensus regarding the effects of Δ^9-THC on memory and learning in that results are often inconsistent and test specific (2 ,6). Hall et al. concluded from clinical observations and cross-cultural studies that chronic marijuana use does not appear to produce severe gross impairment, but rather it may produce subtle cognitive deficits (7). The most frequently mentioned deficits were slower psychomotor performance, poorer perceptual motor coordination, and memory dysfunction. During the past few years, greater attention has been directed toward investigating specific cognitive deficits and relating these effects directly to marijuana use. Whereas THC appears to produce its greatest decrement in free recall or short-term memory, it has been proposed that chronic marijuana use in adolescents may result in long-term memory impairment (6). There are also indications that persons with learning disabilities may be more susceptible to memory deficits (6). Almost all studies have shown that marijuana has no effect on retrieval of already-learned material. THC reliably alters the perception of time, with subjects overestimating elapsed time or experiencing an increase in the subjective rate of time (2). Evidence has emerged from several studies that chronic marijuana use after many years produces subtle cognitive changes, specifically with regard to attention, as well as organization and integration of complex information (8).

These effects on memory have been supported by cannabinoid-induced deficits in several animal models. The naturally occurring cannabinoids as well as a wide range of synthetic compounds have been demonstrated to impair learning and memory in numerous laboratory animal memory tasks. Lichtman and Martin found that many synthetic cannabinoids impaired spatial memory in rats, as assessed by the eight-arm radial maze, and retarded completion time (9). Direct injection of cannabinoids into the hippocampus impaired memory, and this appeared specific to cognition because no other pharmacologic effects were produced (10).

Tolerance and Dependence

There is convincing evidence for the development of tolerance to THC in humans (11). Tolerance developed to a variety of THC’s effects, after oral administration, including cannabinoid-induced decreases in cardiovascular and autonomic functions, increases in intraocular pressure, sleep disturbances, and mood changes (11). High doses of Δ^9-THC were required for a sustained period of time to achieve behavioral tolerance. If the doses of Δ^9-THC were sufficiently small and infrequent, little behavioral tolerance seemed to develop.

Although it is well known that cessation of chronic marijuana exposure does not result in severe withdrawal symptoms, numerous case reports attest to the development of dependence (12). Several early reports came from countries where potent marijuana was used for long periods of time. On deprivation of marijuana, users experienced auditory and visual hallucinations and irritability. The development of tolerance and dependence has been studied under rigorous and controlled conditions (12 ,13). In one study, high doses of marijuana extract or Δ^9-THC were administered for up to 21 days, and the most prominent subjective symptoms were increased irritability and restlessness. Other symptoms included insomnia, anorexia, increased sweating, and mild nausea, although they were variable. Objective symptoms were increased body temperature, weight loss, and hand tremor. Readministration of a marijuana cigarette or oral Δ^9-THC alleviated the objective and subjective effects, a finding suggesting the establishment of a withdrawal symptom. Similar findings were reported by Georgotas and

Zeidenberg in abstinent subjects who had smoked high quantities of marijuana on a long-term basis (13). One study (14) found that lower doses of THC (80 and 120 mg/day, orally, each for 4 days) initially produced ratings of “high,” increased food intake over baseline by 35% to 45%, and decreased verbal interaction among participants (14). Tolerance developed to the subjective effects of THC but not to its effects on food intake or social behavior. Abstinence from THC produced anxious, depressed, and irritable symptoms, decreased the quantity and quality of sleep, and decreased food intake (14). A similar study conducted with marijuana cigarettes resulted in similar effects and led to the conclusion that abstinence symptoms may play a role in maintaining daily marijuana use, even at levels of use that do not produce tolerance (15).

Epidemiologic data support marijuana dependence as reviewed by Hall et al. (7). There are numerous cases in which persons seek treatment for dependence of which marijuana is the primary cause (16). These patients typically complained of being unable to stop or to decrease their marijuana use despite experiencing sleepiness, depression, inability to concentrate, and memorization difficulties that they directly attributed to marijuana exposure. Kandel and Davies reported similar problems in daily users of marijuana (17). Several groups of investigators have used DSM-III-R and DSM-IV criteria to diagnose marijuana dependence (18 ,19 and 20). With regard to prevalence of marijuana abuse and dependence, the strongest evidence was provided by the Epidemiological Catchment Area study involving 20,000 persons in five geographic areas of the United States (21). Approximately 4.4% of the population were diagnosed for marijuana abuse or dependence, and three-fifths of these met the criteria for dependence. After an extensive review of the literature, Hall et al. concluded that the risk of developing marijuana dependence was probably similar to that of alcohol, and daily use over a period of weeks to months results in the greatest risk of dependence development (7). Kandel and Davies estimated that the risk of dependence in near-daily marijuana users was one in three (17). Hall et al. estimated that the risk of developing dependence is 10% for those who ever used marijuana, with the risk rising to 20% to 30% for those using more than five times (7). Factors that have been associated with marijuana dependence include poor academic achievement, deviant behavior, rebelliousness, maladjustment, difficult parental relations, early initiation of drug use, and family history of drug use (7). The major complaints by marijuana-dependent persons are loss of control over drug use, cognitive and motivational impairments, lowered self-esteem, depression, and spousal discord. The risk of cannabis abuse and dependence was found to increase with the frequency of smoking occasions and slightly decreased with age (22). More severe comorbidity was associated with dependence compared with abuse, a finding suggesting that cannabis may be used to self-medicate major depression.

More recently, Budney et al. reported that most marijuana users seeking treatment for marijuana dependence had experienced symptoms consistent with either moderate or severe dependence (23). These investigators also reported that marijuana-dependent persons exhibit substantial problems (24). Comparison of marijuana- and cocaine-dependent patients revealed comparable substance-use histories and a range of impairments in both groups. However, the marijuana-dependent patients showed less severe dependence. The marijuana group was more ambivalent and less confident about stopping their marijuana use than the cocaine group was about stopping their cocaine use. The authors concluded that treatment-seeking, marijuana-dependent persons exhibit substantial problems and urged development of effective treatments for this population (24).

Some predisposing factors may contribute to marijuana dependence in some persons. Crowley et al. reported that juveniles diagnosed with both substance abuse and conduct disorders have serious problems related to cannabis, and most met standard adult criteria for cannabis dependence (25). Two-thirds of these cannabis-dependent patients reported withdrawal. The data indicate that for adolescents with conduct problems, cannabis use is not benign. Genetic risk factors may also contribute. Kendler et al. examined a large female twin population for lifetime cannabis use, heavy use, abuse, and dependence as defined by DSM-IV criteria (26). These investigators concluded that in women, genetic risk factors have a moderate impact on the probability of ever using cannabis and a strong impact on the liability to heavy use, abuse, and, probably, dependence. By contrast, the family and social environment substantially influences risk of ever using cannabis but plays little role in the probability of developing heavy cannabis use or abuse (26).

One of the difficulties in establishing the presence of cannabinoid dependence was the lack of a reliable animal model. Early attempts to demonstrate spontaneous withdrawal after cessation of chronic marijuana or THC treatment resulted in equivocal findings. However, the development of a specific cannabinoid antagonist, SR 141617A, made it possible to precipitate withdrawal in rats (27 ,28), mice (29), and dogs (30) treated chronically with THC. The physical withdrawal syndrome for cannabinoids and opioids in rodents shares many of the same characteristics. It is also clear that, in humans, THC is an essential reinforcing component in marijuana (31). Contrary to most drugs abused by humans, it has been difficult to train animals to self-administer cannabinoids. Although the physical characteristics of cannabinoids probably contribute to this difficulty, the general opinion persists that cannabinoids lack rewarding effects and therefore are devoid of dependence liability. However, Martelotta et al. demonstrated that the synthetic cannabinoid agonist WIN 55,212-2 was intravenously self-administered by mice in a concentration-dependent manner according to an inverted U-shaped curve (32).

Therefore, it appears that WIN 55,212-2 elicits both rewarding and aversive effects, depending on the concentration used. It may well be that these dual properties have hindered the development of a THC model of self-administration. Nevertheless, these studies clearly demonstrate that cannabinoid self-administration is not confined to humans.

ENDOGENOUS CANNABINOID SYSTEM

Part of "106 - Marijuana "

Cannabinoid Receptors

It is now widely recognized that most of the neurobehavioral and peripheral actions of marijuana and THC result from activation of selective receptors, two of which, named CB_1 and CB_2, have been cloned and characterized (33 ,34). The development of transgenic mice lacking the genes encoding for either of these two receptors, the CB_1 and CB_2-receptor knockout mice (35 ,36 and 37), have provided conclusive evidence that the effects of THC on motor behavior, body temperature, cardiovascular function, and nociception, on the one hand, and on some immunologic responses, on the other hand, are mediated by CB_1 and CB_2 receptors, respectively.

CB_1 receptors are widely distributed throughout mammalian tissues and have been found not only in the central and peripheral nervous systems, but also in both male and female reproductive organs, immune cells, the gastrointestinal tract, the liver, and the heart (38). In the central nervous system, CB_1 receptors are most abundant in the hippocampus (i.e., the dentate gyrus and CA pyramidal cells), the basal ganglia (namely, the substantia nigra pars reticulata, globus pallidus, caudate putamen), and the cerebellum and the olfactory bulb (39), findings in agreement with the inhibitory actions of THC on memory and cognitive functions, spontaneous activity, locomotion, motor coordination, and posture. Lower density of CB_1 receptors is present in discrete nuclei of other brain regions such as the hypothalamus, brainstem, thalamus, and limbic forebrain, thus possibly accounting for THC activity on body temperature, appetite, supraspinal mechanisms of pain perception, sensory perception, and mood or reward. CB_1 receptors are associated with nerve fibers and axon terminals, but not in the neuronal soma. This pattern is consistent with the presynaptic inhibitory effects of cannabinoids on neurotransmitter release in the brain (see ref. 40 for review). CB_1-expressing cells in mouse forebrain can be divided into distinct neuronal subpopulations. Most of the cells that highly express CB_1 are GABAergic neurons belonging mainly to the cholecystokinin-positive type of interneurons (basket cells). In the hippocampus, amygdala, and entorhinal cortex area, CB_1 mRNA is present at low but significant levels in many non-GABAergic cells that can be considered as projecting principal neurons. These data are in good agreement with the observation that cannabinoids act on principal glutamatergic circuits as well as modulate local GABAergic inhibitory circuits by inhibiting glutamate and GABA release. Virtually all striatal projection neurons contain CB_1 mRNA, which is also expressed in putative GABAergic interneurons that enable functional interactions between the direct and indirect striatal output pathways (41). CB_1 mRNA is found in striatonigral neurons that contain dynorphin and substance P and striatopallidal neurons that contain enkephalin, whereas local circuit neurons in striatum that contain somatostatin or acetylcholine do not synthesize CB_1 mRNA.

The presence of CB_1 receptors in sensory (42) and autonomic peripheral fibers (43 ,44) has been reported. CB_1 receptors seem to be mostly restricted to spinal interneurons, rather than at the axonal level (45), thus possibly accounting for spinal mechanisms of pain control ascribed to psychotropic cannabinoids. However, indirect evidence also exists for the presence of CB_1 receptors in peripheral sensory afferents (46), a finding thus supporting the concept that cannabinoids may also exert analgesia at the peripheral level. The presence of CB_1 receptors in parasympathetic and sympathetic fibers, on the other hand, may be at the basis of the vascular and smooth muscle–relaxing activity of THC through inhibition of norepinephrine and acetylcholine release, respectively (43 ,44). There is no evidence for the presence of CB_2 receptors in the central nervous system, except for their expression in microglia. Clearly, given that CB_2 receptors seem to be mostly confined to cells of the immune system (34), it would not be surprising to find these proteins only in those central nervous system cells deputed to immune defense, such as the microglia and resident mast cells. This finding may explain some of the neuroprotective activities of cannabinoids *in vivo* (see later).

Transduction Mechanisms

Studies have revealed that activation of the α subunits of G_i/G_o proteins, with subsequent inhibition of adenylate cyclase through both CB_1 and CB_2 receptors (47), blockade of voltage-activated calcium (Ca^{2+}) channels of the N- and P/Q-type through CB_1 receptors (48), and activation of inwardly rectifying potassium channels through CB_1 receptors (49), may not be the sole intracellular signaling messages delivered by psychoactive cannabinoids. There is now evidence for the coupling of CB_1, but not CB_2 receptors, to G_s proteins, with consequent activation of adenylyl cyclase. It is not clear yet whether this effect may explain the biphasic nature of cannabinoid effects on behavior in several tests. Another typical consequence of CB_1-receptor activation is the modulation of nitric oxide (NO) release. In neurons, THC and synthetic and endogenous cannabinoids can either stimulate (50) or inhibit (51) NO formation. The former effect results in inhibition of dopamine release from invertebrate ganglia, whereas the inhibition of NO release in granule cerebellar cells seems to result from inhibition of voltage-activated Ca^{2+} channels. In any case, modulation of NO levels may result in changes in cyclic guanosine

monophosphate intracellular concentrations. Finally, protein phosphorylation catalyzed by mitogen-activated protein kinase is coupled to both CB_1- and CB_2-receptor stimulation (52). This intracellular effect, together with inhibition of the cyclic adenosine monophosphate (cAMP)–dependent protein kinase A, is at the basis of cannabinoid action on the expression of several genes such as *krox-24* in HL60 cells (52) or the prolactin receptor and the high-affinity receptors *trk* for the nerve growth factor in human breast cancer cells (53). Mitogen-activated protein kinase activation by cannabinoids may occur independently from inhibition of protein kinase A (52), or it may result in part from inhibition of cAMP formation (53). Likewise, CB_1-induced activation of focal adhesion kinase in hippocampal slices, an effect suggested to lead to modulation by cannabinoids of synaptic plasticity and learning, results from inhibition of adenylate cyclase and protein kinase A.

Endogenous Ligands (Endocannabinoids)

Since the mid-1990s, several fatty acid derivatives have been isolated from mammalian tissues and have been shown to mimic the pharmacologic properties of THC. Not all these substances, however, can displace high-affinity cannabinoid ligands from selective binding sites in membrane preparations containing the CB_1 or the CB_2 receptor. Anandamide (arachidonoylethanolamide, AEA), the amide of arachidonic acid with ethanolamine, was the first of such compounds to be isolated (54). The other prominent endogenous ligand is a glycerol ester, 2-arachidonoyl glycerol (2-AG) (55). These compounds share the ability to bind to and to activate CB_1 and (particularly in the case of 2-AG) CB_2 receptors. Therefore, they induce a series of pharmacologic effects *in vitro* and *in vivo* that are, to some extent, similar to those exerted by THC. Other fatty acid derivatives, such as palmitoylethanolamide and oleamide, do not have high affinity for either CB_1 or CB_2 receptors, and yet they exhibit pharmacologic actions that in some cases resemble those of THC (56). The molecular mode of action of these latter compounds is still a subject for investigation.

Several structure-activity relationship studies have been carried out on AEA and have revealed that this compound does share with THC some of the structural prerequisites necessary for interaction with the CB_1 receptor. This relationship can be best appreciated with a successful conformational model (57), in which AEA may assume a low-energy conformation resulting in the superimposition of its *n*-pentyl chain, carbonyl amide group and ethanolamine hydroxyl group, respectively, with the *n*-pentyl chain, the phenolic hydroxyl group and the C-9 hydroxyl group in 9-*nor*-9ß-OH-hexahydrocannabinol, a potent THC analogue. Structure-activity relationship studies for the interaction with CB_2 receptors have not been performed yet, the sole exception being the article by Sugiura et al. (58) on the structural moieties necessary to 2-AG analogues to induce Ca^{2+} transients in HL60 cells through these receptors. Interestingly, in this study, AEA was shown to be a very weak and partial agonist at CB_2 receptors. Whatever its role as an endogenous agonist at CB_2 receptors, AEA, and much more so its metabolically stable analogues (*R*)-methanandamide and 2′-fluoro-2-methyl-arachidonoyl-ethanolamide, act as relatively potent (K_i between 12 and 100 nM) and selective CB_1-receptor agonists, and thus can be considered useful pharmacologic tools for studies on the bioactivity of endocannabinoids.

AEA shows, in some cases, effects qualitatively and quantitatively different from those of classic cannabinoids, possibly in part because of the rapid metabolism of this compound both *in vitro* and *in vivo* (59), and because AEA is a partial agonist in some functional assays of CB_1 activity (60). In the brain, AEA was shown to exert inhibitory actions on learning and memory (61) and to modulate the extrapyramidal control of motor behavior (62). These effects probably result from the capability of AEA to induce, by activation of CB_1 receptors, modulation of neurotransmitter (e.g., glutamate, GABA, dopamine) release, action or reuptake through intracellular signaling events similar to those described earlier for THC (40). This neuromodulatory action may also underlie AEA regulation of hormone release at the level of the hypothalamus-pituitary-adrenal axis, as well as the antinociceptive effects of the compound through both spinal and supraspinal mechanisms (63).

Endocannabinoid levels in tissues and cells can be modulated through the regulation of either their biosynthesis or inactivation. It is commonly accepted that the AEA and 2-AG are not stored as such in cells, but rather are synthesized and are directly released by cells “on demand,” after Ca^{2+} influx into the cell (such as that occurring in neurons on depolarization or in mast cells after IgE-mediated activation) and the hydrolysis of phospholipid precursors (40). For, example, AEA is produced in neurons and leukocytes together with other *N*-acyl-ethanolamines (NAEs) from the hydrolysis of the corresponding *N*-acyl-phosphatidyl-ethanolamines (NAPEs) (64). This reaction is catalyzed by a Ca^{2+}-independent phospholipase D, whereas a Ca^{2+}-dependent *trans*-acylase catalyzes the formation of NAPEs form phosphatidylethanolamine and the fatty acids on the *sn*-1 position of phosphoglycerides. Several mechanisms for the inactivation of endocannabinoids have been identified in neuronal and blood cells. A membrane-bound intracellular hydrolase catalyzes AEA hydrolysis after its diffusion into neuronal cells and leukocytes (64). A mechanism for the facilitated diffusion of AEA into cells according to its concentration gradient across the plasma membrane has been partially characterized as a saturable, temperature-sensitive, selective and sodium-independent “carrier” (64). This “carrier,” probably a protein, may be used for both the reuptake by and the release from cells of AEA. The major enzyme responsible for AEA hydrolysis is the fatty acid amide hydrolase

(FAAH), cloned so far in four different mammalian species (65).

Because the biosynthetic precursors for AEA and 2-AG, by being products of membrane phospholipid remodeling, are likely to occur in most animal tissues, the two endocannabinoids are probably to be found, at least in minute amounts, as ubiquitous metabolites. However, for these compounds to work as endogenous agonists of CB_1 and CB_2 receptors, their tissue concentrations need to be increased up to at least 50 to 100 nM after cell stimulation (e.g., neuronal depolarization, immune challenge) and subsequent activation of the proteins involved in their biosynthesis and release. Furthermore, the inactivation of endocannabinoids may be subject to regulation. In agreement with possible regulation of endocannabinoid levels under physiologic and pathologic conditions, the amounts of AEA or 2-AG have been found to vary during brain development, to be higher in some of the brain regions with the highest density of CB_1 receptors, such as the basal ganglia and the hippocampus (66), to decrease and increase in the striatum and limbic forebrain, respectively, of rats after chronic treatment with THC (67), to be inversely correlated with spontaneous activity in the globus pallidus of reserpine-treated rats (68), to vary during pregnancy in mouse uterus, levels of these agents being maximal when the uterus is least receptive to embryo implantation (69), and to be enhanced during septic or hemorrhagic shock in rat macrophages and platelets (70 ,71). Indeed, several possible regulative mechanisms have been reported for both the biosynthesis and inactivation of AEA and 2-AG in isolated, intact cells.

PHYSIOLOGIC ROLE OF ENDOGENOUS SYSTEM

Part of "106 - Marijuana "

The finding of variations in AEA and 2-AG levels during physiologic or pathologic conditions, together with observations of their pharmacological activity *in vivo* and *in vitro*, provide useful information on the possible biological role played by these compounds. Additionally, *in vivo* pharmacologic studies carried out by administering selective cannabinoid receptor antagonists may reveal a possible endocannabinoid-induced “tone” of CB_1- and CB_2-receptor activation during certain conditions, although the capability of the antagonists so far developed to behave as inverse agonists as well (72) should always be taken into account.

Pain

Extensive studies (see refs. 63 and 73 for review) have been carried out demonstrating the involvement of endocannabinoids in the control of nociception and, in particular, chronic and inflammatory pain. Electrical stimulation of the periaqueductal gray was shown to induce CB_1-mediated analgesia while leading to the release of AEA in microdialysates from this region of the brainstem (74). Moreover, the injection of formalin into the paw induced a nociceptive response concomitantly to the release of AEA from the periaqueductal gray and thereby established an correlation between supraspinal nociception and endocannabinoid release. In fact, an earlier investigation had suggested that an endocannabinoid tone may down-modulate pain perception through CB_1 receptors in another region of the brainstem, the rostral ventromedial medulla, through the same circuit previously shown to contribute to the pain-suppressing effects of morphine (75). However, other studies have shown that blockade of the action or expression of spinal CB_1 receptors by SR141716A or a CB_1-receptor antisense oligonucleotide, respectively, leads to hyperalgesia (76), a finding thus suggesting the existence of an endocannabinoid tone down-modulating nociceptive response also at the spinal level. The same group gained evidence for the presence of CB_1 receptors in peripheral sensory afferents in the skin, and for their involvement in the control of inflammatory pain (77). It may well be that an endocannabinoid and CB_1/CB_2 receptors mediate tone controlling pain at the peripheral level, because local administration of the antagonist for each receptor subtype leads to hyperalgesia and exogenous AEA blocks the painful response of mice to formalin injection. Several studies, taken together with that by Meng et al. (75) and Walker et al. (74), suggest that if endocannabinoids tonically modulate inflammatory pain perception, they may do so at sites different from those of inflammation.

There has been considerable interest in determining what role, if any, opioids play in cannabinoid-induced antinociception. In one study, marijuana produced significant dose-dependent antinociception (increased finger withdrawal latency) and behavioral effects. Naltrexone did not significantly influence marijuana dose-effect curves, a finding suggesting no role of endogenous opiates in marijuana-induced antinociception under these conditions (78). Conversely, it has been shown that cannabinoids stimulate release of endogenous opioids that contribute to cannabinoid antinociception (79). Meng et al. showed that the rostral ventromedial medulla that contributes to the pain-suppressing effects of morphine is also required for the analgesic effects of cannabinoids (75). Although cannabinoids produce analgesia by modulating rostral ventromedial medulla neuronal activity in a manner similar to morphine, their actions are not identical. They also show that endogenous cannabinoids tonically regulate pain thresholds in part through the modulation of rostral ventromedial medulla neuronal activity. These authors concluded that analgesia produced by cannabinoids and opioids involves similar brainstem circuitry and that cannabinoids are indeed centrally acting analgesics with a new mechanism of action.

Memory

The effects of cannabinoids on memory in rats are also blocked by a specific cannabinoid antagonist, SR 141716A, a finding providing strong evidence that these effects are mediated through cannabinoid receptors in the brain (9). Mallet and Beninger used a two-component instrumental discrimination task, consisting of a conditional discrimination, and a non–match-to-position to assess recent or working memory (61). These investigators found that both THC and anandamide impaired performance, an effect that could be attenuated with the administration of the CB_1-receptor antagonist SR141716A. These results suggest that anandamide-induced memory disruption is mediated by CB receptors. Studies have shown that THC produces memory deficits similar to those produced by neurochemical lesions of the hippocampus. A possible role for cannabinoid receptors and endogenous cannabinoids may be to regulate the storage and retrieval of information (80).

As discussed previously, cannabinoid receptors are highly expressed in the hippocampus, a brain region that is believed to play an important role in certain forms of learning and memory. The notion that endocannabinoids are involved in the control of learning and memory processes at the level of the hippocampus is supported by several different types of observations. First, both AEA and 2-AG inhibit hippocampal long-term potentiation (81) and modulate the release of glutamate or acetylcholine from hippocampal slices (40). Second, AEA modulates both short-term and long-term memory (61). Third, SR141716A enhances long-term potentiation, a finding thus suggesting a CB_1-receptor tone in the control of this process. Fourth, CB_1-receptor knockout mice exhibit enhancement of memory as well as of long-term potentiation (82). Finally, CB_1 receptors, AEA, and FAAH are found in high levels in the hippocampus of humans, rats, and mice (66). These findings suggest that constitutive activation of CB_1 receptors in this brain region leads to inhibition of learning and memory processes.

There is evidence that memory deficits induced by cannabinoids may be mediated through cholinergic and dopaminergic systems (83). The systemic administration of THC reduced hippocampal extracellular acetylcholine concentrations while impairing working memory in rats. Both effects were blocked by the CB_1 cannabinoid and D2 dopamine receptor antagonists and potentiated by the D2 dopamine receptor agonist quinpirole. The inhibition of hippocampal extracellular acetylcholine concentration and working memory produced by the combination of (-)- quinpirole and THC was suppressed by either CB_1 cannabinoid and D2 dopamine receptor antagonists. These researchers concluded that cannabinoid impairment of working memory and inhibition of hippocampal extracellular acetylcholine concentration are mediated by the concomitant activation of D2 and CB_1 receptors.

Movement

Central cannabinoid receptors are densely located in the output nuclei of the basal ganglia (globus pallidus, substantia nigra pars reticulata), a finding suggesting their involvement in the regulation of motor activity. However, different approaches have not managed to give a precise role of endocannabinoids in the inhibition of spontaneous activity and induction of catalepsy in rodents, typical of all CB_1-receptor agonists (84). In fact, although CB_1-receptor knockout mice seem to have different baseline locomotor activity than wild-type mice, it is not clear whether deletion of the CB_1-receptor gene in these transgenic animals leads to hypermotility (35) or hypomotility (36). However, an endogenous cannabinoid tone negatively controlling spontaneous activity and motor behavior is supported by the finding that AEA, but not 2-AG, is released in microdialysates from the dorsal striatum of freely moving rats (85), and the levels of AEA are also very high in the substantia nigra and external layer of the globus pallidus (68).

In this latter brain region, endocannabinoid levels are *inversely* correlated with spontaneous activity in the reserpine-treated rat, an animal model of Parkinson disease, in which dopamine and other catecholamines in the striatum are depleted (68). AEA levels in the striatum of normal rats are increased by selective stimulation of D2 dopamine receptors by quinpirole, whereas the CB_1 antagonist SR141716A strongly enhances quinpirole-induced movement in both normal and reserpine-treated rats (68 ,85). These data suggest that the endocannabinoid system may act as a brake on dopaminergic stimulation of movement in the basal ganglia, and an exaggerated endocannabinoid tone in this region may produce (or at least contribute to) parkinsonian symptoms in rats (68). Further evidence for such suggestions has been provided by the finding that tolerance to the motor inhibitory actions of THC in rats chronically treated with the cannabinoid is accompanied not only by down-regulation of cannabinoid receptors in the striatum, but also by a significant decrease of endocannabinoid levels in this brain area (67).

Craving, Appetite Stimulation, and Reward

The finding of CB_1 receptors in the arcuate nucleus and the medial preoptic area of the hypothalamus, the presence of endocannabinoids and their biosynthetic precursors in the hypothalamus and pituitary, and the effect of endocannabinoids on body temperature, food intake, and pituitary hormone release suggest a role for endocannabinoids in the control of hypothalamic functions, and in particular on appetite and hormone release. Indeed, the CB_1-receptor–selective antagonist, SR141716A, inhibits palatable food intake in rodents (86). It has not been established whether

this effect results from the inverse agonist properties of SR141716A (72) or from its blockade of a food-intake stimulatory tone by endocannabinoids. Another brain region possibly involved in the control of appetite and craving is the limbic forebrain and, more particularly, the nucleus accumbens. In this brain area, cannabinoids, by enhancing the release of dopamine from dopaminergic terminals originating in the ventral tegmental area, may exert reinforcing actions on the effects of other drugs of abuse or, under more physiologic conditions, may participate in the regulation of feelings of craving and reward (87). Furthermore, it was found that chronic treatment of rats with THC causes an almost fourfold increase of AEA levels (and no down-regulation of cannabinoid receptors) in the limbic forebrain (67). It is possible that dopamine released in the nucleus accumbens on chronic treatment with THC triggers AEA formation, as previously shown for the dorsal striatum (85). Conversely, dopamine may be released in this region after the activation of CB_1 receptors by AEA. Indeed, studies carried out with CB_1-receptor knockout mice showed reduced opioid dependence (35), as well as lack of morphine-induced dopamine release in the nucleus accumbens of these transgenic animals (88). Thus, contrary to the basal ganglia, endocannabinoids released in the nucleus accumbens may act to enhance the action or release of dopamine, thereby participating in reward, craving, and pleasure or in the reinforcement of drug of abuse effects. There are indications that withdrawal from chronic cannabinoid administration is associated with reduced dopaminergic transmission in the limbic system, similar to that observed with other addictive drugs, a finding consistent with a role in drug craving and relapse into drug addiction or in the reinforcing effects of drugs of abuse.

Neuroprotection

The possibility that endocannabinoids may play a role in diminishing cellular or neuronal damage is of particular relevance to neurodegenerative disorders. The suggestion that endocannabinoids may have a neuroprotective function during cell injury stems from the finding that a similar role was proposed also for other ethanolamide of fatty acids (89), as well as for both psychoactive and nonpsychoactive cannabinoids. This hypothesis is supported by the finding that stimuli leading to high intracellular Ca^{2+} concentrations (e.g., glutamate-induced excitotoxicity) and noxious agents such as ethanol and sodium azide lead to increased synthesis of AEA and related compounds in neuronal cells. Cannabinoid receptors do not appear to be involved in this elevation of Ca^{2+}. In fact AEA, via CB_1 receptors, produces the opposite effect. It inhibits Ca^{2+} influx into neurons through voltage-gated Ca^{2+} channels and counteracts membrane permeability to Ca^{2+} through *N*-methyl-D-aspartate receptor–coupled channels. Therefore, endocannabinoids should be able to inhibit glutamate-induced excitotoxicity (or other pathologic conditions arising from high intracellular Ca^{2+} concentrations) by acting at CB_1 receptors, particularly because they do not share the antioxidant effects of some synthetic cannabinoids. In conclusion, further studies are necessary to assess whether and through what mechanisms AEA and 2-AG prevent neuronal damage.

MEDICAL MARIJUANA

Part of "106 - Marijuana "

Scientific Justification

The nonmedical use of marijuana has a very long history, primarily for its mind-altering effects and the sense of well-being that it can provide. Therefore, the potential use of marijuana for diseases of the brain is a logical extension of the popularity of the use of the material in producing mood-altering effects. The initial therapeutic uses proposed for marijuana included the treatment of mental disorders and pain. As more information about the pharmacologic effects of the plant material emerged, other potential therapeutic uses became apparent. Since the 1970s, investigators have proposed many different therapeutic uses for marijuana including, but not limited to, nausea and vomiting induced by cancer chemotherapeutic agents, the wasting syndrome accompanying AIDS, mental illness, convulsions, glaucoma, cognition disorders, muscle spasticity, and neuropathic pain (90).

The consequences of the social use of marijuana, both real and as often exaggerated by opponents and proponents, have caused increased anxiety on both sides of the controversy over the medical use of marijuana. Strong proponents of the use of smoked marijuana for the treatment of various syndromes and diseases argue that smoking marijuana has produced beneficial effects in at least one disease state that could not be achieved by the oral administration THC or by any other treatment modality. Opponents are concerned with the deleterious effects of the smoked marijuana, especially the prolonged use of this plant material. The issue is further complicated by the fact that many strong proponents of the use of marijuana in medicine also advocate for its legal recreational use. Conversely, those who are opposed to its use, especially by adolescents and young adults who may be especially vulnerable to problems of abuse, effects on energy, memory, and acquisition of interpersonal skills, have not always considered the possible benefits with the same degree of objectivity as would be afforded other potential therapeutic agents. One of the major problems contributing to this dilemma is the lack of well-controlled studies attesting to the efficacy and the safety of marijuana in humans. Such studies require a reasonable hypothesis to be tested and an appropriate investigation under conditions that completely eliminate the possibility of subjectivity in the measurements. This controversy is not likely to be resolved until such studies are forthcoming.

The need for an approved medical use of marijuana itself

is questioned by some investigators because of the availability of marijuana's active constituent. THC was approved for the treatment of the nausea and vomiting associated with cancer chemotherapy in the 1980s and for the treatment of the wasting syndrome in AIDS patients in the early 1990s. It has been moved from Schedule II to Schedule III. Marijuana proponents counter that the therapeutic benefits derived from smoked marijuana are the result of many chemicals in the plant, not solely THC. There are, however, toxic pulmonary consequences to marijuana smoking. Further, the smoking route of administration used for marijuana has advantages over the oral route used for the administration of THC. The onset of action is faster while at the same time allowing the smoker to titrate blood levels better. The heightened interest in medical marijuana has not yet been translated into a satisfactory resolution of the differences. There is no doubt of the severity of the conditions and diseases for which marijuana or THC has been proposed. The availability of alternative delivery systems (e.g., aerosol) and alternative synthetics that have the desired therapeutic effect with minimal intoxicating effects are needed to resolve the controversy. Even then, however, it is likely that some will argue that it is the intoxicating effect combined with the other effects that makes marijuana particularly useful.

Medication with Plant Material

Active Constituents

Hundreds of compounds have been isolated and identified from the marijuana plant (91). Most have been shown to have minimal pharmacologic activity, and most exist in the plant in very small quantities. The major active constituent in marijuana has been shown to be THC. The pharmacologic profile of THC is essentially the same as that of smoked marijuana, and the evidence is now overwhelming that the predominant effects of marijuana on the brain result from this compound. Other cannabinoids such as Δ^8-THC, cannabidiol, and cannabinol have been studied and have been shown to have interesting pharmacologic profiles. The Δ^8-THC isomer produces many of the same effects as the Δ^9- isomer (THC), but it is generally less potent, and the quantity of Δ^8-THC in the plant material usually is less than that of THC. Cannabidiol and cannabinol are of interest because they exist in reasonable quantities in the plant and produce interesting pharmacologic effects in some systems but also are considerably devoid of activity on the central nervous system, especially in relation to mental health, memory, and cognition. The lack of effects on the central nervous system is an advantage in the potential use of one of these agents or an analogue to treat a disease or a condition with a locus of action outside the brain. The search for other cannabinoids that could have therapeutic potential has shifted almost totally to the synthetic chemistry process.

THC Content

The identification of THC as the active agent in marijuana stimulated a concentrated effort to quantitate the amount of this material in various samples of the cannabis plant. The initially reported concentrations of THC in confiscated marijuana were approximately 2% but have increased to more than 4% during the past few years (92). It was found that by altering the soil conditions and the environment, the concentration could be increased several fold. As one would expect, the pharmacologic effects of smoking marijuana are directly related to the concentration of THC. Advances in biogenetic engineering as applied to agriculture suggest that manipulations could be made to increase the concentration of THC even higher. Concentrations of more than 20% have been reported in some marijuana grown under artificial conditions in the Netherlands. How available increased-potency marijuana is in the United states remains unclear.

Consistency

As described earlier, the concentration of the active constituent in marijuana can vary over a large range. This variation clearly complicates the delivery of a consistent dose of medication. It would not be practical to quantitate the concentration of the active ingredient in each cigarette before its consumption. Most of the proposed indications for the medical use of marijuana require chronic administration, which magnifies the problem of inconsistent dosing. Administration of any drug through smoking presents an additional problem when a standard procedure does not exist for preparing cigarettes with a constant quantity of plant material in each cigarette. Further, the variability from individual to individual in the size and the rate of puffing produces another variable for the consumption of drugs by this route of administration. Even in the same patient, the volume of smoke inhaled can often differ from time to time.

Unwanted Side Products

The administration of a drug by smoking plant material causes other problems because many substances are being taken in along with the active ingredient. Each of these substances has its pharmacologic and toxicologic effects. Further, these other substances may either potentiate or interfere with the effects of the active ingredient. It is abundantly clear from the vast literature on the smoking of other products, predominantly tobacco, that numerous compounds are produced in the burning process. These pyrolysis products also have their own pharmacologic and toxicologic profile, and as with the other ingredients in the plant material, these pyrolysis products have the potential to alter the effects of the active ingredient. The major problem is the inability to control the exposure of the patient carefully to

a consistent and correct dose of the active ingredient and to establish a treatment regimen that is devoid of interference from other chemicals in the preparation or made during the administration process.

Optimal Delivery

One of the major concerns with the potential use of marijuana, and, for that matter with any of the cannabinoids, is the observation that these agents produce a multiplicity of effects, and they all seem to occur at similar doses. One possible way to overcome this problem is to develop a delivery mechanism that limits the distribution of the drug to the desired site of action. This is particularly difficult when the site of action is in the brain, as it is with the cannabinoids and their potential usefulness in treating symptoms of mental illness. One potential therapeutic use of the cannabinoids is in the treatment of glaucoma. Local administration directly into the eye is the preferred mechanism of drug delivery. One of the problems of using cannabinoids in this fashion is that they are insoluble and must be administered in a vehicle, which may have deleterious effects when it is placed in the eye. An optimal delivery for this indication would be to have a water-soluble cannabinoid (93) with good efficacy in lowering intraocular pressure that can be applied directly to the eye and not be irritating. The direct application of such a drug would provide the intended therapeutic effect and would not produce the undesirable side effects that would be observed if the drug were absorbed into the general circulation.

The reemergence on the debate of the use of marijuana for medicinal purposes has also been the impetus for developing an acceptable delivery form of aerosolized cannabinoids. A nebulizer was used to generate an aerosol with sufficiently small particle size such that exposure to rodents produced pharmacologic effects, These results demonstrate that the development of an aerosolized form of cannabinoids for human medicinal use is feasible (94).

Future Developments

Selective Receptor Agonists and Selective Pharmacologic Profiles

Because the cannabinoids have multiple effects on so many different body functions at approximately the same dose, considerable effort continues to be directed toward identifying the portion of the cannabinoid molecule that is most responsible for each unique pharmacologic effect. The identification of multiple cannabinoid receptors and the observation that certain cannabinoids have selectivity for one type of receptor over the other are encouraging (95 ,96). Further research into different receptor types and the identification of specific endogenous ligands for each of these receptor types will provide guidance for the medicinal chemist synthesizing new and, one hopes, more selective cannabinoids. The hypothesis that each receptor subtype has its own specific ligand, as in the case of AEA preference for the CB_1 receptor, is a reasonable approach. These investigations will be guided by the continued progress in the efforts of researchers to identify and understand the cellular and molecular effects of the cannabinoids. Mechanisms may be found that will help to provide selectivity resulting from the effects of the cannabinoids on an intracellular site of action. The more detail we know about the genetic influences and the structural makeup of the receptors and other cellular elements, the better we will be able to design cannabinoids with selective activity.

Another approach to identifying cannabinoids with more selectivity of action is to investigate the interactions of exogenously administered substances with AEA and other endogenous cannabinoids. If one were to hypothesize that an altered tonic activity of the endogenous cannabinoid system is at the basis of some neurologic disorders, there are again two approaches that could be taken in the search for more selective agents. One would be to alter either the synthesis or degradation of the endocannabinoids, and the other is to modulate their actions. An example of the latter approach would be to compete with or block the receptor or to interfere with the signaling events through which the endogenous compound produces its effect. The advantage of these approaches would be that drugs interfering with endocannabinoid metabolism or action would exhibit higher effects in those tissues where the levels and activity of endocannabinoids are pathologically altered.

Selective Transport Blockers and Enzyme Inhibitors

Several blockers of AEA-facilitated transport have been developed so far, but only two were shown to enhance the actions of AEA *in vivo* or *in vitro*. The most widely used one, AM404, potentiates AEA analgesic effects in the hot plate test and inhibition of adenylate cyclase (97), although it also activates vanilloid receptors. Inhibitors of FAAH have also been developed (98), a potent and selective one of which, palmitylsulfonylfluoride (AM347), acts as a covalent inhibitor (99). The most potent irreversible FAAH inhibitor developed so far is methylarachidonoylfluorophosphonate (100), which unfortunately also binds to CB_1 receptors. Future studies will have to establish whether defective biosynthesis or exaggerated metabolism of endocannabinoids contributes to pathologic conditions, and, therefore, whether therapeutically useful drugs can be developed using these or more selective inhibitors of AEA degradation. These compounds are likely to act most efficaciously only at the site where AEA and 2-AG levels are pathologically altered, and, also for this reason, they will be devoid of undesired psychotropic side effects.

Receptor Antagonists

Ideally, the development of an antagonist to any new drug would be an asset to provide protection against an accidental overdose or to reverse the effects in a hypersensitive individual. Clearly, an antagonist is a very useful tool for studies directed toward elucidation of the pharmacologic profile of the agonist. The development of an antagonist as a new therapeutic agent requires the same demonstration of efficacy and safety needed for a cannabinoid agonist. As discussed earlier, a cannabinoid agonist may have potential therapeutic uses for disorders that have characteristics opposite to the pharmacologic effects of THC. Conversely, cannabinoid antagonists may have potential in treating memory impairment, obesity, and perhaps certain psychiatric disorders associated with marijuana use. The basis of the therapeutic usefulness of a cannabinoid antagonists rests on the premise that the endogenous cannabinoid system is under tonic control, and the antagonist can either block the actions of the endogenous ligand or exert inverse agonist effects by interactions with the cannabinoid receptor.

REFERENCES

1. Johnston LD, O'Malley PM, Bachman JG. *Monitoring the future: national results on adolescent drug use*. Washington, DC: U.S. Department of Health and Human Services, 2000.

2. Chait LD, Pierri J. Effects of smoked marijuana on human performance: a critical review. In: Murphy L, Bartke A, eds. *Marijuana/cannabinoids: neurobiology and neurophysiology*. Boca Raton, FL: CRC, 1992:387–423.

3. Liguori A, Gatto CP, Robinson JH. Effects of marijuana on equilibrium, psychomotor performance, and simulated driving. *Behav Pharmacol* 1998;9:599–609.

4. Fant RV, Heishman SJ, Bunker EB, et al. Acute and residual effects of marijuana in humans. *Pharmacol Biochem Behav* 1998;60:777–784.

5. Kurzthaler I, Hummer M, Miller C, et al. Effect of cannabis use on cognitive functions and driving ability. *J Clin Psychiatry* 1999;60:395–399.

6. Schwartz RH. Chronic marihuana smoking and short-term memory impairment. In: Nahas GG, Latour C, eds. *Cannabis: physiopathology, epidemiology, detection*. Boca Raton, FL: CRC, 1993:61–71.

7. Hall W, Solowij N, Lemon J. *The health and psychological consequences of cannabis use*. National Drug Strategy monograph series no. 25. Canberra: Australian Government Publication Service, 1994.

8. Pope H, Yurgelun-Todd D. The residual cognitive effects of heavy marijuana use in college students. *JAMA* 1996;275:521–527.

9. Lichtman AH, Martin BR. Δ^9-Tetrahydrocannabinol impairs spatial memory through a cannabinoid receptor mechanism. *Psychopharmacology (Berl)* 1996;126:125–131.

10. Lichtman AH, Dimen KR, Martin BR. Systemic or intrahippocampal cannabinoid administration impairs spatial memory in rats. *Psychopharmacology (Berl)* 1995;119:282–290.

11. Jones RT, Benowitz N, Bachman J. Clinical studies of cannabis tolerance and dependence. *Ann NY Acad Sci* 1976;282:221–239.

12. Jones RT. Cannabis tolerance and dependence. In: Fehr KO, Kalant H, eds. *Cannabis and health hazards*. Toronto: Addiction Research Foundation, 1983:617–689.

13. Georgotas A, Zeidenberg P. Observations on the effects of four weeks of heavy marihuana smoking on group interaction and individual behavior. *Compr Psychiatry* 1979;20:427–432.

14. Haney M, Ward AS, Comer SD, et al. Abstinence symptoms following oral THC administration to humans. *Psychopharmacology (Berl)* 1999;141:385–394.

15. Haney M, Ward AS, Comer SD, et al. Abstinence symptoms following smoked marijuana in humans. *Psychopharmacology (Berl)* 1999;141:395–404.

16. Jones RT. Marijuana: health and treatment issues. *Psychiatr Clin North Am* 1984;7:703–712.

17. Kandel DB, Davies M. Progression to regular marijuana involvement: phenomenology and risk factors for near-daily use. In: Glantz M, Pickens R, eds. *Vulnerability to drug abuse*. Washington, DC: American Psychological Association, 1992.

18. Kosten TR, Roundsville BJ, Babor TF, et al. Substance-use disorders in DSM-III-R. *Br J Psychiatry* 1987;151:834–843.

19. Newcombe MD. Understanding the mulidimensional nature of drug use and abuse: the role of consumption, risk factors and protective factors. In: Glantz M, Pickens R, eds. *Vulnerability to drug abuse*. Washington, DC: American Psychological Association, 1992.

20. Rounsaville BJ, Bryant K, Babor T, et al. Cross-system agreement for substance use disorders: DSM-III-R, DSM-IV and ICD-10. *Addiction* 1993;88:337–348.

21. Robins LN, Regier DA. *Psychiatric disorders in America*. New York: Free Press, MacMillan, 1991.

22. Grant BF, Pickering R. The relationship between cannabis use and DSM-IV cannabis abuse and dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. *J Subst Abuse* 1998;10:255–264.

23. Budney AJ, Novy PL, Hughes JR. Marijuana withdrawal among adults seeking treatment for marijuana dependence. *Addiction* 1999;94:1311–1322.

24. Budney AJ, Radonovich KJ, Higgins ST, et al. Adults seeking treatment for marijuana dependence: a comparison with cocaine-dependent treatment seekers. *Exp Clin Psychopharmacol* 1998;6:419–426.

25. Crowley TJ, Macdonald MJ, Whitmore EA, et al. Cannabis dependence, withdrawal, and reinforcing effects among adolescents with conduct symptoms and substance use disorders. *Drug Alcohol Depend* 1998;50:27–37.

26. Kendler KS, Prescott CA. Cannabis use, abuse, and dependence in a population-based sample of female twins. *Am J Psychiatry* 1998;155:1016–1022.

27. Tsou K, Patrick S, Walker JM. Physical withdrawal in rats tolerant to Δ^9-tetrahydrocannabinol precipated by a cannabinoid receptor antagonist. *Eur J Pharmacol* 1995;280:R13–R15.

28. Aceto M, Scates S, Lowe J, et al. Cannabinoid precipitated withdrawal by the selective cannabinoid receptor antagonist, SR 141716A. *Eur J Pharmacol* 1995;282:R1–R2.

29. Cook SA, Lowe JA, Martin BR. CB1 receptor antagonist precipitates withdrawal in mice exposed to Delta9-tetrahydrocannabinol. *J Pharmacol Exp Ther* 1998;285:1150–1156.

30. Lichtman AH, Wiley JL, LaVecchia KL, et al. Acute and chronic cannabinoid effects: characterization of precipitated withdrawal in dogs. *Eur J Pharmacol* 1998;357:139–148.

31. Haney M, Comer SD, Ward AS, et al. Factors influencing marijuana self-administration by humans. *Behav Pharmacol* 1997;8:101–112.

32. Martellotta MC, Cossu G, Fattore L, et al. Self-administrarion of the cannabinoid receptor agonist WIN 55,212-2 In drug-naive mice. *Neuroscience* 1998;85:327–330.

33. Matsuda LA, Lolait SJ, Brownstein MJ, et al. Structure of a cannabinoid receptor and functional expression of the cloned cDNA. *Nature* 1990;346:561–564.

34. Munro S, Thomas KL, Abu-Shaar M. Molecular characterization of a peripheral receptor for cannabinoids. *Nature* 1993;365:61–64.

35. Ledent C, Valverde O, Cossu G, et al. Unresponsiveness to cannabinoids and reduced addictive effects of opiates in CB1 receptor knockout mice. *Science* 1999;283:401–404.

36. Zimmer A, Zimmer AM, Hohmann AG, et al. Increased mortality, hypoactivity, and hypoalgesia in cannabinoid CB1 receptor knockout mice. *Proc Natl Acad Sci USA* 1999;96:5780–5785.

37. Buckley NE, McCoy KL, Mezey E, et al. Immunomodulation by cannabinoids is absent in mice deficient for the cannabinoid CB(2) receptor. *Eur J Pharmacol* 2000;396:141–149.

38. Herkenham M. Localization of cannabinoid receptors in brain and periphery. *Neuropschychopharmacology* 1994;10:143S-143S.

39. Herkenham M, Lynn AB, Little MD, et al. Cannabinoid receptor localization in the brain. *Proc Natl Acad Sci USA* 1990;87:1932–1936.

40. Di Marzo V, Melck D, Bisogno T, et al. Endocannabinoids: endogenous cannabinoid receptor ligands with neuromodulatory action. *Trends Neurosci* 1998;21:521–528.

41. Hohmann AG, Herkenham M. Localization of cannabinoid CB(1) receptor mRNA in neuronal subpopulations of rat striatum: a double-label *in situ* hybridization study. *Synapse* 2000;37:71–80.

42. Hohmann AG, Herkenham M. Localization of central cannabinoid CB1 receptor messenger RNA in neuronal subpopulations of rat dorsal root ganglia: a double-label *in situ* hybridization study. *Neuroscience* 1999;90:923–931.

43. Pertwee RG. Pharmacology of cannabinoid CB_1 and CB_2 receptors. *Pharmacol Ther* 1997;74:129–180.

44. Kunos G, Jarai Z, Varga K, et al. Cardiovascular effects of endocannabinoids: the plot thickens. *Prostaglandins Other Lipid Mediat* 2000;61:71–84.

45. Farquhar-Smith WP, Egertova M, Bradbury EJ, et al. Cannabinoid CB(1) receptor expression in rat spinal cord *Mol Cell Neurosci* 2000;15:510–521.

46. Hohmann AG, Herkenham M. Cannabinoid receptors undergo axonal flow in sensory nerves. *Neuroscience* 1999;92:1171–1175.

47. Howlett A. Pharmacology of cannabinoid receptors. *Annu Rev Pharmacol Toxicol* 1995;35:607–634.

48. Mackie K, Devane W, Hille B. Anandamide, an endogenous cannabinoid, inhibits calcium currents as a partial agonist in N18 neuroblastoma cells. *Mol Pharmacol* 1993;44:498–503.

49. Mackie K, Lai Y, Westenbroek R, et al. Cannabinoids activate an inwardly rectifying potassium conductance and inhibit Q-type calcium currents in AtT20 cells transfected with rat brain cannabinoid receptor. *J Neurosci* 1995;15:6552–6561.

50. Stefano GB, Liu Y, Goligorsky MS. Cannabinoid receptors are coupled to nitric oxide release in inverebrate immunocytes, microglia, and humanb monocytes. *J Biol Chem* 1996;271:19238–19242.

51. Hillard CJ, Muthian S, Kearn CS. Effects of CB(1) cannabinoid receptor activation on cerebellar granule cell nitric oxide synthase activity. *FEBS Lett* 1999;459:277–281.

52. Bouaboula M, Poinot-Chazel C, Marchand J, et al. Signaling pathway associated with stimulation of CB2 peripheral cannabinoid receptor involvement of both mitogen-activated protein kinase and induction of *krox*-24 expression. *Eur J Biochem* 1996;237:704–711.

53. Melck D, Rueda D, Galve-Roperh I, et al. Involvement of the cAMP/protein kinase A pathway and of mitogen-activated protein kinase in the anti-proliferative effects of anandamide in human breast cancer cells. *FEBS Lett* 1999;463:235–240.

54. Devane WA, Hanus L, Breuer A, et al. Isolation and structure of a brain constituent that binds to the cannabinoid receptor. *Science* 1992;258:1946–1949.

55. Mechoulam R, Ben-Shabat S, Hanus L, et al. Identification of an endogenous 2-monoglyceride, present in canine gut, that binds to cannabinoid receptors. *Biochem Pharmacol* 1995;50:83–90.

56. Lambert DM, Di Marzo V. The palmitoylethanolamide and oleamide enigmas: are these two fatty acid amides cannabimimetic? *Curr Med Chem* 1999;6:757–773.

57. Tong W, Collantes ER, Welsh WJ, et al. Derivation of a pharmacophore model for anandamide using constrained conformational searching and comparative molecular field analysis. *J Med Chem* 1998;41:4207–4215.

58. Sugiura T, Kondo S, Kishimoto S, et al. Evidence that 2-arachidonoylglycerol but not N-palmitoylethanolamine or anandamide is the physiological ligand for the cannabinoid CB2 receptor: comparison of the agonistic activities of various cannabinoid receptor ligands in HL-60 cells. *J Biol Chem* 2000;275:605–612.

59. Deutsch DG, Chin SA. Enzymatic synthesis and degradation of anandamide, a cannabinoid receptor agonist. *Biochem Pharmacol* 1993;46:791–796.

60. Breivogel CS, Selley DE, Childers SR. Cannabinoid receptor agonist efficacy for stimulating [^{35}S]GTPgS binding to rat cerebellar membranes correlates with agonist-induced decreases in GDP affinity. *J Biol Chem* 1998;273:16865–16873.

61. Mallet PE, Beninger RJ. The cannabinoid CB1 receptor antagonist SR141716A attenuates the memory impairment produced by delta-9-tetrahydrocannabinol or anandamide. *Psychopharmacology (Berl)* 1998;140:11–19.

62. Romero J, de Miguel R, García-Palomero E, et al. Time-course of the effects of anandamide, the putative endogenous cannabinoid receptor ligand, on extrapyramidal function. *Brain Res* 1995;694:223–232.

63. Martin BR, Lichtman AH. Cannabinoid transmission and pain perception. *Neurobiol Dis* 1998;5:447–461.

64. Di Marzo V, Fontana A, Cadas H, et al. Formation and inactivation of endogenous cannabinoid anandamide in central neurons. *Nature* 1994;372:686–691.

65. Cravatt BF, Giang DK, Mayfield SP, et al. Molecular characterization of an enzyme that degrades neuromodulatory fatty-acid amides. *Nature* 1996;384:83–87.

66. Bisogno T, Berrendero F, Ambrosino G, et al. Brain regional distribution of endocannabinoids: implications for their biosynthesis and biological function. *Biochem Biophys Res Commun* 1999;256:377–380.

67. Di Marzo V, Berrendero F, Bisogno T, et al. Enhancement of anandamide formation in the limbic forebrain and reduction of endocannabinoid contents in the striatum of delta-9- tetrahydrocannabinol-tolerant rats. *J Neurochem* 2000;74:1627–1635.

68. Di Marzo V, Hill MP, Bisogno T, et al. Enhanced levels of endogenous cannabinoids in the globus pallidus are associated with a reduction in movement in an animal model of Parkinson's disease. *FASEB J* 2000;14:1432–1438.

69. Schmid P, Paria BC, Krebsbach R, et al. Changes in anandamide levels in mouse uterus are associated with uterine receptivity for embryo implantation. *Proc Natl Acad Sci USA* 1997;94:4188–4192.

70. Wagner JA, Varga K, Ellis EF, et al. Activation of peripheral CB_1 receptors by an endogenous cannabinoid contributes to hypotenxion in haemorrhagic shock. *Nature* 1997;390:518–521.

71. Varga K, Wagner JA, Bridgen DT, et al. Platelet- and macrophage-derived endogenous cannabinoids are involved in endotoxin-induced hypotension. *FASEB J* 1998;12:1035–1044.

72. Landsman RS, Burkey TH, Consroe P, et al. SR141716A is an inverse agonist at the human cannabinoid CB1 receptor. *Eur J Pharmacol* 1997;334:R1–R2.

73. Walker JM, Hohmann AG, Martin WJ, et al. The neurobiology of cannabinoid analgesia. *Life Sci* 1999;65:665–673.

74. Walker JM, Huang SM, Strangman NM, et al. Pain modulation by release of the endogenous cannabinoid anandamide. *Proc Natl Acad Sci USA* 1999;96:12198–12203.

75. Meng ID, Manning BH, Martin WJ, et al. An analgesia circuit activated by cannabinoids. *Nature* 1998;395:381–383.

76. Richardson JD, Aanonsen L, Hargreaves KM. SR 141716A, a cannabinoid receptor antagonist, produces hyperalgesia in untreated mice. *Eur J Pharmacol* 1997;319:R3–R4.

77. Richardson JD, Kilo S, Hargreaves KM. Cannabinoids reduce hyperalgesia and inflammation via interaction with peripheral CB1 receptors. *Pain* 1998;75:111–119.

78. Greenwald MK, Stitzer ML. Antinociceptive, subjective and behavioral effects of smoked marijuana in humans. *Drug Alcohol Depend* 2000;59:261–275.

79. Houser SJ, Eads M, Embrey JP, et al. Dynorphin B and spinal analgesia: induction of antinociception by the cannabinoids CP55,940, Delta(9)-THC and anandamide. *Brain Res* 2000;857:337–342.

80. Hampson RE, Deadwyler SA. Role of cannabinoid receptors in memory storage. *Neurobiol Dis* 1998;5:474–482.

81. Terranova J-P, Michaud J-C, Le Fur G, et al. Inhibition of long-term potentiation in rat hippocampal slices by anandamide and WIN55212-2: reversal by SR141716A, a selective antagonist of CB1 cannabinoid receptors. *Nauyn Schmiedebergs Arch Pharmacol* 1995;352:576–579.

82. Bohme GA, Laville M, Ledent C, et al. Enhanced long-term potentiation in mice lacking cannabinoid CB1 receptors. *Neuroscience* 2000;95:5–7.

83. Nava F, Carta G, Battasi AM, et al. D(2) dopamine receptors enable delta(9)-tetrahydrocannabinol induced memory impairment and reduction of hippocampal extracellular acetylcholine concentration. *Br J Pharmacol* 2000;130:1201–1210.

84. Fride E, Mechoulam R. Pharmacological activity of the cannabinoid receptor agonist, anandamide, a brain constituent. *Eur J Pharmacol* 1993;231:313–314.

85. Giuffrida A, Parsons LH, Kerr TM, et al. Dopamine activation of endogenous cannabinoid signaling in dorsal striatum [see Comments]. *Nat Neurosci* 1999;2:358–363.

86. Arnone M, Maruani J, Chaperon F, et al. Selective inhibition of sucrose and ethanol intake by SR 141716, an antagonist of central cannabinoid (CB1) receptors. *Psychopharmacology (Berl)* 1997;132:104–106.

87. Gardner EL, Vorel SR. Cannabinoid transmission and reward-related events. *Neurobiol Dis* 1998;5:502–533.

88. Mascia MS, Obinu MC, Ledent C, et al. Lack of morphine-induced dopamine release in the nucleus accumbens of cannabinoid CB(1) receptor knockout mice. *Eur J Pharmacol* 1999;383:R1–2.

89. Hansen H, Lauritzen L, Moesgaard B, et al. Formation of N-acyl-phosphatidylethanolamines and N-acylethanolamines. *Biochem Pharmacol* 1998;55:719–725.

90. Joy JE, Watson SJ, Benson J, et al. *Marijuana and medicine*. Washington, DC: National Academy Press, 1999

91. Turner CE, Elsohly MA, Boeren EG. Constituents of *Cannabis sativa L*. XVII. A review of the natural constituents. *J Nat Prod* 1980;43:169–234.

92. El Sohly MA, Ross SA, Mehmedic Z, et al. Potency trends of delta9-THC and other cannabinoids in confiscated marijuana from 1980–1997. *J Forensic Sci* 2000;45:24–30.

93. Pertwee RG, Gibson TM, Stevenson LA, et al. O-1057, a potent water-soluble cannabinoid receptor agonist with antinociceptive properties. *Br J Pharmacol* 2000;129:1577–1584.

94. Lichtman AH, Peart J, Poklis JL, et al. Pharmacological evaluation of aerosolized cannabinoids in mice. *Eur J Pharmacol* 2000;399:141–149.

95. Huffman JW, Liddle J, Yu S, et al. 3-(1′,1′-dimethylbutyl)-1-deoxy-Δ8-THC and related compounds: synthesis of selective ligands for the CB2 receptor. *Bioorg Med Chem* 1999;7:2905–2914.

96. Hillard CJ, Manna S, Greenberg MJ, et al. Synthesis and characterization of potent and selective agonists of the neuronal cannabinoid receptor (CB1). *J Pharmacol Exp Ther* 1999;289:1427–1433.

97. Beltramo M, Stella N, Calignano A, et al. Functional role of high-affinity anandamide transport, as revealed by selective inhibition. *Science* 1997;277:1094–1097.

98. Di Marzo V, Deutsch DG. Biochemistry of the endogenous ligands of cannabinoid receptors. *Neurobiol Dis* 1998;5:386–404.

99. Deutsch DG, Lin S, Hill WAG, et al. Fatty acid sulfonyl fluorides inhibit anandamide metabolism and bind to the cannabinoid receptor. *Biochem Pharmacol* 1997;53:1–16.

100. Deutsch DG, Omeir R, Arreaza G, et al. Methyl arachidonyl fluorophosphonate: a potent irreversible inhibitor of anandamide amidase. *Biochem Pharmacol* 1997;53:255–260.

107

Therapeutics for Nicotine Addiction

Reese T. Jones

Neal L. Benowitz

Reese T. Jones: Department of Psychiatry, University of California, San Francisco, California.

Neal L. Benowitz: Departments of Medicine, Psychiatry, and Biopharmaceutical Sciences, University of California, San Francisco; Division of Clinical Pharmacology and Experimental Therapeutics, Medical Service, San Francisco General Hospital Medical Center, San Francisco, California.

About 25% of adults in the United States smoke tobacco cigarettes. Most continue smoking because they are addicted to nicotine. That nicotine is central to maintaining tobacco use is well established (49). When asked, 70% of cigarette smokers report that they would like to quit. Each year, less than 1% will actually succeed without any therapeutic interventions. Few other conditions in medicine present nicotine addiction's mix of lethality, prevalence, cost, and relative therapeutic neglect, despite effective and readily available treatment interventions. Health care providers too often fail to assess or treat tobacco addiction despite substantial evidence that even brief therapeutic interventions are effective (2).

Worldwide potential benefits of prevention and adequate treatment are staggering (96). More than 1.2 billion people regularly smoke tobacco. During the twentieth century, only approximately 0.1 billion people died of tobacco use–related illnesses. If current smoking patterns continue, 1 billion additional people will die of smoking-related illness during this century. Half will die during middle age. About 4 million people died of tobacco-related disease in 1998. Projections indicate 10 million tobacco-related deaths yearly by the year 2030, with 70% of those deaths in developing countries. Reducing the number of current smokers by 50% would avoid 25 million premature deaths in the first quarter of this century and about 150 million more by midcentury (96).

Understanding the role of nicotine in sustaining tobacco addiction offers a basis for optimal and rational treatments for preventing or stopping smoking (49). Nicotine addiction has much in common with other addictions, so consideration of therapeutics should help in development of therapies for less common addictions to stimulants such as cocaine and amphetamines or to other drugs. Nicotine addiction resulting from tobacco cigarette smoking is emphasized. Other routes for nicotine delivery, chewing tobacco, buccal and nasal snuff, and smoked pipes and cigars, deliver substantial amounts of nicotine but with different pharmacokinetics, although the pharmacology is otherwise similar. Nicotine is addicting when it is delivered by any route, but special attributes and the ubiquity of cigarette delivery systems warrant special attention.

WHY DO PEOPLE SMOKE?

Part of "107 - Therapeutics for Nicotine Addiction "

Although the pharmacologic effects of nicotine are essential to sustaining tobacco smoking, the beginnings of tobacco addiction result from nonpharmacologic learned or conditioned factors, social settings, personality, and genetics (4). Taking nicotine enhances an addicted smoker's mood and performance. Nicotine is rewarding. Smoking is an extremely effective way of rapidly and conveniently delivering concentrated doses of nicotine to the brain (4 ,5). Nicotine smokers appear able to discriminate small, rewarding effects from each individual puff and to titrate nicotine dose from each cigarette. From the typical 10 puffs per cigarette, a one pack-per-day smoker receives 73,000 distinct drug reinforcements per year. Although nicotine can enhance mood directly, what is even more important for understanding nicotine addiction therapeutics is that when nicotine is taken by an addicted smoker, the negative consequences of prior nicotine use are diminished. A nicotine withdrawal syndrome is relieved (4).

Most smokers say they want to stop smoking (2 ,13). Most adults have made several attempts to quit and require four or more attempts before quitting permanently. Relapse typically occurs because of disrupted emotional state, work performance, enjoyment of leisure activities, and interpersonal relationships. Life-threatening smoking-related illness should motivate one to quit, yet 50% of smokers after a

myocardial infarction continue to smoke (13 ,49). Smokers with lung or throat cancer or those suffering from chronic obstructive lung disease behave similarly.

Tobacco-taking behavior is made more likely to recur, reinforced by the pharmacologic actions of nicotine (49). With each successive cigarette, a beginning smoker, usually an adolescent, learns to associate certain moods, situations, and environmental factors with the rewarding effects of nicotine. Associations between cues associated with smoking, anticipated nicotine effects and the resulting urge to use tobacco (craving) become all important in maintaining smoking.

Smoking is more likely in certain situations: after a meal, with coffee or alcohol, and with friends who smoke (49). Associations between events and smoking repeated thousands of times make for powerful cues facilitating an urge to smoke. Manipulation of smoking paraphernalia, taste, smell, and sensations from smoke in upper airways become associated with pleasurable effects. Unpleasant or dysphoric moods come to serve as conditioned cues for smoking. For example, an adolescent smoker, usually within the first year of smoking, learns that not having a cigarette available is associated with feelings of irritability and learns that just a few puffs from a cigarette diminish irritability and other dysphoric nicotine withdrawal symptoms. After hundreds of repeated experiences, irritability from any source serves as a cue for smoking.

Left to nature, it is unlikely that many people would make or find a cigarette, light it, and smoke it (49). Conditioning and learning linking nicotine pharmacology and environmental contingencies are facilitated by advertising encouraging, often in subtle ways, the use of tobacco. In the beginning, teenage smokers teach each other. Quickly, links between the pharmacologic actions of nicotine and associated behaviors become powerful (7). Conditioning loses its power only gradually without nicotine delivered in the right dose and context. Conditioning is a major factor in relapse to nicotine use after quitting. Dealing with it is important in any therapeutics for nicotine addiction.

Many smokers report that smoking improves concentration and elevates mood. Cigarette smoking or nicotine administration improves attention, reaction time, and problem-solving, particularly in recently abstinent smokers (55 ,74). Smokers typically report enhanced pleasure and reduced anger, tension, depression, and stress after a cigarette. Whether enhanced performance and improved mood after smoking are mostly or entirely the result of the relief of abstinence symptoms or rather are intrinsic effects of nicotine on the brain remains unclear (49). Improvement in the performance of nonsmokers after nicotine suggests at least some direct enhancement (8).

NICOTINE PHARMACOKINETICS AND METABOLISM

Part of "107 - Therapeutics for Nicotine Addiction "

Some special attributes of smoked nicotine delivery are important for understanding mechanisms and therapeutics of tobacco addiction (5 ,49). Nicotine, a tertiary amine structurally similar to acetylcholine, binds to nicotinic cholinergic receptors in the brain and elsewhere. During smoking nicotine, steam distilled from the burning tobacco is inhaled into the small airways and alveoli on small droplets of tar, buffered to a physiologic pH, absorbed rapidly into the pulmonary capillaries, and thence into systemic arterial blood. Initial arterial blood levels of nicotine are two to six times greater than venous levels (11). Within 10 to 20 seconds after each puff, relatively high levels of nicotine reach the brain. Nicotine levels in plasma and in brain tissue then decline rapidly because of rapid distribution into peripheral tissues.

During a typical smoker’s day, peak and trough plasma and brain nicotine levels vary considerably before and after each cigarette, but nicotine gradually accumulates over 6 to 8 hours of repeated smoking because of nicotine’s 2-hour half-life (5). By midafternoon, relatively constant, steady-state, venous plasma levels, 20 to 40 ng/mL, are reached, but with transient 50-ng/mL increments in arterial and brain levels after each cigarette. During sleep, plasma concentration of nicotine falls progressively but is still measurable on awakening when the first cigarette of the next day is smoked, typically within 30 minutes of awakening. Thus, smoking results in exposure of brain to nicotine 24 hours of each day but with regular brain level perturbations after each puff and each cigarette (5 ,12).

Smokers regulate smoked nicotine intake to maintain their preferred range of concentrations by varying puff and inhalation timing, volume, and number (49). Nicotine intake and resulting plasma levels vary. Smokers can compensate for differing machine-determined nicotine yields to obtain a preferred dose of nicotine whether smoking a high- or low-yield brand (39). Nicotine delivered by cigarettes offers smokers individualized control of nicotine dose unattainable by other nicotine delivery systems (49). The special attributes of smoked nicotine dosimetry are relevant when designing animal experiments to model human tobacco dependence properly (53) and when considering nicotine replacement therapies (NRTs). In contrast to smoking, chewing tobacco and snuff deliver nicotine through oral or nasal mucosa. Plasma and brain nicotine concentrations rise more gradually, reach plateau levels after about 30 minutes, and then decline slowly over the next few hours (5).

NICOTINE RECEPTOR-BASED NEURAL MECHANISMS RELEVANT TO THERAPEUTICS

Part of "107 - Therapeutics for Nicotine Addiction "

Nicotine binds stereoselectively to a diverse family of nicotinic cholinergic receptors widely distributed in brain, autonomic ganglia, adrenal medulla, and neuromuscular junctions (15 ,16). Nicotine’s effects on nicotinic cholinergic receptors in the brain enhance release of an array of neurotransmitters; dopamine, norepinephrine, acetylcholine, serotonin,

vasopressin, ß-endorphin, glutamate, γ-aminobutyric acid, and others (12 ,49). Nicotinic cholinergic receptors have varied functional characteristics, different chemical conductances for sodium and calcium, and variable sensitivity to different nicotinic agonists (17 ,92). Receptor diversity probably accounts for the diverse effects of nicotine experienced by smokers (19). The undoubtedly complex relationships between specific nicotinic cholinergic receptor subtypes and release of specific neurotransmitters are still to be fully characterized (92). Neurotransmitter release is assumed to mediate nicotine effects such as arousal, relaxation, cognitive enhancement, relief of stress, and depression.

A brain nicotinic cholinergic receptor is a ligand-gated ion channel, with five subunits. Most brain nicotinic cholinergic receptors are composed of α and ß subunits. The α subunits are responsible for ligand binding. The ß subunits mediate other aspects of receptor function (29). The nicotinic cholinergic receptor, consisting of α-4 and ß-2 subunits, accounts for 90% of high-affinity nicotine binding in rat brain and may play a critical role in stimulant and rewarding effects (21). The ß-2 subunit is critical for dopamine release, judging from studies of knockout mice lacking that subunit who have less nicotine-induced dopamine release and do not self-administer nicotine as do wild-type mice (76).

When nicotine binds to nicotine receptors, allosteric changes lead to different functional states including a resting state, an activated state (channel open), and two desensitized states (channel closed) (10). Receptor change to the desensitized state probably accounts for tolerance and for the observation that tolerance to nicotine is associated with increased numbers of nicotinic cholinergic receptors in animals during chronic nicotine treatment and in brains of human smokers (24 ,25 ,26 and 27).

Nicotine's effects on brain dopaminergic and noradrenergic systems are important in reinforcing self-administration (49). The mesolimbic dopamine system is assumed to mediate pleasurable and other rewards from nicotine as with other drugs of abuse. Nicotinic receptors are on the nerve terminal membranes in the nucleus accumbens and on membranes of the dopamine-secreting neurons innervating nucleus accumbens located in the midbrain. Unlike cocaine and amphetamine, which exert effects by binding to presynaptic dopamine transporters on nerve terminal membranes, nicotine's effects depend more on modulating the flow of impulses to the terminal field (17). As happens after repeated exposure to other stimulants, repeated exposure to nicotine results in sensitization of its effects on dopamine release in the accumbens. There appears to be co-stimulation of *N*-methyl-D-aspartate (NMDA) receptors for glutamate because the development and the expression of the sensitized dopamine response is attenuated or blocked by the administration of NMDA-receptor antagonists. In this respect, some consequences of repeated nicotine exposure on these pathways are similar to those of other stimulant drugs. The consequences of nicotine's modulating effects on multiple neuronal systems remains to be determined (49).

Sustained exposure to nicotine desensitizes some but not all nicotinic cholinergic receptors and results in a state in which nicotine is needed to maintain normal neurotransmission. As nicotine levels decrease, diminished neurotransmitter release or altered modulation of neurotransmitter systems (17) contributes to a relative deficiency state and in humans, symptoms of lethargy, irritability, restlessness, inability to concentrate, depressed mood, and other symptoms making up the nicotine withdrawal syndrome. Plasma concentrations of nicotine in smokers are sufficient to desensitize mesolimbic dopamine neuron nicotinic receptors reinforcing nicotine self-administration. Thus, self-administration of eight to ten nicotine bolus doses (puffs) during the smoking of each cigarette would cause gradually decreasing dopamine release in the nucleus accumbens. With each successive cigarette and gradually rising levels of brain nicotine, desensitization would increase. If so, tobacco smokers continue to smoke during the latter half of each smoking day under conditions in which nicotine is less likely to stimulate neurotransmitter release than while smoking the first cigarettes of the day. Thus, other mechanisms likely contribute to the rewarding properties of nicotine in the latter portion of the daily cycle of smoking (49).

Nicotine increases or decreases brain serotonin levels, depending on concentration and pattern of exposure (16). A possible role for serotonin release in reward mechanisms is suggested by selective serotonin ($5\text{-}HT_3$) antagonists that reduce nicotine reinforcing effects. Chronic exposure to nicotine results in reduced capacity to synthesize 5-HT in serotonergic terminals. Postmortem human studies indicate that tobacco smoking is associated with reductions in hippocampal 5-HT and 5-hydroxyindole acetic acid (16). Functional consequences of the nicotine-induced changes in 5-HT remain to be established but could partially explain anxiety reduction commonly reported by smokers. Increased 5-HT release could result in anxiety and related symptoms common during the early stages of nicotine withdrawal (49).

Nicotine-mediated release of norepinephrine plays a role in the release of adrenocorticotropic hormone (ACTH) and cortisol. Nicotine, acting on α-7 cholinergic receptors, releases glutamate, enhances fast excitatory synaptic transmission possibly contributing to improved learning and memory (28 ,36), and regulates dopaminergic function. Activation of the locus ceruleus produces behavioral arousal with nicotine-increased burst firing an adaptive reaction to stressful situations (49). Activation of nicotinic cholinergic receptors in the adrenal medulla releases epinephrine and perhaps ß-endorphin, a factor contributing to nicotine's systemic actions (12).

In addition to brain receptor-mediated effects, nicotine activates afferent nerves, an effect possibly accounting for the importance of sensory phenomena in cigarette smoking satisfaction and important in shaping conditioned aspects of smoking behaviors. For example, intravenous nicotine

produces burst firing of locus ceruleus neurons before injected nicotine reaches the brain (47). After an initial rapid onset, brief activation that can be blocked by a peripheral nicotine antagonist, a second longer-lasting activation, mediated by central nicotinic receptors, occurs (31).

NATURAL HISTORY OF NICOTINE DEPENDENCE

Part of "107 - Therapeutics for Nicotine Addiction "

Most nicotine addicts begin smoking during adolescence. Adolescent smoking has been increasing since the 1990s. In the United States, about 3 million adolescents smoke. Each day, 6,000 more begin. Most perceive themselves to be dependent on nicotine within their first year of smoking. Adolescent daily smokers appear to inhale doses of nicotine similar to doses inhaled by adults. When asked, about 50% report wanting to quit, and 71% report having tried and failed (49). Adolescents report withdrawal symptoms similar to those reported by adults (32).

Without treatment interventions, smoking quitting rates in adolescent smokers in the United States are comparable to those of addicted adult smokers. Young, still experimenting smokers are likely to become regular smokers; however, the proportion of adolescents who go on to regular smoking and what influences the progression remain obscure. The first symptoms of nicotine dependence occur within weeks of the onset of occasional use, often before daily smoking begins (7). As many as one-third to one-half of adolescents experimenting with cigarettes become regular smokers.

Interventions to prevent progression to tobacco addiction in adolescents are less effective than in adult smokers (33). Adolescents have less interest in treatment, high treatment dropout rates, and low quitting rates (20). Reviews of adolescent tobacco smoking conclude that better characterization of nicotine dependence (35) and assessment of pharmacotherapies are needed, given the almost epidemic proportions of smoking in adolescents.

RISK FACTORS

Part of "107 - Therapeutics for Nicotine Addiction "

Comorbidity

Some smokers report that smoking helps relieve their depression and other mood disorders. Others become severely depressed when they stop smoking (16 ,52). Smokers are more likely to have experienced major depression, and those who have are less likely to quit smoking (37). Several mechanisms may link smoking and depression (16). Depression sensitizes patients to the dysphoric effects of stressful stimuli. Smokers exposed to stressful stimuli become conditioned to nicotine's diminishing of the adverse effects. Nicotine-related decreases in 5-HT formation and release in the hippocampus could be a factor. Stopping addicting drugs, including tobacco, has been hypothesized to result in a negative affect state with dysphoria, malaise, and inability to experience pleasure that has been termed *hedonic dysregulation* (84). Smokers may be protected from such consequences by the antidepressant properties of nicotine.

Consistent with a notion that nicotine may be self-administered by some smokers to manage affective disorders is an uncontrolled study reporting that transdermal nicotine lessened depression in nonsmokers with major depression (56). Another intriguing connection is that cigarette smoking inhibits activity of brain monoamine oxidase (MAO) A and B as measured in the brains of smokers and nonsmokers by positron emission tomography using MAO ligands (40 ,41). Smokers have a 30% to 40% suppression of brain MAO A and B activity. Medications that inhibit MAO sometimes have antidepressant activity. Conceivably, cigarette smoking could have similar effects. Finally, some researchers suggest that links between depression and cigarette smoking result from a common genetic predisposition (42).

Schizophrenia is a risk factor for nicotine addiction; approximately 80% of patients with schizophrenia smoke (43). An abnormality in neuronal nicotinic acetylcholine receptor expression or function may be involved in the neuropathophysiology of schizophrenia. Nonschizophrenic smokers have increased nicotinic receptor binding in postmortem brain hippocampus, cortex, and caudate with increasing tobacco use. In contrast, schizophrenic smokers have reduced nicotinic receptor levels, a finding suggesting abnormal regulation of high-affinity neuronal nicotinic receptors after nicotine use (43). One theory linking schizophrenia and susceptibility to nicotine addiction comes from observations that schizophrenic patients often have abnormalities in auditory sensory gating. Sensory gating is mediated by functions of the α-7 nicotinic cholinergic receptor (44). Cigarette smoking and nicotine improve abnormal sensory gating in humans and animals. The abnormality of sensory gating in schizophrenia has been linked to the gene also encoding the α-7 subunit (45).

Dependence on alcohol, heroin, cocaine, and other drugs frequently coexists with nicotine addiction (4). Alcohol and nicotine addiction have common heritability (46 ,60). Stimulant drug exposure may cross-sensitize to neurochemical effects of other stimulants, so nicotine and other stimulants enhance the effects of one another (14 ,48). Because all addicting drugs release dopamine in the mesolimbic system, drugs may be interacting or substituting for one another to produce common changes in dopamine-related reinforcement.

Genetics

Nicotine addiction involves multiple genetic and environmental factors (50 ,51). Genetic factors account for a significant proportion in the variation in the use of tobacco in twin studies, with heritabilities estimated to be as high as

84% and 82% for liability to lifetime and current tobacco use, respectively (9), influencing both initiation and maintenance of tobacco smoking (60). Another twin pair study (51) reported a similar pattern, with genetic factors appearing more important than environment (78% versus 22%) in smoking initiation and in development of dependence (72% versus 28%). Genetic factors also appeared important for the appearance of alcohol dependence (55% versus 45%), consistent with a common genetic vulnerability and showing that nicotine and alcohol dependence occur together (60). Environmental factors more strongly determined age of first use of tobacco and alcohol, whereas latency between first use and patterns of regular use were more genetically determined (54). A major genetic influence accounting for about 70% of the variance in risk in a group of Vietnam era twin pairs is consistent with other studies suggesting that heritable traits such as sensitivity to nicotine are relevant to smoking prevention and treatment (18).

Genetically determined dopamine receptor functional differences and genetic variation in hepatic enzyme activity important in metabolizing nicotine suggest possible mechanisms. Individuals with TaqIA alleles (A1 and A2) and TaqIB (B1 and B2) of the D2 dopamine receptor gene had earlier onset of smoking, smoked more, and made fewer attempts to quit (58). Specific gene mutations, including those associated with dopamine D2 receptors (23) and dopamine transporter proteins (95), have been implicated as possible determinants for nicotine addiction. People lacking a fully functional genetically variable enzyme CYP2A6 important in the metabolism of nicotine to cotinine are slow nicotine metabolizers (30). This genotype may be important in protecting against tobacco dependence because of impaired nicotine metabolism and may be important as well in determining dose and response to NRTs.

Tobacco and Nicotine Exposure During Pregnancy

In the United States, about 25% of pregnant women smoke cigarettes, and so each year about 1 million babies are exposed *in utero* to tobacco smoke (89). That so many pregnant women smoke has important implications both for the determinants of tobacco addiction and its therapeutics. Tobacco smoking has long been known to present considerable fetal risks (59). Less well appreciated is that nicotine itself is a neuroteratogen (89). Nicotine given to rats during gestation or adolescence at levels assumed to be consistent with those in human smokers alters gene expression and produces long-lasting central nervous system cellular damage, by reducing cell number and impairing synaptic activity and cell signaling (62). Developing brain cells appear particularly vulnerable. In adult rats, similar exposure stimulates nicotinic cholinergic receptors without lasting cellular changes. Some of nicotine’s effects on cell numbers continue post partum, well after termination of nicotine exposure. The alterations in synaptic function plausibly would account for associated behavioral disruptions evident in humans (98) and in animal models (1 ,62 ,89).

Nicotine doses that do not produce growth retardation still produce central nervous system cell damage, loss, and synaptic dysfunction. The fetal effects of nicotine may be greater during the later stages of pregnancy, a finding suggesting the first trimester is the most desirable period for NRT during pregnancy. Based on the rat data, it may be preferable to introduce NRTs early in pregnancy to try to reduce the fetal exposure to nicotine before the second or third trimester. In the rat models, episodic nicotine delivery, as happens with smoking, is associated with less nicotine exposure to the fetus than continuous exposure from a nicotine patch. Of course, fetal exposure to tobacco smoke presents a host of other toxins besides nicotine.

Maternal smoking during pregnancy has long-term behavioral consequences in humans (89 ,98). Cognitive deficits, behavioral problems in childhood, particularly attention-deficit/hyperactivity disorders, conduct disorders, and substance abuse in the exposed children are associated with maternal smoking. Children whose mothers smoked ten or more cigarettes daily during pregnancy had a fourfold increased risk of prepubertal-onset conduct disorder in boys and a fivefold risk in adolescent-onset drug dependence in girls (98). The outcomes are not explained by other risk factors. Maternal prenatal smoking appeared to be related to future criminal behavior in male children, with a dose–response relationship between intensity of third trimester smoking and arrest history of 34-year-old men whose mothers smoked during pregnancy (63). Although such studies are limited by retrospective maternal reports of smoking behaviors during pregnancy, there is a consistency with outcomes in studies directly assessing maternal smoking.

MANAGEMENT AND THERAPEUTICS OF NICOTINE ADDICTION

Part of "107 - Therapeutics for Nicotine Addiction "

Ideally, therapeutics for nicotine addiction should be available for the 80% of the world’s smokers who live in low- and middle-income countries. Within those countries, smokers have the lowest income, are the least educated, and have the poorest access to health care. Thus, from a world view, cost of therapeutics and access become important considerations. Prevention is obviously an important strategy, but strategies to prevent tobacco addiction must deal with a politically powerful and wealthy multinational industry promoting use of tobacco (64). The tobacco industry in the United States alone spent 6 billion dollars in 1998 to market cigarettes, about 18 million dollars each day. More is spent promoting tobacco use elsewhere in the world. Although successful prevention strategies exist (65), well-funded competition encouraging tobacco use will remain (64).

More intensive therapeutics typically include behavioral interventions combined with NRT delivered over a series of sessions. Individual or group behavioral treatments appear almost equally effective. Intensive treatment programs are effective in assisting even very dependent smokers to stop for a few months. However, as with other addictions, relapse is a major problem. Initial quitting rates of 50% to 60% at 1 month typically decrease to 20% to 30% at 1 year. Various relapse-prevention procedures have been tried. None has proven clearly effective. Most tobacco addicts repeat the quitting process on average every 3.5 years and try three or four times before finally stopping forever (66). In that respect, stopping smoking is similar to overcoming addictions to other psychoactive drugs. Tobacco addiction treatment programs are cost-effective. Average treatment costs per year of life saved are $1,000 to $2,000 per year for brief counseling alone and $2,000 to $4,000 per year of life saved with more intensive counseling and pharmacotherapy to aid in smoking cessation (34 ,67). Smoking cessation treatments are less costly per year of life saved than are generally accepted therapies for hypertension, hypercholesterolemia, and other chronic disorders.

Therapeutics: Clinical Guidelines

Guidelines for treating tobacco dependence were published in 2000 by the United States Public Health Service (2 ,13). The detailed report resulted from critical review of approximately 6,000 peer-reviewed articles on tobacco addiction therapeutics and 50 metaanalyses based on that literature (69).

The major general conclusions were as follows:

1. Tobacco dependence is a chronic condition warranting repeated treatment until abstinence is achieved.
2. Effective treatments for tobacco dependence exist. All tobacco users should be offered treatment.
3. Clinicians and health care systems must institutionalize consistent identification, documentation, and treatment of every tobacco user at every visit.
4. Brief tobacco dependence treatment is effective. Every tobacco user should be offered at least brief treatment.
5. There is a strong relationship between the intensity of tobacco dependence counseling and effectiveness.
6. Three types of counseling are especially effective: practical counseling, social support as a part of treatment, and social support outside of treatment.
7. Five pharmacotherapies for tobacco dependence are effective: nicotine gum, nicotine inhaler, nicotine nasal spray, nicotine patch, and bupropion. At least one of these medications should be prescribed in the absence of contraindications.
8. Tobacco dependence treatments are cost-effective when compared with other medical and disease prevention interventions. Health insurance plans should include as a benefit the counseling and pharmacotherapies identified as effective in the guideline (2).

Contemporaneous reviews of tobacco addiction therapeutics (59 ,70 ,71 ,72 and 73) and an extensive report on tobacco addiction pharmacology and therapeutics from the Royal College of Physicians (49) offered similar conclusions. A summary review from the Cochrane Tobacco Addiction Review Group identified and summarized evidence of efficacy for tobacco addiction therapeutics (91). Details of the 20 systematic reviews are available on the Internet in the Cochrane Library (75). The reviews used a similar strategy and reviewed much the same literature on tobacco addiction therapeutics as did the Public Health Service review.

The Cochrane reviews considered the results from randomized controlled trials having at least 6 months of follow-up (91). Sustained abstinence or point prevalence quit rates were used in the metaanalysis as necessary. Simple advice from physicians presented during routine care was studied in 31 trials that included 26,000 smokers in a variety of clinical settings. Brief advice increased quit rate more than no advice (odds ratio, 1.69; 95% confidence interval, 1.5 to 1.98). Individual counseling was better than brief advice or usual care. Group therapy was more effective than self-help materials alone but not consistently better than interventions with more personal contact. Self-help informational material and printed descriptions of behavioral strategies had a small treatment effect (75 ,91).

Nicotine Replacement Therapeutics

NRT decreases the discomfort of nicotine withdrawal. The relatively stable brain nicotine levels resulting from NRT should facilitate a desensitized state for some nicotinic cholinergic receptors. Because some nicotinic receptor subtypes are more desensitized than others, both nicotine agonistic and desensitization mechanisms could operate together in NRT. In a nicotine-induced desensitized state, norepinephrine release normally stimulated by endogenous acetylcholine would be diminished. Other neurotransmitter release normally stimulated by endogenous acetylcholine could be diminished as well. The resulting NRT modulation of mood states in itself could be rewarding. In addition, some blunting of the reinforcing effects of cigarettes smoked during cessation lapses is likely during NRT. However, the mechanisms of NRT still remain uncertain because the intensity of withdrawal alone is not a good predictor of success for ultimately stopping smoking (3). Even though withdrawal symptoms can be diminished by NRT, other mechanisms, learning coping skills, and replacing some of the positive effects of nicotine are important as well. Whatever the mechanisms, NRT is clearly effective and safe for helping smokers to quit (70).

The Cochrane review of clinical trials with nicotine gum, transdermal nicotine patches, nasal spray, and inhalers concluded

that NRT enhanced early cessation and reduced early relapse when compared with placebo (75 ,91). All products enhanced quitting smoking about twofold. Quitting rates, depending on intensity of concurrent behavioral interventions, ranged from 10% to 30% of patients with a 1-year follow-up. Higher nicotine doses were more effective, although the dose–response function is shallow. NRT did not appear to have significant dependence potential or to cause significant harm (70). Characteristics of long-term NRT users resembled those of treatment failures. It appeared many would be smoking or smoking more if NRT were not available.

However, 70% to 90% of addicted smokers fail to stop smoking despite NRT. Why (77)? Most studies included only nicotine-addicted smokers, so the usefulness of NRT for less addicted smokers remains uncertain. Although recommendations have been made for use of combinations of NRT products, for example, patch plus spray, patch plus gum, or higher-dose NRT, too few trials preclude clear evidence of effectiveness. Long-term reduction in smoking by concomitant use of NRT while smoking continues is being investigated (78). Nicotine inhalers and skin patches have been used safely and with sustained reduction in smoking for up to 30 months (79 ,80).

Particularly for highly dependent smokers, nicotine replacement from patches and gum probably delivers nicotine to the brain too gradually and without the transient but rewarding brief surges in brain nicotine levels from puffing on a cigarette (5 ,19). Nicotine nasal sprays or inhalers more closely approximate smoking in this respect, but only partially so, and clinically they do not offer advantages to patch failures (77). An inhaled nicotine aerosol would, in principle, be an ideal substitute nicotine delivery system, but despite many attempts, a practical inhaled aerosol system providing the control over dose offered by a tobacco cigarette has not been brought to market.

Non–Nicotine Replacement Pharmacotherapies

The consequences of neuroadaptive changes in brain function associated with chronic nicotine exposure should, in principle, be modified by appropriate neurochemical interventions (71 ,81 ,82). Pharmacotherapies mimicking nicotine’s neurochemical effects by increasing or modulating brain levels of dopamine, epinephrine, serotonin, and other neurotransmitters should correct the neurochemical deficiency states associated with nicotine withdrawal. Pharmacotherapies may also mimic some of nicotine’s actions on brain reward systems. Nicotinic receptor antagonism offers an additional strategy. Although treatment with anxiolytics did not improve outcome, antidepressants, bupropion, and nortriptyline increased quit rates (2 ,83 ,91). The mechanisms by which antidepressant drugs benefit smoking cessation are yet to be determined. The neurochemical consequences of chronic nicotine exposure have similarities to the effects of some antidepressants (16 ,52 ,84).

Bupropion

As with many pharmacotherapies, recognition that bupropion could be useful for treating tobacco addiction resulted from serendipitous observations. Smokers being treated with bupropion for depression reported less desire to smoke or greater success in stopping smoking. Bupropion is structurally related to phenethylamines resembling an anorectic drug diethylproprion and is believed to assist smoking cessation by blocking neuronal uptake of dopamine and norepinephrine and possibly by decreasing firing of the locus ceruleus (71). Bupropion and some other antidepressants functionally antagonize some nicotinic cholinergic receptors in muscle and autonomic ganglia and reduce receptor response to nicotine (85). Whether antidepressant drugs similarly antagonize brain nicotine receptors is undetermined.

Bupropion was effective judging from two large trials and two smaller unpublished ones (86). Bupropion alone or combined with a nicotine patch was more effective than the patch alone (87). Although the drug caused dry mouth and insomnia, serious side effects were uncommon. Bupropion was as effective in patients with a history of depression as with those without such a history (88). When given to a group of smokers not trying to quit permanently, bupropion decreased some withdrawal symptoms but had no effect on craving (86). Bupropion was a more cost-effective therapeutic agent for tobacco addiction than NRT (68).

Other Therapies With and Without Utility

Clonidine shares some pharmacologic effects of bupropion and tricyclic antidepressants. The Cochrane review of six clinical trials of clonidine found increased smoking quit rates, but side effects of sedation and postural hypotension posed problems for many patients (75 ,91).

Sensory stimulants mimic mouth and airway sensory responses to smoking that become associated with the pharmacologic effects of nicotine and thus become reinforcers. Ascorbic acid aerosols and citric acid inhalers evaluated in cessation trials reduce craving and some withdrawal symptoms over the short term (71).

The effects of opiate antagonists on cigarette smoking have been studied to examine how opioid systems modulate smoking behavior and to determine whether opioid antagonists could be useful to aid in smoking cessation (22). Naloxone precipitates opiate withdrawal-like symptoms and increases desire to smoke (38). The effects of naloxone or naltrexone on *ad libitum* smoking over brief periods in a laboratory were inconsistent, but some smokers smoked less. A clinical trial compared naltrexone, 50 mg daily for 12 weeks, or placebo, with or without transdermal nicotine (61). Only transdermal nicotine increased abstinence rates.

Naltrexone had no effect on cessation rates. Transdermal nicotine reduced craving and cigarette smoking in smokers who did not quit. Naltrexone had no such effects. Another 4-week trial of naltrexone or placebo found no difference in smoking 6 months later (93). Thus, the clinical trial data indicate no useful role for opioid antagonists in smoking cessation therapy despite the suggestive laboratory results.

Nicotinic receptor antagonism offers another possible strategy. A nicotine antagonist mecamylamine has been investigated as a cessation aid both alone and in combination with NRT (59 ,71). Mecamylamine started before quitting smoking and continued afterwards appeared useful in two studies (94). Combined use of mecamylamine and nicotine patch increased quit rate more than nicotine alone, a finding leading to consideration of mecamylamine blockade of nicotine's rewarding effects (97).

Circulating antibodies binding nicotine in blood and preventing its reaching the brain would be functionally equivalent to a receptor antagonist's preventing nicotine receptor access. Antibodies have been induced by immunization of rats with nicotine linked to an immunogen (68). Immunized animals had reduced brain nicotine concentrations and reduced behavioral and cardiovascular effects after intravenous nicotine (68). Whether immunization alters the reinforcing effects of nicotine remains to be determined.

Lobeline, a nonpyridine alkaloid and partial nicotine receptor agonist from the Indian tobacco plant *(Lobelia inflata),* has long been used in proprietary smoking treatments (59 ,71). Although no longer marketed in the United States, lobeline is available elsewhere. No clinical trials had more than a 6-month follow-up. The drug was judged unproven by the Cochrane review (75 ,91).

ACTH has been suggested to aid smoking cessation, based on the notion that nicotine increases ACTH and cortisol release and that during nicotine withdrawal, there may be a state of hypoadrenocorticism. Uncontrolled trials with small numbers of smokers given a few ACTH injections during the first week after quitting reported high quit rates or decreased smoking, but without controlled clinical trials, ACTH must still be considered unproven (71).

Silver acetate has long been available as an over-the-counter smoking deterrent in the form of chewing gum, lozenges, and spray. A reaction with cigarette smoke produces an unpleasant metallic taste, the basis for this aversive therapy. Several clinical trials reported short-term efficacy, particularly in less addicted smokers (71). Whenever the urge to smoke is great, it is easy to stop silver acetate use, so it does not appear an effective therapeutic for severe nicotine addiction.

The effectiveness of other aversion therapies, acupuncture, hypnotherapy, and exercise was at best considered uncertain (75 ,91).

FUTURE RESEARCH

Part of "107 - Therapeutics for Nicotine Addiction "

As nicotinic cholinergic receptor subtype–specific agonists and antagonists are developed, more specific treatments for nicotine addiction in subtypes of smokers should result (49). Possibly, medications useful for treating affective disorders, schizophrenia, Alzheimer disease, and other brain disorders may result as well from such research. (15 ,16 ,43).

Adolescent smoking initiation rates remain high (49). Beginning smoking among young adults is common (90) and is perhaps increasing (57). Although we understand much about nicotine addiction (2), we still do not know enough to prevent people from becoming addicted or how best to treat highly dependent tobacco users (77).

Research on vulnerability to nicotine addiction should include linking of phenotypes of nicotine response and other aspects of tobacco dependence to genotypes related to specific receptors and other proteins involved in nicotine addiction (49). Study of the mechanisms of nicotine reinforcing effects in people with affective disorders (16), schizophrenia (43), and other drug dependence is important. Study of hormonal and psychosocial mechanisms will help us to understand gender differences in nicotine addiction (99). Longitudinal studies of children of mothers who smoke, with better measures of smoke exposure *in utero* and assessments of behavior through childhood, will better define the natural history of nicotine addiction and will lead to better strategies for prevention and cessation interventions (1 ,49 ,62 ,89 ,98).

CONCLUSION

Part of "107 - Therapeutics for Nicotine Addiction "

Research since the early 1980s has expanded our understanding of nicotine addiction. Now the challenge is to translate knowledge of the biology of nicotine addiction into pharmacotherapies and other therapeutics addressing individual differences in addicted smokers. To do so will require the development of new drugs, better understanding of existing ones, and wisdom needed to match optimal therapies to individual smokers. Much the same could be said about therapeutics for all drug addictions.

ACKNOWLEDGMENTS

Part of "107 - Therapeutics for Nicotine Addiction "

Preparation of this manuscript was supported in part by US Public Health Service grant nos. DA02277, DA12393, and DA00053 from the National Institute on Drug Abuse, National Institutes of Health.

REFERENCES

1. Trauth JA, Seidler FJ, McCook EC, et al. Adolescent nicotine exposure causes persistent upregulation of nicotinic cholinergic receptors in rat brain regions. *Brain Res* 1999;851:9–19.

2. Fiore MC, Bailey WC, Cohen SJ, et al. A clinical practice guideline for treating tobacco use and dependence: a US Public Health Service report. *JAMA* 2000;283:3244–3254.

3. Piasecki TM, Niaura R, Shadel WG, et al. Smoking withdrawal dynamics in unaided quitters. *J Abnorm Psychol* 2000;109:74–86.

4. Department of Health and Human Services, Public Health Service. *The health consequences of smoking: nicotine addiction: a report of the surgeon general*. DHHS (CDC) publication no. 88–8406. Washington, DC: Government Printing Office, 1988.

5. Benowitz NL, Jacob P III. Pharmacokinetics and metabolism of nicotine and related alkaloids. In: Arneric SP, Brioni JD, eds. *Neuronal nicotinic receptors*. New York: Wiley-Liss, 1999:213–234.

6. Shiffman S, Johnston JA, Khayrallah M, et al. The effect of bupropion on nicotine craving and withdrawal. *Psychopharmacology (Berl)* 2000;148:33–40.

7. DiFranza JR, Rigotti NA, McNeill AD, et al. Initial symptoms of nicotine dependence in adolescents. *Tobacco Control* 2000;9:313–319.

8. LeHouezec J, Halliday R, Benowitz NL, et al. A low dose of subcutaneous nicotine improves information processing in nonsmokers. *Psychopharmacology (Berl)* 1994;114:628–634.

9. Maes HH, Woodard CE, Murrelle L, et al. Tobacco, alcohol and drug use in eight- to sixteen-year-old twins: the Virginia Twin Study of Adolescent Behavioral Development. *J Stud Alcohol* 1999;60:293–305.

10. Lena C, Changeux JP. Allosteric modulations of the nicotinic acetylcholine receptor. *Trends Neurosci* 1993;16:181–186.

11. Gourlay SG, Benowitz NL. Arteriovenous differences in plasma concentration of nicotine and catecholamines and related cardiovascular effects after smoking, nicotine nasal spray, and intravenous nicotine. *Clin Pharmacol Ther* 1997;62:453–463.

12. Benowitz NL. Nicotine addiction. *Prim Care* 1999;26:611–631.

13. United States Public Health Service. *Treating tobacco use and dependence: clinical practice guideline*. http://www.surgeongeneral.gov/tobacco/2000.

14. Suemaru K, Gomita Y, Furuno K, et al. Chronic nicotine treatment potentiates behavioral responses to dopaminergic drugs in rats. *Pharmacol Biochem Behav* 1993;46:135–139.

15. Balfour DJK, Fagerstrom KO. Pharmacology of nicotine and its therapeutic use in smoking cessation and neurodegenerative disorders. *Pharmacol Ther* 1996;72:51–81.

16. Balfour DJK, Ridley DL. The effects of nicotine on neural pathways implicated in depression: a factor in nicotine addiction? *Pharmacol Biochem Behav* 2000;66:79–85.

17. Wonnacott S. Presynaptic nicotinic ACh receptors. *Trends Neurosci* 1997;20:92–98.

18. True WR, Heath AC, Scherrer JF, et al. Genetic and environmental contributions to smoking. *Addiction* 1997;92:1277–1287.

19. Fattinger K, Verotta D, Benowitz NL. Pharmacodynamics of acute tolerance to multiple nicotinic effects in humans. *J Pharmacol Exp Ther* 1997;281:1238–1246.

20. Moolchan ET, Ernst M, Henningfield JE. A review of tobacco smoking in adolescents: treatment implications. *J Am Acad Child Adolesc Psychiatry* 2000;39:682–693.

21. Grottick AJ, Trube G, Corrigall WA, et al. Evidence that nicotinic alpha (7) receptors are not involved in the hyperlocomotor and rewarding effects of nicotine. *J Pharmacol Exp Ther* 2000;294:1112–1119.

22. Pomerleau OF. Endogenous opioids and smoking: a review of progress and problems. *Psychoneuroendocrinology* 1998;23:115–130.

23. Lerman C, Caporaso NE, Audrain J, et al. Evidence suggesting the role of specific genetic factors in cigarette smoking. *Health Psychol* 1999;18:14–20.

24. Benwell MEM, Balfour DJK, Anderson JM. Evidence that tobacco smoking increases the density of (-)-[^{3}H]nicotine binding sites in human brain. *J Neurochem* 1988;50:1243–1247.

25. Breese CR, Marks MJ, Logel J, et al. Effect of smoking history on (^{3}H)nicotine binding in human postmortem brain. *J Pharmacol Exp Ther* 1997;282:7–13.

26. Collins AC, Luo Y, Selvaag S, et al. Sensitivity to nicotine and brain nicotinic receptors are altered by chronic nicotine and mecamylamine infusion. *J Pharmacol Exp Ther* 1994;271:125–133.

27. Court JA, Piggott MA, Lloyd S, et al. Nicotine binding in human striatum: elevation in schizophrenia and reductions in dementia with Lewy bodies, Parkinson's disease and Alzheimer's disease and in relation to neuroleptic medication. *Neuroscience* 2000;98:79–87.

28. Gray R, Rajan AS, Radcliffe KA, et al. Hippocampal synaptic transmission enhanced by low concentrations of nicotine. *Nature* 1996;383:713–716.

29. McGehee DS, Role LW. Physiological diversity of nicotinic acetylcholine receptors expressed by vertebrate neurons. *Annu Rev Physiol* 1995;57:521–546.

30. Tyndale RF, Pianezza ML, Sellers EM. A common genetic defect in nicotine metabolism decreases risk for dependence and lowers cigarette consumption. *Nicotine Tobacco Res* 1999;1:S63–S67.

31. Engberg G, Hajos M. Nicotine-induced activation of locus coeruleus neurons: an analysis of peripheral versus central induction. *Naunyn Schmiedebergs Arch Pharmacol* 1994;349:443–446.

32. Colby SM, Tiffany ST, Shiffman S, et al. Are adolescent smokers dependent on nicotine? A review of the evidence. *Drug Alcohol Depend* 2000;59:S83–S95.

33. Hurt RD, Croghan GA, Beede SD, et al. Nicotine patch therapy in 101 adolescent smokers: efficacy, withdrawal symptom relief, and carbon monoxide and plasma cotinine levels. *Arch Pediatr Adolesc Med* 2000;154:31–37.

34. Nielsen K, Fiore MC. Cost-benefit analysis of sustained-release bupropion, nicotine patch, or both for smoking cessation. *Prev Med* 2000;30:209–216.

35. Colby SM, Tiffany ST, Shiffman S, et al. Measuring nicotine dependence among youth: a review of available approaches and instruments. *Drug Alcohol Depend* 2000;59:S23–S39.

36. McGehee DS, Heath MJS, Gelber S, et al. Nicotine enhancement of fast excitatory synaptic transmission in CNS by presynaptic receptors. *Science* 1995;269:1692–1696.

37. Covey LS. Tobacco cessation among patients with depression. *Prim Care* 1999;26:691–706.

38. Krishnan-Sarin S, Rosen MI, O'Malley SS. Naloxone challenge in smokers: preliminary evidence of an opioid component in nicotine dependence. *Arch Gen Psychiatry* 1999;56:663–668.

39. Scherer G. Smoking behaviour and compensation: a review of the literature. *Psychopharmacology (Berl)* 1999;145:1–20.

40. Fowler JS, Volkow ND, Wang GJ, et al. Inhibition of monoamine oxidase B in the brains of smokers. *Nature* 1996;379:733–737.

41. Fowler JS, Volkow ND, Wang GJ, et al. Wolf. Brain monoamine oxidase A inhibition in cigarette smokers. *Proc Natl Acad Sci USA* 1996;93:14065–14069.

42. Kendler KS, Neale MC, MacLean CJ, et al. Smoking and major depression: a causal analysis. *Arch Gen Psychiatry* 1993;50:36–43.

43. Breese CR, Lee MJ, Adams CE, et al. Abnormal regulation of high affinity nicotinic receptors in subjects with schizophrenia. *Neuropsychopharmacology* 2000;23:351–364.

44. Dalack GW, Healy DJ, Meador-Woodruff JH. Nicotine dependence in schizophrenia: clinical phenomena and laboratory findings. *Am J Psychiatry* 1998;155:1490–1501.

45. Freedman R, Coon H, Myles-Worsley M, et al. Linkage of a neurophysiological deficit in schizophrenia to a chromosome 15 locus. *Proc Natl Acad Sci USA* 1997;94:587–592.

46. Hughes JR. Genetics of smoking: a brief review. *Behav Ther* 1986;17:335–345.

47. Tung CS, Ugedo L, Grenhoff J, et al. Peripheral induction of burst firing in locus coeruleus neurons by nicotine mediated via excitatory amino acids. *Synapse* 1989;4:313–318.

48. Horger BA, Giles MK, Schenk S. Preexposure to amphetamine and nicotine predisposes rats to self-administer a low dose of cocaine. *Psychopharmacology (Berl)* 1992;107:271–276.

49. Tobacco Advisory Group, Royal College of Physicians. *Nicotine addiction in Britain: a report of the Tobacco Advisory Group of the Royal College of Physicians*. London: Royal College of Physicians, 2000.

50. Clarke PBS. Tobacco smoking genes, and dopamine. *Lancet* 1998;352:84–85.

51. Kendler KS, Neale MC, Sullivan P, et al. A population-based twin study in women of smoking initiation and nicotine dependence. *Psychol Med* 1999;29:299–308.

52. Markou A, Kosten TR, Koob GK. Neurobiological similarities in depression and drug dependence: a self-medication hypothesis. *Neuropsychopharmacology* 1998;18:135–174.

53. Stolerman IP. Inter-species consistency in the behavioural pharmacology of nicotine dependence. *Behav Pharmacol* 1999;10:559–580.

54. Stallings MC, Hewitt JK, Beresford T, et al. A twin study of drinking and smoking onset and latencies from first use to regular use. *Behav Genet* 1999;29:409–421.

55. Levin ED, Conners CK, Silva D, et al. Transdermal nicotine effects on attention. *Psychopharmacology (Berl)* 1998;140:135–141.

56. Salin-Pascual RJ, Rosas M, Jimenez-Genchi A, et al. Antidepressant effect of transdermal nicotine patches in nonsmoking patients with major depression. *J Clin Psychiatry* 1996;57:387–389.

57. Wechsler H, Rigotti NA, Gledhill-Hoyt J, et al. Increased levels of cigarette use among college students: a cause for national concern. *JAMA* 1998;280:1673–1678.

58. Spitz MR, Shi H, Yang F, et al. Case-control study of the D2 dopamine receptor gene and smoking status in lung cancer patients. *J Natl Cancer Inst* 1998;90:358–363.

59. Haustein KO. Pharmacotherapy of nicotine dependence. *Int J Clin Pharmacol Ther* 2000;38:273–290.

60. True WR, Xian H, Scherrer JF, et al. Common genetic vulnerability for nicotine and alcohol dependence in men. *Arch Gen Psychiatry* 1999;56:655–661.

61. Wong GY, Wolter TD, Croghan GA, et al. A randomized trial of naltrexone for smoking cessation. *Addiction* 1999;94:1227–1237.

62. Trauth JA, Seidler FJ, Slotkin TA. An animal model of adolescent nicotine exposure: effects on gene expression and macromolecular constituents in rat brain regions. *Brain Res* 2000;867:29–39.

63. Brennan PA, Grekin ER, Mednick SA. Maternal smoking during pregnancy and adult male criminal outcomes. *Arch Gen Psychiatry* 1999;56:215–219.

64. Anderson P, Hughes JR. Policy interventions to reduce the harm from smoking. *Addiction* 2000;95[Suppl 1]S9–S11.

65. Biener L, Harris JE, Hamilton W. Impact of the Massachusetts tobacco control programme: population based trend analysis. *BMJ* 2000;321:351–354.

66. Hughes JR. Four beliefs that may impede progress in the treatment of smoking. *Tobacco Control* 1999;8:323–326.

67. Cromwell J, Bartosch WJ, Fiore MC, et al. Cost-effectiveness of the clinical practice recommendations in the AHCPR guideline for smoking cessation: Agency for Health Care Policy and Research [see Comments]. *JAMA* 1997;278:1759–1766.

68. Pentel PR, Malin DH, Ennifar S, et al. A nicotine conjugate vaccine reduces nicotine distribution to brain and attenuates its behavioral and cardiovascular effects in rats. *Pharmacol Biochem Behav* 2000;65:191–198.

69. Treating Tobacco Use and Dependence. Clinical Practice Guideline, Meta-Analysis. http://www.surgeongeneral.gov/tobacco/meta.htm 2000.

70. Balfour D, Benowitz N, Fagerstrom K, et al. Diagnosis and treatment of nicotine dependence with emphasis on nicotine replacement therapy: a status report. *Eur Heart J* 2000;21:438–445.

71. Benowitz NL, Peng MW. Non-nicotine pharmacotherapy for smoking cessation: mechanisms and prospects. *CNS Drugs* 2000;13:265–285.

72. Davis RM. Moving tobacco control beyond “the tipping point.” *BMJ* 2000;321:309–310.

73. Hughes JR, Goldstein MG, Hurt RD, et al. Recent advances in the pharmacotherapy of smoking [see Comments]. *JAMA* 1999;281:72–76.

74. LeHouezec J, Benowitz NL. Basic and clinical psychopharmacology of nicotine. *Clin Chest Med* 1991;12:681–699.

75. Cochrane Tobacco Addiction Group, Cochrane Library. http://www.update-software.com/ccweb/cochrane/revabstr/g160index.htm 2000.

76. Picciotto MR, Zoll M, Rimondini R, et al. Acetylcholine receptors containing the b2 subunit are involved in the reinforcing properties of nicotine. *Nature* 1998;391:173–177.

77. Hughes JR, Grass JA, Pillitteri JL. Treatment resistant smokers: a pilot study of nicotine nasal spray and inhaler. *J Addict Dis* 2000;19:95–100.

78. Benowitz NL, Zevin S, Jacob P. Suppression of nicotine intake during ad libitum cigarette smoking by high-dose transdermal nicotine. *J Pharmacol Exp Ther* 1998;287:958–962.

79. Bolliger CT, Zellweger JP, Danielsson T, et al. Smoking reduction with oral nicotine inhalers: double blind, randomised clinical trial of efficacy and safety. *BMJ* 2000;321:329–333.

80. Hughes JR. Reduced smoking: an introduction and review of the evidence. *Addiction* 2000;95[Suppl 1]:S3–S7.

81. Ferry LH. Non-nicotine pharmacotherapy for smoking cessation. *Prim Care* 1999;26:653–669.

82. Hughes JR. New treatments for smoking cessation. *CA Cancer J Clin* 2000;50:143–151.

83. Hughes JR, Stead LF, Lancaster T. Anxiolytics and antidepressants for smoking cessation. http://www.update-software.com/ccweb/cochrane/revabstr/ab000031.htm 2001.

84. Koob GF, Le Moal M. Drug abuse: hedonic homeostatic dysregulation. *Science* 1997;278:52–58.

85. Fryer JD, Lukas RJ. Noncompetitive functional inhibition at diverse, human nicotinic acetylcholine receptor subtypes by bupropion, phencyclidine, and ibogaine. *J Pharmacol Exp Ther* 1999;288:88–92.

86. Hurt RD, Sachs DPL, Glover ED, et al. A comparison of sustained-release bupropion and placebo for smoking cessation. *N Engl J Med* 1997;337:1195–1202.

87. Jorenby DE, Leischow SJ, Nides MA, et al. A controlled trial of sustained-release bupropion, a nicotine patch, or both for smoking cessation. *N Engl J Med* 1999;340:685–691.

88. Hayford KE, Patten CA, Rummans TA, et al. Efficacy of bupropion for smoking cessation in smokers with a former history of major depression or alcoholism. *Br J Psychiatry* 1999;174:173–178.

89. Slotkin TA. Fetal nicotine or cocaine exposure: which one is worse? *J Pharmacol Exp Ther* 1998;285:931–945.

90. Rigotti NA, Lee JE, Wechsler H. US college students' use of tobacco products: results of a national survey. *JAMA* 2000;284:699–705.

91. Lancaster T, Stead L, Silagy C, et al. Regular review: effectiveness of interventions to help people stop smoking. Findings from the Cochrane Library. *BMJ* 2000;321:355–358.

92. Wonnacott S. The paradox of nicotinic acetylcholine receptor upregulation by nicotine. *Trends Pharmacol Sci* 1990;11:216–219.

93. Covey LS, Glassman AH, Stetner F. Naltrexone effects on short-term and long-term smoking cessation. *J Addict Dis* 1999;18:31–40.

94. Rose JE, Behm FM, Westman EC. Brand-switching and gender effects in mecamylamine/nicotine smoking cessation treatment. *Nicotine Tobacco Res* 1999;1:286–287.

95. Sabol SZ, Nelson ML, Fisher C, et al. A genetic association for cigarette smoking behavior. *Health Psychol* 1999;18:7–13.

96. Peto RL, Lopez AD. The future worldwide health effects of current smoking patterns. In: Koop CE, Pearson CE, Schwarz MR, eds. *Critical issues in global health*. San Francisco: Jossey-Bass, 2000.

97. Rose JE, Behm FM, Westman EC. Nicotine-mecamylamine treatment for smoking cessation: the role of pre-cessation therapy. *Exp Clin Psychopharmacol* 1998;6:331–343.

98. Weissman MM, Warner V, Wickramaratne PJ, et al. Maternal smoking during pregnancy and psychopathology in offspring followed to adulthood. *J Am Acad Child Adolesc Psychiatry* 1999;38:892–899.

99. Benowitz NL, Hatsukami D. Gender differences in the pharmacology of nicotine addiction. *Addict Biol* 1998;3:383–404.

108

Psychedelic Drugs

Henry David Abraham

Una D. McCann

George A. Ricaurte

Henry David Abraham: Department of Psychiatry, Harvard Medical School, Cambridge, Massachusetts.

Una D. McCann: Department of Psychiatry and Behavioral Sciences, The Johns Hopkins School of Medicine, Baltimore, Maryland.

George A. Ricaurte: Department of Neurology, The Johns Hopkins School of Medicine, Baltimore, Maryland.

As defined in this chapter, the term *psychedelic drugs* includes both classic hallucinogens [i.e., indolalkylamines and phenylalkylamines, such as lysergic acid diethylamide (LSD) and mescaline, respectively], "dissociative" drugs [i.e., arylcyclohexamines, such as phencyclidine (PCP) and ketamine], and substituted amphetamine analogues [i.e., phenylpropanolamines, such as 3,4-methylenedioxymethamphetamine (MDMA, "ecstasy")]. The use of psychedelic drugs dates from the dawn of recorded history and continues today. Indeed, in Western culture, their use appears to be on the rise. Despite the longstanding popularity of psychedelic drugs, controlled research evaluating their effects in humans has been surprisingly scant, and data from preclinical studies have been largely limited to the last several decades. This chapter reviews preclinical and clinical research involving indolalkylamines, arylcyclohexamines, and substituted amphetamines, for which LSD, PCP, and MDMA are used as the prototypes, respectively. Significant recent advances are highlighted, and promising areas toward which future research should be directed are identified.

INDOLALKYLAMINES

Part of "108 - Psychedelic Drugs "

Epidemiology

Surveys in the United States and Western Europe reveal an increased use of indolalkylamine hallucinogens. For example, trend data in the United States, gathered from 15,000 high school seniors, showed a rise in prevalence of lifetime hallucinogen use from 6% to 13.7% between 1986 and 1999 (1 ,2). Similarly, in Great Britain, the use of LSD rose from 7% to 11% between 1989 and 1993. Among German drug abusers, the prevalence of LSD use was reported at 14.1%, and 7.2% of Danes reported the use of hallucinogenic mushrooms (3).

In the United States, a survey of 633 undergraduates found that 23.8% had experimented with hallucinogenic mushrooms, and 16.3% had had experience with LSD. Among LSD users, 6.6% reported problems associated with LSD (Abraham and Koob, *unpublished data*). Of this group, 46.9% reported symptoms of hallucinogen persisting perception disorder (HPPD), 37.5% described alcohol dependence, 25% major depression, 18.8% persisting delusions, 15.6% panic attacks, and 12.5% auditory hallucinations. LSD use is most likely to occur between the ages of 18 and 25. Use is more common in male Caucasians and Hispanics. Of note is that although the parents of LSD users tend to be of a higher socioeconomic status, the users themselves exhibit an inverse relationship between LSD use and educational achievement (4).

Early Neurophysiologic Studies

Work in the 1950s intimated that hallucinogens simultaneously activate and depress neural systems in mammals. In 1953, Gaddum (5) reported that LSD antagonizes the effects of serotonin (5-HT). In the visual system, LSD decreased by 80% the amplitude of the postsynaptic response in the lateral geniculate nucleus of the cat following stimulation of the optic nerve (6). Pentobarbital was found to sensitize the cells to LSD, and asphyxia transiently overcame the LSD effect. These observations were among the first to suggest that in the visual system, LSD is inhibitory, like γ-aminobutyric acid (GABA), and is antagonized by excitatory amino acids released during hypoxia.

Neurophysiologic studies in animals and humans indicate that hallucinogens produce arousal (7). Multiple EEG studies of LSD in rabbits, cats, and humans have documented an increasing shift of alpha frequencies to low voltage, fast rhythms, and alpha disappearance (8). In studies of evoked sensory potentials in cats, a low dose of LSD facilitated both auditory and visual primary responses, whereas high doses depressed auditory responses while continuing

to facilitate visual responses (9). Thus, LSD appears to affect the midbrain and cerebral cortex, particularly the visual cortex, and its effects both stimulate and inhibit, depending on the system studied.

Behavioral Studies

A variety of behavioral models in animals have been employed to study psychedelics. The strength of such models over human studies is that ethical concerns are mitigated, experimental controls are more comprehensive, tissue is available for *in vitro* assessment, and genetic studies are possible with the use of knockout, mutagenesis, and antisense nucleotide strategies. The weakness of animal models is that they cannot provide a direct, reliable method to determine if or when an animal is hallucinating. Despite this limitation, drug discrimination paradigms have been useful in establishing comparative benchmarks between LSD, mescaline, and other hallucinogens, associating potency data with binding at specific receptor types, correlating animal potencies with human data, and describing structure–activity relationships (10). Sophisticated behavioral studies by Geyer et al. (11) suggest that LSD disrupts two fundamental mechanisms of filtering of sensory information, habituation and prepulse inhibition.

Neuropharmacology

The mechanism of action of the hallucinogens is one of the compelling questions in pharmacology, the answer to which promises insights into the mechanisms of perception, mood, and psychosis. Early studies of LSD in peripheral tissue implicated serotoninergic receptors in the mechanism of hallucinogenic activity. Freedman (12) found that LSD decreases brain 5-HT turnover. This effect correlated with behavioral changes and the plasma half-life of LSD, was limited to hallucinogens, and was replicated in several species. Hirschhorn and Winter (13) showed that rats can discriminate LSD and mescaline from saline solution. Discrimination fell in the presence of serotonin antagonists, supporting a 5-HT-agonist mechanism for the action of hallucinogens.

In intracellular recordings from serotoninergic dorsal raphe neurons of the rat brain *in vivo*, LSD directly inhibited firing, but other hallucinogens did not (14). In 1979, it was shown that the effects of LSD on cat behavior are dissociated from raphe responses and involve postsynaptic serotonin activity (15). The same year, Peroutka and Snyder (16) reported the discovery of multiple serotonin receptor types. A high density of 5-HT_{1A} autoreceptors was found on raphe neurons, which explained the direct inhibition of this system by LSD (17). Based on the ability of receptor antagonists to block hallucinogen discrimination in animals, it was proposed that hallucinogens act as agonists at postsynaptic 5-HT_2 receptors (18). Hallucinogen potency in animals was found to correlate with affinity at the 5-HT_2 receptor (19).

Chemistry

Considerable work has been directed at structure–activity relationships of the ergoline hallucinogens (20 ,21). Substitution at the N(1) position of LSD abolishes activity, as does substitution at the C(2) position with a halogen. (*R*)-stereochemistries are essential at both C(5) and C(8) for activity. Reduction of the double bond at the 9,10 position abolishes hallucinogenic activity. Hydroxylation of C(13), which may occur *in vivo*, confers a high level of dopaminergic potency on ergolines (21). Most interesting is that ethylation of LSD at N(6) enhances potency, as determined in both animal and human studies. A monoalkyl amide, a diastereomer of chlorobutyl LSD, is at least 50% more potent than LSD. In ligand binding at 5-HT_2, 5-HT_{1A}, D1, and D2 receptors, the (*R*)-2-butylamide substituent is likewise more potent. Cloning of the 5-HT-$_2$ receptor permitted replacement of aspartate 120 in second transmembrane domain with asparagine. This resulted in a significant decrease in affinity for LSD and abolished phosphatidylinositol turnover. Additionally, aspartate 155 is required for agonist and antagonist binding (22). Second messenger systems in hallucinogen-responsive receptors represent another promising avenue to unraveling the mechanism of hallucinogens. 5-HT_2 receptors are coupled to at least three transduction systems: potassium channels, cationic I_h channels, and phosphoinositide hydrolysis. The close correlation between hallucinogen affinities for the 5-HT_2 and 5-HT_{1C} receptors raises the possibility that the latter may play an independent or complementary role in hallucinogenic activity. This is supported by the fact that LSD is an agonist at 5-HT_{1C} receptors, as determined by phosphoinositide hydrolysis (23).

Recent Neurophysiologic Studies

More recent electrophysiologic studies of hallucinogens in animal models support the involvement of postsynaptic 5-HT_2 and 5-HT_{1C} receptors in hallucinogen activity. The locus ceruleus, considered a sensory novelty detector in the pons, projects widely throughout the brain. Hallucinogens indirectly decrease spontaneous activity in the locus ceruleus by activating $GABA_A$ inputs, and they enhance sensory responses of the locus ceruleus by activating excitatory inputs via *N*-methyl-D-aspartate (NMDA) receptors (24). 5-HT_{2A}-receptor antagonists block these effects. In rat piriform cortex, both 5-HT and hallucinogens at 5-HT_{2A} receptors excite GABAergic interneurons, which then induce inhibitory postsynaptic potentials (25). In prefrontal cortex, the opposite occurs, where the drugs release glutamate and increase excitatory potentials (26). Both 5-HT_{2A} and glutamatergic antagonists block this effect. Direct studies of neocortical

cells suggest that 5-HT_{2A} receptors induce glutamate release by a focal mechanism, not by impulse flow. Recently, it has been suggested that hallucinogens act at 5-HT_{2A} cortical receptors by promoting late, asynchronous excitatory potentials. In such a model, 5-HT itself would antagonize hallucinogens by activating 5-HT_1 receptors (14). This model explains the clinical observation that selective serotonin reuptake inhibitors blunt the effects of LSD, whereas serotonin depletion enhances them.

Although the dominant hypothesis of hallucinogenic activity currently is that it results from partial agonism at the 5-HT_{2A} receptor, similar affinity and agonism data exist for the 5-HT_{2C} receptor (27). Finally, functional interaction is likely to occur between receptor types and subtypes.

Recent Human Studies of Indolalkylamine Hallucinogens

The extraordinary mental effects of LSD described in 1943 by Hofmann prompted hope in the following two decades that a powerful therapeutic tool was at hand. The drug was used experimentally to treat neuroses, childhood schizophrenia, sociopathy, and alcoholism, and as a comfort to the terminally ill (28). Methodologies were inadequate by contemporary standards, and no treatment stands unambiguously as effective. In recent years, renewed interest in hallucinogen research has been sparked by the emergence of positron emission tomography (PET), single-photon emission computed tomography (SPECT), and magnetic resonance imaging (MRI) technologies. For example, PET studies by Vollenweider and colleagues (29) have shown that psilocybin, another hallucinogen, increases frontal glucose metabolism in healthy volunteers, which suggests that the behavioral effects of psilocybin involve the frontal cortex (Fig. 108.1). Similar imaging work has been done with the phenethylamine mescaline.

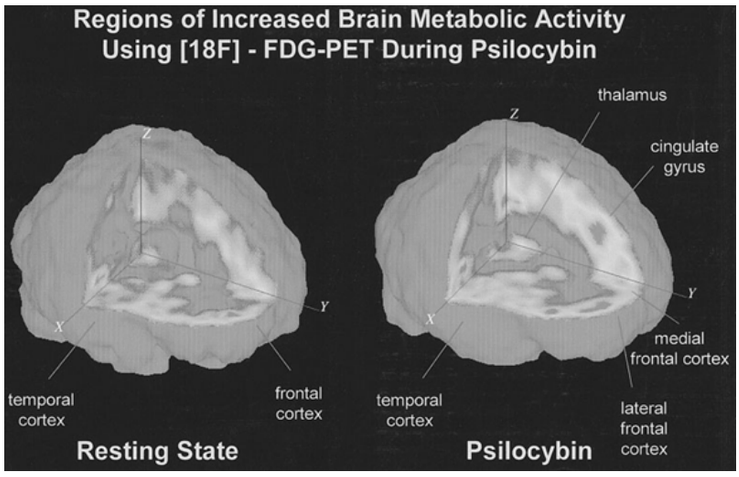


FIGURE 108.1. Positron emission tomography with [^{18}F]-fluorodeoxyglucose before and after a 15- or 20-mg dose of psilocybin in healthy volunteers. Psychotomimetic doses of psilocybin were found to produce a global increase in the cerebral metabolic rate of glucose, with significant and most marked increases in the frontomedial, frontolateral, anterior cingulate, and temporomedial cortex. The increase correlated positively with psychotic symptoms. (Modified from Vollenweider FX, et al. Positron emission tomography and fluorodeoxyglucose studies of metabolic hyperfrontality and psychopathology in the psilocybin model of psychosis. *Neuropsychopharmacology* 1997;16:257–272, with permission.) See color version of figure.

Human dose–response studies of hallucinogens since 1965 have been rare. Shulgin et al. (30) synthesized 179 phenethylamines and informally screened them in human volunteers for hallucinogenic potency. Strassman and Qualls (31), using DMT (N,N-dimethyltryptamine) in carefully screened and supervised subjects, were able to develop a hallucinogen rating scale and measure a number of dose-dependent neuroendocrine responses to the drug. Equally importantly, this work demonstrated that hallucinogen experimentation could be safe as well as informative.

Acute Adverse Psychiatric Effects of Hallucinogens

Clinically, the flow of thoughts, feelings, and perceptions that constitute a hallucinogenic experience can, on occasion, result in panic. Thus, a man who was using LSD while driving tried to crash the vehicle when he saw his companion turn into a giant lizard (32). The treatment for hallucinogen-induced panic is an oral benzodiazepine. The utter efficacy and rapidity of the response to this class of medications implicates GABA receptors as the neuromodulators of this hallucinogenic experience in humans.

Hallucinogen Persisting Perception Disorder

Hallucinogens sometimes appear to alter psychological functions years after drug use (32). Elkes et al. (33) originally noted recurrences of drug experiences, flashbacks, to occur episodically. Surveys among college students reveal that more than 40% of those using LSD report minor spontaneous visual experiences weeks to months after LSD use (32). Less common are patients who report persistent, continuous visual disturbances following LSD use. These people

are affected by a variety of disturbances, such as after-imagery, geometric pseudohallucinations, halos around objects, and the trailing of visual images as they move through the visual field (32).

Hallucinogen persisting perception disorder appears to be a permanent, or slowly reversible, disorder of disinhibition of visual processing, which suggests the defective sensory gating described by Braff and Geyer (34). Evidence for this comes from psychophysical experiments in which visual signals in subjects with HPPD persisted significantly longer than in LSD-naïve controls (35). Quantitative electrophysiology (qEEG) in this population shows abnormalities in visually evoked potentials as long as 26 years after last LSD use, consistent with visual disinhibition (36). Thus, these studies are consistent with others showing that the visual system is especially sensitive to the effects of LSD. Second, LSD hallucinations involve the cerebral cortex (37). Third, inhibitory systems appear, at least in certain circumstances, to be involved in LSD effects and probably LSD aftereffects. Fourth, flashbacks may in certain cases become long-lived, continuous, and probably permanent. And fifth, HPPD is associated with cortical disinhibition.

Recently, it has been found that the GABA agonist midazolam rapidly reduces experimentally induced afterimages in persons with HPPD to approximate the responses from controls without the disorder (Abraham, *unpublished data*). Clinically, GABA agonists are known to reduce, but not resolve, symptoms of HPPD symptoms. This modulation of after-imagery suggests that the visual dysregulation of HPPD may be related to a permanent loss of GABA-mediated inhibition. Risperidone, a 5-HT_2 antagonist, has been found to exacerbate HPPD in persons with the disorder (38). Nefazodone, also a 5-HT_2 antagonist, is associated with visual trailing phenomena. One suggestion regarding the etiology of HPPD is that LSD in vulnerable persons reduces the population of 5-HT_2 inhibitory interneurons modulating visual processing by excitotoxicity, thus reducing GABA efferents to glutamatergic neurons. Treatment for HPPD remains empiric and palliative. It may include benzodiazepines, sertraline, naltrexone, and clonidine. Nonaddictive agents are preferred in patients with histories of addiction.

Psychosis

Studies of LSD administration to research subjects report an incidence of subsequent psychosis in 0.08% to 4.6% of the samples. Psychiatric patient status appears to be a risk factor for the development of psychosis (39). Case histories tend to support phenomenology of a schizoaffective presentation with the added feature of visual disturbances. Positive symptoms of schizophrenia tend to be present. Effective treatments include neuroleptic, lithium, and electroconvulsive therapies.

ARYLCYCLOHEXAMINES

Part of "108 - Psychedelic Drugs "

The arylcyclohexamine PCP (“angel dust,” “peace pill”) can be considered the prototypal dissociative anesthetic. Other drugs in this class include ketamine and dizocilpine maleate (MK-801). PCP was first synthesized in the 1950s, when it was marketed as a surgical anesthetic under the trade name Sernyl. Initially widely used in surgical settings, it was withdrawn in 1965 because of its association with a variety of behavioral disturbances, including agitation, dysphoria, delirium, hallucinations, paranoia, rage, and violence (40). In approximately half of patients who received PCP, a psychotic syndrome developed that sometimes persisted for more than a week (41). Today, the psychotic syndrome produced by PCP or ketamine is considered a leading drug model of schizophrenia (42).

Chemistry

Phencyclidine and other dissociative anesthetics consist of a phenyl ring, a piperidine group, and a cyclohexyl ring. The two conformations of drugs in this class are categorized according to the cyclohexyl spine and subsequent location of the phenyl ring. In particular, if the phenyl ring is located in the axial plane, the drug is active, whereas location of the phenyl group in the equatorial plane renders the drug inactive. Ring number and substitutions can significantly alter the potency of drugs in this class (43).

Epidemiology

Until recently, typical PCP users were white, blue collar men with a high school (or partial high school) education working in unskilled or semiskilled jobs (40). However, dissociative anesthetics such as ketamine and PCP are increasingly used in the growing “club” or “dance” culture, and it appears that the popularity of these drugs has risen in young adults. Data from the Monitoring the Future Study indicate that lifetime prevalence rates of PCP use increased from 2.4% to 2.7% in young men ages 18 to 29 from 1997 to 1998, whereas use among twelfth graders remained at 3.9% in 1997 and 1998 and fell to 3.4% in 1999 (2).

Neuropharmacology

The mechanism of action of dissociative anesthetics is both unique and complex, involving a number of distinct neurotransmitter and neuromodulator systems. PCP behaves as a cholinergic antagonist at both central and peripheral sites, acting at both nicotinic and muscarinic receptors. (44). PCP is also a D2-receptor antagonist, and actions at this receptor are believed to underlie many of the behavioral symptoms that follow drug administration (40). In addition to D2 blockade, PCP increases the rate of dopamine release from synaptic vesicles and prevents dopamine reuptake inactivation.

Amphetamine-like activating effects of PCP are believed to involve not only dopamine uptake blockade but also actions of PCP in the frontal cortex and consequent neuromodulatory effects of the frontal cortex on the basal ganglia (45). Actions at σ and μ opiate receptors are thought to underlie the anesthetic effect of PCP, whereas actions at serotonin receptors may underlie its hallucinogenic effects (40). Notably, cross-tolerance occurs between PCP and the classic hallucinogens LSD, mescaline, and psilocybin (46), and PCP substitutes for LSD or mescaline in two-level drug discrimination studies in rats. PCP also binds to two specific PCP sites in the brain. One PCP receptor, located within the NMDA receptor-gated ion channel, is stimulated by NMDA-receptor agonists such as L-glutamate and can be modulated by a variety of modulatory agents, such as glycine-like amino acids and polyamines (47). A second, lower-affinity PCP receptor has been identified but is less well characterized (40).

Behavioral Effects in Humans

Phencyclidine produces a mixture of stimulant, depressant, anesthetic, and hallucinogenic effects, with the particular presentation dependent, in part, on dosage. In particular, low doses are associated with anticholinergic symptoms (red and dry skin, nystagmus, amnesia, conceptual disorganization); moderate doses are more likely to be associated with opiate receptor activity (anesthesia, dreamlike states); at high doses, dopaminergic symptoms predominate (hallucinations, paranoia). However, this rule of thumb should not be considered diagnostic. The mnemonic RED DANES was coined by Giannini and colleagues (48 ,49) to characterize eight acute symptoms of PCP intoxication that may be seen at any dose: rage, erythema, dilated pupils, delusions, amnesia, nystagmus in the horizontal plane, excitation, and skin dry. It is important to note that the toxic effects of PCP may persist for days because the half-life of PCP after overdose may be as long as 3 days (50).

In addition to acute toxicity, a number of researchers have reported persistent cognitive deficits in long-term PCP users, particularly in short-term memory function (51 ,52 ,53 and 54). Also, abrupt lapses into confusional states occurring weeks or months after PCP ingestion have been reported.

Phencyclidine Neurotoxicity

Olney and colleagues (55) were the first to report that single doses of PCP and related compounds (MK-801 and ketamine) lead to neurotoxic damage of neurons in layers III and IV of the posterior cingulate and retrosplenial cortex in rats. These cells display abnormal cytoplasmic vacuolization that is directly correlated with the potency of noncompetitive NMDA blockade. Initially, these were believed to be short-term changes, but higher doses of MK-801 were observed to cause necrotic changes persisting at least 48 hours after drug administration. Subsequently, other researchers reported a number of observations suggesting that PCP is neurotoxic. In particular, following PCP administration, vacuolization of neurons in hippocampal fields CA1 and CA3 and the subiculum has been demonstrated (56). PCP induces a microglial response and a 70-kilodalton heat shock protein in cerebellar Purkinje cells (57); most recently, it has been found that PCP induces apoptosis in striatopallidal cells in rats (58). The mechanisms for the actions of PCP at these various anatomic sites are likely to differ, with cortical injury involving activity of cholinergic, GABAergic, and adrenergic neuronal systems (59) and apoptotic changes observed in striatopallidal cells involving excess corticosteroids (58). Research is needed to determine whether PCP-induced neurotoxicity underlies the memory deficits seen in some PCP users.

SUBSTITUTED AMPHETAMINES (MDMA, “ECSTASY”)

Part of "108 - Psychedelic Drugs "

Chemistry

3,4-Methylenedioxymethamphetamine bears structural similarity to both the psychomotor stimulant amphetamine and the hallucinogen mescaline. Of the two optical isomers of MDMA, the dextrorotatory isomer exhibits more potent central nervous system activity (60). In contrast, most potent hallucinogenic amphetamines are more potent in their levorotatory forms (61). The aromatic methylenedioxy substituent of MDMA is similar to the substance found in oils of the natural products safrole and myristicin, once proposed to be the intoxicants of sassafras and nutmeg (62).

Epidemiology

Data from the most recent Monitoring the Future Study indicate that MDMA use has continued to rise since 1989 (1 ,2). For example, annual use of MDMA among college students rose from 2.4% in 1997 to 3.9% in 1998, with lifetime figures reflecting similar increases, from 4.7% in 1997 to 6.8% in 1998 (1). Notably, figures from 1996 to 1999 indicate that approximately 3% of eighth graders and approximately 8% of twelfth graders have experimented with MDMA in their lifetime (2), which suggests that in the United States MDMA use begins at an early age.

Patterns of Use

At present, MDMA is used primarily for recreational purposes, although some still advocate the use of MDMA for psychotherapeutic purposes (63). During the last decade, the most frequently reported use of MDMA has been in the context of large, organized social events known as “raves,” often held in warehouses or dance clubs. Festively

dressed “ravers” use MDMA as their drug of choice and typically dance through the night to music accompanied by computer-generated videos and laser light shows. The amount of MDMA typically used during raves varies widely, with doses ranging from 75 to 1,250 mg over several hours.

Acute Neurochemical Effects

The most pronounced acute biochemical effect of MDMA is increased 5-HT neurotransmission, brought about by a calcium-independent release of 5-HT from nerve endings (64). MDMA-induced 5-HT release involves both vesicular and plasma membrane monoamine transporter (65). Actions at the serotonin transporter are also thought to lead to reuptake inactivation (66). MDMA also appears to release dopamine, but this effect is less pronounced than those on serotonin neurons (66). Unlike the actions of classic hallucinogens, the acute neurochemical actions of MDMA are primarily indirect rather than mediated directly at postsynaptic 5-HT receptors, for which MDMA has s low affinity (67).

The binding potential of MDMA at a number of postsynaptic receptor sites and reuptake sites has been evaluated (68 ,69). The affinity of racemic MDMA for receptors was initially found to be greatest for the serotonin transporter (SERT), followed in turn by the α_2-adrenergic receptor, the 5-HT_2receptor, the histamine H1 receptor, and the muscarinic M1 receptor (70). In a subsequent study by Pierce and Peroutka (69), in which a more selective 5-HT_{2A}-receptor agonist, 2,5-dimethoxy-4-^{77}Br-amphetamine (DOB), was used, the binding potency of MDMA at 5-HT_{2A} receptors was greater than that at α_2-adrenergic receptors.

Behavioral Effects in Animals

The administration of MDMA in animals leads to typical signs of mild sympathomimetic stimulation; these include increased locomotor activity, heart rate, and body temperature in rats (71) and mydriasis, salivation, piloerection, and hyperthermia in dogs and monkeys (72 ,73). Locomotor studies suggest that MDMA can be distinguished from amphetamine, and in some behavioral paradigms, it appears to have a greater similarity to hallucinogens than to amphetamine (74).

In drug discrimination studies, MDMA substitutes for D-amphetamine in rats (75), pigeons (76), and monkeys (77). In contrast, despite structural similarities to mescaline, responses to MDMA differ from those to the hallucinogen DOB (61), but they are similar to those for the indolalkylamine α-methyltryptamine (78).

Animal studies investigating the abuse potential of MDMA are consistent with epidemiologic studies and abuse patterns previously described in humans. In particular, baboons self-administer MDMA (28). Rhesus monkeys trained to self-administer cocaine prefer MDMA to vehicle, and they sometimes administer MDMA at a higher rate than cocaine (79). In rats, MDMA lowers the electric threshold for self-stimulation in the medial forebrain bundle (80). Thus, in three different behavioral paradigms, MDMA appears to have significant potential for abuse.

Human Studies with MDMA

As would be predicted from studies in animals, MDMA exhibits both stimulant and hallucinogen-like activity. The stimulant effects of MDMA, typically noted shortly after drug ingestion, include increased heart rate, increased blood pressure, decreased appetite, increased alertness, and euphoria (81). Data regarding the effects of MDMA in humans come from both retrospective, uncontrolled studies and controlled, laboratory-based research. These studies are described below.

Greer and Tolbert (82) summarized experiences from 29 separate clinical therapy sessions during which MDMA was utilized as a psychotherapeutic adjunct. Patients received doses of MDMA ranging between 75 and 150 mg after a 6-hour fast (one subject, at his request, received a higher dose). A second dose of 50 or 75 mg was offered when the effects of the first dose began to subside. The 21 patients who were engaging in couples therapy reported increased closeness or enhanced communication with their partner, and all 29 patients reported positive attitudinal and emotional changes. Of the 29 patients, 22 reported “cognitive” benefits, such as “an expanded mental perspective, insight into problems, and issue resolution.” All patients reported adverse effects, including fatigue, jaw clenching, nausea, transient gait disturbance, and sympathomimetic symptoms.

In the first double-blinded, randomized study involving the prospective administration of MDMA to humans (83), subjects received MDMA orally at doses ranging from 0.25 to 1.0 mg/kg (17.5 to 70 mg in a 70-kg adult). These doses were associated with increased heart rate and blood pressure and positive psychological effects. In a second double-blinded, placebo-controlled study (84), the effects of MDMA (1.7 mg/kg; 119 mg in a 70-kg person) were evaluated in 13 MDMA-naïve healthy volunteers. MDMA was reported to enhance mood, a sense of well-being, and emotional sensitivity. Some subjects reported anxiety. Other symptoms reported included mild depersonalization and derealization, altered time perception, moderate thought disorder, poor coordination, heightened sensory awareness, and increased energy. A hypertensive reaction developed in one subject. Adverse subjective somatic effects of MDMA included jaw clenching, anorexia, impaired gait, and restless legs. After 24 hours, subjects’ complaints included poor energy and appetite, restlessness, insomnia, trismus, poor concentration, and brooding. In the most recent prospective, double-blinded study of MDMA administration in humans (85), the effects of 75 and 125 mg of MDMA were compared with those of 40 mg of amphetamine and placebo.

Both doses of MDMA led to significant increases in blood pressure (increases in systolic blood pressure averaging 40 mm Hg), heart rate (increases averaging 30 beats/min), and pupillary diameter in comparison with placebo. No hallucinations were reported following any drug. All active drugs led to increases in euphoria that were greatest following 125 mg of MDMA. MDMA was also reported to produce altered visual and auditory perception.

Neuroendocrine Effects

In rats, the systemic administration of MDMA leads to a pronounced elevation in levels of corticosterone and prolactin, accompanied by an elevation in temperature (86 ,87). These effects appear to be mediated by 5-HT receptors because they are attenuated or completely blocked by pretreatment with the 5-HT neurotoxin *p*-chlorophenylalanine (86). MDMA-induced increases in corticosterone levels and temperature are blocked by 5-HT_2-receptor antagonists but not by 5-HT_{1A}-receptor antagonists or nonspecific 5-HT-receptor antagonists. In contrast, MDMA-induced prolactin responses are not attenuated by either 5-HT_{1A}-receptor or 5-HT_2-receptor antagonists, which suggests that the two MDMA-induced neuroendocrine responses involve different 5-HT receptors.

Several studies have evaluated the neuroendocrine effects of MDMA in humans. MDMA doses of up to 75 mg are associated with increases in cortisol, and higher doses lead to increases in both cortisol and prolactin (83 ,85). Notably, evidence in both animals and humans is increasing that previous exposure to MDMA leads to alterations in neuroendocrine responses (87 ,88 ,89 ,90 ,91 and 92), possibly as a consequence of long-term effects on brain 5-HT neurons.

Biodisposition in Animals

The metabolic pathways of MDMA have been well characterized in several animal species. *In vivo* studies in rats have shown that MDMA is metabolized via *N*-demethylation, *O*-dealkylation, deamination, and conjugation (*O*-methylation, *O*-glucuronidation, and *O*-sulfation) (93). The (*S*)-(+)-MDMA isomer of MDMA appears to be metabolized more rapidly (94) and extensively (95) than the (*R*)-(-)-MDMA isomer, with half-life estimates being 73.8 and 100.7 minutes for (*S*)-(+)- and (*R*)-(-)-MDMA, respectively (94). Nonconjugated metabolites of MDMA are present in blood, brain, liver, feces, and urine for a 24-hour period following drug administration, with the exception of the *O*-dealkylated catechol metabolite, which is found only in brain tissue (93). This latter pathway, mediated via constitutive cytochrome P-450 isozymes, is a primary route of metabolism in rat brain microsomes.

Biodisposition in Humans

Three studies have evaluated the biodisposition of MDMA in humans (85 ,96 ,97). In the neuroendocrine study by Mas et al. (85), maximum concentrations of MDMA and elimination half-lives were evaluated for 75- and 125-mg doses of MDMA in healthy men. Maximum plasma concentrations were 130.9 and 236.4 ng/mL for the 75- and 125-mg doses respectively and reached peak at 1.8 and 2.4 hours following drug ingestion, respectively. Elimination half-life was 7.7 hours for the 75-mg dose of MDMA and 8.6 hours for the 125-mg dose. Plasma concentrations of (*R*)-(-)-MDMA exceed those of the (*S*)-(+) enantiomer (96). Most recently, de la Torre and colleagues (97) found that relatively small increases in MDMA doses are translated to disproportionate rises in MDMA plasma concentrations, even in persons with high levels of CYP2D6 activity (i.e., extensive metabolizers).

Clinically Reported Adverse Effects

Acute adverse medical effects of MDMA have been reviewed extensively elsewhere (98 ,99). These effects, which are undoubtedly related to the sympathomimetic and serotoninergic properties of MDMA, include nausea, vomiting, jaw clenching, bruxism, hypertension, palpitations, headaches, hyperreflexia, difficulty walking, urinary urgency, diaphoresis, anorexia, muscle aches or tension, hot and cold flashes, nystagmus, blurred vision, insomnia, and dry mouth.

Aside from one report of an acute hypertensive crisis in a prospective study (84), serious acute medical complications of MDMA use have appeared in the literature as case reports or reports from poison centers and coroners. Among the serious problems that have been associated with MDMA use are cerebrovascular incidents (100) and arrhythmias (101), likely related to the potent sympathomimetic and vasoconstrictive effects of MDMA. Electrolyte imbalance or the syndrome of inappropriate secretion of antidiuretic hormone, sometimes associated with cerebral edema or seizures, has been reported by numerous authors (102 ,103).

Numerous reports of chronic medical sequelae of MDMA have also been published, and readers are referred elsewhere for a more comprehensive review of this topic (98 ,99). One serious adverse medical event associated with MDMA, multiple organ system failure, appears to be directly related to the use of MDMA in raves, where users become hot and dehydrated in crowded conditions. In this setting, MDMA is associated with a life-threatening syndrome involving dehydration, hyperthermia, seizures, rhabdomyolysis, disseminated intravascular coagulation, renal failure, and death (104 ,105 and 106). This is reminiscent of the phenomenon of aggregation toxicity in animals (107), in which the lethality of amphetamines is greatly potentiated by crowded housing conditions. Reports of hepatotoxicity, aplastic anemia, and toxic leukoencephalopathy in MDMA users may be related to contaminants in MDMA synthesis or represent idiopathic drug reactions (108 and 109).

Adverse neuropsychiatric effects have also been associated with MDMA. Acute psychiatric complications of MDMA include panic attacks (110), psychosis (111), delirium (112),

and impulsive irrational behavior with subsequent severe medical consequences or death (101 ,113). Chronic neuropsychiatric syndromes reported in MDMA users include panic disorder (114), psychosis (115), aggressive outbursts (116), flashbacks (111), major depressive disorder (117), and cognitive disturbances (117).

Serotonin Neurotoxicity

Like its structural relative methylenedioxyamphetamine (118), MDMA is a well-documented serotonin neurotoxin in a variety of animal species (119 ,120 ,121 and 122). In nonhuman primates, MDMA-induced brain serotonin neurotoxicity is long-lasting and possibly permanent (123 ,124).

The administration of MDMA in animals leads to the persistent loss of a variety of markers specific to brain serotonin neurons. These include brain 5-HT itself (121); 5-hydroxyindolacetic acid (5-HIAA), the major metabolite of serotonin (125); tryptophan hydroxylase, the rate-limiting enzyme in serotonin synthesis (126); and the SERT, a structural protein on the 5-HT nerve terminal (119). Anatomic evidence also indicates a persistent loss of brain serotonin axons and axon terminals. For example, following MDMA administration, quantitative autoradiographic studies with radioligands that bind to the SERT, and immunocytochemical studies in which antibodies are directed at either serotonin or the SERT, show pronounced, long-lasting reductions of the SERT and reduced density of serotonin axons with sparing of serotonin cell bodies (127). These selective serotonin deficits have been observed up to 7 years after drug discontinuation in nonhuman primates (123).

Efforts to determine whether selective serotonin neurotoxicity develops in human MDMA users, as in animals exposed to MDMA, have been limited by the paucity of available methods for assessing the status of central nervous system serotonin structure and function in living humans. At present, two methods for detecting MDMA-induced brain 5-HT neurotoxicity in living humans have been validated. These include measurement of spinal fluid 5-HIAA and PET neuroimaging of the SERT. Both of these methods have demonstrated capability for detecting MDMAinduced neurotoxic injury in nonhuman primates (128 ,129). With these methods, two studies have shown decrements in human cerebrospinal fluid 5-HIAA that are similar to those seen in monkeys with known MDMA-induced 5-HT neurotoxic damage (92 ,130). Similarly, imaging studies with PET have revealed reductions in brain SERT binding in MDMA users that are similar to those seen in baboons with demonstrated MDMA-induced 5-HT damage (49). Further, reductions in the SERT could be correlated with the extent of previous MDMA use.

Studies attempting to identify the functional consequences of MDMA neurotoxicity in humans suggest that brain serotonin function is abnormal in human MDMA users. In particular, as previously described, abnormal neuroendocrine responses to the serotonin-releasing drugs fenfluramine and *m*-chlorophenylpiperazine (m-CPP) have been demonstrated in MDMA users. In the case of *m*-CPP, MDMA users also differ in their behavioral responses to drug. Several research groups have found cognitive impairments in MDMA users in comparison with controls, including decrements of visual and verbal memory, attention, and verbal reasoning (92 ,131 ,132 ,133 and 134). MDMA users have also been found to score higher on measures of impulsivity (88,135,136, but not 130), consistent with work showing an inverse relationship between 5-HT markers and impulsivity (137).

FUTURE RESEARCH DIRECTIONS

Part of "108 - Psychedelic Drugs "

Since Hofmann's discovery of LSD in 1943, significant progress has been made toward understanding the mechanism of action of LSD and other drugs in its class (Fig. 108.2). Despite advances in understanding the mechanism of hallucinogenic action, many questions remain unanswered. During the next decade, it should be possible to refine further the 5-$HT_{2A/1C}$ hypothesis of psychedelic activity, to characterize better the neuroanatomy of the pharmacologic action of LSD, and to use modern neuroimaging techniques to compare and contrast the effects of LSD with those of idiopathic psychiatric illnesses in which hallucinations are a feature. Similarly, future studies of PCP may elucidate certain aspects of idiopathic psychotic illnesses. Clinical studies in PCP users, like those previously conducted in MDMA users, should be directed toward determining whether humans, like rodents, are susceptible to PCP neurotoxic injury and defining the functional consequences of such injury if it occurs.

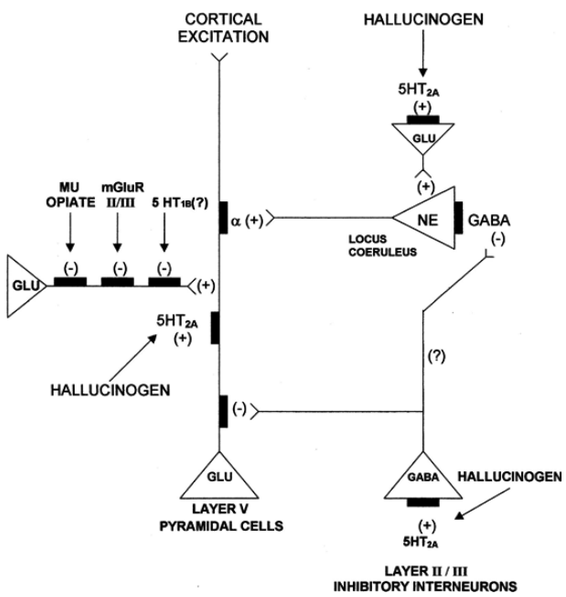


FIGURE 108.2. Schematic diagram of putative electrophysiologic mechanism of action of hallucinogenic drugs. Depicted are serotoninergic hallucinogenic inputs at the raphe nuclei and locus ceruleus projecting to the vicinity of apical dendrites of layer V pyramidal cells in the neocortex. Hallucinogens, acting as partial agonists at 5-hydroxytryptamine subtype 2A (5-HT_{2A}) receptors, induce the release of glutamate from excitatory nerve terminals. Also shown are inhibitory modulators of 5-HT_{2A}-induced glutamate release: γ-aminobutyric acid, μ opiate, group II and III metabotropic glutamate, and possibly 5-HT_{1B} receptors. NE, noradrenergic input; alpha$_1$, α_1-adrenergic receptor; mGluR II/III, group II and III metabotropic glutamate receptor; GABA, γ-aminobutyrate. (Modified from Aghajanian GK, Marek GJ. Serotonin and hallucinogens. *Neuropsychopharmacology* 1999;21:16S–23S, with permission.)

MDMA research during the next decade should also yield significant advances. Preclinical studies aimed at determining the mechanism of MDMA-induced 5-HT neurotoxicity may not only increase our understanding of serotonin neuronal function but also provide insight into idiopathic neurodegenerative illnesses and neuronal responses to injury. Long-term studies in nonhuman primates and humans will be essential to learn whether recovery from MDMA-induced 5-HT neurotoxicity can occur (and if so, under what conditions), and will be useful in defining the functional consequences of MDMA-induced neurotoxicity. It may be possible, by using information derived from preclinical studies, to design treatments for persons in whom chronic MDMA-related neuropsychiatric illnesses develop. Increased efforts should be directed toward identifying those at greatest risk for the development of MDMA-related neuropsychiatric illnesses. Finally, cost-effective methods should be devised to detect MDMA-induced neurotoxicity, so as to identify those who may benefit from alternative, science-based treatment strategies.

REFERENCES

1. Johnston LD, O'Malley PMJ, Bachman JG. *National survey results on drug use from the Monitoring the Future Study, 1975–1998*. NIH Publication No 99-4661. Rockville, MD: National Institute on Drug Abuse, 1999.

2. Johnston L, Bachman J, O'Malley P. *The Monitoring the Future national results on adolescent drug use: overview of key findings, 1999*. NIH Publication No 00-4690. Rockville, MD: National Institute on Drug Abuse, 2000.

3. Lassen JF, Lassen NF, Skov J. Consumption of psilocybin-containing hallucinogenic mushrooms by young people [in Danish]. *Ugeskrift for Laeger* 1992;154:2678–2681.

4. Johnston LJ, O'Malley P, Brachman JG. *Monitoring the future: national survey results on drug use, 1975–1999*. Vols. 1 and 2. (NIH Publication Nos. 00–4802 and 00–4803.) Bethesda, Maryland: National Institute on Drug Abuse, 2001.

5. Gaddum JH. Antagonism between lysergic acid diethylamide and 5-hydroxytryptamine. *J Physiol* 1953;121:15.

6. Evarts EV, et al. Some effects of lysergic acid diethylamide and bufotenine on electrical activity in the cat's visual system. *Am J Physiol* 1955;182: 594–598.

7. Hofmann A. Chemical, pharmacological and medical aspects of psychotomimetics. *J Exp Med Sci* 1961;5:31–51.

8. Rinkel M, et al. Experimental schizophrenia-like symptoms. *Am J Psychiatry* 1952;108:572–578.

9. Purpura DP. Electrophysiological analysis of psychotogenic drug action I. *AMA Arch Neurol Psychiatry* 1956;75:122–131.

10. Glennon RA. Classical hallucinogens: an introductory overview. In: Lin GC, Glennon RA, eds. *Hallucinogens: an update*. Rockville, MD: National Institute on Drug Abuse, 1994:4–32.

11. Geyer MA, Braff DL, Swerdlow NR. Startle-response measures of information processing in animals: relevance to schizophrenia. In: Haug M, Whalen RE, eds. *Animal models of human emotion and cognition*. Washington, DC: American Psychiatric Association, 1999:103–116.

12. Freedman DX. Effects of LSD-25 on brain serotonin. *J Pharmacol Exp Ther* 1961;134:160–166.

13. Hirschhorn ID, Winter JC. Mescaline and lysergic acid diethylamide (LSD) as discriminative stimuli. *Psychopharmacologia* 1971;22:64–71.

14. Aghajanian GK, Marek GJ. Serotonin and hallucinogens. *Neuropsychopharmacology* 1999;21:16S–23S.

15. Trulson ME. Dissociations between the effects of hallucinogens on behavior and raphe unit activity in behaving cats. *Pharmacol Biochem Behavior* 1986;24:351–357.

16. Peroutka SJ, Snyder SH. Multiple serotonin receptors: differential binging of [^{3}H]5-hydroxytryptamine, [^{3}H]lysergic acid diethylamide and [^{3}H]spiroperidol. *Mol Pharmacol* 1979;16:687–699.

17. Aghajanian GK. Electrophysiology of serotonin receptor subtypes and signal transduction mechanisms. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995;451–460.

18. Glennon RA, Young R, Rosecrans JA. Antagonism of the stimulus effects of the hallucinogen DOM and the purported serotonin agonist quipazine by 5-HT$_2$ antagonists. *Eur J Pharmacol* 1983;91:189–192.

19. Titeler M, Lyon RA, Glennon RA. Radioligand binding evidence implicates the brain 5-HT$_2$ receptor as a site of action for LSD and phenylisopropylamine hallucinogens. *Psychopharmacology* 1988;94:213–216.

20. Brimblecombe RW, Pinder RM. *Hallucinogenic agents*. Bristol, UK: Wright-Scientechnica, 1975.

21. Pfaff, RC, et al. Lysergamides revisited. In: Lin GC, Glennon RA, eds. *Hallucinogens: an update*. Rockville, MD: National Institutes of Health, 1994:52–73.

22. Shih JC, Chen K, Gallaher TK. Structure and function of serotonin 5-HT_2 receptors. In: Lin GC, Glennon RA, eds. *Hallucinogens: an update*. Rockville, MD: National Institute on Drug Abuse, 1994:284–297.

23. Burris, KD, Breeding M, Sanders-Bush E. (+)Lysergic acid diethylamide, but not its nonhallucinogenic congeners, is a potent serotonin $5HT_{1C}$ receptor agonist. *J Pharmacol Exp Ther* 1991;258:891–896.

24. Chiang C, Aston-Jones G. A 5-hydroxytryptamine$_2$ agonist augments gamma-butyric acid and excitatory amino acid inputs to noradrenergic locus coeruleus neurons. *Neuroscience* 1993;54:409–420.

25. Sheldon PW, Aghajanian GK. Serotonin (5-HT) induces IPSPs in pyramidal layer cells of rat piriform cortex: evidence for the involvement of a 5-HT_2-activated interneuron. *Brain Res* 1990;506:62–69.

26. Aghajanian GK, Marek GJ. Serotonin induces excitatory postsynaptic potentials in apical dendrites of neocortical pyramidal cells. *Neuropharmacology* 1997;36:589–599.

27. Nichols DE. Role of serotonergic neurons and 5-HT receptors in the action of hallucinogens. In: Baumgarten HG, Gothert M, eds. *Serotonergic neurons and 5-HT receptors in the CNS*. Berlin: Springer-Verlag, 1997:563–585.

28. Lamb RJ, Griffiths RR. Self-injection of 3,4-methylenedioxymethamphetamine (MDMA) in the baboon. *Psychopharmacology* 1987;91:268–272.

29. Vollenweider FX, et al. Positron emission tomography and fluorodeoxyglucose studies of metabolic hyperfrontality and psychopathology in the psilocybin model of psychosis. *Neuropsychopharmacology* 1997;16:257–272.

30. Shulgin A, Shulgin A. *Pihkal: A chemical love story*. Berkeley. CA: Transform Press, 1991.

31. Strassman RJ, Qualls CR. Dose–response study of *N,N*-dimethyltryptamine in humans. I. Neuroendocrine, autonomic, and cardiovascular effects. *Arch Gen Psychiatry* 1994;51: 85–97.

32. Abraham HD. Hallucinogen-related disorders. In: Sadock BJ, Sadock VA, eds. *Kaplan and Sadock's comprehensive textbook of psychiatry*, seventh edition. Philadelphia: Lippincott, Williams, & Wilkins, 2000:1015–1024.

33. Elkes C, Elkes J, Mayer-Gross W. Hallucinogenic drugs. *Lancet* 1955;268:719.

34. Braff DL, Geyer MA. Sensorimotor gating and schizophrenia. *Arch Gen Psychiatry* 1990;47:181–188.

35. Abraham HD, Wolf E. Visual function in past users of LSD: psychophysical findings. *J Abnorm Psychol* 1988;97:443–447.

36. Abraham HD, Duffy FH. Stable quantitative EEG difference in post-LSD visual disorder by split-half analysis: evidence for disinhibition. *Psychiatr Res Neuroimaging* 1996;67:173–187.

37. Krill AE, Alpert HJ, Ostfield AM. Effects of a hallucinogenic agent in totally blind subjects. *Arch Ophthalmol* 1963;69:180–185.

38. Abraham HD, Mamen A. LSD-like panic from risperidone in post-LSD visual disorder. *J Clin Psychopharmacol* 1996;16:238–241.

39. Abraham HD, Aldridge AM, Gogia P. The psychopharmacology of hallucinogens. *Neuropsychopharmacology* 1996;14:285–298.

40. Giannini AJ. Phencyclidine. In: Tarter RE, Ammerman RT, Ott PJ, eds. *Handbook of substance abuse: neurobehavioral pharmacology*. New York, Plenum Press, 1998.

41. Johnstone M, Evans V, Baigel S. Sernyl (CI-395) in clinical anesthesia. *Br J Anaesthesia* 1958;31:433–439

42. Olney JW, Newcomer JW, Farber NB. NMDA receptor hypofunction model of schizophrenia. *J Psychiatr Res* 1999;33:523–533.

43. Giannini AJ, Loiselle RH, Malone DM. Treatment of PHP psychosis with haloperidol. *J Toxicol Clin Toxicol* 1985;23:185–189.

44. Albuquerque TX, Aguayo LG, Warnick JE, et al. Interactions of phencyclidine with ion channels of nerve and muscle: behavioral implications. *Fed Proc* 1983;42:2852–2859.

45. Jentsch JD, Tran A, Taylor JR, et al. Prefrontal cortical involvement in phencyclidine-induced activation of the mesolimbic dopamine system: behavioral and neurochemical evidence. *Psychopharmacology* 1998;138:89–95.

46. Giannini AJ, Loiselle RH, Graham BH, et al. Behavioral response to buspirone in cocaine and phencyclidine withdrawal. *J Subst Abuse Treat* 1993;10:523–527.

47. Javitt DC, Frusciante MJ, Zukin SR. Activation-related and activation-independent effects of polyamines on phencyclidine receptor binding within the *N*-methyl-D-aspartate receptor complex. *J Pharmacol Exp Ther* 1994;270:604–613.

48. Giannini AJ, Loiselle RH, Price WA. Antidotal strategies in phencyclidine intoxication. *Int J Psychiatry Med* 1984;4:513–518.

49. Giannini AJ, Loiselle RH, Giannini MC, et al. The dissociatives. *Med Psychiatry* 1987;3:197.

50. McCarron MM. *Phencyclidine intoxication. P 209*. NIDA Research Monograph Series. DHHS Publ No (ADM) 86-1371. Washington, DC: US Government Printing Office, 1986.

51. Aniline O, Pitts FN. Phencyclidine (PCP): a review and perspectives. *Crit Rev Toxicol* 1982;10:1045.

52. Lerner SE, Burns RS. Phencyclidine use among youth: history, epidemiology and acute and chronic intoxication. In: Peterson RC, Stillman RC, eds. *PCP phencyclidine abuse: an appraisal*. NIDA Research Monograph No 21. Rockville, MD: National Institute on Drug Abuse, 1978:66–118.

53. Weibel SL, Wolf HH. Opiate modification of intracranial self-stimulation in the rat. *Pharmacol Biochem Behav* 1979;10:71–78.

54. Cosgrove J, Newell T. Recovery of neuropsychological functions during reduction in use of phencyclidine. *J Clin Psychol* 1991;47:159–169.

55. Olney JW, Labruyere J, Price MT. Pathological changes induced in cerebrocortical neurons by phencyclidine and related drugs. *Science* 1989;244:1360–1362.

56. Gao X, Shirakawa F, Tamminga C. Delayed regional metabolic actions of phencyclidine. *Eur J Pharmacol* 1993;241:7–25.

57. Nakki R, Koistinaho J, Sharp F, et al. Cerebellar toxicity of phencyclidine. *J Neurosci* 1995;15:2097–2108.

58. Griffiths MR, Cooper AJ, Barber DJ, et al. Pharmacological mechanisms mediating phencyclidine-induced apoptosis of striatopallidal neurons: the roles of glutamate, dopamine, acetylcholine and corticosteroids. *Brain Res* 2000;855:1–10.

59. Kim SH, Price MT, Olney JW, et al. Excessive cerebrocortical release of acetylcholine induced by NMDA antagonists is reduced by GABAergic and alpha$_2$-adrenergic agonists. *Mol Psychiatry* 1999;4:344–352.

60. Anderson GM, Braun G, Braun U, et al. *Absolute configuration and psychomimetic phenylalkylamines*. NIDA Research Monograph No 22. Rockville, MD: National Institute on Drug Abuse, 1978.

61. Nichols DE, Hoffman AJ, Oberlender RA, et al. (1986) Derivatives of 1-(1,3-benzodioxole-5-yl)-2-butanamine: representatives of a novel therapeutic class. *J Med Chem* 1986;29:2009–2015.

62. Shulgin AT, Sargent T, Naranjo C. The chemistry and psychopharmacology of nutmeg and of several related phenylisopropylamines. In: Effron D, ed. *Ethnological search for psychoactive drugs*. Washington, DC: US Government Printing Office, 1967:202–214.

63. Greer GR, Tolbert R. A method of conducting therapeutic sessions with MDMA. *J Psychoactive Drugs* 1998;30:371–379.

64. Johnson MP, Hoffman AH, Nichols DE. Effects of the enantiomers of MDA, MDMA and related analogues on [^{3}H]-serotonin and [^{3}H]-dopamine release from superfused rat brain slices. *Eur J Pharmacol* 1986;132:269–276.

65. Rudnick G, Wall SC. The molecular mechanism of “ecstasy” [3,4-methylenedioxymethamphetamine (MDMA)]: serotonin transporters are targets for MDMA-induced serotonin release. *Proc Natl Acad Sci U S A* 1992;89:1817–1821.

66. Steele TD, Nichols DE, Yim GKW. Stereochemical effects of 3,4-methylenedioxymethamphetamine (MDMA) and related amphetamine derivatives on inhibition of uptake of [^{3}H]-monoamines into synaptosomes from different regions of rat brain. *Biochem Pharmacol* 1987;36:2297–2303.

67. Titeler M, Lyon RA, Glennon RA. Radioligand binding evidence implicates the brain 5-HT_2 receptor as a site of action for LSD and central phenylisopropylamine hallucinogens, *Psychopharmacology* 1988;94:213–216.

68. Battaglia G, Brooks BP, Kulsakdinum C, et al. Pharmacologic profile of MDMA (3,4-methylenedioxymethamphetamine) at various brain recognition sites. *Eur J Pharmacol* 1988;149:159–163.

69. Pierce PA, Peroutka SJ. Ring-substituted amphetamine interactions with neurotransmitter receptor binding sites in human cortex. *Neurosci Lett* 1988;95:208–212.

70. Battaglia G, Brooks BP, Kulsakdinum C, et al. Pharmacologic profile of MDMA (3,4-methylenedioxymethamphetamine) at various brain recognition sites. *Eur J Pharmacol* 1988;149:159–163.

71. Gordon CJ, Watkinson WP, O’Callaghan JP, et al. Effects of 3,4-methylenedioxymethamphetamine on autonomic thermoregulatory responses of the rat. *Pharmacol Biochem Behav* 1991;38:339–344.

72. Hardman HF, Haavik CO, Seevers MH. Relationship of the structure of mescaline and seven analogs to toxicity and behavior in five species of laboratory animals. *Toxicol Appl Pharmacol* 1973;25:299–309.

73. Frith CH, Chang LW, Lattin DL, et al. Toxicity of 3,4-methylenedioxymethamphetamine in the dog and rat. *Fundam Appl Toxicol* 1987;9:110–119.

74. Gold LH, Koob GF, Geyer, MA. Stimulant and hallucinogenic profiles of 3,4-methylenedioxymethamphetamine and *N*-ethyl-3,4-methylenedioxyamphetamine in rats. *J Pharmacol Exp Ther* 1988;247:547–555.

75. Glennon RA, Young RY. Further investigations of the discriminative stimulus properties of MDA. *Pharmacol Biochem Behav* 1984;20:501–505.

76. Evans SM, Johanson CE. Discriminative stimulus properties of (+)-3,4-methylenedioxymethamphetamine and (+)-3,4-methylenedioxyamphetamine in pigeons. *Drug Alcohol Depend* 1986;18:159–164.

77. Kamien JB, Johanson CE, Schuster CR, et al. The effects of (+)-methylenedioxymethamphetamine and (+)-methylenedioxyamphetamine in monkeys trained to discriminate (+)-amphetamine from saline. *Drug Alcohol Depend* 1986;18:139–147.

78. Glennon RA. MDMA-like stimulus effects of alpha-ethyltryptamine and the alpha-ethyl homolog of DOM. *Pharmacol Biochem Behav* 1993;46:459–462.

79. Beardsley PM, Balster RL, Harris LS. Self-administration of methylenedioxymethamphetamine (MDMA) by rhesus monkeys, *Drug Alcohol Depend* 1986;18:149–157.

80. Hubner CB, Bird M, Rassnick S, et al. The threshold lowering effects of MDMA (ecstasy) on brain-stimulation reward, *Psychopharmacology* 1988;95:49–51.

81. Downing J. The psychological and physiological effects of MDMA on normal volunteers. *J Psychoactive Drugs* 1986;18: 335–340.

82. Greer G, Tolbert P. Subjective reports of the effects of MDMA in a clinical setting, *J Psychoactive Drugs* 1986;18:319–327.

83. Grob CS, Poland RE, Chang L, et al. Psychobiologic effects of 3,4-methylenedioxymethamphetamine in humans: methodological considerations and preliminary observations. *Behav Brain Res* 1996;73:103–107.

84. Vollenweider FX, Gamma A, Liechti M, et al. Psychological and cardiovascular effects and short-term sequelae of MDMA (“ecstasy”) in MDMA-naïve healthy volunteers. *Neuropsychopharmacology* 1998;19:241–251.

85. Mas M, Farre M, de la Torre R, et al. Cardiovascular and neuroendocrine effects and pharmacokinetics of 3, 4-methylenedioxymethamphetamine in humans. *J Pharmacol Exp Ther* 1999;290:136–145.

86. Nash JF Jr, Meltzer HY, Gudelsky GA. Elevation of serum prolactin and corticosterone concentrations in the rat after the administration of 3,4-methylenedioxymethamphetamine. *J Pharmacol Exp Ther* 1988;245:873–879.

87. Poland RE, Lutchmansingh P, McCracken JT, et al. Abnormal ACTH and prolactin responses to fenfluramine in rats exposed to single and multiple doses of MDMA. *Psychopharmacology* 1997;131:411–419.

88. Gerra G, Zaiovic A, Giucastro G, et al. Serotonergic function after ecstasy in humans. *Int Clin Psychopharmacol* 1998;13:1–9.

89. Gerra G, Zaimovic A, Ferri M, et al. Long-lasting effects of (+/-)3,4-methylenedioxymethamphetamine (ecstasy) on serotonin system function in humans. *Biol Psychiatry* 2000;47:127–136.

90. Poland RE. Diminished corticotropin and enhanced prolactin responses to 8-hydroxy-2(di-*n*-propylamino)tetralin in methylenedioxymethamphetamine pretreated rats. *Neuropharmacology* 1990;29:1099–1101.

91. Series HG, Cowen PJ, Sharp T. *p*-Chloroamphetamine (PCA), 3,4-methylenedioxy-methamphetamine (MDMA) and *d*-fenfluramine pretreatment attenuates *d*-fenfluramine-evoked release of 5-HT *in vivo*. *Psychopharmacology (Berl)* 1994;116:508–514.

92. McCann UD, Eligulashvili V, Mertl M, et al. Altered neuroendocrine and behavioral responses to *m*-chlorophenylpiperazine in 3,4-methylenedioxymethamphetamine (MDMA) users. *Psychopharmacology (Berl)* 1999;147:56–65.

93. Lim HK, Foltz RL. *In vivo* and *in vitro* metabolism of 3,4-(methylenedioxy)methamphetamine in the rat: identification of metabolites using an ion trap detector. *Chem Res Toxicol* 1988;1:370–378.

94. Cho AK, Hiramatsu M, DiStefano EW, et al. Stereochemical differences in the metabolism of 3,4-methylenedioxymethamphetamine *in vivo* and *in vitro*: a pharmacokinetic analysis, *Drug Metab Dispos* 1990;18:686–691.

95. Fitzgerald RL, Blanke RV, Rosecrans JA, et al. Stereochemistry of the metabolism of MDMA to MDA. *Life Sci* 1989;45:295–301.

96. Fallon JK, Kicman AT, Henry JA, et al. Stereospecific analysis and enantiomeric disposition of 3,4-methylenedioxymethamphetamine (ecstasy) in humans. *Clin Chem* 1999;45:1058–1069.

97. de la Torre R, Farre M, Ortuno J, et al. Non-linear pharmacokinetics of MDMA (“ecstasy”) in humans. *Br J Clin Pharmacol* 2000;49:104–109.

98. McCann UD, Slate SO, Ricaurte GA. Adverse reactions with alpha-3,4-methylenedioxymethamphetamine (MDMA, “ecstasy”). *Drug Safety* 1996;15:107–115.

99. McCann UD, Mertl MM, Ricaurte GA. Methylenedioxymethamphetamine (MDMA, “ecstasy”’). In: Tarter RE, Ammerman RT, Ott PJ, eds. *Sourcebook on substance abuse: etiology, methodology, and intervention*. New York: Allyn and Bacon, 1998:567–577.

100. Hughes JC, McCabe M, Evans RJ. Intracranial haemorrhage associated with ingestion of “ecstasy.” *Arch Emerg Med* 1993;10:372–374.

101. Dowling GP, McDonough ET, Bost RO. “Eve” and “ecstasy.” A report of five deaths associated with the use of MDEA and MDMA. *JAMA* 1987;257:1615–1617.

102. O’Connor A, Cluroe A, Couch R, et al. Death from hyponatraemia-induced cerebral oedema associated with MDMA (“ecstasy”) use. *N Z Med J* 1999;112:255–256.

103. Holmes SB, Banerjee AK, Alexander WD. Hyponatraemia and seizures after ecstasy use. *Postgrad Med J* 1999;75:32–33.

104. Brown C, Osterloh J. Multiple complications from recreational ingestion of MDMA (“ecstasy”). *JAMA* 1987;258:780–781.

105. Chadwick IS, Linsley A, Freemont AJ, et al. Ecstasy, 3-4 methylenedioxymethamphetamine (MDMA), a fatality associated with coagulopathy and hyperthermia. *J R Soc Med* 1991;84:371.

106. Walubo A, Seger D. Fatal multi-organ failure after suicidal overdose with MDMA, “ecstasy”: case report and review of the literature. *Hum Exp Toxicol* 1999;18:119–125.

107. Chance, M. Aggregation as a factor influencing the toxicity of sympathomimetic amines in mice. *J Pharmacol Exp Ther* 1946;198:214–219.

108. Jones AL, Simpson KJ. Review article: mechanisms and management of hepatotoxicity in ecstasy (MDMA) and amphetamine intoxications. *Aliment Pharmacol Ther* 1999;13:129–133.

109. Clark AD, Butt N. Ecstasy-induced very severe aplastic anaemia complicated by invasive pulmonary mucormycosis treated with allogeneic peripheral blood progenitor cell transplant. *Clin Lab Haematol* 1997;19:279–281.

110. Whitaker-Azmitia PM, Aronson TA. “Ecstasy” (MDMA)induced pain. *Am J Psychiatry* 1989;146:119.

111. Creighton FJ, Black DL, Hyde CE. “Ecstasy” psychosis and flashbacks. *Br J Psychiatry* 1991;159:713–715.

112. Alciati A, Scaramelli B, Fusi A, et al. Three cases of delirium after “ecstasy” ingestion. *J Psychoactive Drugs* 1999;31:411–419.

113. Crifasi J, Long C. Traffic fatality related to the use of methylenedioxymethamphetamine. *J Forensic Sci* 1996;41:1082–1084.

114. McCann UD, Ricaurte GA. MDMA (“ecstasy”) and panic disorder: induction by a single dose. *Biol Psychiatry* 1992;32:950–953.

115. McGuire P, Fahy T. Chronic paranoid psychosis after misuse of MDMA (“ecstasy”). *Br Med J* 1991;302:607.

116. Schifano F, Magni G. MDMA (“ecstasy”) abuse: psychopathological features and craving for chocolate: a case series. *Soc Biol Psychiatry* 1994;36:763–767.

117. McCann UD, Ricaurte GA. Lasting neuropsychiatric sequelae of α-methylenedioxymethamphetamine (“ecstasy”) in recreational users. *J Clin Psychopharmacol* 1991;11:302–305.

118. Ricaurte G, Bryan G, Strauss L, et al. Hallucinogenic amphetamine selectively destroys brain serotonin nerve terminals. *Science* 1985;229:986–988.

119. Battaglia G, Yeh SY, O’Hearn E, et al. 3,4-Methylenedioxymethamphetamine and 3,4-methylenedioxyamphetamine destroy serotonin terminals in rat brain: quantification of neurodegeneration by measurement of [^{3}H]paroxetine-labeled serotonin uptake sites. *J Pharmacol Exp Ther* 1987;242:911–916.

120. Commins DL, Vosmer G, Virus R, et al. Biochemical and histological evidence that methylenedioxymethylamphetamine (MDMA) is toxic to neurons in the rat brain.*J Pharmacol Exp Ther* 1987;241:338–345.

121. Stone DM, Hanson GL, Gibb JW. Differences in the central serotoninergic effects of 3,4-methylenedioxymethamphetamine (MDMA) in mice and rats. *Neuropharmacology* 1987;26:1657–1661.

122. Ricaurte GA, DeLanney LE, Irwin I, et al. Toxic effects of MDMA on central serotoninergic neurons in the primate: importance of route and frequency of drug administration. *Brain Res* 1988;446:165–168.

123. Hatzidimitriou G, McCann UD, Ricaurte GA. Aberrant serotonin innervation in the forebrain of monkeys exposed to MDMA seven years previously: Factors influencing abnormal recovery. *J Neurosci* 1999;19:5096–5107.

124. Ricaurte GA, Katz JL, Martello MB. Lasting effects of alpha-3,4-methylenedioxymethamphetamine on central serotoninergic neurons in nonhuman primates. *J Pharmacol Exp Ther* 1992;261:616–622.

125. Schmidt CJ, Wu L, Lovenberg W. Methylenedioxymethamphetamine: a potentially neurotoxic amphetamine analog. *Eur J Pharmacol* 1986;124:175–178.

126. Schmidt CJ, Taylor VL. Depression of rat brain tryptophan hydroxylase activity following the acute administration of methylenedioxymethamphetamine. *Biochem Pharmacol* 1987;36:4095–4102.

127. O’Hearn EG, Battaglia G, De Souza EB, et al. Methylenedioxyamphetamine (MDA) and methylenedioxymethamphetamine (MDMA) cause selective ablation of serotonergic axon terminals in forebrain: immunocytochemical evidence for neurotoxicity. *J Neurosci* 1988;8:2788–2803.

128. Ricaurte GA, Delanney LE, Wiener SG, et al. 5-Hydroxyindoleacetic acid in cerebrospinal fluid reflects serotoninergic damage induced by 3,4-methylenedioxymethamphetamine in CNS of nonhuman primates. *Brain Res* 1988;474:359–363.

129. Scheffel U, Szabo Z, Mathews WB, et al. *In vivo* detection of short- and long-term MDMA neurotoxicity—a positron emission tomography study in the living baboon brain. *Synapse* 1998;29:183–192.

130. McCann UD, Ridenour A, Shaham Y, et al. Brain serotoninergic neurotoxicity after MDMA (“ecstasy”): a controlled study in humans. *Neuropsychopharmacology* 1994;10:129.

131. Krystal JH, Price LH, Opsahl C, et al. Chronic 3,4- methylenedioxymethamphetamine (MDMA) use: effects on mood and neuropsychological function. *Am J Drug Alcohol Abuse* 1992;18:331–341.

132. Parrot AC, Lees A, Granham NJ, et al. Cognitive performance in recreational users of MDMA or “ecstasy”: evidence for memory deficits. *J Psychopharmacol* 1998;12:79–83.

133. Parrott AC, Lasky J. Ecstasy (MDMA) effects upon mood and cognition: before, during and after a Saturday night dance. *Psychopharmacology* 1998;139:261–268.

134. Bolla KI, McCann UD, Ricaurte GA. Impaired memory function in abstinent MDMA (“ecstasy”) users. *Neurology* 1998;51:1532–1537.

135. Morgan MM. Memory deficits associated with recreational use of “ecstasy” (MDMA). *Psychopharmacology* 1999;141:30–36.

136. Ricaurte GA, McCann UD. Experimental studies on 3,4-methylenedioxymethamphetamine (MDMA, “Ecstasy”) and its potential to damage brain serotonin neurons. *Neurotoxicity Res* 2001;3:85–99.

137. Brown GL, Goodwin FK, Ballenger JC, et al. Aggression in humans correlates with cerebrospinal fluid metabolites. *Psychiatry Res* 1979;1:131–139.

109

Epidemiology of Drug Dependence

James C. Anthony

James C. Anthony: Johns Hopkins University, School Of Hygeine And Public Health, Baltimore, Maryland.

If one judges solely by the cumulative table of contents of the official journal of the American College of Neuropsychopharmacology, *Neuropsychopharmacology*, the intersection where epidemiology meets neuropsychopharmacology is quite empty. When one looks elsewhere, the traffic becomes visible, with a scope that encompasses topics such as the characteristics of incarcerated drug users, adolescent drug use, epidemics of drug taking, the "overmedication" of American society, and postmarketing surveillance of new neuropsychopharmacologic drug products (e.g., see refs. 1 , 2 , 3 and 4).

One of the possible reasons that epidemiologic research articles have been seen rarely in *Neuropsychopharmacology* is the generally nonexperimental and observational character of the studies. In this respect, epidemiology shares features of astronomy, geology, ecology, and other sciences in which the evidence comes mainly from field studies, without the benefit of maximal control over experimental error and sometimes with heavy reliance on retrospection. This reliance on retrospection has been a source of considerable criticism, and a countervailing trend has developed toward prospective, longitudinal, and even randomized experimental studies in epidemiology. Nonetheless, the scale and environment of epidemiologic research introduce constraints not seen elsewhere in human biology and the biomedical sciences, even when epidemiology harnesses the power of a randomized trial.

Against a such a background, the primary goal of this chapter is to describe the focus of epidemiologic research in drug dependence and to elucidate the contributions that such research can make when connected with that in other areas of neuropsychopharmacology. A secondary goal is to aid neuropsychopharmacologists who may wish to know more about what can be learned by collaborating with epidemiologists.

For focus, this overview concentrates on the clinical syndromes of drug dependence, as defined in recent diagnostic and statistical manuals of the American Psychiatric Association (e.g., DSM-III, DSM-IIIR, and DSM-IV) and the tenth revision of the World Health Organization International Classification of Disease (ICD-10). The chapter is organized in relation to five main rubrics or subheadings for the subject matter of epidemiologic research. Under each rubric is included a selection of recent examples of epidemiologic evidence regarding drug dependence.

THE FIVE MAIN RUBRICS OF EPIDEMIOLOGY

Part of "109 - Epidemiology of Drug Dependence "

Morris (5) described seven "uses" of epidemiology, which can be simplified in relation to the five "rubrics" or main subheadings of epidemiology listed in Table 109.1 . These five rubrics offer an easily remembered way of organizing the central research questions and subject matter of this branch of biomedical science (6). Each rubric corresponds to a research question, and each research question demonstrates the substantive research focus of epidemiology and creates an opportunity to explain some of the concepts, principles, and methods that are used to make progress in epidemiology.

The Rubrics	General Issues	Research Questions Associated with Each Rubric
1. Quantity (Prevalence and incidence)	How many?	"In the population, how many are becoming new cases of drug dependence?" "How many already have become drug-dependent?"
2. Location (variation)	Where?	"In the population, does the frequency or occurrence of drug dependence cases vary from place to place, from time to time, or in relation to individual-level characteristics, conditions, or processes?"
3. Causes (Etiology)	Why?	"In the population, why do some people become drug-dependent while others are spared?"
4. Mechanisms	How?	"What sequences of circumstances, conditions, and processes lead to the development of drug dependence?"
5. Prevention and Control	What can be done?	"What can be done to prevent, reduce, or ameliorate the adverse impact of drug dependence?"

Adapted from Anthony JC, Van Etten ML. Epidemiology and its rubrics. In: Bellock AS, Hersen M, eds. *Comprehensive clinical psychology*, first ed. New York: Pergamon, 1998, with permission.

TABLE 109.1. THE MAIN RUBRICS AND RESEARCH QUESTIONS OF EPIDEMIOLOGY, AS APPLIED TO CLINICAL SYNDROMES OF DRUG DEPENDENCE

To some extent, the progress of an individual epidemiologic investigator can be plotted in relation to a mastery of the concepts, principles, and methods that fall under each rubric listed in Table 109.1 . In time, it may prove useful to plot the progress of epidemiology over generations of scientists in terms of the relative balance of attention to the more advanced rubrics. To some extent, progress may be represented by increased attention to issues addressed under the last three rubrics: causal inference, causal mechanisms, and means of prevention and control. As progress is made in future generations, the attention given to estimating how many people are affected and describing how cases are distributed within a population, from place to place or during successive seasons or years, may be correspondingly reduced.

RUBRIC 1, QUANTITY: "IN THE POPULATION, HOW MANY ARE BECOMING CASES?"

Part of "109 - Epidemiology of Drug Dependence "

Concepts and History

The first and most basic of the rubrics of epidemiology involves quantification of the disease burden. Generally, in epidemiologic research on disease states or health events, the main research questions under the rubric of quantity are these: “In the population of interest, how many people are affected?” and “How many people are becoming affected?” Expressed as a proportion of the total population size, the first question concerns the *prevalence* of the condition. Expressed as a rate, the second question concerns the *incidence* of the condition.

As a concept at the level of individuals within a population, the prevalence of a disease can be discriminated from its incidence. Prevalence relates to “an individual’s probability of *being* a case” at some point in time or during a specified interval, whereas incidence concerns “the individual’s probability or risk of *becoming* a case for the first time.” Accordingly, an *incident case* is one that has just become a case (6).

Examples of Epidemiologic Evidence under the Rubric of Quantity

Preclinical research describes a broad range of species that self-administer psychoactive drugs, sometimes to a point of maladaptation and self-harm. These studies have also demonstrated substantial within-species individual differences in predisposition to initiate or sustain drug-taking behavior. Clinical studies under controlled laboratory conditions have clarified that drug self-administration can be shaped by manipulating the profiles of available reinforcers, and by increasing the availability of nondrug reinforcers. Nevertheless, these laboratory studies have not been able to characterize the likelihood of becoming drug-dependent in free-living human populations. At the group level, with population-averaged estimates, this task has been accomplished by means of epidemiologic research in the community. Consider the group of internationally regulated, controlled drugs such as cannabis, cocaine, and heroin, and consider a clinical syndrome defined by the co-occurrence of sustained use of one or more of these drugs with features such as tolerance or withdrawal, with or without signs and symptoms of secondary harm (e.g., loss of a job, recurrent infection or abscess, drug overdose), as encompassed by the DSM-III concept of “psychoactive drug use disorders.” The first research to estimate the risk of becoming a DSMdefined case of “drug use disorder” was a coordinated set of prospective follow-up studies conducted as part of the National Institute of Mental Health Epidemiologic Catchment Area Program. Case ascertainment was via the diagnostic interview schedule method. Based on field survey evidence from these prospective studies of community-dwelling adults, most never treated for drug problems and studied between 1980 and 1985, the risk for becoming a case of “drug use disorder” was estimated at 1.1% per year for a community-dwelling adult in the United States (standard error, 0.4%). In other words, of the literally thousands of adults who did not have drug use disorder at the start of the follow-up interval, drug dependence or a related drug use disorder developed during the 1-year follow-up interval in just over 1% (7).

Ethanol was treated as a separate drug, with “alcohol use disorder” defined in terms of sustained use, tolerance or withdrawal, and secondary harms. Based on the Epidemiologic Catchment Area evidence, for a community-dwelling adult in the United States, the risk for becoming a case of alcohol use disorder was estimated at 1.8% per year (standard error, 0.4%), a risk some 70% greater than that for the development of dependence or a related disorder involving an internationally regulated psychoactive drug (7).

Roughly 10 years after the Epidemiology Catchment Area field studies, the National Comorbidity Survey provided new epidemiologic evidence to complement these estimates of the risk for becoming a case of drug use disorder. Although entirely retrospective and cross-sectional in character and lacking the prospective features of the Epidemiologic Catchment Area studies, the National Comorbidity Survey produced useful information necessary to estimate how many users of various classes of drugs had acquired a clinical syndrome of drug dependence, with the syndrome defined and made operational in relation to the DSM-IIIR criteria (8). Based on its nationally representative sample of community-dwelling Americans between 15 and 54 years of age in the early 1990s, the National Comorbidity Survey estimated how many persons had started taking each of several different drugs (e.g., alcohol, cannabis, cocaine), and also how many of them had become dependent on each drug (i.e., alcohol dependence, cannabis dependence, cocaine dependence). On this basis, it was possible to derive a population-average estimate for each drug; once someone had started taking a drug, how likely was it that he or she would have become drug-dependent?

From epidemiologic data derived retrospectively and cross-sectionally in the National Comorbidity Survey, it was determined that for persons who had consumed tobacco on at least once occasion, the probability of having become tobacco-dependent was an estimated 33%. Among persons who had consumed heroin, DSM-IIIR heroin dependence had developed in about 23% (standard error, 5.6%). Among those who had taken cocaine, cocaine dependence had developed in an estimated 16% to 17% (standard error, 1.5%), a value not too distant from that observed for alcohol dependence, 15% (standard error, 0.7%) (8).

The estimated probability that a clinical syndrome of dependence had developed was somewhat lower for users of cannabis, the psychostimulant drugs, anxiolytic–sedative–hypnotic drugs, hallucinogens such as lysergic acid diethylamide (LSD), and inhalant drugs (e.g., glue, gasoline). For example, among stimulant users, the estimate was about 1 in 9 (11%; standard error, 1.6%). For cannabis users, it was 1 in 11 (9%; standard error, 0.7%). Figure 109.1 shows these and other epidemiologic estimates based on the National Comorbidity Survey data (8).

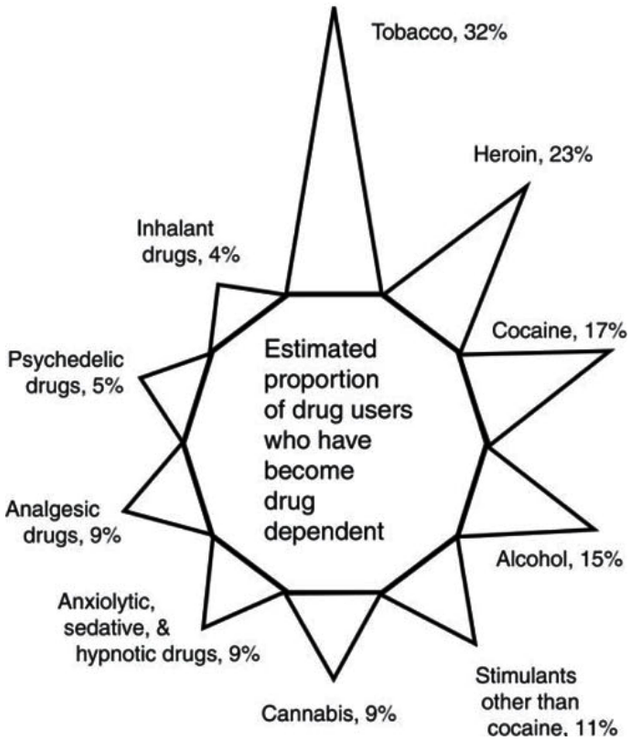


FIGURE 109.1. Estimated probability of drug dependence among drug users, by drug group. (From Anthony JC, Warner LA, Kessler RC. Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalents: basic findings from the National Comorbidity Survey. *Exp Clin Psychopharmacol* 1994;2:244–268, with permission.)

The interpretation of epidemiologic estimates of this type can be tricky. These estimates certainly do not reflect which drugs are associated with a greater potential for dependence than others. In the community at large, exogenous factors, such as the relative availability of a drug (e.g., tobacco vs. cocaine), influence whether drug dependence has a chance to develop once drug use is initiated. In addition, some drug users do not survive from the time of first use to the time of field survey assessment, either dying or disappearing from the sampling frame of the epidemiologic survey before a diagnostic assessment can be completed (see ref. 8). Despite limitations such as these, estimates of this type draw attention to the variability in response to drugs such as nicotine and cocaine, even when laboratory studies demonstrate the robust reinforcing functions served by these drugs. In counterpoint, the laboratory studies demonstrate more limited reinforcing functions served by cannabis and LSD. Nonetheless, DSM-IIIR dependence syndromes appear to have developed in a substantial proportion of alcohol, cannabis, and hallucinogen users (Fig. 109.1).

A slightly different, and more complex, epidemiologic estimate has been derived by dividing the number of currently dependent drug users by the number of currently active drug users (see refs. 9 , 10 , 11 , 12 , 13 , 14 , 15 and 16). The complexity starts in estimating the numerator of the ratio; here, it is necessary to mix the probability of becoming dependent with the probability of continuing to be dependent. Complexity is sustained in estimating the denominator of the ratio. To focus on currently active drug users, it is necessary to mix the probability of starting to use the drug with the probability of continuing to use the drug. An additional complexity enters the picture because drug dependence, as a process, becomes one of the determinants of whether a person continues to use a drug once drug use has been initiated. Hence, the force of persisting drug dependence is exerted not only in the numerator of this ratio but also in its denominator. In

consequence, the resulting estimate cannot be interpreted as a risk for becoming dependent, much less as an indication of relative dependence liability. At best, this estimated ratio reflects the proportion of active drug users who may, in theory, require drug dependence treatment services—that is, it is an indicator of burden. This kind of statistic may be helpful in planning services. Its utility in etiologic studies is compromised by its complexity.

Table 109.2 presents the most recently published drug-specific estimates for the proportion of active drug users who have currently active drug dependence, based on the 1998 National Household Survey on Drug Abuse in the United States (14). Reading the table, one can see that 709 recently active cocaine users were included in the nationally representative survey sample of community-dwelling respondents ages 12 years and older. According to the population estimates, among active cocaine users in the study population, 38% reported at least one of seven active clinical signs or symptoms of cocaine dependence, and an estimated 18% reported at least three active clinical features. Applied in an estimate of burden in the general population, these values indicate that about 0.7% of the study population have a cocaine-related problem and about 0.3% have three or more cocaine-related problems, perhaps meriting treatment or intervention services. By comparison, the corresponding estimates for cannabis, based on 3,444 active cannabis users in the sample, indicate that 42% of active cannabis users reported at least one cannabis dependence problem and 17% reported three or more clinical features of cannabis dependence. In terms of population burden, an estimated 3.6% of the study population have an active cannabis problem and an estimated 1.5% have three or more active features of cannabis dependence. Values for alcohol and for tobacco cigarettes are included in Table 109.2 for comparison with the values for cocaine and cannabis.

Drug or Drug Group	Number of Active Drug Users in the Sample	Estimated Proportion with One or More Clinical Features of Active Drug Dependence (%)	Estimated Proportion with Three or More Clinical Features of Active Drug Dependence (%)
Cocaine	709	38	18
Cannabis	3,444	42	17
Alcohol	14,596	23	8
Tobacco	8,187	60	34

Data from *National household survey on drug abuse: main findings, 1998.* DHHS publication No. (SMA) 00-3381. Rockville, MD: Department of Health and Human Services, Substance abuse and Mental Health Services administration, 2000.

TABLE 109.2. ESTIMATED PROPORTION OF ACTIVE DRUG USERS WHO REPORT FEATURES OF ACTIVE DRUG DEPENDENCE

An increasing number of epidemiologic studies have started to produce estimates of this type, helping to quantify the number of affected cases in various parts of the world and for selected subgroups of the population, such as young adults. For example, Grant (15 ,16) estimated that alcohol dependence developed in about 20% of drinkers, that drug dependence developed in about 19% of persons initiating illicit drug use, and that 16% of active illicit drug users were dependent on illicit drugs. In addition to other recent U.S. survey estimates for the number of active dependence cases among active drug users (10 ,11 ,13 ,14 ,17), estimates have now been made for national populations or subpopulations in Australia, the United Kingdom, Germany, and other countries (18 ,19 and 20).

Against the background of rapidly accumulating prevalence estimates based on cross-sectional epidemiologic surveys, prospective studies and incidence estimates for the drug dependence syndromes have progressed much more slowly. Although prospective studies are much more difficult to complete, they cannot be omitted if we are to understand the force of drug-related morbidity, and distinguish the separate conditions and processes that promote the initiation of drug dependence, as distinct from the conditions and processes that sustain drug dependence once the syndrome has started. In this respect, it is unfortunate that the Epidemiologic Catchment Area estimates, now more than 15 years old, are currently our most authoritative values for the risk for the development of alcohol or other drug dependence in the U.S. adult population (7). Elsewhere in the world, prospectively gathered data on the incidence of clinically defined syndromes of alcohol or other drug dependence (21), sometimes obtained with rigorous methods in quite isolated populations (e.g., 22), are very limited.

Much of the postmarketing surveillance of a population’s actual experience with newly distributed medicines falls under the first rubric of epidemiology. Recent efforts to monitor the abuse potential of tramadol (Ultram) demonstrate the utility of epidemiologic concepts, principles, and methods at the intersection of epidemiology with neuropsychopharmacology (23).

Before we leave the rubric of quantity, it may be useful to note that several generations of epidemiologically oriented

scientists have attempted to estimate the number of “hard-core” drug users in the United States by extrapolating from the number of cases seen in treatment, law enforcement, or other facilities. In the earliest reports, the approach involved guessing the number of untreated or nonincarcerated drug users for every case registered in treatment or known by law enforcement authorities. Since the 1960s, sophisticated mathematical estimation procedures have been used, with advanced statistical treatments such as projections from truncated Poisson distributions and capture–recapture methods (see ref. 24).

This work is at the margin of the scientific enterprise. It may be understood best for its enduring political popularity. As a source of authoritative scientific evidence, it is of dubious value and based on assumptions that are not well tested and may never be testable. In a brief, pithy article, Newman and Cates (25) made this point 30 years ago, quoting from the classic study by Terry and Pellens (26): “As a matter of fact, it is not necessary to know the exact number of users or even the minimal extent, to realize that there are a large number [of addicts] and that the problem is serious.” Newman and Cates also summarized an observation made by the very talented epidemiologist Leon G. Hunt, whose work is mentioned in the next section. Hunt was quoted as saying, “The question is not whether there are three or four million [addicts], but that the number is several million rather than only several hundred thousand” (25).

The view espoused by Newman and Cates (25) actually was much more harsh. They wrote, “The great disparity of the findings of studies using different methods of enumeration generally ensures that data will be found to support any position, and contradictory information is simply ignored.” They concluded, “Objectives of studies that are intended to measure the incidence and prevalence of addiction must be reassessed in terms of experience. It is necessary to ask candidly what impact such research has had in the past and to question the premise that knowledge, for its own sake, is sufficient justification [to undertake these studies.]” This sentiment is especially appropriate in an era of dramatically increased investment by the U.S. government in surveys to estimate the number of active drug users in populations at the state level, with sample sizes for the U.S. National Household Survey on Drug Abuse growing from under 10,000 respondents per year in the early 1990s to more than 70,000 respondents per year in the early twenty-first century.

RUBRIC 2, LOCATION (VARIATION): “IN THE POPULATION, IS THERE ANY VARIATION IN THE OCCURRENCE OF CASES?”

Part of "109 - Epidemiology of Drug Dependence "

Concepts and History

The best epidemiologic studies to quantify the occurrence of drug dependence also have had a more general purpose of studying variations in prevalence or incidence in relation to characteristics of place (e.g., geographic variation), time (e.g., from year to year), or person (e.g., male vs. female drug users). Often, the analyses to disclose variation are not intended to produce links in a chain of causal inferences. Rather, the purpose of these analyses is description, as in Fig. 109.2 , or they may be a necessary step of clarifying variation before anyone undertakes a more probing causal analysis or new investigation (6).

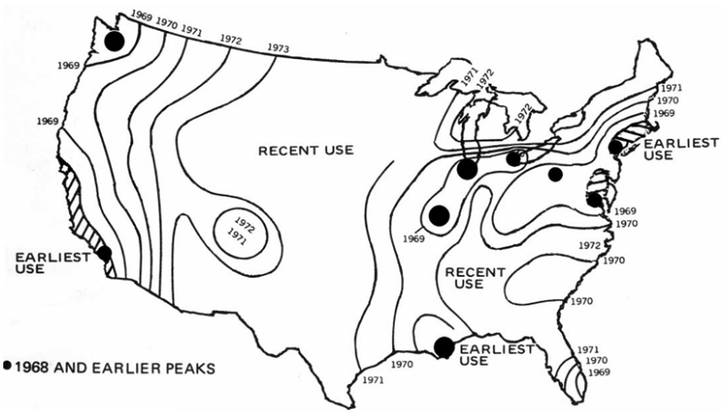


FIGURE 109.2. Retrospectively constructed geographic spread of the epidemic of heroin use in the United States during the Vietnam War era. Data from age at onset of cases admitted to treatment facilities. (From Greene MH, Kozel NJ, Hunt LG, et al. An assessment of the diffusion of heroin abuse to medium-sized American cities. Washington, DC: Special Action Office for Drug Abuse Prevention, 1974, with permission.)

Examples of Epidemiologic Evidence under the Rubric of Location

Two of the most robust findings from epidemiologic studies on the location of drug dependence cases within population subgroups are a male excess and an excess in the age group 15 to 44 years old, disclosed by both prevalence differences and relative risk estimates (see ref. 7). To be sure, the male excess in the occurrence of drug dependence can be contradicted with certain evidence involving some specific drugs, such as those in the group of anxiolytic, sedative, and hypnotic medicines. In addition, in some places, the use of opium derivatives is commonplace in persons in the later years of middle age and among the elderly, and several studies have noted a slight upturn in the risk for alcohol dependence during the last decades of life, at least among men (27). Nonetheless, these exceptions help prove the more general rule.

A recent intriguing discovery about male–female differences in drug use within the United States is that a male excess is found at the earliest stages of drug involvement; boys are more likely than girls to be exposed to opportunities to try illicit drugs. However, once presented with the opportunity, they are equally likely to try a drug (28). Furthermore, once drug use has started, women are almost as likely as men to become drug-dependent (12.6% vs. 16.4%), alcohol and cannabis being two noteworthy exceptions (8).

When cases are located in relation to time, the past 35 years have seen a marked increase in the prevalence of illicit drug use, mainly between 1965 and 1980, as illustrated for the United States in Table 109.3 . Analyses of retrospective data from the Epidemiologic Catchment Area studies and the National Comorbidity Survey have highlighted corresponding differences in the experiences of successive birth cohorts born during the first two-thirds of the twentieth century. According to this evidence, the risk for drug dependence has been markedly greater in persons born since World War II than in prior birth cohorts (17 ,29 ,30 and 31).

TABLE 109.3. ESTIMATED PREVALENCE OF ILLICIT DRUG USE IN THE UNITED STATES

Survey Year	Number of Survey Respondents	Estimated Proportion with a History of Illicit Drug Use (%)	Estimated Prevalence of Recent Illicit Drug Use (%)
1971–72	3,760[a]	15–22[a]	N/A[a]
1979	7,224	31.3	17.5
1985	8,021	34.4	16.3
1991	32,594	34.1	11.1
1992	28,832	33.3	9.7
1993	26,489	34.2	10.3
1994	17,809	34.4	10.8
1995	17,747	34.2	10.7
1996	18,269	34.8	10.8
1997	24,505	35.6	11.2
1998	25,500	35.8	10.6

[a]Data from the *National household servey on drug abuse: main findings, 1998.* DHHS publication No. (SMA) 00-3381. Rockville, MD: Department of Health and Human Services, substance abuse and Mental Health Services administration, 2000.

The location of cases in relation to geography also has been scrutinized in descriptive epidemiologic studies. Figure 109.2 provides a cartoon summary of an apparently epidemic spread of heroin use in the United States during the Vietnam War era, based on the analyses of Hunt (24) of heroin-dependent persons entering treatment and their age at onset of heroin use. Figure 109.3 pertains to the more recent outbreaks of cocaine involvement among young people

living in Chile and shows a substantially greater prevalence of coca paste (“pasta base”) smoking in areas in the north of the country, near the borders with coca-producing countries, than in the south of the country (32).

FIGURE 109.3. Estimated geographic distribution of coca paste smoking among Chilean youth, 1999. (From Dormitzer C, Caris L, Anthony JC. Parental attention and risk of coca paste smoking in Chile: preliminary data from the 1999 national school survey in Chile. Rockville, MD: Department of Health and Human Services, 2000, with permission.)

What is common to all these epidemiologic observations is that they describe, but do not explain or account for, the observed variation. We know that men have been more likely to become drug-dependent than women, but the studies that produced solid evidence on this variation have not helped us explain why this is so. The same is true for the epidemiologic observations of birth cohort differences, which prompt speculation about the greater availability of illicit drugs but then beg the question of what prompted the greater availability. The patterns of epidemic spread of heroin use in the United States remain unexplained, and it is possible that the exclusive attention to treated cases in the study of Hunt (24) may have produced a biased impression of the temporal sequencing of spread through different areas. Finally, at one level, the north–south pattern of coca paste smoking in Chile may be interpreted as a manifestation of proximity to the coca-producing countries, but no probing analysis has confirmed the impression that coca paste is substantially more available in the north. More probing epidemiologic analyses are required to confirm the impression left by initial descriptive observations of this type.

RUBRIC 3, CAUSES: “IN THE POPULATION, WHY DO SOME BECOME AFFECTED WHILE OTHERS ARE SPARED?”

Part of "109 - Epidemiology of Drug Dependence "

Concepts and History

What differentiates the rubric of “causes” from the rubric of “location” is the degree to which the analysis is oriented toward explaining and accounting for the observed phenomena, rather than merely describing the patterns of occurrence. To the extent that the search for causes can lead us toward more effective intervention maneuvers, work under this rubric merits a special status; many regard this search as one of the highest callings of epidemiology (33). Nonetheless, numerous examples show that epidemiologic research can have a considerable impact on the health of a population even before the search for causes is complete. John Snow’s effective demonstration that proper water sanitation can reduce or prevent outbreaks of cholera anticipated Robert Koch’s identification of *Vibrio cholerae* by several decades. Epidemiologic evidence plotting an offspring’s risk for Down syndrome by age of the mother at the time of delivery created one pathway toward effective prevention of trisomy 21 and associated conditions, although we still do not know the causes of the trisomies. HIV prevention efforts directed toward the unsafe sex practices of gay men in the United States helped to change the dynamics of the HIV/AIDS epidemic in the early 1980s, when many believed AIDS to be caused by inhalant drug use (“poppers”) and before isolation and identification of the AIDS-causing virus (6).

In some instances, weak links in the chains of disease causation have been spotted by epidemiologists working with basic quantitative methods, such as cross-tabulation or plotting of incidence estimates, as was done for maternal age and risk for Down syndrome. Complex diseases and conditions such as the drug dependence syndromes do not yield so readily; what might seem to be an apparently simple “chain” of causation actually turns out to be a complex “web” of causation of multifactorial origin.

As in the other sciences allied with neuropsychopharmacology, epidemiology sometimes can turn to the power of randomized, controlled trials and multiple replications for definitive evidence of the web of causation. However, a great many of the important questions in the intersection between epidemiology and neuropsychopharmacology cannot be answered with randomized, controlled trials; in some instances of “natural experiments,” the concept of replication leaves much to be desired.

With respect to a “natural experiment” that may never be repeated, we have the experience of members of the U.S. Armed Forces who served in Vietnam. Virtually all Vietnam veterans were exposed to the opportunity to try heroin within a span of a few months in-country; many (but not all) tried heroin when the opportunity arose (34). When diagnostic interview schedule field study methods were used to assess a large representative sample of Vietnam returnees, almost 20% of the study sample qualified as cases of active heroin dependence during the tour of duty. Nevertheless, no more than a fraction continued to use heroin or remained heroin-dependent once they returned stateside to home. O'Brien et al. (35) have suggested that heroin availability had much to do with this situation. Studying Vietnam veterans who came home to urban areas known for heroin availability, they found substantially higher fractions returning to heroin use. Nonetheless, even in this study sample, a great many heroin users did not come back from Vietnam to the United States and return to heroin dependence once they had settled in urban areas. These results challenge conventional notions of any inherent addictive quality, abuse potential, or dependence liability of heroin as a chemical substance. During an era of discovery that genetic predispositions are prominent among causes of drug dependence, these “subject as own control” results from once-in-a-lifetime natural experiments demonstrate that environmental contingencies also are important (36 ,37).

The Vietnam era research also highlights the necessarily observational character of much epidemiologic work; no responsible investigator would undertake a randomized, controlled trial of exposure to heroin in otherwise drug-naïve young adults. Recent evidence on the early onset of drug use raises similar issues. Since the 1960s, epidemiologic evidence from observational studies has accumulated about

age at first illicit drug use and the subsequent risk for drug problems, including drug dependence, as illustrated in Fig. 109.4 . Very simple cross-tabulations of study data were enough to bring this association to light (38 ,39 ,40 ,41 and 42).

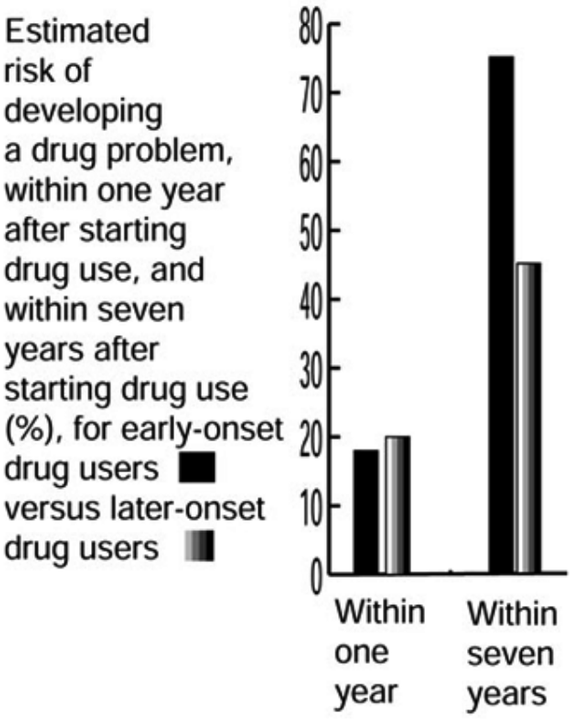


FIGURE 109.4. Estimated risk for the development of a drug problem 1 year after the start of illicit drug use and 7 years after the start of illicit drug use, by age at onset of illicit drug use. Data from National Institute of Mental Health Epidemiologic Catchment Area Program, 1980–1984. (From Anthony JC, Petronis KR. Early-onset drug use and risk of later drug problems. *Drug Alcohol Depend* 1995;40:9–15, with permission.)

Many observers have been convinced by this cross-tabular evidence and have inferred that drug dependence can be prevented by delaying the onset of illicit drug use. Nonetheless, some observers have noted that persons who start using drugs at an early age have more time to experience the hazards of drug use (38) or have other characteristics that make it seem that early age at onset is to blame for the excess when alternative explanations are more plausible (43 ,44). In addition, basic cross-tabulations typically cannot rule out the possibility of an underlying common predisposition, laid down at conception or later in development, that manifests itself not only in an earlier age at onset of drug use but also in a greater risk for becoming drug-dependent. If some predisposition regulates both age at onset and risk for drug dependence (e.g., something linked to a predisposition toward anxiety disorders or antisocial behavior), then efforts to delay age at onset of drug use may have a limited effect on risk for drug dependence.

In this type of context, a direct randomized, controlled trial is out of the question; no one would deliberately expose children to illicit drugs for the sake of experimental confirmation of the observational evidence. Instead, epidemiologists have turned to alternative approaches. First, the life table method developed by Edmund Halley (of comet fame) was used to clarify that the risk for drug problems is not simply a consequence of the fact that early-onset drug users have more time to experience drug problems (Fig. 109.4). Then, “survival analyses” were completed for different strata of the population (e.g., males vs. females) to rule out some of the worrisome predispositions that might lead to a distorted and imperfect view of causal relationships (38). Whereas a randomized, controlled trial may call on randomization to bring distorting predispositions into balance, epidemiologists seek an approximation of this balance via stratified analyses of observational study data or by “matching.” Just as a clinical researcher may match subjects by age or sex, epidemiologists create more homogeneity within “risk sets” by matching subjects on measured variables thought to represent confounding predispositions. Some of these confounding predispositions are genetic and can be matched within monozygotic twin pairs, or they are environmental and can be matched within neighborhoods of residence (6). Second, over and above stratification and matching, multiple regression methods and other forms of the generalized linear model have been used to make statistical adjustments for an array of suspected confounding variables, too many to control via matching or stratification. Finally, indirect randomized, controlled trials are being undertaken in a final push to challenge the belief that the early onset of drug use causes drug dependence. These indirect trials involve the random allocation of preventive interventions intended to delay the onset of drug use; subsequent post-intervention follow-up tests whether delayed onset of drug use is followed by a reduced risk for drug dependence.

In summary, whereas trialists may use a direct randomized, controlled trial to probe suspected causal relationships in the setting of a laboratory or experimental clinic, in the context of population research, epidemiology turns to tools such as stratification, matching, statistical modeling, and the indirect randomized, controlled trial, in which the causal factor is a proximal target for intervention and the condition to be prevented is a more distal outcome (6). Of course, many trialists also make use of stratification, matching, and statistical models—for example, when they anticipate that randomization will not yield completely balanced distributions or when randomization has failed to bring suspected confounding variables into balance. In this sense, a methodologic intersection exists between epidemiology and neuropsychopharmacology, the epidemiologist typically working with larger and less restricted samples at the population level rather than the smaller samples of patients seeking help seen in most clinical trials.

Examples of Epidemiologic Evidence under the Rubric of Causes

Challenging epidemiologic problems have surfaced in research on the suspected hazards of illicit drug use and neuropsychopharmacologic drug products. For example, during

the height of the cocaine epidemic in the late twentieth century in the United States, ethnographic studies of cocaine users and a limited number of case reports from psychiatrists drew attention to symptoms of panic anxiety, experienced not only during episodes of cocaine intoxication but also afterward. Because of some resonance with pharmacology and neurobiological theory, several clinicians inferred that cocaine use was precipitating panic attacks and panic disorder in persons who had not experienced them previously (45 ,46). Nonetheless, the ethnographic study samples were small (e.g., see ref. 47), and the research was based on relatively uncontrolled study designs (e.g., assessors were not blinded with respect to the suspected causal hypothesis).

In addition, it may not come as a surprise that cocaine users visiting psychiatrists had an apparent excess of a psychiatric condition. The well-known Berkson's bias (48) can account for a false appearance of comorbidity when samples are clinical rather than epidemiologic.

With Berkson's bias in mind, several epidemiologists investigated the suspected causal link of cocaine use to panic attack and panic disorder, each with the strengths of study methods that involved double-blinding with respect to the causal hypotheses (i.e., neither the clinical assessors nor the study participants knew that the hypotheses would be tested). One community sample study of young adults was oriented to anxiety in general rather than to discrete panic attacks specifically, and its general evidence about anxiety did not support the published clinical observations (49). However, drawing prospectively ascertained incident cases of first-time panic attack from within the Epidemiologic Catchment Area study sample, and using multiple regression methods to constrain a range of suspected confounding variables, another epidemiologic study produced statistically robust evidence that cocaine users in the community are about three times more likely to experience panic attacks than are age- and neighborhood-matched nonusers (46). A third study, which applied a new case-crossover research design for epidemiology to data from the National Household Survey on Drug Abuse, produced an estimated relative risk not too distant from the one observed in the prospective Epidemiologic Catchment Area studies (RR = 3) (50).

This example of a suspected psychiatric hazard of cocaine use illustrates how epidemiology makes use of study procedures such as standardized diagnostic assessment, double-blinding, matching, and multiple regression to strengthen the basis for causal inference from evidence based on nonrandomized observational studies. Nevertheless, more work has actually been done and the history of epidemiologic research is longer on the "other side" of co-occurring psychiatric disturbances and drug dependence, where the observed "comorbidity" is thought to arise because the psychiatric disturbance leads to drug dependence. Perhaps the oldest tradition of this type of comorbidity research started with clinical observations about sociopathy and criminal backgrounds in drug dependence cases sent to federal narcotics farms. For example, clinical investigators such as Kolb and Pescor (51) estimated that as many as 50% to 60% of incarcerated cases qualified as antisocial or socially maladapted, with evidence of social maladaption generally predating drug use.

The potential for a type of Berkson's bias is ripe in this context; one should not be too surprised to find an excess of socially maladapted persons among incarcerated prisoners of any stripe. Even so, it was more than 25 years after the initial observations before a proper epidemiologic investigation of this relationship was undertaken. This investigation was a "nonconcurrent prospective study" of children seen in child guidance clinics of the 1920s, some with record-based evidence of childhood rule breaking and deviance, and others without such evidence. The research team secured old child guidance records in the 1950s, and by the early 1960s they had successfully traced, reengaged, and used standardized diagnostic survey methods to assess the vast majority of the sample, who were then well into adulthood, mostly within or beyond the end of the effective period of risk for the development of drug dependence. To the extent that any single study can do so, this classic epidemiologic investigation set to rest most of the concerns about Berkson's bias and showed that prior childhood deviance is linked to a subsequent risk for dependence on heroin and other "narcotic" drugs, with the evidence tending to support a causal inference linking earlier deviance with later drug problems (52 ,53).

Numerous subsequent observational studies followed along this path, with the strength of community samples outside clinics and prisons, but generally with cross-sectional and retrospective research designs (53). One noteworthy exception was the Woodlawn Project "concurrent prospective study" of a large sample of first-graders recruited in the mid-1960s and followed later as teenagers and then as young adults. The teenage follow-up study produced evidence that resonated with the nonconcurrent prospective evidence from the child guidance study sample. Namely, an excess occurrence of "heavy" drug use was noted among teenage boys whose first-grade teachers had rated them as "aggressive" rule-breakers in the classroom (54; M. Ensminger, *personal communication*). Subsequent follow-up, when the teens had matured to adulthood and entered the fourth decade of their lives, showed an excess occurrence of cocaine use in association with teacher-rated aggression measured more than 20 years beforehand.

It is difficult to imagine a direct randomized, controlled trial to test the causal influence of antisocial, deviant, or aggressive behavior on later risk for drug dependence. Regrettably, the observational studies described to this point do not rule out the possibility of an underlying diathesis or predisposition that gives rise both to unruly behavior and to drug dependence, the first coming developmentally earlier in the expression of the diathesis, and the second developmentally

later (see ref. 55 for a recent discussion). This, of course, is one of the main problems of causal inference in contemporary "psychiatric comorbidity" research. Other problems include "shared methods covariation" resulting from heavy reliance on self-report and recall in the measurement of drug dependence and other psychiatric disturbances, and problems associated with an assumption that coarse-grained age-of-onset data can illuminate which condition came first, even when the prodrome of the conditions is known to develop insidiously, often over a span of years (56).

Fortunately, some twin and adoption paradigm research has clarified the issue of shared genetic and environmental predispositions toward antisocial behavior and the problems associated with drug dependence (57 ,58). In addition, an indirect randomized, controlled trial has been completed to evaluate a new developmentally sensitive modality of drug prevention programming and to shed new light on the suspected causal influence of early deviance, aggression, and rule breaking. A detailed description of the indirect randomized, controlled trial is beyond the scope of this chapter and can be read elsewhere (59 ,60). Nevertheless, in brief, the research design involved a repetition of the Woodlawn Project recruitment of a large sample of first-graders (*n* = 2,311), but with random assignment of the children to a "good behavior game" condition, which involved a teacher-led classroom-based intervention designed to improve the behavior and rule abidance of children and promote their social interactions, versus control conditions (either the standard curriculum or a "mastery learning" curriculum designed to improve reading achievement). The children assigned to the experimental interventions were kept in the same primary school classrooms and the same conditions for 2 years, during which they received increasing "doses" of the behavior and reading curriculum. The children assigned to the "standard-setting" or control classrooms also were kept together for 2 years and received just the usual and customary curriculum of the local public school system.

Follow-up assessment of the children who grew up and went to school in this urban public school system occurred on an annual basis from grades 3 to 4 of primary school to grades 7 to 8 of middle school. These assessments involved private face-to-face interviews with each child, in which standardized survey research methods with blinding (assessors did not know which children had received which intervention) and teacher ratings were used. Once the children were old enough, they were allowed to mark their answers on an answer sheet that could not be read by the assessor and was sealed in an envelope for later data entry. The assessments also involved ratings by the teachers in these later grades; the middle school teachers knew that the children had been in a prevention experiment, but they did not know which of the three conditions the child might have received during the first 2 years of primary school.

Life table and regression analyses of the follow-up teacher ratings and self-reported age at first use of tobacco provided evidence consistent with the preventive hypotheses; (a) boys who had received the good behavior game intervention were rated as better-behaved than their counterparts in the other study conditions ($p < .05$), and (b) the risk of starting to smoke tobacco by age 13 or 14 years was substantially greater for boys in the "standard-setting" control classrooms than in those who had spent first and second grades in the "good behavior game" classrooms (RR = 2.0; $p < .05$). Consistent with the observational evidence suggesting that deviant or aggressive behavior is a stronger determinant of drug use for boys than for girls, the "good behavior game" effect was less pronounced among the girls (59). Current continuing follow-up of these study participants into their young adult years, supported by the National Institute of Mental Health and the National Institute of Drug Abuse, will reveal whether the apparent intervention effects are long-lasting and influence the risk for the use of other drugs and the development of drug dependence.

Whereas an indirect randomized, controlled trial of this type is a challenge and requires follow-up over spans of time that exceed typical durations of National Institutes of Health grant awards, this type of experimental investigation of the suspected causal link of early deviance or aggression to later drug use and dependence is indispensable when the task is causal inference. Random assignment of the children to different intervention conditions helps to bring into balance an array of suspected confounding variables. Thereafter, careful measurement and regression modeling help to constrain what randomization has not constrained. Although it is not possible to make a random assignment of children to higher or lower levels of rule breaking and aggressive behavior, it is possible to assign them at random to interventions designed to reduce these levels.

Of course, no single indirect randomized, controlled trial will settle the outstanding issues about this form of psychiatric comorbidity. The case for causal inference will depend on completion of more indirect trials along these lines, with each replication adding strength to the chain of inference and web of causation.

Other forms of "comorbidity" with drug dependence have come under scrutiny in epidemiologic research, most recently an observed co-occurrence involving the anxiety disorders, especially phobic disorders (61). One can expect a more rapid acceleration of epidemiologic attention to the link of anxiety disorders to alcohol or other drug dependence, with time from the first nonexperimental observations to the first indirect randomized, controlled trial measured in years rather than in decades, as was the case for the link of childhood deviance and antisocial behavior to drug dependence. In this context, it may be important to note another feature of the Woodlawn Project findings, which drew attention to the combination of shyness and aggression

or rule breaking among boys. The Woodlawn Project report noted an interaction of shyness and aggression, directing attention mainly to the excess risk for heavy drug use among boys who were rated in first grade as being both shy and aggressive. However, careful inspection of the Woodlawn Project study data indicated that the observed interaction depended heavily on a quite low occurrence of drug use among the boys who were shy but not aggressive. In other words, in the absence of aggression, being shy and not having many friends may help protect an inner city youth against the risks of illicit drug use. The link to current anxiety disorders and comorbidity research involves the prominence of phobias in the observed association with alcohol and other drug dependence, and the Woodlawn Project measurement of shyness as a trait that encompasses not having very many friends, being socially withdrawn, and staying on the fringes of or outside social groups (54).

Before we leave the rubric of causes, several recent studies merit special mention because of their pertinence in the genetic epidemiology of drug dependence. The Vietnam Veteran twin study is especially noteworthy because it is seeking to partition genetic, shared and nonshared environmental influences across a sequence of transitions leading to drug dependence. Unique to this study is its attention to the transition from before to after the first exposure to opportunities to try drugs (62). The importance of this transition in etiologic research on drug dependence is discussed in the next section, under the heading of causal mechanisms.

The work of Kendler and colleagues (63 ,64) is noteworthy for its initial focus on female twins and its spotlight on gene–environment interactions. This work is leading us to a better understanding of how genetic predispositions may have an important influence on entry into risk-laden environments, where exposure to drugs and drug taking becomes more likely, over and above any influence of inherited characteristics on responses to drug exposure. In a related line of work on parent–child interactions, Kendler et al. (65) have recently added new evidence that parental coldness or aloofness may affect the occurrence of alcohol or other drug dependence, but the evidence is not generally supportive of an influence of active parenting styles (e.g., authoritarianism). Of course, evidence to the contrary exists, including some new evidence on how children shape the parenting behaviors displayed by their mothers and fathers (66). Soon, results will be available from indirect randomized, controlled trials in which interventions have been used to increase the aspects of parenting behavior suspected of being most influential in early drug involvement (e.g., supervision and monitoring). Here, the causal inference is supported by a fairly solid body of observational evidence contributed by many different research groups (67 ,68 ,69 ,70 and 71). Nonetheless, as in the link between sociopathy and drug dependence, replications from indirect randomized, controlled trials are apt to provide the most definitive evidence regarding these issues of causal inference. In time, indirect randomized, controlled trials with large epidemiologic samples will probably be performed, possibly with specific targets identified and characterized through elaborations of the human genome project. For example, one can imagine research on parental influences on drug taking that includes measurement of parenting behaviors in addition to inherited determinants of persistent drug use, such as the alleles controlling the cytochrome P-450 enzyme, which is important in nicotine metabolism (72).

RUBRIC 4, MECHANISMS: “HOW DO SEQUENCES OF CIRCUMSTANCES, CONDITIONS, AND PROCESSES LEAD TO DISEASE?”

Part of "109 - Epidemiology of Drug Dependence "

Epidemiology as a discipline places emphasis on studies of the “natural history” of disease, in part because of its early confluence with clinical medicine, bacteriology, and virology. Here, natural history may be understood as the outward manifestations of an evolving causal process and the expression of causal mechanisms that lead toward the fatal or nonfatal resolution of the condition under study. In the study of diseases, “clinical course” can be differentiated from “natural history” once clinical attention can make a fundamental difference. Before then, what we see is natural history. Once effective treatment maneuvers have been started, what we see is the clinical course, or natural history modified by clinical attention.

Until recently, when the concept of natural history was applied in the epidemiology of drug dependence, most attention was given to observable “stages” and “developmental sequences.” For example, Robins (73) and Winick (74) advocated decomposition of the addiction process into stages. They specified a pre-initiation stage that involved first exposure to an opportunity to try a drug. Thereafter, some presented with an opportunity go on to try the drug, whereas others do not. Among those who actually try the drug, the drug-using stage may or may not be followed by another stage—transition into drug dependence. Some users actually “mature out” of stages of very serious drug use (e.g., see ref. 74).

In the work of Kandel and Davies (75), the natural history of drug involvement is conveyed as a stage developmental sequence in which different drugs are tried, first legally available beer or wine, then hard liquor or tobacco, then marijuana as the first “illicit” drug in the sequence, then other illicit drugs. The last development in Kandel’s sequence is use of prescription psychotherapeutic medicines (Fig. 109.5) (85); others have confirmed this position for prescription drug use (86).

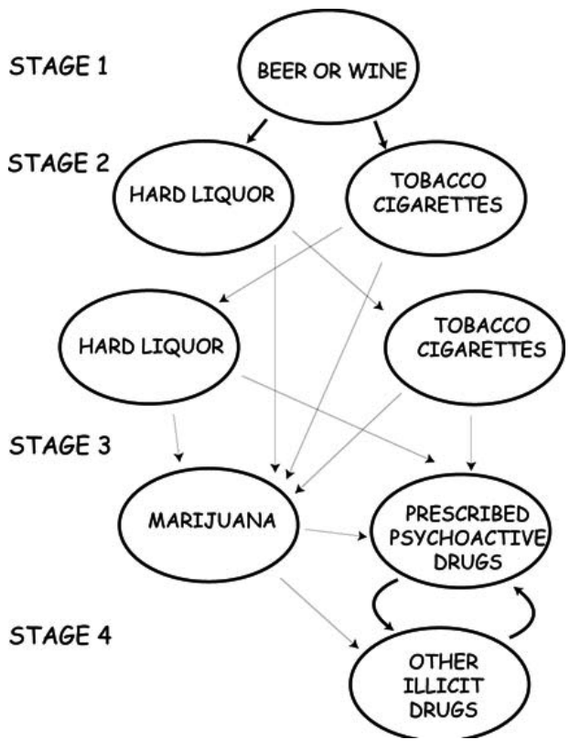


FIGURE 109.5. Stages in the developmental sequence of adolescent drug involvement. (Adapted from Kandel DB, Yamaguchi K, Chen K. Stages of progression in drug involvement from adolescence to adulthood: further evidence for the gateway theory. *J Stud Alcohol* 1992;53:447–57, with permission.)

Nonetheless, some observers have argued that this “gateway description” of sequences from drug to drug may rest solely on different levels of availability or opportunity to use different drugs. Other investigators have challenged the stage transition concept as applied to drug dependence and youthful tobacco smoking (77). They have advocated an

analysis of levels of drug involvement in terms of a hybrid concept that allows for discrete stage transitions, but with dimensional movement within each stage (e.g., see ref. 27). Statistical methods for studying this hybrid transition progression model are being developed for epidemiologic research (e.g., see refs. 78 , 79).

A recent development has been epidemiologic research on the natural history and clinical course of the various clinical features of alcohol dependence (e.g., see refs. 80 , 81 and 82). Separate lines of clinical and epidemiologic research on the natural history of dependence on drugs other than alcohol have also been initiated (e.g., see refs. 83 , 84).

RUBRIC 5, PREVENTION AND CONTROL: “WHAT CAN BE DONE TO PREVENT, REDUCE, OR AMELIORATE THE ADVERSE IMPACT?”

Part of "109 - Epidemiology of Drug Dependence "

The central position of prevention in epidemiology already has been mentioned in this chapter, although many epidemiologists’ careers are devoted to observational studies, with little attention to intervention research. “Control” is also a key concept in epidemiology, referring to maneuvers such as quarantine or the effective treatment of active cases to limit spread to other persons.

During the evolution of epidemiology in the nineteenth century, a new type of professional emerged—a public health officer equipped with newly found knowledge of epidemiology and armed with police powers necessary to protect the larger population from the threat of infectious diseases. In twentieth century efforts to mount an effective societal response to drug dependence, the police authority was split from the public health authority. As a result, when most people now think of the prevention of drug dependence, what comes to mind are health education classes for young people of school age or mass media campaigns to publicize the hazards one faces once drug use starts. We do not tend to think of the international, federal, state, and local laws or police actions as societal instruments for prevention. Nor do we tend to think of early interventions for drug-dependent cases, tracing of secondary contacts who may be sources of sustained outbreaks, or effective treatment of active cases as a means of preventing new cases. Indeed, in some quarters, the opinion has been expressed that concepts of epidemiology and public health should not be applied to drug dependence because these concepts are tied inherently to coercive actions, such as quarantine (85).

Notwithstanding these concerns, during the past quarter-century, some epidemiologists have directed attention to the evaluation of laws and regulatory activities thought to prevent and control drug dependence and associated hazards. Starting in the 1960s, de Alarcon (86) and Hughes et al. (87) refined methods of tracing secondary cases and of street outreach to curb urban outbreaks of heroin dependence. Figure 109.6 shows the pattern of spread of heroin injection that was central to de Alarcon’s work on epidemiology and the prevention of heroin epidemics.

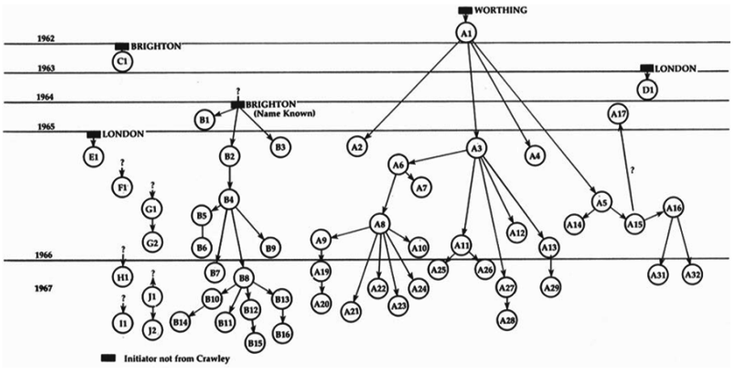


FIGURE 109.6. The person-to-person spread of heroin injection across time and space in the United Kingdom in the 1960s. (From De Alarcon R. The spread of heroin abuse in a community. *Community Health (Bristol)* 1969;1:155–161, with permission.)

It may come as a surprise that epidemiologists have not been the ones to sustain this work or build on it. For example, the most advanced efforts to evaluate drug policy have come from systems research models that make use of epidemiologic data but are based more on econometrics and operations research than on epidemiologic principles, concepts, and methods (e.g., see refs. 88 , 89 and 90). Indeed, more epidemiologic attention and evaluative research have been devoted to community mobilization to prevent HIV infection and AIDS than to the prevention and control of drug dependence, although a new impetus for community mobilization is coming from drug treatment researchers (e.g., see refs. 91 , 92).

Noteworthy exceptions to this generalization about drug prevention research do exist. As examples, the work of Pentz and Perry and their colleagues (93 ,94 ,95 ,96 and 97) involves the mobilization of communities to shape policy and procedures, with a core of interventions directed toward young persons of school age. Gutman and Clayton (98) have recently urged that greater attention be paid to “upstream” prevention maneuvers that affect large aggregations of communities, such as state, federal, and even international policy initiatives. In alcohol and tobacco research, some noteworthy examples can be found of the evaluation of “upstream” interventions

(e.g., see refs. 99 , 100 , 101 and 102), in addition to evaluations of “downstream” maneuvers, such as limiting tobacco sales to minors (103).

Recently, interest has been renewed in “multilevel” statistical models that take into account different levels of organization, from the community at large to the local neighborhood to the household or individual, and in models of “dependent happenings,” such as are seen when innovations (e.g., drug use) diffuse from one person or group to the next (104 ,105 and 106). A carryover into the domain of prevention research has been expressed in recent articles and a textbook (107). These developments, coupled with a greater appreciation of gene–environment transactions or reciprocities, rather than gene–environment competition, promise to transform and sharpen the focus of prevention research during future decades as the human genome project yields new targets (108 ,109 and 110).

CONCLUSION AND FORECAST

Part of "109 - Epidemiology of Drug Dependence "

It is possible to make an optimistic forecast regarding the application of epidemiology to the study of drug dependence. Under the rubric of “quantity,” sustained growth in the number of cross-sectional “prevalence surveys” that estimate the frequency of drug dependence in various populations and subpopulations of the world is apparent. Diagnostically oriented national surveys, such as the National Comorbidity Survey and the National Household Survey on Drug Abuse in the United States, will continue to be at center stage. The “World Mental Health 2000” surveys, organized by Professor Ronald Kessler of Harvard University and Dr. Bedirhan Ustun of the World Health Organization, will enlarge these national perspectives and offer epidemiologic data on the prevalence of drug dependence in more than 15 different participating countries. Because of the greater difficulty and complexity of the Epidemiologic Catchment Area studies, it is less likely that we will see similar growth in prospectively derived estimates of the incidence of drug dependence and the risk for becoming drug-dependent. Most likely, we will have to make do with approximate estimates of risk based on retrospective data from the cross-sectional surveys.

The sustained attention given to determining the prevalence of drug dependence within the context of more general surveys of psychiatric disturbances essentially guarantees a raft of new findings on the location of cases and “psychiatric comorbidity” within the populations of the world. We are likely to see more and more data on the male excess in drug dependence cases, although in some countries, because of the use of psychotherapeutic medicines, a female excess may be shown for some drug categories. Similarly, the excess occurrence among 15- to 44-year-olds in comparison with other age groups may prove to be a general rule via exceptions,

such as the high prevalence of heroin or opium dependence among elderly persons living in the opium-growing regions of the world. Nonetheless, it seems that new findings from these cross-sectional surveys will be most useful in confirming past observations. One may hope for transformative evidence, but the work is not likely to be ground-breaking.

Under the rubric of “causes,” the intersection of epidemiology intersects with the human genome project provides a basis for optimism. As discussed elsewhere, epidemiology has a special capacity to discover environmental circumstances, conditions, and processes that modify inherited predispositions. To the extent that epidemiologic studies are able to incorporate measurements of genetic polymorphisms and to characterize participants as heterozygotes and homozygotes, they will disclose variations in the expression of risk. These variations, linked to environmental conditions or processes, will clarify the webs of causation leading to drug dependence.

The capacity of epidemiology to yield definitive evidence regarding macrosocial causes of drug dependence, such living within an inner city community or being of low socioeconomic status (e.g., see ref. 111), is less of a reason for optimism. The definitive quality of research on these topics will remain limited without a truly massive investment in prospective studies within urban areas, and without levels of subject cooperation and participation far in excess of what is now achieved in these areas.

Under the rubric of “mechanisms,” the above-mentioned statistical advances will bear fruit once investments have been made in longitudinal studies designed to make the measurements required to characterize the hybrid sequences of transitions and progressions. Linked with advances in human genetics and the measurement of environmental conditions and processes, these longitudinal studies promise advances in our understanding, but as a “basic science” initiative, the clinical application of this new understanding is not immediately clear.

Under the rubric of “prevention and control,” we will begin to see long-term results from rigorous drug prevention research during the first decade of the twenty-first century. This evidence should help us to clarify central issues, such as whether preventing the onset of illicit drug use in the early teenage years will be followed by a reduced risk for drug dependence in later adolescence and early adulthood. One may hope for an intersection of etiologic research and prevention research, but any new gains in understanding the genetics of drug dependence may not yield practical preventive interventions for a half-century or more.

The recently developed systems research models suffer mainly from inadequate epidemiologic and law enforcement data (e.g., see ref. 90). However, having forged these systems research models, the policy analysts should provoke epidemiologists to improve the data, and it is hoped that the societal investment in improved data will lead to more compelling evaluations of drug policies and societal responses to drug dependence and illicit drug use.

Under this same rubric, a ray of light has begun to shine forth from the National Institute of Drug Abuse in the United States, where a new unit has been established to promote research on community mobilization and efforts at the community level to curb outbreaks of drug taking and drug dependence. Coupled with continuing progress in the ongoing evaluation of school-based prevention programs and mass media campaigns, this initiative represents an important step in the next generation of progress in epidemiologic research on drug dependence.

ACKNOWLEDGMENT

Part of "109 - Epidemiology of Drug Dependence "

Some of the material in this chapter overlaps with material and ideas presented in other review articles and chapters written by the author. For example, the concepts associated with the rubrics of epidemiology originally were presented in a chapter by Anthony and Van Etten (6); concepts on the hybrid transition–progression model were presented by Anthony and Helzer (27). The appropriate work has been cited, and the editors have been notified of the circumstances.

REFERENCES

1. Leighton AH, Clausen JA, Wilson RN. *Explorations in social psychiatry*. New York: Basic Books, 1958.

2. Ball JC, Chambers CD. *The epidemiology of opiate addiction in the United States*. Springfield, IL: Charles C Thomas Publisher, 1970.

3. Terry CE, Pellens M. *The opium problem*. New York: Bureau of Social Hygiene, 1928.

4. Parry H, Balter MB, Mellinger GD, et al. National patterns of psychotherapeutic drug use. *Arch Gen Psychiatry* 1973;28:18–74.

5. Morris JN. *Uses of epidemiology*. London: Livingstone, 1957.

6. Anthony JC, Van Etten ML. Epidemiology and its rubrics. In: Bellack AS, Hersen M, eds. *Comprehensive clinical psychology*, first ed. New York: Pergamon, 1998.

7. Eaton WW, Kramer M, Anthony JC, et al. The incidence of specific DIS/DSM-III mental disorders: data from the NIMH Epidemiologic Catchment Area Program. *Acta Psychiatr Scand* 1989;79:163–178.

8. Anthony JC, Warner LA, Kessler RC. Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalents: basic findings from the National Comorbidity Survey. *Exp Clin Psychopharmacol* 1994;2:244–268.

9. Kessler RC, Crum RM, Warner LA, et al. Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Arch Gen Psychiatry* 1997;54:313–321.

10. Kandel D, Chen K, Warner LA, et al. Prevalence and demographic correlates of symptoms of last year dependence on alcohol, nicotine, marijuana and cocaine in the U.S. population. *Drug Alcohol Depend* 1997;44:11–29.

11. Chen K, Kandel DB, Davies M. Relationships between frequency and quantity of marijuana use and last year proxy dependence among adolescents and adults in the United States. *Drug Alcohol Depend* 1997;46:53–67.

12. Hall W, Teesson M, Lynskey M, et al. The 12-month prevalence of substance use and ICD-10 substance use disorders in Australian adults: findings from the National Survey of Mental Health and Well-Being. *Addiction* 1999;94:1541–1550.

13. Caetano R, Tam T, Greenfield T, et al. DSM-IV alcohol dependence and drinking in the U.S. population: a risk analysis. *Ann Epidemiol* 1997;7:542–549.

14. *National household survey on drug abuse: main findings, 1998*. DHHS publication No. (SMA) 00-3381. Rockville, MD: Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, 2000.

15. Grant BF. Prevalence and correlates of drug use and DSM-IV drug dependence in the United States: results of the National Longitudinal Alcohol Epidemiologic Survey. *J Subst Abuse* 1996;8:195–210.

16. Grant BF. The relationship between ethanol intake and DSM-III-R alcohol dependence: results of a national survey. *J Subst Abuse* 1993;5:257–267.

17. Warner LA, Kessler RC, Hughes M, et al. Prevalence and correlates of drug use and dependence in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1995;52:219–229.

18. Hall W, Teesson M, Lynskey M, et al. The 12-month prevalence of substance use and ICD-10 substance use disorders in Australian adults: findings from the National Survey of Mental Health and Well-Being. *Addiction* 1999;94:1541–1550.

19. Meltzer H, Jenkins R. The national survey of psychiatric morbidity in Great Britain. *Int Rev Psychiatry* 1994;6:349–356.

20. Nelson CB, Wittchen HU. DSM-IV alcohol disorders in a general population sample of adolescents and young adults. *Addiction* 1998;93:1065–1077.

21. Hagnell O, Lanke J, Rorsman B, et al. Predictors of alcoholism in the Lundby Study. I. Material and methods. *Eur Arch Psychiatry Neurol Sci* 1986;235:187–191.

22. Chen WJ, Cheng AT. Incidence of first onset alcoholism among Taiwanese aborigines. *Psychol Med* 1997;27:1363–1371.

23. Cicero TJ, Adams EH, Geller A, et al. A postmarketing surveillance program to monitor Ultram (tramadol hydrochloride) abuse in the United States. *Drug Alcohol Depend* 1999;57:7–22.

24. Hunt LG. Prevalence of active heroin use in the United States. In: Rittenhouse JD, ed. *Report of the task force on the epidemiology of heroin and other narcotics*. Menlo Park, CA: Stanford Research Institute, 1976:35–52.

25. Newman RG, Cates MS. Estimating the incidence and prevalence of addiction: Why? In: Rittenhouse JD, ed. *Report of the task force on the epidemiology of heroin and other narcotics*. Menlo Park, CA: Stanford Research Institute, 1976:71–72.

26. Terry CE, Pellens M. *The opium problem*. New York: Bureau of Social Hygiene, 1928.

27. Anthony JC, Helzer JE. Epidemiology of drug dependence. In: Tsuang TZ, ed. *Textbook of epidemiology*. New York: Wiley-Liss, 1995:361–406.

28. Van Etten ML, Neumark YD, Anthony JC. Male–female differences in the earliest stages of drug involvement. *Addiction* 1999;94:1413–1419.

29. Anthony JC, Helzer JE. Syndromes of drug abuse and dependence. In: Robins LN, Regier DA, eds. *Psychiatric disorders in America: the Epidemiologic Catchment Study*. New York: The Free Press, 1991:116–154.

30. Johnson RA, Gerstein DR. Initiation of use of alcohol, cigarettes, marijuana, cocaine, and other substances in U.S. birth cohorts since 1919. *Am J Public Health* 1998;88:27–33.

31. Johnson RA, Gerstein DR. Age, period, and cohort effects in marijuana and alcohol incidence: United States females and males, 1961–1990. *Subst Use Misuse* 2000;35:925–948.

32. Dormitzer C, Caris L, Anthony JC. Parental attention and risk of coca paste smoking in Chile: preliminary data from the 1999 national school survey in Chile. 2000.

33. Morris JN. *Uses of epidemiology*. London: Livingstone, 1957.

34. Robins LN. Estimating addiction rates and locating target populations. How decomposition into stages helps. *NIDA Res Monogr* 1977;16:25–39.

35. O'Brien CP, Nace EP, Mintz J, et al. Follow-up of Vietnam veterans. I. Relapse to drug use after Vietnam service. *Drug Alcohol Depend* 1980;5:333–340.

36. Robins LN, Davis DH, Goodwin DW. Drug use by U.S. Army enlisted men in Vietnam: a follow-up on their return home. *Am J Epidemiol* 1974;99:235–249.

37. Robins LN, Helzer JE, Davis DH. Narcotic use in southeast Asia and afterward. An interview study of 898 Vietnam returnees. *Arch Gen Psychiatry* 1975;32:955–961.

38. Anthony JC, Petronis KR. Early-onset drug use and risk of later drug problems. *Drug Alcohol Depend* 1995;40:9–15.

39. Tennant FS Jr, Detels R, Clark V. Some childhood antecedents of drug and alcohol abuse. *Am J Epidemiol* 1975;102:377–385.

40. Margulies RZ, Kessler RC, Kandel DB. A longitudinal study of onset of drinking among high-school students. *J Stud Alcohol* 1977;38:897–912.

41. Robins LN, Przybeck TR. Age of onset of drug use as a factor in drug and other disorders. In: Jones CL, Battjes RL, eds. *Etiology of drug abuse: implications for prevention*. NIDA Research Monograph No 56. Rockville, MD: National Institute on Drug Abuse, 1985:178–192..

42. Grant BF, Dawson DA. Age of onset of drug use and its association with DSM-IV drug abuse and dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. *J Subst Abuse* 1998;10:163–173.

43. Kandel DB, Chen K. Types of marijuana users by longitudinal course. *J Stud Alcohol* 2000;61:367–378.

44. Prescott CA, Kendler KS. Age at first drink and risk for alcoholism: a noncausal association. *Alcohol Clin Exp Res* 1999;23:101–107.

45. Aronson TA, Craig TJ. Cocaine precipitation of panic disorder. *Am J Psychiatry* 1986;143:643–645.

46. Anthony JC, Tien AY, Petronis KR. Epidemiologic evidence on cocaine use and panic attacks. *Am J Epidemiol* 1989;129:543–549.

47. Chitwood DD. Patterns and consequences of cocaine use. *NIDA Res Monogr* 1985;61:111–129.

48. Berkson J. Limitations of the application of fourfold table analysis to hospital data. *Biometrics* 1946;2:47–53.

49. Newcomb MD, Bentler PM, Fahy B. Cocaine use and psychopathology: associations among young adults. *Int J Addict* 1987;22:1167–1188.

50. Wu LT, Anthony JC. The use of the case-crossover design in studying illicit drug use. *Subst Use Misuse* 2000;35:1035–1050.

51. Pescor MJ. The Kolb classification of drug addicts. *Public Health Rep* 1939; [Suppl 55].

52. Robins LN. *Deviant children grown up*. Baltimore: Williams & Wilkins, 1966.

53. Robins LN. The intimate connection between antisocial personality and substance abuse. *Soc Psychiatry Psychiatr Epidemiol* 1998;33:393–399.

54. Kellam SG, Brown CH, Rubin BR, et al. Paths leading to teenage psychiatric symptoms and substance use: developmental epidemiologic studies in Woodlawn. In: Guze SB, Earls FJ, Barrett JE, eds. *Childhood psychopathology and development*. New York: Raven Press, 1983:17–51.

55. Limosin F, Ades J, Gorwood P. Relationships between antisocial personality and alcoholism: genetic hypotheses. *Eur Psychiatry* 2000;15:123–128.

56. Wu LT, Anthony JC. Tobacco smoking and depressed mood in late childhood and early adolescence. *Am J Public Health* 1999;89:1837–1840.

57. Cadoret RJ, Yates WR, Troughton E, et al. Adoption study demonstrating two genetic pathways to drug abuse. *Arch Gen Psychiatry* 1995;52:42–52.

58. van den Bree MB, Svikis DS, Pickens RW. Genetic influences in antisocial personality and drug use disorders. *Drug Alcohol Depend* 1998;49:177–187.

59. Kellam SG, Ling X, Merisca R, et al. The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Dev Psychopathol* 1998;10:165–185.

60. Kellam SG, Anthony JC. Targeting early antecedents to prevent tobacco smoking: findings from an epidemiologically based randomized field trial. *Am J Public Health* 1998;88:1490–1495.

61. Merikangas KR, Risch NJ, Weissman MM. Comorbidity and co-transmission of alcoholism, anxiety and depression. *Psychol Med* 1994;24:69–80.

62. Tsuang MT, Lyons MJ, Harley RM, et al. Genetic and environmental influences on transitions in drug use. *Behav Genet* 1999;29:473–479.

63. Kendler KS, Prescott CA. Cocaine use, abuse and dependence in a population-based sample of female twins. *Br J Psychiatry* 1998;173:345–350.

64. Straub RE, Sullivan PF, Ma Y, et al. Susceptibility genes for nicotine dependence: a genome scan and follow-up in an independent sample suggest that regions on chromosomes 2, 4, 10, 16, 17 and 18 merit further study. *Mol Psychiatry* 1999;4:129–144.

65. Kendler KS, Myers J, Prescott CA. Parenting and adult mood, anxiety and substance use disorders in female twins: an epidemiological, multi-informant, retrospective study. *Psychol Med* 2000;30:281–294.

66. Collins WA, Maccoby EE, Steinberg L, et al. Contemporary research on parenting. The case for nature and nurture. *Am Psychol* 2000;55:218–232.

67. Patterson GR, Stouthamer-Loeber M. The correlation of family management practices and delinquency. *Child Dev* 1984;55:1299–1307.

68. Curran PJ, Chassin L. A longitudinal study of parenting as a protective factor for children of alcoholics. *J Stud Alcohol* 1996;57:305–313.

69. Chilcoat HD, Dishion TJ, Anthony JC. Parent monitoring and the incidence of drug sampling in urban elementary school children. *Am J Epidemiol* 1995;141:25–31.

70. Chilcoat HD, Anthony JC. Impact of parent monitoring on initiation of drug use through late childhood. *J Am Acad Child Adolesc Psychiatry* 1996;35:91–100.

71. Chilcoat HD, Breslau N, Anthony JC. Potential barriers to parent monitoring: social disadvantage, marital status, and maternal psychiatric disorder. *J Am Acad Child Adolesc Psychiatry* 1996;35:1673–1682.

72. Pianezza ML, Sellers EM, Tyndale RF. Nicotine metabolism defect reduces smoking. *Nature* 1998;393:750.

73. Robins LN, Helzer JE, Davis DH. Narcotic use in southeast Asia and afterward. An interview study of 898 Vietnam returnees. *Arch Gen Psychiatry* 1975;32:955–961.

74. Winick CS. Maturing out of narcotic addiction. *Bull Narc* 1962;14:1–7.

75. Kandel DB, Davies M. Progression to regular marijuana involvement: phenomenology and risk factors for near-daily use. In: Glantz M, Pickens R, eds. *Vulnerability to drug abuse*. Washington, DC: American Psychological Association, 1992:211–254.

76. Trinkoff AM, Anthony JC, Munoz A. Predictors of the initiation of psychotherapeutic medicine use. *Am J Public Health* 1990;80:61–65.

77. Mayhew KP, Flay BR, Mott JA. Stages in the development of adolescent smoking. *Drug Alcohol Depend* 2000;59[Suppl 1]:S61–S81.

78. Muthen B, Muthen LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res* 2000;24:882–891.

79. Reboussin BA, Anthony JC. Latent class marginal regression models for modelling youthful drug involvement and its suspected influences. *Stat Med* 2001;20:623–639

80. Langenbucher JW, Chung T. Onset and staging of DSM-IV alcohol dependence using mean age and survival-hazard methods. *J Abnorm Psychol* 1995;104:346–354.

81. Schuckit MA, Anthenelli RM, Bucholz KK, et al. The time course of development of alcohol-related problems in men and women. *J Stud Alcohol* 1995;56:218–225.

82. Schuckit MA, Daeppen JB, Tipp JE, et al. The clinical course of alcohol-related problems in alcohol dependent and nonalcohol dependent drinking women and men. *J Stud Alcohol* 1998;59:581–590.

83. Anglin MD, Hser YI, McGlothlin WH. Sex differences in addict careers. 2. Becoming addicted. *Am J Drug Alcohol Abuse* 1987;13:59–71.

84. Rosenberg M, Anthony JC. Early clinical manifestations of cannabis dependence in a community sample. *Drug Alcohol Depend* 2001;64:123–131.

85. Sidel VW, Drucker E. Further comments on the communicable disease model of heroin addiction. *Am J Drug Alcohol Abuse* 1976;3:369–372.

86. De Alarcon R. The spread of heroin abuse in a community. *Community Health (Bristol)* 1969;1:155–161.

87. Hughes PH, Barker NW, Crawford GA, et al. The natural history of a heroin epidemic. *Am J Public Health* 1972;62:995–1001.

88. Rydell CP, Everingham SS. *Controlling cocaine: supply versus demand programs*. Santa Monica, CA: Drug Policy Research Center: RAND, 1994.

89. Behrens DA, Caulkins JP, Tragler G, et al. A dynamic model of drug initiation: implications for treatment and drug control. *Math Biosci* 1999;159:1–20.

90. Manski CF, Pepper JV, Thomas YF. *Assesment of two cost-effectiveness studies on cocaine control policy*. Washington, DC: National Academy Press, 1999.

91. Belding MA, Iguchi MY, Morral AR, et al. Assessing the helping alliance and its impact in the treatment of opiate dependence. *Drug Alcohol Depend* 1997;48:51–59.

92. Abbott PJ, Weller SB, Delaney HD, et al. Community reinforcement approach in the treatment of opiate addicts. *Am J Drug Alcohol Abuse* 1998;24:17–30.

93. Pentz MA, MacKinnon DP, Flay BR, et al. Primary prevention of chronic diseases in adolescence: effects of the Midwestern Prevention Project on tobacco use. *Am J Epidemiol* 1989;130:713–724.

94. Chou CP, Montgomery S, Pentz MA, et al. Effects of a community-based prevention program on decreasing drug use in high-risk adolescents. *Am J Public Health* 1998;88:944–948.

95. Perry CL, Williams CL, Veblen-Mortenson S, et al. Project Northland: outcomes of a community-wide alcohol use prevention program during early adolescence. *Am J Public Health* 1996;86:956–965. [Comment appears in *Am J Public Health* 1996;86:923–924.]

96. Perry CL, Williams CL, Komro KA, et al. Project Northland high school interventions: community action to reduce adolescent alcohol use. *Health Educ Behav* 2000;27:29–49.

97. Wynn SR, Schulenberg J, Maggs JL, et al. Preventing alcohol misuse: the impact of refusal skills and norms. *Psychol Addict Behav* 2000;14:36–47.

98. Gutman M, Clayton R. Treatment and prevention of use and abuse of illegal drugs: progress on interventions and future directions. *Am J Health Promot* 1999;14:92–97.

99. Allamani A, Casswell S, Graham K, et al. Introduction: community action research and the prevention of alcohol problems at the local level. *Subst Use Misuse* 2000;35:1–10.

100. Toomey TL, Wagenaar AC. Policy options for prevention: the case of alcohol. *J Public Health Policy* 1999;20:192–213.

101. Lantz PM, Jacobson PD, Warner KE, et al. Investing in youth tobacco control: a review of smoking prevention and control strategies. *Tobacco Control* 2000;9:47–63.

102. Gratias EJ, Krowchuk DP, Lawless MR, et al. Middle school students' sources of acquiring cigarettes and requests for proof of age. *J Adolesc Health* 1999;25:276–283.

103. Minozzi S, Grilli R. The systematic review of studies on the efficacy of interventions for the primary prevention of alcohol abuse among adolescents. *Epidemiol Prev* 1997;21:180–188.

104. Duncan TE, Duncan SC, Hops H. Latent variable modeling of longitudinal and multilevel alcohol use data. *J Stud Alcohol* 1998;59:399–408.

105. Bobashev GV, Anthony JC. Clusters of marijuana use in the United States. *Am J Epidemiol* 1998;148:1168–1174.

106. Petronis KR, Anthony JC. Perceived risk of cocaine use and experience with cocaine: do they cluster within U.S. neighborhoods and cities? *Drug Alcohol Depend* 2000;57:183–192.

107. Murray DM. *Design and analysis of group-randomized trials*. New York: Oxford University Press, 1998.

108. Neiderhiser JM, Reiss D, Hetherington EM, et al. Relationships between parenting and adolescent adjustment over time: genetic and environmental contributions. *Dev Psychol* 1999;35:680–692.

109. Anthony JC. The promise of psychiatric enviromics. *Br J Psychiatry* 2001;178(Suppl 40): S8–S11.

110. O'Connor TG, Deater-Deckard K, Fulker D, et al. Genotype-environment correlations in late childhood and early adolescence: antisocial behavioral problems and coercive parenting. *Dev Psychol* 1998;34:970–981.

111. Dohrenwend BP, Levav I, Shrout PE, et al. Socioeconomic status and psychiatric disorders: the causation–selection issue. *Science* 1992;255:946–952.

110

Neuroimaging of Cocaine Craving States: Cessation, Stimulant Administration, and Drug Cue Paradigms

A. R. Childress

Teresa R. Franklin

J. Listerud

Paul D. Acton

Charles P. O'Brien

A. R. Childress, Teresa R. Franklin, J. Listerud, Paul D. Acton, and Charles P. O'Brien: Addiction Treatment Research Center, Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania.

Cocaine craving is a cardinal feature of addiction to the drug and is clinically significant because of its potential link to relapse. Addicted patients describe the craving as a powerful, “must-have” pull that causes them to risk, and sometimes lose, their relationships, families, money, possessions, jobs, and even their lives. “Anticraving” medications are greatly needed, but despite intensive research efforts since the mid-1980s, such agents have remained elusive (1 ,2). This dilemma is in part a consequence of our inability to define or measure cocaine “craving” clearly. Diversity in measurement may well account for some of the variability in the data collected, as described below. In part, the problem is our nascent understanding of which brain substrate(s) an “anticraving” medication should address. Until recently, the activity of the brain during cocaine craving was a matter of inference rather than direct observations. The increased availability of powerful tools for brain imaging *in vivo* has thrust research on drug craving forward into a new era. Several laboratories have begun to measure brain activity during cocaine-craving states directly. This chapter reviews current findings, offering a framework for the results and a discussion of their theoretic and treatment implications.

IS THERE MORE THAN ONE KIND OF COCAINE CRAVING?

Part of "110 - Neuroimaging of Cocaine Craving States: Cessation, Stimulant Administration, and Drug Cue Paradigms "

Despite sophisticated tools for brain imaging, the study of cocaine craving is complicated by the fact that desire for the drug can be reported under dramatically different conditions, and measurement is carried out with a range of different scales. For example, craving may be reported in association with *cocaine cessation* (3), but also in association with *cocaine administration* (4) and *cues signaling cocaine* (5 ,6). Brain substrates may differ across these conditions. Interestingly, changes in the mesolimbic dopamine (DA) systems of the brain have been implicated in all three conditions, but the direction of change is not the same in all three. A possible (tonic) decrease in mesolimbic (nucleus accumbens) DA has been proposed in the case of cessation/functional depletion (7 ,8 and 9), whereas an (episodic) *increase* in DA has been hypothesized in the case of cocaine administration (10) and response to cocaine cues (11 ,12). Of course, these differing DA-related states are not mutually exclusive and may interact in important ways; for example, a dose of cocaine or a cocaine-related cue may have a different brain impact in early versus late cessation, depending on the dynamic re-regulation of the affected substrates. The possible “duality” (too much DA vs. too little DA) of cocaine craving has implications not only for the design of imaging studies but also for the treatment of cocaine craving (2). If the *same* brain substrates are responsible for the craving associated with DA deficiency and the craving associated with DA excess, how can either of these craving states be treated without worsening the other? We will return to this question after a review of the neuroimaging evidence for craving across the conditions.

Neurotransmitter systems other than DA are likely to be involved in cocaine reward and motivation (13 ,14). For example, serotonin (15 ,16), glutamate (17), corticosteroids (18 ,19 and 20), and opioids (21 ,22) have each received substantial cocaine-related research attention. However, in the *neuroimaging*

of cocaine craving, many of the studies have used a DA-related framework for their hypotheses, interpretations, or both. The emphasis on a dopaminergic system reflects not only a long literature on DA involvement in stimulant reward and motivation (23 ,24 and 25) but also the availability of DA-related neuroimaging tools for humans (26). Imaging studies testing the role of other neurotransmitters and neuromodulators in cocaine craving will be a welcome addition to the field. The only currently available study in this category is included in the current review (27).

The neuroimaging studies of cocaine craving discussed below are categorized according to whether their primary focus is *cessation, stimulant administration*, or *drug cue paradigms*. In studies in which the paradigms may overlap (e.g., patients in early cessation viewing cocaine cues), this is noted, as it may be helpful in the final integration of craving results across studies and paradigms.

IMAGING OF CRAVING DURING COCAINE CESSATION

Part of "110 - Neuroimaging of Cocaine Craving States: Cessation, Stimulant Administration, and Drug Cue Paradigms "

Hypotheses

Early in the cocaine epidemic, investigators hypothesized that the long-term use of cocaine resulted in a functional depletion of brain DA (9) in regions critical for the regulation of mood, motivation, thought, and concentration. Acute cocaine administration to laboratory animals or to humans elevates levels of synaptic DA (and other monoamines) by reuptake blockade (28 ,29). However, brain measurements in long-term users (or laboratory animals maintained on cocaine) have revealed differences from controls, such as reduced DA synthesis (30), reduced cocaine uptake (31), down-regulation of postsynaptic DA sites (32 ,33), and altered responses in the endogenous opioid system (34 ,35 and 36), differences that may reflect the homeostatic attempts of the brain to cope with a cocaine-induced flood of synaptic DA. DA transporters (DATs) in the long-term user of cocaine may also be dysregulated, although the direction of observed change varies across studies (35 ,37 ,38 ,39 and 40), potentially reflecting the oscillating nature of re-regulation (see also refs. 31 , 41). The DA dysregulation observed in cocaine users following cessation is supported by numerous animal studies. These studies have found alterations in accumbens DA levels (7), the threshold for rewarding brain self-stimulation (42), the metabolism of reward-relevant regions (43), DA synthesis (44), and postsynaptic DA receptors (44). Thus, on cocaine cessation, the DA systems of some cocaine users may be in a neurologically adapted, dysregulated state. The lowered mood, energy, and concentration experienced by some cocaine patients during abstinence may be related, at least in part, to a dysfunction in brain DA systems. According to this general view, the *craving* that arises during abstinence may also reflect DA system dysregulation (8). Do neuroimaging studies support this hypothesis?

Data

The results of several neuroimaging studies are consistent with the notion of DA dysregulation in cocaine cessation; we review here the studies that have also included craving measures.

Early Cessation (1 Week or Less post Cocaine)

Volkow et al. (45) found a higher rate of metabolism by positron emission tomography (PET) with ^{18}F-fluorodeoxyglucose in cocaine users in early cessation (≤1 week post cocaine) than in controls in two major regions of DA projection, the orbitofrontal cortex and the basal ganglia. In this study, craving (“none,” “mild,” “moderate,” or “severe”) during the week before the scan was positively correlated with (both relative and absolute) metabolic rates in the orbitofrontal cortex, and with (absolute but not relative) metabolic rates in another major mesocortical DA projection region, the prefrontal cortex.

Malison et al. (39) studied DATs with single-photon emission computed tomography (SPECT) and B-CIT, a tracer that binds to dopamine and serotonin transporters, in patients abstinent from cocaine for 96 hours or less to determine whether they exhibited the transporter elevations predicted by a neuroadaptive hypothesis. DAT increases of approximately 20% in comparison with controls were indeed detectable, although they were more modest than the DAT increases found in postmortem studies of cocaine users (37 ,38). An inverse relationship was noted between DAT level and depression scores, but no relationship of DAT level to craving scores (on the Cocaine Craving Scale) was found (3).

Early and Later Cessation

Long-term cocaine use can alter μ-opioid binding (34 ,35 and 36), which indicates an interaction of DA and endogenous opioid systems. Given the role of endogenous opioids in mood modulation, up-regulated μ-opioid receptors may contribute to dysphoria or craving in cocaine cessation. Zubieta et al. (27) used PET and ^{11}C-carfentanil (a high-affinity μ-opioid agonist tracer) to image μ-opioid-receptor binding in cocaine patients at 1 to 4 days, and then at 4 weeks post cocaine. In comparison with controls, the cocaine patients showed up-regulation of μ-opioid receptors in the caudate, thalamus, anterior cingulate, frontal, and temporal brain regions; these changes persisted for 4 weeks in all but the temporal region. On the early cessation scan, craving (Minnesota Craving Scale on the prior evening, 100-mm visual analogue scale just before the PET, or both) was positively correlated with μ-opioid binding in the amygdala and the anterior cingulate, frontal, and temporal cortex. These four regions all receive significant DA projections, consistent with DA–opioid interactions. Scores on the Beck Depression

Inventory did not correlate with μ-opioid-receptor availability, which indicates that the early cessation craving was not simply a depressed mood. None of the correlations of craving with μ-opioid binding were significant for the later, 4-week scan.

Later Cessation

In abstinence extending beyond a week (imaging at 1 to 6 weeks, and again at 3 to 4 months post cocaine), neuroimaging studies from the laboratory of Volkow et al. (46) have shown that frontal metabolism is decreased in the brains of cocaine users in comparison with that in controls. Although craving (averaging 3 ± 1 on a scale of 0 to 7) was assessed as a subject characteristic in this study, no correlations with metabolic rates were reported. In a further study of the same patient sample, reductions in orbitofrontal cortex and cingulate metabolism were particularly profound, and these reductions were correlated with reductions in (striatal) DA D2-receptor availability (33). However, craving (on a scale of 0 to 10) during the week of the study did not correlate with striatal D2-receptor availability; correlations of craving with metabolic rates were not reported.

Paralleling the metabolic findings of Volkow, Childress et al. found the resting regional cerebral blood flow (rCBF) to be significantly lower in the anterior cingulate (47) and left medial orbitofrontal cortex (gyrus rectus) (48) of cocaine patients at an average of 13.5 days after cessation in comparison with controls. However, neither baseline cocaine craving nor withdrawal (self-rated on a scale of 0 to 9) at the time of the scan correlated with rCBF in these regions.

Summary of Cessation Studies

The neuroimaging data clearly show a number of differences between the brains of cocaine patients undergoing cessation in comparison with controls not using drugs. The observed differences often are clearly linked to brain DA systems. However, the relationship of these brain indices to “cessation craving” has been variable. At this time, correlative evidence from studies in early (≤1 week) cessation, but not from those conducted at longer intervals post drug, indicates a relationship between craving and brain responses (orbitofrontal and prefrontal cortex metabolism, μ-opioid binding).

Because the “early” and “late” cessation studies were often conducted in separate populations, and with important differences across studies (e.g., monitoring of abstinence, inpatient vs. outpatient population), any conclusions must be drawn with caution. Powerful within-subject designs, including frequent scans and subjective measures across a period of prolonged abstinence, would be helpful in clarifying cessation–brain–craving relationships. If such longitudinal studies were to confirm a lack of a relationship between craving and later resting *hypo*activity (by metabolism and rCBF) in DA projection areas, such findings would have implications for both theory and treatment. For example, if low DA tone is not associated with craving, then enhancing DA activity might not help craving, and could even be problematic (49).

Even if unrelated to craving, the differences between cocaine patients and controls that are evident at both early and later time points (e.g., in DA D2 and μ-opioid receptors) may still be very important. They may reflect other functional consequences of cocaine use, or group differences that possibly predate cocaine use or even predispose subjects to such use. The latter possibility has received very recent support. In a study of Volkow et al. (50), initial liking for intravenous methylphenidate (a stimulant also acting by inhibition of reuptake of DA) in normal persons was *inversely* related to D2-receptor availability. The D2-receptor levels for normal subjects who liked methylphenidate were significantly lower than those for normal subjects who did not like the drug, and they were strikingly similar to those of long-term cocaine users in earlier studies by the same investigators (51). Reduced D2 receptors may thus be a marker for vulnerability to stimulant misuse in addition to, or perhaps even instead of, a possible consequence of misuse.

IMAGING OF CRAVING DURING STIMULANT ADMINISTRATION

Part of "110 - Neuroimaging of Cocaine Craving States: Cessation, Stimulant Administration, and Drug Cue Paradigms "

Hypotheses

Administering cocaine in the laboratory is a reliable and robust trigger of cocaine desire (4), and cocaine users complain that the first dose (“taste”) of cocaine in a binge rapidly elicits profound craving, drug-seeking behavior, and a second dose. Almost two decades ago, Eikelboom and Stewart (52) modeled this behavior in rats, showing that small drug “primes” could motivate drug seeking and reinstate extinguished responding for drug. According to the priming hypothesis, the initial drug effect always precedes the full drug effect and comes (through classic conditioning) to trigger a druglike brain state. The state has powerful positive incentive properties, “pulling” the organism back to the drug.

Koob (53) has proposed a different way in which the first dose of cocaine might lead to the next. In this “opponent process” view, the brain responds to cocaine with homeostatic processes, some of which are the hedonic opposite of the direct effect of the drug, euphoria. The opponent response (i.e., dysphoria) emerges as the direct effects of the drug dissipate and motivates drug seeking to reduce discomfort. Clinically, patients do complain about the jittery offset of the cocaine high, and they recognize that taking another dose of the drug will alleviate this state.

So, is the craving associated with human cocaine administration more closely related to a brain state that occurs at the *onset* or at the *offset* of the drug response? Although the question is posed as a choice, these possibilities (unfortunately

for the task of developing medications) are not mutually exclusive. Craving of the positive, appetitive, “primed” variety may map onto brain responses associated with the initial effect of the drug and be followed shortly thereafter by the dysphoric craving of offset, which may map onto a later set of brain responses that are *opposite* in direction to those of drug. The exquisite temporal sensitivity of functional magnetic resonance imaging (fMRI) allows these possibilities to be examined directly, as discussed below.

What are the likely neuroanatomic and neurochemical features of the craving state(s) associated with cocaine administration? More than two decades of animal research (see refs. 25 , 54 , 55 and 56 for review) support the involvement of the mesolimbic and mesocortical DA systems of the brain in cocaine reinforcement and motivation. Thus, *a priori* neuroanatomic predictions include the familiar projections of the DA cells in the ventral tegmental area of the midbrain to the ventral striatum (nucleus accumbens), amygdala, basal forebrain, orbitofrontal cortex, and medial prefrontal/anterior cingulate cortex. And, although recent animal (13) and human (57) research leaves room for the contribution of other brain systems, DA neuronal elements (DATs, postsynaptic D2 receptors) have also been initial neuroimaging targets.

Data

Several laboratories have now imaged cocaine users during stimulant administration, but some of these studies either did not obtain a craving measure (58 ,59 ,60 and 61), did not analyze the craving item (62), or analyzed a craving item but did not attempt to relate it to the brain measure under study (63). The remaining four studies discussed below have been published since 1997.

Breiter et al. (64) used temporally sensitive fMRI technology with a BOLD (blood oxygen level-dependent) scan to map the brain circuitry activated during a period 5 minutes before, and 13 minutes after, cocaine (0.6 mg/kg, maximum dose of 40 mg) versus saline infusion in cocaine subjects abstinent a minimum of 18 hours. Subjective ratings (“rush,” “high,” “craving,” and “low”) were taken each minute throughout the experiment. “Rush” and “craving” (defined as wanting more cocaine) ratings were later correlated with the group-averaged temporal pattern of signals from each brain region meeting specified threshold and extent criteria for differential activation by cocaine. “Craving” ratings increased steadily from the onset of the cocaine infusion, peaking at approximately 12 minutes post cocaine. No activity in any single brain region precisely echoed the onset and late peak of “craving” ratings. However, significant positive correlations were obtained with regions having early-onset (during euphoria) but *sustained* activations. These included the nucleus accumbens/subcallosal cortex and some paralimbic sites (a section of the parahippocampal gyrus, a section of the posterior insula, and a section of the anterior cingulate). Signal change in the amygdala during cocaine administration was initially reported as heterogeneous (some patients showed increases and others showed decreases), not correlated with rush, and negatively correlated with craving ratings. However, in a follow-up study with cardiac gating of the fMRI signal (see below), the direction of signal change in amygdala was positive for all subjects.

In contrast to “craving” ratings, “rush” ratings peaked quickly (at 3 minutes) after infusion and then declined rapidly. Although the brain regions that correlated with “craving” and “rush” overlapped substantially, a clear dissociation was also noted. “Rush,” but not “craving,” correlated with early maximal, short-duration signals from the ventral tegmental area and basal forebrain (and sections of the cingulate). On the other hand, “craving,” but not “rush,” correlated with an early-onset but sustained signal from the nucleus accumbens/subcallosal cortex. All the activated regions showed early onset to cocaine. Thus, the primary difference between “craving” and “rush” (euphoria) substrates was not a matter of which regions were activated, but of how long. Put another way, a full orchestra is playing from the outset of cocaine administration. As the “rush” wanes, some instruments drop out. “Craving” corresponds to the strains of those that play on.

How do these data fit with the “priming” and “opponent process” hypotheses of craving in response to cocaine? Unfortunately, the fit is not completely straightforward for either view. “Craving” mapped onto the activity of structures (nucleus accumbens/subcallosal cortex) with an early but sustained signal. At the first level of examination, this finding seems consistent with a priming effect; the signal occurs very early and therefore looks like a direct drug effect. However, according to the priming hypothesis, “craving” should map better onto the brain correlates (ventral tegmental area and basal forebrain) for the direct effects of “rush” and euphoria. In terms of clear evidence for a simple opponent process view, no later-occurring activations *opposite* to the direction of the “direct” drug effects in ventral tegmental area/basal forebrain, nucleus accumbens/subcallosal cortex, or other brain regions were identified. However, because the direct drug effect was a positive signal change in virtually all brain areas, detecting opposite direction effects would necessitate the detection and interpretation of fMRI signal decreases (“negative signal change”). Because of the physiologic basis of the BOLD signal, the meaning of “negative signal change” is an ongoing research challenge for fMRI.* Until this issue is resolved, PET in combination

with a ^{15}O bolus performed during and after cocaine administration offers sufficient temporal resolution (^{15}O has a half-life of 128 seconds) that it can be used to sort out “early/direct” from any “later/opposed” effects of cocaine. Another possible complication in detecting “opposed” drug effects is that opponent processes can be conditioned; during the course of thousands of cocaine administrations, a response may “move back in time” so that it becomes nearly engaged at drug onset and persists. In experienced users (the only subjects who can be given cocaine in human studies), distinguishing “direct” from “opposed” effects could be very difficult. Animal research mapping the temporal correlates of brain response to cocaine and its signals during the course of initial and repeated administrations could clarify these relationships; of course, such experiments cannot ethically be conducted in humans.

The other three neuroimaging experiments involving stimulant administration to cocaine users and measures of craving have been conducted by the Volkow team at Brookhaven Laboratories. In one of these experiments, cocaine was used as the stimulant probe; in the other two, methylphenidate was used. In the cocaine study, cocaine users were given intravenous cocaine in doses of 0.3 to 0.6 mg/kg (a dose range known to induce euphoria), administered together with a tracer dose of ^{11}C-cocaine to measure DAT occupancy (29). Subjective self-ratings were taken every minute for the first 20 minutes post infusion, and then at 10-minute intervals for the next 40 minutes. The ratings of cocaine “high” and “rush” were positively correlated with DAT occupancy, but “craving” (the desire for cocaine, rated on a scale from 1 to 10) and “restlessness” were not. Thus, as in the earlier study of Breiter et al. (64), “craving” did not map onto precisely the same substrate as “rush” and “high.” This variability among studies may be a consequence of differences in the way craving was measured in each study.

The results from studies of methylphenidate administration suggest that a simple DA hypothesis of stimulant effects (including craving) may be insufficient. One study compared the subjective and brain DA response of cocaine users (3 to 6 weeks post cocaine) with that of controls after an injection of methylphenidate, which (like cocaine) blocks DA reuptake (51). An intravenous injection of 0.5 mg of methylphenidate per kilogram was followed by an injection of ^{11}C-raclopride, a dopamine D2 ligand sensitive to competition by endogenous DA. Regions of interest were the striatum, thalamus, and cerebellum (as a comparison region devoid of D2 receptors). Subjective measures were taken 5 minutes before and 27 minutes after methylphenidate. “Craving” in response to methylphenidate was much greater in the cocaine users than the controls and was correlated with an enhanced response (reduced raclopride binding) in the thalamus. “High” did not correlate with either thalamic or striatal raclopride binding, but a possible correlation may have been compromised by taking the subjective measures at 27 minutes, when the high had waned. Interestingly, both the “high” and the DA response to methylphenidate (measured by raclopride binding) in the striatum were greater in the *controls* than in the cocaine users, as though the cocaine users’ response to methylphenidate had been blunted.

In a second study of cocaine users (abstinent on average for 14 days), two sequential doses (0.5 and 0.25 mg/kg) of methylphenidate were administered 90 minutes apart, after which metabolism (measured by PET and ^{18}F-fluorodeoxyglucose), D2-receptor availability (measured by ^{11}C-raclopride), and subjective responses (27 minutes after each infusion) were determined (66). The actions of methylphenidate on brain metabolism showed substantial variability across subjects that correlated with striatal D2-receptor availability; the stimulant increased metabolism in subjects with a high level of D2-receptor availability and reduced it in subjects with a low level of D2-receptor availability. Although methylphenidate induced metabolic increases in several areas (cingulate, thalamus, cerebellum), it increased right orbitofrontal and right striatal metabolism *only* in the subjects who experienced cocaine craving (“desire for cocaine and perception of loss of control over cocaine”). This observation indicates a possible (lateralized) role for these regions in stimulant-induced craving, but it also shows that an increase in DA secondary to reuptake blockade is not in itself sufficient to induce the metabolic increases in the frontal regions. As in the prior study, ratings of “high” at 27 minutes post methylphenidate did not correlate with the effects of the drug on metabolism or D2-receptor availability.

Summary

The only fMRI study performed during cocaine administration showed widespread brain activation, including activation of the classic mesolimbic–mesocortical circuitry often implicated in the reinforcing and motivational effects of cocaine. Although several brain regions were commonly activated by both craving and rush, rush-associated ventral tegmental area/basal forebrain signals rapidly declined while the craving-associated signals in the nucleus accumbens/subcallosal cortex persisted. This partial dissociation suggests at least some independence of substrates for these two states. The early onset of brain activation in cocaine-induced craving is consistent with the priming hypothesis; the partial dissociation with rush/euphoria is not. The lack of apparent “offset” activations correlated with craving is inconsistent with a simple opponent process view, but conditioned opponent effects could occur very early in highly experienced users, so that their origins would be obscured.

Studies that use a priming dose of cocaine necessarily image both the direct effects of the drug on the system being measured (e.g., DA system) and simultaneously any cognitive–affective–craving brain activity related to the cocaine cue. This dual time course of effects may account for some of the variability or inconsistency in effects reported.

With regard to which dopaminergic neuronal elements

are involved in the craving response, the results are mixed. Cocaine-induced craving was unrelated to DAT occupancy. However, methylphenidate-induced cocaine craving was correlated with right striatal and right orbitofrontal increases in metabolism and with an enhanced response (indexed by ^{11}C-raclopride) in the thalamus in comparison with controls. These findings could support a postulated DA dysfunction in the striatal–thalamic–orbitofrontal circuit in cocaine addiction (33 ,67). One similarity between the craving findings in early cocaine cessation (reviewed earlier) and the craving results in stimulant administration is a positive correlation with orbitofrontal activation. As discussed in the final section, the orbitofrontal cortex is thus far the only region linked to craving in all three paradigms: (early) cessation, stimulant administration, and exposure to drug-related cues.

Based on the neuroimaging data during stimulant administration, a role of transmitters other than DA in methylphenidate-induced effects (both high and craving) seems likely. Methylphenidate alone was insufficient to increase frontal metabolism, and in other studies by the Brookhaven team, significant levels of DAT occupancy by intravenous methylphenidate did not always result in a subjective high in healthy controls (craving was not probed) (68). The studies needed to determine whether these variable effects are also likely for cocaine (namely, giving pharmacologic doses of cocaine to healthy controls) cannot be performed; primate studies offer an important alternative.

IMAGING OF CRAVING DURING COCAINE CUES

Part of "110 - Neuroimaging of Cocaine Craving States: Cessation, Stimulant Administration, and Drug Cue Paradigms "

Hypotheses

A guiding hypothesis of much of the research in cue reactivity is that the powerful craving and arousal responses to drug-related cues are based on simple classic conditioning (5 ,6). According to this view, cocaine cues trigger cocaine-related subjective and physiologic responses, including craving, because they reliably predict the arrival of the direct effects of the drug. As both the animal and human literature has shown, however, “simple” classic conditioning is often anything but simple. Drug-related conditioning can result in both responses similar to those produced by the drug itself (“druglike”) and responses *opposite* to those of the drug (“drug-opposite”) (69), likely reflecting a conditioned compensatory response to drug onslaught. Both kinds of responses may be of motivational significance. Druglike responses to cues reminiscent of the drug (“Wow… It’s like I’m feeling it already… and I haven’t even had any yet! I can’t wait!!”) may act as a powerful positive incentive, pulling the user back to the drug. If drug-opposite responses to the cues are uncomfortable (in opiate users, these include tearing, yawning, sweating, and nausea), they may also prompt drug seeking.

Is cue-elicited cocaine craving more “druglike” or “drug-opposite”? Do the multiple cues surrounding the cocaine experience become linked predominantly to the brief, intense, orgasmic euphoria… or to the jittery, sometimes uncomfortable, offset effects of this very short-acting drug? Both links seem possible, but they would yield opposite predictions for neuroimaging and pharmacotherapies. If craving to common external (paraphernalia, other users, drug-buying locations, drug talk) or internal (e.g., drug dreams, recurrent memories of “high”) cues is predominantly druglike, we would expect that some elements of the mesolimbic–mesocortical DA system activated by cocaine itself are also activated during cue-induced craving. In animals, cues for cocaine can indeed trigger mesolimbic DA overflow in nucleus accumbens and amygdala (70) and can activate c-*fos* (an immediate early gene) in the cingulate (71 ,72), amygdala (72), and nucleus accumbens (71). Patients often describe cocaine-like effects in response to cues, including heart pounding, ear ringing, head buzzing, stomach “flipping,” mild euphoria, a “taste” of cocaine in the back of the throat, even the “smell” of cocaine in the room… and, of course, profound desire. But what do the brains of cocaine users say about the nature of cue-induced cocaine craving?

Data

As shown in Table 110.1 , the first neuroimaging study in which drug cues were used to induce craving was presented in 1992 (73 ,74). It tested whether video-induced cocaine craving might increase endogenous DA, as indexed by competition with a D2 ligand, ^{123}I-iodobenzamide, in SPECT. The low signal-to-noise ratio of ^{123}I-iodobenzamide, the low resolution of SPECT, and timing of the cue activation (after uptake of the tracer) likely undermined the ability of this early study to detect increased DA release in response to cocaine cues. After this initial effort, imaging studies addressed the neuroanatomic rather than the neurochemical substrates of cue-induced craving.

Laboratory	Imaging Technology	Cocaine Population	Days of Cessation	Cue Description	Results[a]
U. Pennsylvania					
Childress et al., 1992 (73,74)	SPECT, ^{123}I-IBZM competition	Treatment-seeking (n = 10; 10 controls)	Range, 6–51	15′ personalized audio followed by 10″ video (separate days for drug/nondrug)	^{123}I-IBZM not displaced from striatum
Childress et al., 1994–1999 (47,79)	PET, ^{15}O bolus (ROI)	Treatment-seeking (n = 14; 6 controls)	Average, 13.5	25′ narrative nondrug video, then 25′ cocaine video (both with soundtrack)	Limbic: amygdala + a. cingulate + OFC 0 hippocampus 0 Comparison: striatum— DLPFC 0 cerebellum 0 visual cortex 0
Childress et al., 1999–2000 (103)	PET, ^{15}O bolus (ROI)	Treatment-seeking (n = 7)	Range, 7–22	(as above)	$GABA_B$ agonist baclofen may blunt limbic activation
Listerud et al., 2000 (80)	fMRI, BOLD	Treatment-seeking (n = 7; 12 controls)	Average, 14.5 Range, 3–38	15′ nondrug video, then 15′ cocaine video	amygdala + a. cingulate +
NIDA ARC					
Grant et al., 1996 (83)	PET, FDG (ROI)	Nontreatment (n = 13; 5 controls)	36–48 h (verified)	10+ repetitions of a 2.5′ silent video plus paraphernalia; snort option (separate days for drug/nondrug cues)	DLPFC +, ^ VLPFC + m. OFC + m. temporal (amygdala) + ^ retrosplenial + temporal/ parietal + temporal + extrastriate/ striate + peristriate + cerebellum ^ r.a. OFC ^(–)
Harvard–McLean					
Maas et al., 1998 (86)	fMRI, BOLD (ROI)	Nontreatment (n = 6; 6 controls)		Alternating 2.5″ drug/ nondrug video clips from Childress tapes (faces blurred)	a. cingulate +, ^ DLPFC +, ^ cerebellum 0
Emory University					
Kilts et al., 1996–2000 (82)	PET, ^{15}O bolus (SPM)	Treatment-seeking (n = 8)	Range, 7–17	1′ guided imagery (by audiotapes of nondrug, cocaine, and angry scenarios, two trials each, same order)	r. amygdala + l. a. cingulate + l. a. insula +, ^(–) r. Nac/SCC +, ^(–) OFC 0 DLPFC 0 Cerebellum 0
Medical College of Wisconsin					
Garavan et al., 1998–2000 (84)	fMRI (AFNI)	Nontreatment (n = 17; 14 controls)		4′ nature video, followed by 4′ cocaine/sex videos; 5″ distraction task in between	amygdala + l. cingulate + caudate + parietal + frontal +
Brookhaven Laboratories					
Wang et al., 1999 (85)	PET, FDG (ROI)	Nontreatment (n = 13)	Average, 7 ± 9 (self-report)	30′ family genogram interview 30′ "cocaine preparation ritual" interview with paraphernalia (separate days for drug/nondrug cues)	OFC + l. insula + r. insula cerebellum + DLPFC 0 a. cingulate
Harvard–Massachussets General					
Breiter et al., 1997–1998 (64)	fMRI, BOLD	Nontreatment (n = 4, retest subgroup)	18 h minimum	saline infusion in an fMRI magnet where cocaine had previously been administered	NAc/SCC + r. insula + OFC +

[a]Results legend: +, activated; 0, no difference; —, deactivated; ^, positive correlation with cocaine craving; ^(–), negative correlation with cocaine craving. When controls were studied, the summary reflects significant difference between groups for the drug cue vs. nondrug cue conditions. When no controls were studied, the effects are only for drug cue vs. nondrug cue condition in cocaine users. Because of space constraints, results are summarized and "no difference" regions are presented only as relevant to discussion in the text. Please refer to the original articles for a complete listing of neuroanatomic regions studied.
Abbreviations: SPECT, single-photon emission computed tomography; ^{123}I-IBZM, iodobenzamide, a D2-receptor ligand; PET, positron emission tomography; ROI, region of interest analysis; OFC, orbitofrontal cortex; DLPFC, dorsolateral prefrontal cortax; VLPFC, ventrolateral prefrontal cortex; fMRI, functional magnetic resonance image; BOLD, blood oxygen level-dependent technique; SPM, statistical parametric mapping technique; FDG, ^{12}F-fluorodeoxyglucose; NAc/SCC, nucleus accumbens/subcallosal cortex; l., left; r., right; a., anterior, m., medial.

TABLE 110.1. BRAIN IMAGING OF CUE-INDUCED COCAINE CRAVING

At least six additional laboratories (Childress et al., Volkow et al., Kilts et al., Grant et al., Garavan et al., Wang et al., Maas et al.) have conducted imaging studies with cocaine cues since 1992. These studies cover a range of imaging technologies (PET with ^{15}O bolus, PET with ^{18}F-fluorodeoxyglucose, fMRI) and include several variations on the method of presenting drug cues to induce craving. The variations are useful because results obtained in only one laboratory, or with one method of cue presentation, may be more related to the specifics of the laboratory or stimuli than to the psychological state of interest (cueinduced cocaine craving). Conversely, any replication and convergence of findings across multiple laboratories and methods are very encouraging.
A survey of the data in Table 110.1 reveals several convergent findings and suggests that brain activation in response

to cocaine cues often occurs in a particular subset of regions activated by cocaine itself (summarized above). Because of the number of studies, the findings reviewed below are organized according to the anatomic structures that have frequently been activated during cue-induced craving.

Limbic-Related Structures

Several of the structures activated during cue-induced craving are parts of an interconnected rostral limbic system, important in motivation and affective experience. Devinsky et al. (75) describe the rostral limbic system as including "the *amygdala* and septum, and *orbitofrontal*, anterior *insula*, and *anterior cingulate* cortices, the ventral striatum including the *nucleus accumbens*, and several brainstem motor nuclei including the periaqueductal gray." We also include findings for the hippocampus, part of the caudal limbic system in Devinsky's organizational scheme. Several of these limbic structures (e.g., amygdala, anterior cingulate and orbitofrontal cortex) comprise subdivisions of functional significance, but most *in vivo* neuroimaging studies refer to the structures in their entirety because of limited spatial resolution.

Amygdala

The amygdala, located in the medial aspect of the temporal lobe, is interconnected with the other rostral limbic regions and with the hippocampus. The amygdala is critical in signal learning for biologically significant (pleasant or unpleasant) events (76 ,77) and has been activated by cue-induced craving in virtually all the studies able to visualize it. In the first PET study of craving, predicted increases in amygdala rCBF (measured with ^{15}O-water as the perfusion tracer) were found in cocaine patients viewing videos that induced craving (averaging 4.5 on a scale of 0 to 9 for "craving or desire for cocaine") (47 ,78 ,79) versus nature videos. This effect was not evident in controls without a cocaine history. Interestingly, baseline rCBF in the amygdala of cocaine users tended to be lower than in controls, such that increased rCBF in response to the cocaine cues did not exceed the amygdala rCBF in control subjects under the same conditions. Activation of the amygdala in this study was initially documented by a region-of-interest (ROI) analysis and has recently been confirmed by statistical parametric mapping (SPM) of the group data (Fig. 110.1 , *top image*). The amygdala activation to cocaine video cues documented with fMRI (Fig. 110.2) has recently been replicated in an ongoing study (Listerud et al., *unpublished data*). In other PET studies, Schweitzer et al. (81) and Kilts et al. (82) found amygdala activation during memory of cocaine-induced craving ("How strong was the urge to use cocaine, on a scale of 0 to 10?") induced by guided drug imagery, and Grant et al. (83) found a positive correlation of medial temporal lobe activation with video- and paraphernalia-induced craving ("… craving or urge to use cocaine, on a scale of 0 to 10"). A recent fMRI study by Garavan et al. (84) found differential activation of the temporal pole, a region surrounding the ventral amygdala, in response to cocaine versus nature video. The PET camera used in the recent ^{18}F-fluorodeoxyglucose study by Wang et al. (85) did not

permit adequate resolution of the amygdala, and susceptibility artifact in the region of the amygdala prevented reliable imaging in an earlier fMRI study by Maas et al. (86).

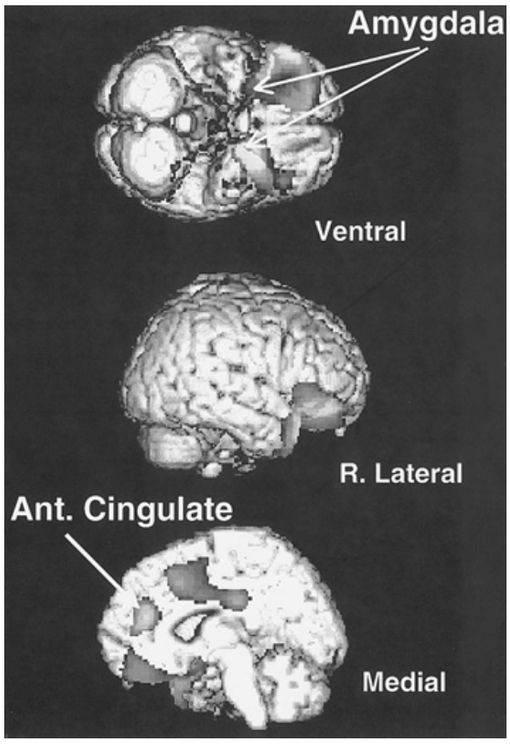


FIGURE 110.1. Amygdala and anterior cingulate activations during craving triggered by cocaine cues. Statistical parametric map (SPM 96) shows differential activation of brain regions by a cocaine video and a nondrug (nature) video; p <.05, corrected. See color version of figure.

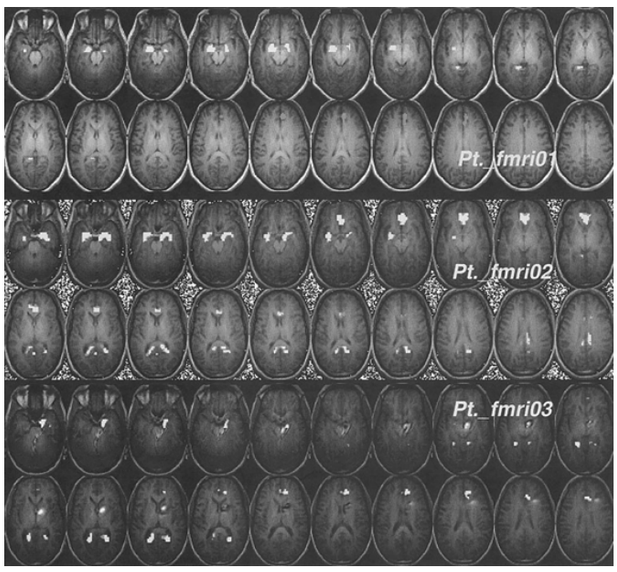


FIGURE 110.2. Functional magnetic resonance imaging of cocaine versus nature video. Individual difference maps show amygdala and anterior cingulate activation in three pilot cocaine patients. See color version of figure.

Anterior Cingulate

The anterior cingulate is located in the dorsomedial prefrontal cortex and is interconnected with other rostral limbic structures, including the amygdala and nucleus accumbens. Multiple roles of the anterior cingulate include selective attention and emotional reactivity to significant stimuli (75 ,87). In parallel with the amygdala findings, the anterior cingulate was differentially activated during video-induced cocaine craving in our initial PET study with ^{15}O bolus (by ROI analysis and later confirmation by SPM analysis of the group data) (Fig. 110.1 , *bottom image*). As with the amygdala effect, the cue-induced rCBF increase in anterior cingulate was from a resting baseline that was hypoactive relative to the baseline of controls. Unpublished fMRI data (Listerud et al.) support the anterior cingulate activation in response to cocaine versus nature cues. Studies by Maas et al. (86), Kilts et al. (82), and Garavan et al. (84) have confirmed anterior cingulate activation during cocaine cues. These studies used fMRI (84 ,86) and PET with ^{15}O bolus (82), imaging techniques that provide good temporal resolution. Studies by Grant et al. (83) and Wang et al. (85) did not detect increased anterior cingulate activation during cocaine cues. They both used PET with ^{18}F-fluorodeoxyglucose; because of its low temporal resolution, this technique may be insensitive to a relatively brief or nonhomogeneous activation of the anterior cingulate.

Nucleus Accumbens/Subcallosal Cortex; Striatum

The nucleus accumbens is located in the forebrain, in the ventral part of the striatum. It is a prominent terminal region for DA cells projecting from the ventral tegmental area, and much animal research points to this mesolimbic–nucleus accumbens pathway as a critical substrate for the reinforcing effects of natural rewards (88), cocaine, and other drugs of abuse (89). In humans, the size of the nucleus accumbens is about 5 mm. Thus, the nucleus accumbens falls at (or often below) the threshold of reliable detection for many early PET cameras, but it often can be localized with the precise anatomic co-registration of fMRI. The 1997 fMRI study of cocaine administration by Breiter et al. (64) contained an "unintentional" cue paradigm (Table 110.1). Subjects given a (double-blinded) infusion of saline solution in the fMRI magnet showed clear activation of the nucleus accumbens if they had previously received an infusion of cocaine in this novel environment. The nucleus accumbens signal to the infusion environment had the same "early-onset, sustained" pattern as the signals to actual cocaine administration. This striking finding suggests that a druglike response to cocaine cues can be established with a single trial.

Other striatal findings in cue paradigms are mixed. In the study performed with PET and ^{15}O bolus, the overall response of the basal ganglia to cocaine versus nature videos was an unpredicted *decrease* in rCBF, but the larger dorsal striatum (which receives primary projections from the substantia nigra) was not parsed from the smaller ventral (nucleus accumbens) portion, which receives projections from the ventral tegmental area (47 ,78 ,79). Garavan et al. (84) found activation of the dorsal striatum (caudate) with fMRI, but did not report on nucleus accumbens. Kilts et al. (82) found nucleus accumbens activation in response to guided cocaine imagery, but the activation was inversely correlated with craving self-reports.

Orbitofrontal Cortex

The orbitofrontal cortex is located in the ventromedial region of the frontal lobes. The orbitofrontal cortex is richly interconnected with DA-related regions involved in reward and stimulus–reward learning (90). Three studies (64 ,83 ,85) found orbitofrontal cortex activation in response to cues; two did not (47 ,82). The remaining two studies, in which fMRI was used, did not report on the orbitofrontal cortex response (84 ,86) (ventral orbital regions are often difficult to image with fMRI because of artifact introduced by air in the sinus cavities). In the three studies finding an orbitofrontal cortex response to cues, the subjects were in early cessation (ranging from 18 hours to 7 days); in the two studies finding no orbitofrontal cortex activation to cues, the patients had been abstinent for longer periods. In the earlier "cessation" section of this chapter, we mentioned that the literature suggests that orbitofrontal cortex hyperactivity in early cocaine cessation is correlated with selfreported craving, whereas (later) orbitofrontal cortex hypoactivity is unrelated to craving. One interpretation that integrates the earlier observations with those in the explicit cue paradigms is that orbitofrontal cortex hyperactivity is associated with enhanced responsivity to cues, whether naturally occurring or presented by a laboratory experiment. Orbitofrontal cortex hypoactivity, on the other hand, clearly does not prevent cue-induced craving and may represent a different vulnerability (see summary below).

Insula

The insular cortex is located interior to the lateral sulcus. It is interconnected with several other regions activated by cocaine cues, including the amygdala and the cingulate and orbitofrontal cortex; it also reflects input from the viscera (autonomic nervous system) and sensory systems. Three laboratories have reported activation of the insula in response to cocaine-related cues, but the effects vary. Wang et al. (85) reported activation of the left insula but found a correlation of craving only with the right insula. Breiter et al. (64) reported activation of the right insula in response to the cocaine infusion environment; no correlation with craving was reported for the small subgroup in this experimental condition. Kilts et al. (82) reported activation of the left insula (in response to guided imagery), but a negative correlation with craving. Given the disparate findings, additional studies will be needed to sort out the nature and direction of cue effects in the insula.

Hippocampus

The hippocampus is located in the medial temporal lobe, posterior to the amygdala. It was not differentially activated by the cocaine videos in our PET study with ^{15}O bolus, and activation has not been reported by the other investigations. This suggests that cue paradigms generally do not make demands on explicit, declarative memory and factual recall, functions closely associated with the hippocampus (91). This is in contrast to the common finding of amygdala activation across several cue studies. Although proximal to the hippocampus and interconnected to it, the amygdala is not activated by explicit memory demands; rather, it supports functions of implicit, emotional memory (92).

Other Structures

Dorsolateral Prefrontal Cortex

The dorsolateral prefrontal cortex, best known for its role in working memory, was not differentially activated by the uninterrupted, narrative cocaine videos in our PET study (47), although craving was robust. Similarly, it was not activated by the paradigm of Kilts et al. (82) with guided imagery, by that of Wang et al. (85) with cocaine theme interviews, or by that of Garavan et al. (84) with an uninterrupted, unrepeated 4-minute video clip.

The paradigm of Maas et al. (86) did activate the dorsolateral

prefrontal cortex; it featured brief alternating exposures to a nondrug and a cocaine video, modified from the tapes of Childress et al. (the faces were blurred to protect the patients' identities). These alternating conditions may have engaged working memory in the cocaine subjects because the same cocaine users reappeared in an ongoing drug scenario that was interrupted by the nondrug video segments. Controls are generally less engaged by cocaine stimuli and therefore would also be expected to show less engagement of working memory. A similar explanation may account for activation of the dorsolateral prefrontal cortex in the paradigm of Grant et al. (83), which featured several repetitions of the same brief cocaine video clip during the period of ^{18}F-fluorodeoxyglucose uptake. In an ongoing fMRI study (Listerud et al., *unpublished data*), activations by uninterrupted versus alternating cocaine videos are being compared within the same cocaine patients. This study will directly test the hypothesis that dorsolateral prefrontal cortex activation in cocaine users is related to the mode of stimulus presentation rather than to cue-induced craving *per se*.

Cerebellum

The cerebellum, important in motor coordination and retention of simple motor schemas, was not activated in our PET study with ^{15}O bolus in which videos were used to induce craving for cocaine (47). Similarly, it was not activated by the cues of Maas et al. (86), Kilts et al. (82), or Garavan et al. (84). The cerebellum was differentially activated in the study of Wang et al. (85), which featured a cocaine theme interview and handling of paraphernalia during the period of ^{18}F-fluorodeoxyglucose uptake. The (highly ritualized and overlearned) handling of cocaine paraphernalia may have triggered motor memories and schemas, a cerebellar function. In support of this notion, the study of Grant et al. (83) also featured handling of paraphernalia as part of the stimulus complex during the period of ^{18}F-fluorodeoxyglucose uptake, and a correlation was found between craving and cerebellar activation. This correlation would occur if handling of paraphernalia acted both as a potent conditioned cue for drug craving and as a trigger for motor memories/highly practiced motor schemas related to cocaine preparation.

Sensory/Association Cortex

In addition to the regions discussed, one imaging study has shown differential activation of visual association areas (peristriate) by cocaine cues (83), and two studies have shown differential activation of the inferior parietal lobe (83 ,84), which is sometimes implicated in working memory. As additional neuroimaging studies accrue, it will be easier to determine whether less common activations such as these reflect a feature of the paradigms used or a feature of the target state.

Summary

The involvement of seven laboratories in neuroimaging cue-induced craving has generated a useful database for drawing preliminary conclusions about the substrates of the state and, importantly, for generating new hypotheses that will help to refine the emerging picture. Despite the variability in imaging techniques, analysis techniques, the abstinence/treatment status of the subjects, and the varied methods used to induce cocaine desire, several convergent findings for regions of activation have been obtained. The most commonly activated regions during cocaine cues, across the laboratories, were the amygdala and anterior cingulate. Two studies that parsed the ventral (nucleus accumbens) from the dorsal striatum showed activation by cocaine cues in the ventral region; the fMRI signal from the nucleus accumbens for a "cocaine-associated environment" (the fMRI magnet!) was strikingly similar to that for cocaine itself. The insula has been activated in at least three cue studies, although the correlations with craving vary in direction. The orbitofrontal cortex was also activated by cues in the (three) studies of cocaine subjects in recent cessation.

The hippocampus was not regularly activated by cocaine cues, which suggests that the cue-induced state does not depend on declarative memory/factual recall. The dorsolateral prefrontal cortex was not activated in most of the cue studies but was activated by cocaine cues that were intermittent or repeated; this activation may be relatively independent of any direct connection to cue-induced craving. The cerebellum was not activated in most cue studies, although it was activated in two paradigms in which paraphernalia handling could have triggered motor memory schemas.

The brain regions activated by cocaine cue paradigms (amygdala, anterior cingulate, nucleus accumbens, insula) do substantially overlap those activated by cocaine itself in the fMRI study of Breiter et al. (64). This is consistent with the druglike phenomena (autonomic arousal, mild euphoria, sense of a cocaine "taste" or "smell") that often accompany cue-induced craving. Most cue paradigms, even those in which fMRI is used, have not yet described the temporal pattern of the signals from these regions during cues; such information would permit a more detailed comparison of cue effects with the multiple effects of cocaine (e.g., "craving" vs. "rush" substrates). No studies have yet examined the ventral tegmental area or basal forebrain in response to cues.

The orbitofrontal cortex deserves special mention in the discussion of craving and drug motivation. Orbitofrontal dysfunction in other disorders has been associated with difficulties in modulating rewarded or punished behavior (e.g., reversing or switching behavior in response to a change in contingencies) (93), impaired somatic/emotional response in anticipation of the consequences of a decision ("future insensitivity") (94), and perseverative, compulsive behaviors (67). Clinically, some of these same difficulties have been

noted in substance abusers, which raises the possibility of a core deficit in some patients (95). Long-term users of amphetamine are impaired on a decision-making task that places demands on ventromedial prefrontal (orbital) function (96), and long-term administration of stimulants clearly erodes orbitostriatal inhibitory function in primates (93). Whether such orbitofrontal cortex deficits predate drug use in humans, predispose to it, or are a consequence of it, the news for long-term cocaine users is not good; they may be at a particular disadvantage in managing their craving for the drug. Interestingly, the reviewed studies link increases in orbitofrontal cortex activity to craving in all three paradigms: (early) cessation, stimulant administration, and response to cues (during early cessation from cocaine). A future challenge is to understand how such activations in a potentially dysfunctional orbitofrontal cortex differ (or not) from normal activation of the orbitofrontal cortex during advantageous decision making (97).

DISCUSSION

Part of "110 - Neuroimaging of Cocaine Craving States: Cessation, Stimulant Administration, and Drug Cue Paradigms "

Theoretic Implications

The neuroimaging of craving states is at a very early stage, and most findings should be replicated before they are taken as either a confirmation or a challenge to theories of addiction and drug motivation. With this caveat kept in mind, some findings from the paradigms reviewed may have implications for current theories.

From the *cessation* paradigms, the preliminary imaging data provide little support for the proposition that craving (drug desire) during cessation arises primarily from a state of prolonged DA depletion or deficiency. Although the DA systems of cocaine patients do differ from those of controls, craving did not show a good relationship to these changes beyond the first week of cessation.

The *drug administration* paradigms show that DA-related regions are activated during stimulant-induced craving, but the temporal pattern of activation fits neither a simple priming (“substrates for craving are the same as the substrates for high”) nor a simple opponent process (“substrates for craving are the opposite of the substrates for high”) view. Studies of craving during methylphenidate administration indicate the possible contribution of non-DA (in addition to DA) systems.

The *cue* paradigms suggest that the regions activated during cue-induced craving often overlap the regions activated by cocaine itself (i.e., the substrate is substantially “druglike”). Of course, these early findings do not preclude “drug-opposite” brain responses or conditioning in response to cocaine cues. The cue data simply suggest that across several laboratory environments (as in patients’ descriptions from the natural environment), cocaine craving with a “druglike” substrate is likely to be evoked.

Interestingly, the current neuroimaging paradigms offer little support for the notion of a “sensitized” substrate, sometimes proposed as a potential mechanism for stimulant drug craving/incentive motivation. Beyond the first week of cessation, cocaine patients often exhibit resting hypoactivity in limbic and frontal regions in comparison with controls. Exposure to cocaine cues or to a stimulant can produce a significant activation in these affected brain regions, but the absolute level of brain response is often no greater than in controls, and may even be less. Of course, the appropriate test for sensitization would require comparing the current responses of cocaine patients with their own initial responses to the drug. This design is unfortunately not feasible. It does, however, highlight a limitation of all our neuroimaging studies in cocaine patients; we usually do not know their baseline before addiction on the brain variables of interest. The discovery of Volkow et al. (50) of striking baseline differences in the availability of D2 receptors in controls (based on liking vs. not liking an initial dose of methylphenidate) is fair warning that the brains of cocaine users who proceed to addiction may differ from those of (some) controls, even before cocaine use begins.

Treatment Implications

Imaging data from the stimulant administration and cue paradigms suggest that craving is often associated with relative increases in the activity of the same brain DA systems that may otherwise be hypoactive in cocaine users. Finding an “anticraving” agent that can modulate the periodic increases in DA in response to cues or drug primes, without worsening symptoms that may be related to chronic hypoactivity (e.g., depressed mood, low levels of energy), poses a particular challenge. Classic DA antagonists (typified by the older antipsychotics) are poor candidates for this modulatory role because they could worsen symptoms related to low levels of DA; they also carry a significant risk for extrapyramidal side effects and tardive dyskinesia.

Partial DA agonists may offer an appealing solution in the future (2 ,98 ,99). These compounds act as agonists under conditions of low DA tone (as may occur in cessation), but as antagonists when the DA concentration increases (as in response to cues or drug primes). The “chameleon-like” nature of partial agonists may possibly offer the cocaine patient a moment-to-moment regulation of the DA system. Unfortunately, for now, no partial agonists (D1, D2, or D3) are approved for humans, although considerable animal research has been done and preliminary safety trials are under way.

Another promising category of DA modulators are the γ-aminobutyric acid type B ($GABA_B$) agonists (100 ,101 ,102 and 103). These compounds may gently modulate the DA system by reducing ventral tegmental area cell firing, thereby reducing DA release in terminal regions. Roberts and colleagues (100 ,101) were the first to demonstrate the blunting of cocaine motivation by the $GABA_B$ agonist baclofen (Fig. 110.3). In

subsequent cocaine-related studies by Dewey et al. (104), the GABA transaminase inhibitor γ-vinyl-GABA also showed promise; its cocaine-related effects are reversible by a $GABA_B$ antagonist. Unpublished preliminary data from Childress et al., testing the ability of baclofen to blunt both subjective and brain responses to cocaine cues by measuring rCBF with PET and ^{15}O bolus, suggest that baclofen, although it has a relatively short half-life, may indeed confer protection against cue-induced craving and the accompanying limbic activation. These data are important because the craving/imaging paradigm is being used to test an "anticraving" medication.

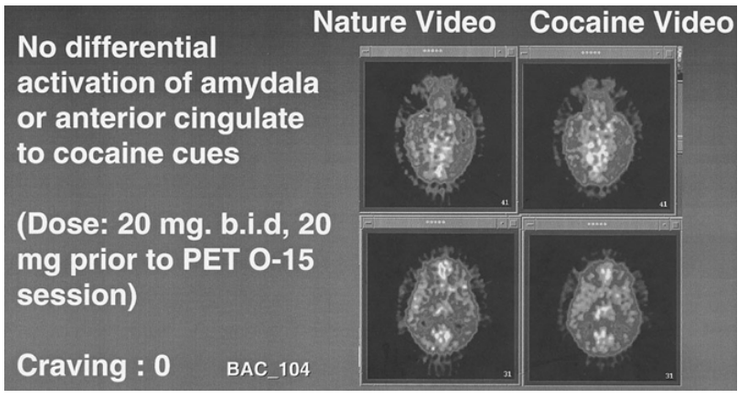


FIGURE 110.3. Baclofen, a γ-aminobutyric acid type B agonist, may blunt limbic activation and craving in response to cocaine cues. See color version of figure.

Future Directions

The neuroimaging studies of cocaine craving reviewed in this chapter will soon be viewed as the "early" stage of our understanding; the imaging field is growing rapidly, as is the sophistication of the tools and their users. Advances in spatial and temporal resolution of imaging devices, and advances in image analysis, will allow the formulation of more precise hypotheses regarding craving substrates. As shown in this review, the future answers are likely to be found within temporal as well as regional patterns of activation. Although DA has played a strong role in shaping the early neurochemical hypotheses, interacting neurotransmitters and neuromodulators will soon be tested as the critical ligands become available. Until then, the combination of pharmacologic probes with neuroanatomic imaging may offer a powerful alternative.

Designs of increased rigor, with attention given to homogeneity of samples (e.g., number of days of cessation, nicotine status, treatment status, urine toxicology status, genetic status, drug use history) and careful characterization of controls, will enhance the replicability of findings across laboratories. Asking for more than one subjective response or "craving item," and asking for these at the optimal time for the paradigm, will ensure a clear test of the relationship between brain activity and subjective state. Characterizing brain activity during other, nondrug states of arousal (e.g., anger, anxiety) will help to determine the specificity of the signature of the brain for cue-induced craving. This is important because the brain structures activated in cueinduced cocaine craving are not "reserved" for this state; rather, they participate in many other states that are not related to cocaine. In this regard, measurement of the brain response during other, nondrug appetitive states (e.g., sexual desire) in subjects who do not use drugs may provide a "positive control" for cocaine craving, which is so often described in sexual terms. Imaging of the craving states for heroin, nicotine, and other drugs of abuse will also provide informative comparisons; these studies have already begun (105).

Although the path ahead is clearly challenging, finding the brain substrates of desire, both for drugs of abuse and for natural rewards, is now a matter of time and effort; the tools are increasingly available. Only a decade ago, and for all of prior human history, brain activity during subjective states was largely a matter of inference. Now, and in the future, these states can be the subject of direct measure. This represents a dramatic paradigm shift, one that enables states such as "desire" and "craving" to be the subject of rigorous scientific research. This research is a critical prerequisite to the rational, and vastly improved, treatment of disorders of desire (i.e., the addictions).

ACKNOWLEDGMENTS

Part of "110 - Neuroimaging of Cocaine Craving States: Cessation, Stimulant Administration, and Drug Cue Paradigms "

We thank V. Mikhalovsky, A. Meyer, and G. Robinson for cyclotron operations and preparation of ^{15}O; S. Riggins, K. Kilroy, D. Dines-Meehan, and C. Herman for PET operations;

W. McElgin for physics consultation; J. Fitzgerald, S. McDonald, J. D. Gray, and A. Fornash for research assistance; and M. Bromwell for editing. This research was supported by research grants NIDA RO1 10241 to Dr. Childress and Core of NIDA P-60 Center to Dr. O'Brien, and by the Medical Research Division of the Philadelphia Department of Veterans Affairs Medical Center.

REFERENCES

1. Kosten TT, McCance E. A review of pharmacotherapies for substance abuse. *Am J Addictions* 1996;5:58–65.

2. Childress AR, O'Brien CP. Dopamine partial agonists may address the duality of cocaine craving. *Trends Pharmacol Sci* 2000.

3. Gawin FH, Kleber HD. Abstinence symptomatology and psychiatric diagnosis in cocaine abusers. *Arch Gen Psychiatry* 1986;43:107–113.

4. Jaffe JJ, et al. Cocaine-induced cocaine craving. *Psychopharmacology* 1989;97:59–64.

5. Childress AR, et al. Conditioned craving and arousal in cocaine addiction: a preliminary report. *NIDA Res Monogr* 1987, 1988;(ADM) 88-1564:74–80.

6. Ehrman RN, et al. Conditioned responses to cocaine-related stimuli in cocaine abuse patients. *Psychopharmacology* 1992;107:523–529.

7. Parsons LH, Smith AD, Justice JB. Basal extracellular dopamine is decreased in the rat nucleus accumbens during abstinence from chronic cocaine. *Synapse* 1991;9:60–65.

8. Markou A, Koob GF. Post-cocaine anhedonia. An animal model of cocaine withdrawal. *Neuropsychopharmacology* 1991;4:17–26.

9. Dackis CA, Gold MS. New concepts in cocaine addiction. The dopamine depletion hypothesis. *Neurosci Biobehav Rev* 1985;9:469–477.

10. Sherer MA, Kumor KM, Jaffe J. Effects of intravenous cocaine are partially attenuated by haloperidol. *Psychiatry Res* 1989;27:117–125.

11. Childress AR, et al. Brain correlates of induced cocaine craving. *Soc Neurosci Abstr* 1993; (abst No. 762.8).

12. Berger SP, et al. Haloperidol antagonism of cue-elicited craving. *Lancet* 1996;347:504–508.

13. Rocha BA, et al. Cocaine self-administration in dopamine transporter knockout mice. *Nat Neurosci* 1998;1:132–137.

14. Sora I, et al. Cocaine reward models: conditioned place preference can be established in dopamine- and in serotonin-knockout mice. *Proc Natl Acad Sci U S A* 1998;95:7699–7704.

15. Glowa J, et al. Phentermine/fenfluramine decreases cocaine self-administration in rhesus monkeys. *Neuroreport* 1997;8:1347–1351.

16. Satel SL, et al. Tryptophan depletion and attenuation of cue-induced craving for cocaine. *Am J Psychiatry* 1995;152:778–783.

17. Kalivas PW, Duffy P. Repeated cocaine administration alters extracellular glutamate in the ventral tegmental area. *J Neurochem* 1998;70:1497–1502.

18. Piazza PV, et al. Stress, glucocorticoids, and mesencephalic dopaminergic neurons: a pathophysiological chain determining vulnerability to psychostimulant abuse. In: *Neurotoxicity and neuropathology associated with cocaine abuse*. NIH publication No. 96-4019, 1994.

19. Goeders N. A neuroendocrine role in cocaine reinforcement. *Psychoneuroendocrinology* 1997;22:237–259.

20. Shaham Y, et al. CP-154,526, a selective, non-peptide antagonist of the corticotropin-releasing factor 1 receptors attenuates stress-induced relapse to drug seeking in cocaine- and heroin-trained rats. *Psychopharmacology* 1998;137:184–190.

21. Spealman RD, Bergman J. Modulation of the discriminative stimulus effects of cocaine by mu and kappa opioids. *J Pharmacol Exp Ther* 1993;261:607–615.

22. Shippenberg TS. Sensitization to the behavioral effects of cocaine: modulation by dynorphin and kappa-opioid receptor agonists. *Pharmacol Biochem Behav* 1997;57:449–455.

23. Roberts DCS, Corcoran ME, Fibiger HC. On the role of ascending catecholamine systems in the self-administration of cocaine. *Pharmacol Biochem Behavior* 1977;6:615–620.

24. Wise RA, Bozarth MA. Brain substrates for reinforcement and drug-self-administration. *Prog Neuropsychopharmacol* 1981;5:467–474.

25. Koob GF, Bloom FE. Cellular and molecular mechanisms of drug dependence. *Science* 1988;242:715–723.

26. Fowler JS, et al. Positron emission studies of dopamine-enhancing drugs. *J Clin Pharmacol* 1999. 13S–16S.

27. Zubieta J-K, et al. Increased mu opioid receptor binding detected by PET in cocaine-dependent men is associated with craving. *Nat Med* 1996;2:1225–1229.

28. Ritz MC, Lamb RJ, et al. Cocaine receptors on dopamine transporters are related to self-administration of cocaine. *Science* 1987;237:1219–1223.

29. Volkow ND, et al. Relationship between subjective effects of cocaine and dopamine transporter occupancy. *Nature* 1997;386:827–830.

30. Baxter LR, Schwartz JM, Phelps M. Localization of neurochemical effects of cocaine and other stimulants in the human brain. *J Clin Psychiatry* 1988;49:23–26.

31. Volkow ND, et al. Cocaine uptake is decreased in the brain of detoxified cocaine abusers. *Neuropsychopharmacology* 1996;14:159–168.

32. Volkow ND, et al. Effects of chronic abuse on postsynaptic dopamine receptors. *Am J Psychiatry* 1990;147:719–724.

33. Volkow ND, et al. Decreased dopamine D2 receptor availability is associated with reduced frontal metabolism in cocaine abusers. *Synapse* 1993;14:169–177.

34. Unterwald EM, et al. Chronic cocaine alters brain mu opioid receptors. *Brain Res* 1992;584:314–318.

35. Hurd YL, Harkenham M. Molecular alterations in the neostriatum of human cocaine addicts. *Synapse* 1993;13:357–369.

36. Hammer RJ. Chronic cocaine alters opiate receptor binding in critical brain reward regions. *Synapse* 1989;8:55–60.

37. Little KY, et al. Cocaine use increases (3-H) WIN 35428 binding sites in human striatum. *Brain Res* 1993;628:17–25.

38. Staley JK, et al. High-affinity cocaine recognition sites on the dopamine transporter are elevated in fatal cocaine overdose victims. *J Pharmacol Exp Ther* 1994;271:1678–1685.

39. Malison RT, et al. Elevated striatal dopamine transporters during acute cocaine abstinence as measured by I-123 B-CIT SPECT. *Am J Psychiatry* 1998;155:832–834.

40. Hitri A, et al. Fewer dopamine transporter receptors in the prefrontal cortex of cocaine users. *Am J Psychiatry* 1994;151:1074–1076.

41. Wilson JM, et al. Striatal dopamine transporter, and vesicular monoamine transporter, in chronic cocaine users. *Ann Neurol* 1996;40:428–439.

42. Markou A, Koob GF. Bromocriptine reverses the elevation in intracranial self-stimulation thresholds observed in a rat model of cocaine withdrawal. *Neuropsychopharmacology* 1992;7:213–224.

43. Hammer RP, et al. Withdrawal following cocaine self-administration decreases regional cerebral metabolic rate in critical brain reward regions. *Synapse* 1993;14:73–80.

44. Trulson ME, Ulissey JJ. Chronic cocaine administration decreases dopamine synthesis rate and increases (H) spiroperidol binding in rat brain. 1987.

45. Volkow ND, et al. Changes in brain glucose metabolism in cocaine dependence and withdrawal. *Am J Psychiatry* 1991;148:621–626.

46. Volkow ND, et al. Long-term frontal brain metabolic changes in cocaine abusers. *Synapse* 1992;11:184–190.

47. Childress AR, et al. Limbic activation during cue-induced cocaine craving. *Am J Psychiatry* 1999;156:11–18.

48. Childress AR, et al. Unpublished data.

49. Gawin F, Riordan C, Kleber HD. Methylphenidate treatment of cocaine abusers without attention deficit disorder: a negative report. *Am J Drug Alcohol Abuse* 1985;11:193–197.

50. Volkow ND, et al. Prediction of reinforcing responses to psychostimulants in humans by brain dopamine D2 receptor levels. *Am J Psychiatry* 1999;156:1440–1443.

51. Volkow ND, et al. Decreased striatal dopaminergic responsiveness in detoxified cocaine abusers. *Nature* 1997;386:830–833.

52. Eikelboom R, Stewart J. Conditioning of drug-induced physiological responses. *Psychol Rev* 1982;89:507–528.

53. Koob GF. Drug addiction: the yin and yang of hedonic homeostasis. *Neuron* 1996;16:893–896.

54. Gardener EL. Brain reward mechanisms. In: Lowinson J, Ruiz P, Millman R, eds. *Comprehensive textbook of substance abuse*. Baltimore: Williams & Wilkins, 1993:70–99.

55. Roberts DCS, Ranaldi R. Effect of dopaminergic drugs on cocaine reinforcement. *Clin Neuropharmacol* 1995;18:S84–S95.

56. Wise RA. Neurobiology of addiction. *Curr Opin Neurobiol* 1996;6:243–251.

57. Volkow ND, et al. Blockade of striatal dopamine transporters by intravenous methylphenidate is not sufficient to induce self-reports of "high." *J Pharmacol Exp Ther* 1999;288:14–20.

58. Pearlson GD, et al. Correlation of acute cocaine-induced changes in local cerebral blood flow with subjective effects. *Am J Psychiatry* 1993;150:495–497.

59. Volkow ND, et al. Is methylphenidate like cocaine? Studies on their pharmacokinetics and distribution in the human brain. *Arch Gen Psychiatry* 1995;52:456–463.

60. Malison RT, et al. Euphorigenic doses of cocaine reduce (^{123}I)ß-CIT SPECT measures of dopamine transporter availability in human cocaine addicts. *Psychopharmacology* 1995;122:358–362.

61. Schlaepfer TE, et al. PET study of competition between intravenous cocaine and C-11 raclopride at dopamine receptors in human subjects. *Arch Gen Psychiatry* 1997;154:1209–1213.

62. London ED, et al. Cocaine-induced reduction of glucose utilization in human brain. *Arch Gen Psychiatry* 1990;47:567–574.

63. Kaufman MJ, et al. Cocaine-induced cerebral vasoconstriction detected in humans with magnetic resonance angiography. *JAMA* 1998;279:376–380.

64. Breiter HC, et al. Acute effects of cocaine on human brain activity and emotion. *Neuron* 1997;19:591–611.

65. Deleted in press.

66. Volkow ND, et al. Association of methylphenidate-induced craving with changes in right striato-orbitofrontal metabolism in cocaine abusers: implications in addiction. *Am J Psychiatry* 1999;156:19–26.

67. Volkow ND, Fowler JS. Addiction, a disease of compulsion and drive: involvement of the orbitofrontal cortex. *Cereb Cortex* 2000;10:318–325.

68. Volkow ND, et al. Relationship between psychostimulant-induced high and dopamine transporter occupancy. *Proc Natl Acad Sci U S A* 1996;93:1–5.

69. Childress AR, et al. Classically conditioned factors in drug dependence. In: Lowinson J, Ruiz P, Millman R, eds. *Comprehensive textbook of substance abuse*. Baltimore: Williams & Wilkins, 1993:56–69.

70. Weiss F, et al. Control of cocaine-seeking behavior by drug-associated stimuli in rats: effects on recovery of extinguished operant-responding and extracellular dopamine levels in amygdala and nucleus accumbens. *Proc Natl Acad Sci* 2000;97:4321–4326.

71. Franklin TR, Druhan JP. Expression of Fos-related antigens in the nucleus accumbens and associated regions following exposure to a cocaine-paired environment. *Eur J Neurosci* 2000;6:2097–2106.

72. Brown EE, Robertson GS, Fibiger HC. Evidence for conditional neuronal activation following exposure to a cocaine-paired environment: role of forebrain limbic structures. *J Neurosci* 1992;12:4112–4121.

73. Childress AR, et al. Brain imaging of withdrawal and craving states in cocaine dependence. In: *Proceedings of the 54th annual meeting of the Committee on Problems of Drug Dependence*, Keystone, Colorado, 1992.

74. Childress AR, et al. Unpublished data.

75. Devinsky O, Morrell MJ, Vogt BA. Contributions of anterior cingulate cortex to behaviour. *Brain* 1995;118:279–306.

76. Everitt BJ, Robbins TW. Amygdala-ventral striatal interactions and reward-related processes. In: Aggleton JP, ed. *The amygdala*. New York: Wiley-Liss, 1992:401–430.

77. LeDoux JE. Emotion and the amygdala. In: Aggleton JP, ed. *The amygdala*. New York: Wiley-Liss, 1992:339–351.

78. Childress AR. Brain imaging during drug-craving states. In: *Symposium on brain imaging in substance abuse*, 56th annual meeting of the Committee on Problems of Drug Dependence, West Palm Beach, FL, 1994.

79. Childress AR, et al. Regional brain blood flow during induced cocaine craving. *Soc Neurosci Abstr* 1994;20:1632(abst No. 666.2).

80. Listerud et al., 2000.

81. Schweitzer J, et al. The neuroanatomy of drug craving in crack cocaine addiction: a PET analysis. *Soc Neurosci Abstr* 1996;22:933(abst No. 365.3).

82. Kilts CD, et al. Functional anatomy of drug craving in human cocaine addiction. *Arch Gen Psychiatry* 2000 *(in press)*.

83. Grant S, et al. Activation of memory circuits during cue-elicited craving. *Proc Natl Acad Sci U S A* 1996;93:12040–12045.

84. Garavan H, et al. Cue-induced cocaine craving: neuroanatomical specificity for drug-users and drug stimuli. *Am J Psychiatry* 2000 *(in press)*.

85. Wang GJ, et al. Regional brain metabolic activation during craving elicited by recall of previous drug experiences. *Life Sci* 1999;64:775–784.

86. Maas LC, et al. Functional MRI of human brain activation during cue-induced cocaine craving. In: *Proceedings of the College on Problems of Drug Dependence*, Nashville, TN, 1996.

87. Vogt BA, Finch DM, Olson CR. Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. *Cereb Cortex* 1992;2:435–443.

88. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science* 1997;275:1593–1599.

89. Koob GF. Neural mechanisms of drug reinforcement. *Ann N Y Acad Sci* 1992;654:171–191.

90. Rolls ET. The orbitofrontal cortex and reward. *Cereb Cortex* 2000;10:284–294.

91. Squire L, et al. Activation of the hippocampus in normal humans: a functional anatomical study of memory. *Proc Natl Acad Sci U S A* 1992;89:1837–1841.

92. LeDoux JE. Emotional memory systems in the brain. *Behav Brain Res* 1993;58:69–79.

93. Jentsch JD, Taylor JR. Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology* 1999;146:373–390.

94. Bechara A, Damasio H, Damasio A. Emotion, decision-making and the orbitofrontal cortex. *Cereb Cortex* 2000;10:295–307.

95. Monterosso. Unpublished data.

96. Rogers RD, et al. Dissociable cognitive deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to the prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. *Neuorpsychopharmacology* 1999;20:323–339.

97. Rogers RD, et al. Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *J Neurosci* 1999;20:9029–9038.

98. Pulvirenti L, Koob GF. Dopamine receptor agonists, partial agonists, and psychostimulant addiction. *Trends Pharmacol Sci* 1994;15:374–379.

99. Pilla M, et al. Selective inhibition of cocaine-seeking behaviour by a partial D3 dopamine receptor agonist. *Nature* 1999;400:371–375.

100. Roberts DCS, Andrews MM, Vickers GJ. Baclofen attenuates the reinforcing effects of cocaine in rats. *Neuropsychopharmacology* 1996;15:417–423.

101. Roberts DCS, Andrews MM. Baclofen suppression of cocaine self-administration: demonstration using a discrete trials procedure. *Psychopharmacology* 1997;131:271–277.

102. Ling W, Shoptaw S, Majewska MD. Baclofen as a cocaine-anti-craving medication: a preliminary clinical study. *Neuropsychopharmacology* 1998;18:403–404.

103. Childress AR, et al. Impact of GABAergics on brain activity during cue-induced cocaine craving. *Soc Neurosci Abstr* 1999;25:815(abst No. 328.14).

104. Dewey SL, et al. A novel strategy for the treatment of cocaine addiction. *Synapse* 1998;30:119–129.

105. Sell LA, et al. Activation of reward circuitry in human opiate addicts. *Eur J Neurosci* 1999;11:1042–1048.

* fMRI is extremely vulnerable to movement artifact. Signals from the amygdala and other structures near the base of the brain can be affected by the slight movement, at each heartbeat, of blood entering the brain through large vessels, and unreliable or uninterpretable data can result. Cardiac gating of the fMRI signal allows the fMRI scanner to be controlled by the heartbeat, and images are collected in the intervals between beats. Under these controlled conditions, the direction of change in the amygdala was positive for all subjects (H. C. Breiter, *personal communication*).

Section XI
Impulsive and Compulsive Disorders

Eric Hollander

Impulsive and Compulsive Disorders - Introduction

The current edition of *Psychopharmacology: the Fifth Generation of Progress* is unique in including a section on impulsive and compulsive disorders. This is genuinely a novel approach to the characterization of a complex, evolving, and rich area of psychopathology and suggests that compulsive and impulsive phenomena may be alternative phenotypic expressions of similar underlying mechanisms. The section focuses on core symptom domains that cut across impulsive and compulsive disorders and links these clinical phenomena to their fundamental neurobiology, as manifested by genetic factors, neurotransmitter/peptide functioning, and neurocircuitry. This approach is of particular clinical interest in leading the clinician to targeted psychopharmacologic approaches to the core symptom domains. In addition, it has important implications for future categoric classification systems such as the Diagnostic and Statistical Manual and the International Classification of Diseases, which eventually must integrate rapidly emerging genotype and neurocircuitry findings with the core phenomenology of impulsive and compulsive disorders.

Rasmussen and Eisen emphasize how the identification of more homogeneous subtypes via factor-analytic, tic-related, or immune phenomena may have important implications for understanding the course of illness and behavioral phenotype. Pauls and colleagues describe how studying genetic marker (and gene product) data together with data characterizing phenotypic expression in the context of specific environments should allow a more complete examination of the simultaneous contribution of genetic and nongenetic factors in obsessive-compulsive disorder. Rosenberg and MacMillan describe how sophisticated brain imaging studies of obsessive-compulsive disorder may help to delineate specific endophenotypes, discriminate between sporadic and familial forms (reducing genetic heterogeneity), and facilitate an understanding of the developmental neurobiologic underpinnings of obsessive-compulsive disorder. Hollander and Pallanti describe the role of targeted pharmacotherapy and other experimental and somatic treatments in modulating the intensity and frequency of repetitive thoughts and behaviors in obsessive-compulsive disorder. Future approaches will need to integrate our understanding of genotype, neurocircuitry, and subtypes and alternative expressions of the behavioral phenotype into experimental therapeutic strategies for obsessive-compulsive disorder.

Smith and Geary describe how greater attention to the microstructure of the behavior of eating has led to advances in the behavioral neuroscience of the controls of eating. Kaye and Walsh describe psychopharmacologic approaches to the clinical eating disorders and their impact on aberrant feeding behaviors and perception of body image. Swerdlow and Leckman describe efforts to link the clinical phenotype of Tourette syndrome to advances in neuropathology, neuroimaging, genetic linkage, and informative animal models, and the use of endophenotypes for a full understanding of the functional relevance of genes for Tourette syndrome.

Olivier and Young describe how new developments in molecular biology, used to generate inducible and brain region-specific mutants, have provided novel tools with which to study the role of genes, the environment, and their interaction in the causation of aggression, and the neural substrates that mediate this behavior. Coccaro and Siever describe how a better understanding of the role of serotonin and other neuromodulators

in the regulation of aggression has led to a more rational approach to the psychopharmacology of impulsive aggression in humans. Potenza and Hollander describe the neurobiology and treatment of the impulse control disorders and pathologic gambling, disorders linked by a failure to resist urges to engage in pleasurable but ultimately self-destructive behaviors. Stein and colleagues explore a broad range of self-injurious behaviors, ranging from anxiety-reducing compulsive to pleasurable impulsive variants. They draw parallels with animal stereotypies and propose neuropsychopharmacologic approaches derived from an understanding of the neural contributions to these behaviors.

111

The Course and Clinical Features of Obsessive-Compulsive Disorder

Steven A. Rasmussen

Jane L. Eisen

Steven A. Rasmussen: Brown University, Butler Hospital, Providence, Rhode Island.

Jane L. Eisen: Department of Psychiatry and Human Behavior, Brown Medical School, Providence, Rhode Island.

Twenty years have passed since the landmark National Epidemiology Catchment Area Survey first demonstrated the prevalence of obsessive-compulsive disorder (OCD) in the general population to be 50 to 100 times greater than had been previously believed (1). This unexpected finding was instrumental in the renewed interest in and rapid growth of our understanding of the clinical features, pathophysiology, and treatment of OCD. Epidemiologic studies in different cultures have confirmed the findings that up to 1% to 2% of the general population worldwide suffer from the disorder at any given time (2). Widespread attention in the media, in addition to growing recognition of the disorder among health care professionals, has resulted in improvements in the diagnosis and treatment of large numbers of patients with OCD who would not even have presented for treatment before 1980.

Knowledge of the clinical features of the disorder has also expanded significantly in the last 10 years. Treatment centers specializing in OCD have succeeded in enrolling large cohorts of patients, so that a more sophisticated analysis of the heterogeneity and comorbidity of OCD and the relationship of these variables to treatment outcome has been possible. Prospective observational studies of the longitudinal course of OCD have contributed further insights into the clinical characteristics and prognosis of the illness (3). Improvements in methodology, including the development of structured interviews with proven reliability and validity, the application of survival analysis and other statistical techniques to assess longitudinal variables, and more sophisticated database management systems, have been instrumental in these advances.

Epidemiologic studies have consistently shown that 2% to 3% of the general population in the United States meet lifetime DSM criteria for OCD (4). In a World Health Organization study that determined the leading causes of mortality and morbidity in developed countries, OCD was found to be the eighth leading cause of disability for any medical or psychiatric condition for ages 15 through 44 (5). Total costs of the disorder in the United States have been estimated at $8 billion in 1990, including $2.1 billion in direct costs and $5.9 billion in indirect costs related to lost productivity (6).

However, despite the increased recognition of the public health significance of OCD during the last decade, surprisingly little is known about the long-term course and prognosis of the disorder. Most studies conducted thus far suggest that OCD is chronic and lifelong. For several reasons, however, questions have been raised about the validity of these findings. Previous studies have been hampered by a number of methodologic limitations, including a lack of standardized assessments, small numbers of subjects, and a sample bias toward more severely ill patients. The introduction of effective treatments for OCD in the last 10 years also raises the question of the relevance of course studies conducted in a pretreatment era.

Obsessive-compulsive disorder spans the life cycle. It has been described in children as young as age 2 (7) and also in the very elderly (8). Evidence supports the hypothesis that OCD is a heterogeneous disorder with multiple causes (9). Neurobiologic studies have demonstrated abnormalities in frontostriatal–basal ganglia circuitry (10). Like any organ system, these neural circuits are susceptible to a variety of pathologic processes, including those associated with autoimmune, infectious, developmental/genetic, and aging processes. Identifying homogeneous subgroups of patients with OCD should help in unraveling its neurobiologic pathogenesis and developing more specific and effective treatment strategies.

This chapter reviews data related to the clinical features

and course of OCD during the lifespan. It focuses on the heterogeneity and comorbidity of the disorder in relation to its course, and points to a new wave of studies that should complement neurobiologic and genetic studies of the pathogenesis of OCD, lead to fuller recognition of its impact on society, and help to measure the effectiveness of behavioral and pharmacologic treatment strategies that have been developed during the past two decades.

SUBTHRESHOLD SYMPTOMS

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

It is generally agreed that it is the frequency of obsessions and compulsions, in addition to the degree with which they interfere with function, that distinguishes normal from abnormal. A patient must have had an hour of obsessive-compulsive symptoms daily for a period of 6 months that interfere with social or occupational function to meet DSM-IV criteria for the disorder (11). This requirement has traditionally been thought to translate to a score of 16 or higher on the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS). Like symptoms of anxiety, obsessive-compulsive symptoms are present to some degree in most people. Rachman and Hodgson (12) found that a high percentage of the normal population report some obsessions and compulsions. Similarly, after screening 861 Israeli military recruits at 16 years of age, Apter et al. (13) concluded that obsessive-compulsive phenomena appear on a continuum, with few symptoms and minimal severity at one end and many symptoms and severe impairment on the other. The receiver operating characteristics that would best distinguish the clinical from the subthreshold syndrome of OCD have yet to be delineated. Using Angst's longitudinal follow-up sample, Degonda et al. (14) found a weighted lifetime prevalence for subthreshold obsessive-compulsive symptoms at age 30 of 5.5%. Goodman (15) screened 958 college students and identified 23 subjects with subclinical OCD. At follow-up 1 year later, 87% continued to have significant symptoms. It has been recognized for many years that most normal children go through developmental stages characterized by obsessive-compulsive or superstitious behavior (16). Determining where the clinical syndrome begins and ends is important for pharmacologic and genetic studies. For example, the multicenter collaborative studies of the selective serotonin reuptake inhibitors (SSRIs) in OCD noted a higher rate of response to placebo in patients with Y-BOCS scores between 16 and 20, a finding that prompting some investigators to suggest that patients with Y-BOCS scores below 20 be excluded from controlled trials (17). Family genetic studies have shown a higher risk for both subthreshold and clinical OCD in OCD probands (18).

Most adult patients who meet DSM criteria for OCD remember subthreshold symptoms in childhood. The clinical significance of subthreshold symptoms in childhood continues to be poorly understood. The risk carried by children of parents with OCD for subsequent development of the disorder is poorly defined. No data are available that would make it possible to predict this transition. Similarly, almost no data are available relating the effect of continuing subthreshold symptoms during a period of remission to the likelihood of relapse in adults. Prospective quantitative longitudinal assessment of probands with subthreshold symptoms is needed in child and adult populations.

DEVELOPMENTAL PSYCHOPATHOLOGY

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

Little systematic study of the developmental antecedents of OCD has been carried out since Janet. In his *Obsessions and Psychasthenia*, Janet (19) postulated that obsessions and compulsions are the most severe stage of an underlying prodromal state that he called *psychasthenia*, a syndrome characterized by feelings of incompleteness and imperfection. He hypothesized that all patients in whom obsessions and compulsions develop pass through a prodromal stage of psychasthenia. His clinical descriptions of the temperamental features of psychasthenics coincide remarkably well with our preliminary findings of the prodromal symptoms of patients with OCD. His description of the patient who "finds on the stairway the word that needed to be said in the parlor" is an astute clinical description and close analogue of the independent variable chosen by Kagan et al. (20) to measure behavioral inhibition (i.e., speech latency in a novel social situation). It is worth noting that Janet included three of the five elements of DSM-III compulsive personality disorder in his description of the psychasthenic state: perfectionism, restricted emotional expression, and indecisiveness. Previous studies have shown that a considerable portion if not the majority of patients with OCD do not meet the DSM-III-R criteria for compulsive personality disorder. The European diagnostic schema for anacastic personality is more directly related to Janet's original definition of psychasthenia and is consistent with the idea of an obsessive spectrum that ranges from normal obsessional behavior through obsessional personality to OCD.

A retrospective study of 90 of our OC probands in which a semistructured format was used was designed to elicit prodromal personality traits or temperamental factors commonly found in OCD (22). During this study, we identified 10 factors commonly found in our adult OC probands as children (Table 111.1). These traits tended to vary minimally during the childhood and adolescent years.

Separation anxiety
Resistance to change or novelty
Risk aversion
Submissiveness (compliance)
Sensitivity
Anacastic
Perfectionism
Hypermorality
Ambivalence
Excess devotion to work

TABLE 111.1. BEHAVIORAL INHIBITION

The developmental antecedents of OCD overlap significantly with the behavioral inhibition syndrome in children that Kagan et al. (20) described. Four of the developmental traits appear to be shared by patients with OCD and those with other major anxiety disorders: separation anxiety, resistance to change or novelty, risk aversion, and submissiveness. Four of the traits are more specific to OCD: perfectionism, ambivalence, excess devotion to work, and

excessive morality. The overlap of the developmental antecedents of panic disorder, social phobia, and OCD is consistent with Janet's original conception of the psychasthenic syndrome and adds credibility to the hypothesis that an element of genetic vulnerability is shared among the anxiety disorders. The relationship of adult personality characteristics and clinical subtypes to developmental antecedents awaits further analysis. It appears that some traits are more commonly seen in particular phenomenologic presentations (e.g., incompleteness in perfectionism and the need for symmetry and precision; abnormal risk assessment in high levels of anxiety). It is probable that temperamental factors such as behavioral inhibition increase the risk for the development of a number of psychiatric syndromes. It would be informative to determine the relative risk for development of each of the major anxiety syndromes by following a group of children with behavioral inhibition longitudinally. The environmental and genetic factors that predispose a given individual to the development of a specific anxiety disorder are unknown. It is also worth noting that a significant minority of patients with OCD do not manifest risk-aversive tendencies as children. Further prospective study of the developmental antecedents of OCD and prospective longitudinal evaluation of children at risk should be an important area for future research.

AGE AT ONSET

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

In most studies of the course of illness, *age at onset* refers to the time that symptoms become severe enough that they meet full DSM criteria for the disorder. The reliability of retrospective recall is an inescapable problem. It is safe to assume that reliability decreases as the years between ascertainment and onset increase. In the Brown cohort drawn from an adult OCD clinic, the mean age at onset of significant OCD symptoms was 20.9 ± 9.6 years, with males having a significantly earlier onset of illness, 19.5 ± 9.2 years, than females, 22.0 ± 9.8 years (p <.003) (212). The illness developed before the age of 25 years in 65% of cases, sometimes as early as 2 years. It developed after age 35 in fewer than 15% of obsessional patients (Fig. 111.1). A significant increase in incidence appeared at puberty. Most adult patients remembered having minor obsessive-compulsive symptoms that did not significantly interfere with their ability to function and that did not cause significant distress before the onset of symptoms meeting DSM-III-R criteria for the disorder. Although male patients noticed minor symptoms earlier than female patients, the difference did not reach statistical significance. Most of the patients described a gradual or insidious onset of illness. Emerging data suggest that a considerable percentage of patients with an early, prepubertal onset have an acute attack followed by an episodic course (22). These patients frequently suffer at the same time from multiple tics and other movement disorders, including choreiform movements and behavioral dysregulation. Swedo et al. (23) systematically characterized 50 children with this cluster of symptoms, which they call *pediatric autoimmune neuropsychiatric disorder* (PANDAS). A diagnosis of PANDAS is made if the following criteria are met: (a) the presence of OCD, a tic disorder, or both; (b) prepubertal onset of symptoms; (c) episodic course with

varying symptom severity; (d) dramatic exacerbation of symptoms following a group A ß-hemolytic streptococcal infection; and (e) association with neurologic abnormalities. In these children, the average age at onset was 6.3 years for tics and 7.4 years for obsessive-compulsive symptoms. The longitudinal course of children with PANDAS and how they differ from patients in whom OCD develops but who do not meet the criteria for PANDAS is unclear.

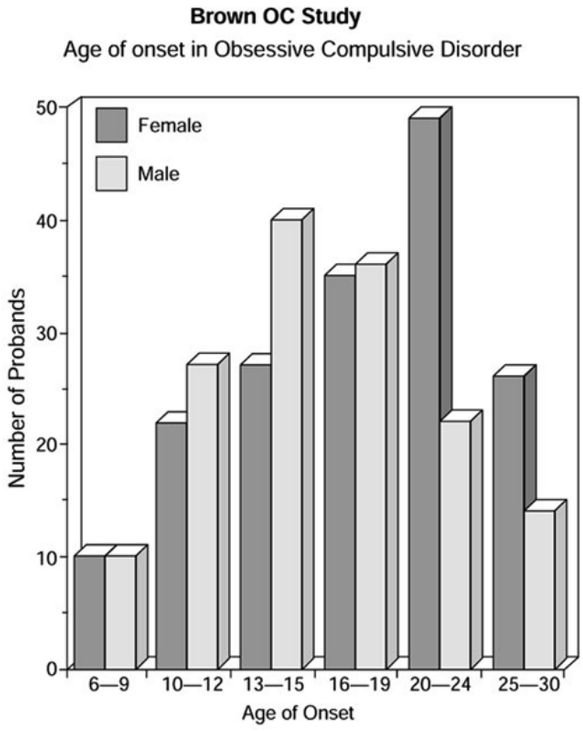


FIGURE 111.1. Age at onset in obsessive-compulsive disorder in the Brown study. (From Jenike MA, Baer L, Minichiello WE. An overview of obsessive-compulsive disorder. In: Jenike MA, Baer L, Minichiello WE, eds. *Obsessive-compulsive disorders practical management, third ed. St. Louis: Mosby, 1998;3–11, with permission*.)

NATURAL HISTORY AND COURSE OF ILLNESS

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

DSM-IV describes the course of OCD as typically chronic with some fluctuation in the severity of symptoms over time. The numerous retrospective and prospective follow-up studies of patients with OCD support this description. However, many of the earlier phenomenologic and follow-up studies of OCD suffered from a number of methodologic limitations, including the following: retrospective study design, small sample size, lack of standardized criteria to determine diagnosis, hospital-based samples not representative of the spectrum of the disorder found in the general population, biases in inclusion and exclusion criteria, chart review rather than personal interview, absence of structured interviews, and lack of consensus regarding the definition of relapse, remission, and recovery. Because of these flaws in design, the earlier studies of OCD may have included subjects who would not meet today's criteria for the diagnosis. In particular, clear distinctions between OCD and compulsive personality disorder were often not made, and obsessions and compulsions occurring in the context of other disorders (e.g., psychosis, eating disorders) may have been included as OCD.

Despite these methodologic shortcomings, several more recent prospective follow-up studies, in which a prospective design, standardized criteria to assess diagnosis, and structured interviews with direct patient contact were used, have also shown that most patients continue to meet either all or some of the criteria for the disorder at follow-up. Relatively few patients experience complete remission. Retrospective and prospective follow-up studies of the course are reviewed in detail below.

Retrospective Follow-up Studies

In retrospective studies, fluctuations in the severity of psychiatric symptoms and impact on functioning over time are ascertained primarily on the basis of subjects' recall. Results of these studies are summarized in Table 111.2 . In most of them, patients were selected based on chart review and were subsequently assessed at the time of the study, either in person or through questionnaire. In the earliest longitudinal study of OCD, a relatively good outcome was observed by Lewis (24), who followed 50 patients with OCD (most of whom received some psychotherapy) at least 5 years after initial assessment; 37% were "quite well," 14% were "much improved," but 46% were minimally improved, unchanged, or worse. Only 10% had had an episodic course marked by later recurrence after remission. Pollitt (25) followed 67

nonleucotomized patients for a mean of 3.4 years; 24% were symptom-free (similar to the results with psychotherapy), 36% had mild symptoms and were functioning well, and 12% were improved but with impaired functioning. Only 25% had symptoms that were unchanged or more severe than at baseline. This study was somewhat unusual because most of the patients were selected from an outpatient practice. The results of this study illustrate how outcome is influenced by the baseline severity of the obsessive-compulsive symptoms of the cohort selected. A longer duration of illness at initial evaluation was associated a with poorer outcome with respect to severity of symptoms at follow-up, as might be expected. Duration of illness was also a predictor of course of illness in a study of 29 inpatients with obsessional symptoms followed for 6 years by Ingram (26). In this study, 72% were minimally improved but functioning poorly, unchanged, or worse, and 21% of the patients were much improved. One conclusion that can be drawn is that chronicity at entry appears to predict chronicity at follow-up.

TABLE 111.2. RESTROSPECTIVE FOLLOW-UP STUDIES OF OBSESSIVE-COMPULSIVE DISORDER

Study (Ref.)	No. Patients	Mean Years of Follow-up	Well (%)	Minimally Improved, Unchanged or Much Improved (%)	Worse (%)	Comments
Lewis, 1936 (24)	50	>5	32	14	44	Episodic course in 10%
Pollitt, 1957 (25)	67	3.4	24	36	37	Mostly outpatients
Ingram, 1961 (26)	29	5.9	7	21[a]	72	Inpatients
Kringlen, 1965 (27)	80	13–20	0	24	76	Inpatients
Grimshaw, 1965	100	5	40	24	35	Inpatients
Coryell, 1981 (31)	44	.5+	8	20	8	Inpatients
Thomsen, 1993 (36)	47	6–22	28	26	46	Childhood OCD
Lo, 1967 (28)	88	3.9	23	50	27	Inpatients and outpatients diagnostic heterogeneity

[a]One patient not leucotomized; five patients leucotomized.
OCD, obsessive-compulsive disorder.

In a study characterized by a long follow-up period, Kringlen (27) found that at 13 to 20 years after initial contact, 42% of patients were unimproved or had worsened symptoms, only 24% were much improved, and 34% described slight improvement in OC symptoms. The patients included in this study were all hospitalized for their first contact, which may contribute to the poorer outcome in this study.

Lo (28) interviewed 88 patients in whom OCD had been diagnosed with a mean follow-up of 3.9 years and found that 23% were symptom-free and 50% had symptoms that were much improved. More than half the patients had distinct obsessions and compulsions. However, 10% had prominent affective symptoms, and 31% were described as having “phobic and ruminative symptoms,” with minimal compulsions. Therefore, some of the patients described as being in remission at follow-up may have had major depression with obsessional or ruminative thinking during their index episode. In reviewing these early follow-up studies, Goodwin et al. (29) concluded that the course of OCD is usually chronic, but variable, with fluctuations in the severity of symptoms.

In follow-up studies conducted since 1980, the course of illness has been evaluated according to criteria different from those used in the earlier studies described above. Patients have been retrospectively assigned to categories of “continuous,” “waxing and waning,” “deteriorative,” and “episodic with full remissions between episodes.” Rasmussen and Tsuang (30) conducted a study in 1986 in which patients were selected based on current enrollment in an outpatient OCD clinic. The course of OCD was described by most patients as chronic or “continuous” (84% of 44 patients); six subjects (14%) had a deteriorating course, and only one (2%) had an episodic course. The average duration of illness at the time of assessment was more than 15 years, which again suggests that the chronicity of the disorder may have been influenced by the sample. Because these subjects were acquired through the process of clinic referral and prospective follow-up was not conducted, no former OCD patients who had already recovered and remained well were included.

Coryell (31) observed some improvement at follow-up in 55.6% of a hospitalized cohort of patients with OCD. However, this cohort was significantly less likely to experience remission after discharge (22%) than the comparison cohort of depressed patients (64%).

Synthesizing methodologically varied studies, some of which present an optimistic picture, others a pessimistic one, may require more careful examination of reported outcomes. It is particularly important to separate the best possible outcome (“full remission” or “symptom-free”) from what is described as “much improved” or “improved,” which may indicate persistent symptoms in the abatement phase of a chronic illness that waxes and wanes. The episodic pattern of full remission (and sometimes later occurrence), when it is clearly identified as such, appears to occur in about 10% to 15% of patients with OCD, although this proportion may increase somewhat as follow-up is extended for several years and may also be greater in childhood OCD (12), in which improvement can be rapid even without treatment (32). In most studies, a smaller proportion of patients (6% to 14%) seem to follow a deteriorating course. Most follow a course marked by chronicity, with some fluctuation of symptoms over time but without clear remissions or deterioration.

Prospective Longitudinal Studies of Course

During the past decade, several prospective longitudinal studies of the course of OCD have been carried out; these are summarized in Table 111.3 . Although studies of adults have supported the hypothesis that OCD is a chronic, lifelong disorder, child and adolescent studies have found a surprisingly high percentage of patients with an episodic course. Flament et al. (33) completed a 2-year follow-up study of 59 adolescents in whom OCD, subclinical OCD, or compulsive personality disorder had been identified in an epidemiologic study of high school students, most of whom had not sought clinical treatment. Of 12 patients who had met the criteria at baseline for OCD, only five still met the full criteria at follow-up. Four patients with subclinical manifestations of OCD at baseline did meet the full criteria for OCD at follow-up. In another 5-year prospective follow-up study, of an OC adolescent cohort seeking treatment at a tertiary clinic, Flament et al. (34) concluded that patterns of course are not easily predicted from baseline variables (34). Some patients with subthreshold symptoms at baseline were severely ill at follow-up, whereas others classified at baseline as severely ill no longer had

clinical levels of symptomatology at follow-up. This finding was substantiated in a recent study by Valleni-Basile et al. (35). When they screened a community sample of 3,283 adolescents with a self-report instrument followed by the Schedule for Affective Disorders and Schizophrenia (SADS), they found 1-year incidence rates of OCD and subthreshold OCD of 0.7% and 8.4%, respectively. Interestingly, transition probabilities demonstrated a pattern of moving from more severe to less severe categories in subsequent years. Of the patients with OCD at baseline, 17% had OCD at follow-up. Only 1.5% of those with subclinical OCD had progressed to OCD that met syndromal criteria. In contrast, in a Danish follow-up study of 23 adolescents presenting with OCD to a community clinic, half of the subjects retained an OCD diagnosis at follow-up. One-third of the subjects had had an episodic course, and two-thirds had had a chronic course (36). Berg et al. (37) reported a 2-year follow-up of 59 high school students with DSM criteria for OCD who were identified as part of an epidemiologic survey. Most of the subjects had never sought treatment. The course of illness was much more variable than had originally been predicted. Some patients with subthreshold symptoms at baseline were severely ill at follow-up, whereas others classified as severely ill at baseline no longer had clinical levels of symptomatology at follow-up. Although these studies have given us our first prospective glimpse of the early course of OCD, they also suffer from significant methodologic limitations. Follow-up was at a single point, an average of 4.8 years from baseline. Interim data about remissions and relapses during the study period were not obtained. However, the evidence suggests that the course of illness may be much more variable and episodic in child and adolescent samples than was previously believed.

Study (Ref.)	Treatments	No. Patients	Mean Follow-up (ys)	Remained in Episode (%)	Partial Remission (%)	Full Remission (%)	Comment
Children and adolescents							
Berg et al., 1989 (37)		12	2	42	17[a]	8	17% had compulsive personality traits
Leonard et al., 1993 (43)	SSRIs, BT, psychotherapy, family therapy	54	3.4	43	46	11[b]	70% on medication at follow-up
Adults							
Orloff et al., 1994 (44)	SSRIs, BT	85	2.1			33	
Eisen et al., 1995	SSRIs, BT	51	2	57	31	12	
Stekette et al., 1996	SSRIs, BT	107	0.5–5	47	31	22	mainly outpatients

[a] Subjects had subclinical OCD at follow-up (i.e., obsessions compulsions were present but not at full criteria).
[b] Three of the six subjects in remission (i.e., symptom-free) were receiving medication.
BT, behavioral therapy; OCD, obsessive-compulsive disorder; SSRIs, selective serotonin reuptake inhibitors; Y-BOCS, Yale–Brown Obsessive-Compulsive Scale.

TABLE 111.3. PROSPECTIVE FOLLOW-UP STUDIES OF OBSESSIVE-COMPULSIVE DISORDER

In a three-site prospective longitudinal study of adult patients with OCD conducted by Eisen et al. (38), data were collected on the course of illness in 78 subjects for 2 years. Two instruments with proven reliability and validity were used to evaluate severity of symptoms: the Y-BOCS (39) and the Psychiatric Rating Scale for OCD (PSR-OC) (40). On the PSR-OC, scores ranged from 6 for patients who were severely symptomatic and unable to function at work or socially to 0 for patients who had no obsessive-compulsive symptoms and used no avoidance. Follow-up measures were obtained at 3, 6, 12, and 24 months after baseline assessment.

The probability of achieving at least partial remission during the 2-year study period was 47%. However, if more stringent criteria were used to define remission, in which patients had only occasional or no obsessions and compulsions for 8 consecutive weeks (PSR-OC score ≤2, which is equivalent to a Y-BOCS score ≤8), the probability of achieving remission was only 12%. Once a patient was in remission, the probability of subsequent relapse (defined as returning to a Y-BOCS score ≥16 and a PSR-OC score >4 for any length of time) was 48%. Of the 22 patients who achieved partial remission, 10 relapsed and 12 remained in partial remission throughout the study.

In another prospective study, 107 clinic patients with OCD were followed for up to 5 years after intake (41). The probability of full remission for at least a 2-month period was 22% at 5 years, and the probability of partial remission was 53%. Although outcome in this study was assessed with

a 3-point rating scale, the results are comparable with those in the study of Eisen et al. (38), in which a 6-point PSR-OC and the Y-BOCS were used.

Skoog and Skoog (42) recently described a 40-year follow-up study of 144 patients with OCD who were identified as inpatients in the late 1940s and early 1950s. Two-thirds were improved within a decade after the onset of OCD, and most of the patients reported an intermittent course, with at least two remissions during that time period. However, a chronic course was more common in the later follow-up period, and 20% showed either no improvement or a deteriorative course during the 40 years. Although the length of follow-up in this study was remarkable, methodologic flaws limit the conclusions that can be drawn. First, the sample consisted of psychiatric inpatients hospitalized in the 1940s. The baseline severity of obsessive-compulsive symptoms was unclear because of the lack of scales with proven reliability and validity. It seems likely that patients with a primary diagnosis of major depression were included. Finally, the study was conducted before the widespread availability of the SSRIs and behavioral treatments.

In summary, only a handful of prospective studies of the course of illness in OCD are available. A significantly greater degree of episodic illness is seen in child and adolescent samples than in the adult population. A number of methodologic considerations may account for some of these inconsistencies. The earlier retrospective studies were completed before the introduction of standardized diagnostic criteria, standardized ratings of symptom severity, and effective pharmacologic and behavioral treatment strategies. In addition, because until recently patients with OCD were reluctant to seek treatment, patients with more debilitating symptoms may have been overrepresented in these earlier studies, so that the results are biased toward a worse prognosis. In our pilot study, patients were followed who were already enrolled in our clinic, a factor that potentially contributed to the chronic course noted in many of the subjects. A prospective longitudinal study of the course of 400 patients with OCD is currently in progress.

Effect of Treatment on Course of Illness

Effective pharmacologic and behavioral treatments for OCD became available in the late 1980s in the United States. A follow-up study of children with OCD was conducted by Leonard et al. (43) to determine outcome after standardized short-term treatment with clomipramine (a medication known to be effective in OCD). Fifty-four children and adolescents were re-interviewed 2 to 7 years after participation in a controlled trial of clomipramine and a variety of interim interventions. Obsessive-compulsive symptoms were more severe in only 10 of the subjects at reassessment, so that as a whole, the cohort had improved at follow-up. However, only three subjects (6%) were considered to be in true remission (defined as no obsessions or compulsions and no medication), and 23 subjects (43%) still met full criteria for OCD. Most of the patients were taking medication at follow-up. It is worth noting that the patients who made up this sample were referred to a tertiary research center and were severely ill with a more chronic course than is seen in most childhood samples of OCD.

The results of a 1994 study conducted by Orloff et al. (44) are a greater cause for optimism than those of the studies described above. Most of the 85 subjects assessed 1 to 3 years after baseline evaluations were much improved at follow-up based on chart review. The mean follow-up Y-BOCS score of 9.3 was in the range of mild to minimal obsessions and compulsions that do not interfere with functioning. This improvement in obsessive-compulsive symptoms in comparison with baseline symptomatology was attributed to the current availability of effective behavioral and pharmacologic treatments for OCD (techniques of exposure–response prevention and SSRIs). In fact, 99% of subjects had received at least a 10-week trial of an SSRI and 45% had received some behavior therapy. Most patients were still taking medication at the time of follow-up. Relapses were common in those patients who discontinued medication, which suggests that continued treatment may be required to maintain an improvement in OC symptoms over time.

The effect of treatment on the course of illness in OCD was also evaluated in the prospective study conducted by Eisen et al. (38), described above, in which 77 adults meeting DSM criteria for OCD were followed with frequent interim assessments for more than 2 years. Pharmacologic data gathered included doses of medications and duration of treatment. Patients had to have received a maximum dose of at least one SSRI for a minimum of 12 weeks to be considered to have received adequate pharmacotherapy for OCD. Information obtained on behavior therapy included amount of time spent in sessions, time spent doing homework, whether the patient practiced exposure–response prevention, and imagined homework. Patients were considered to have received adequate behavior therapy if they reported undergoing behavior therapy with a therapist who used exposure–response prevention and if they spent at least 20 hours practicing exposure–response prevention homework assignments. Fifty-five subjects (84% of the total sample) received an adequate trial of at least one SSRI during the study period, and 12 patients (18%) received adequate behavior therapy. The probability of partial remission for those patients who received an adequate trial of at least one SSRI was 51% during the 2-year study period.

The mean Global Assessment of Function (GAF) and Y-BOCS scores at intake and at 2 years were similar for those subjects who received an adequate trial of an SSRI and those who did not receive adequate pharmacotherapy. However, the mean GAF

scores at intake of patients who subsequently underwent adequate behavior therapy during the course of the study were lower than the mean GAF scores of patients who did not undergo behavior therapy. The change in GAF score at 2 years was significantly greater in the group of patients who received behavior therapy, so that these patients in effect “caught up”; their final GAF scores were similar to the scores of the patients who did not undergo behavior therapy.

Although this study was conducted at a time when current behavioral and pharmacotherapies were available, the results again support the findings that for the majority of patients, the course of illness in OCD is continuous with fluctuations in severity rather than episodic with clear periods of remission between periods of exacerbation of symptoms.

QUALITY OF LIFE

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

No longitudinal follow-up study of OCD has systematically measured psychosocial functioning and quality of life over time. Most treatment outcome studies have primarily focused on symptomatic relief. Also, no attempt has been made to examine the relationship between symptom severity and psychosocial functioning over time. For a significant percentage of OCD patients, impairment in function and quality of life is severe (45). It is the only major psychiatric disorder for which neurosurgery continues to be a treatment option. It will be important in future studies to gather prospective information on levels of psychosocial impairment during periods of remission when subjects no longer meet full criteria for a diagnosis of OCD. In the National Collaborative Study of Depression, even subsyndromal symptoms were associated with significant dysfunction in multiple areas (46). Similarly, preliminary data from Eisen et al. (38) suggest that psychosocial functioning continues to be impaired during partial remission despite symptomatic improvement; for example, after 1 year of follow-up, 29% of subjects in partial remission continued to miss work much of the time or were virtually incapable of carrying out activities at their jobs (38).

PREDICTORS OF LONG-TERM COURSE OF ILLNESS

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

Although a number of studies have examined predictors of outcome in OCD, the results have been inconsistent. Most have focused on identifying predictors of short-term outcome following pharmacologic or behavioral treatment. None of the existing studies has examined predictors of remission or relapse rates. These studies have been methodologically compromised by small sample size, inclusion or exclusion criteria that led to sample bias, and inadequate duration of follow-up. Characteristics such as age at onset of OCD, duration of illness, severity of illness at baseline, and phenomenologic subtype have not been associated with outcome in a consistent way. More recently, emerging data have clarified that OCD is a heterogeneous disorder and have begun to point to the existence of discrete subtypes of illness. It will be important to determine whether these “subtypes” influence the likelihood of remission or relapse. The most likely prediction variables are reviewed below.

One subtype of OCD is associated with a family or lifetime history of tic disorders. Although variation between studies is considerable, it is generally accepted that approximately 20% of patients with OCD have a lifetime history of tics, and that 5% to 10% have a lifetime history of Tourette disorder (47 ,48). Family and genetic studies have demonstrated that OCD patients with a family or lifetime history of multiple tics are more likely to have first-degree family members affected by OCD or Tourette syndrome than are OCD probands without tics (49 ,50) They are also significantly more likely to have onset of illness at an early age. OCD patients with tics appear to be less likely to respond to SSRIs, and their OCD symptoms respond differentially to augmentation of an SSRI with a neuroleptic (51). Certain OCD symptoms have been shown to develop more commonly in this subgroup, including the need for symmetry, ordering, arranging, and hoarding (52). The presence of a tic disorder predicted more severe symptoms of OCD at follow-up in children (47). The predictive power of a personal or family history of multiple tics in regard to remission and relapse rates should be investigated.

The role of personality disorder in outcome also has not been explored prospectively, although the relationship has been investigated in a number of acute treatment studies with inconsistent findings. In the study of Jenike et al. (53), schizotypy was a negative predictor for outcome after pharmacologic and behavioral treatment. In a study by Baer et al. (54), the presence of any single personality disorder was not related to improvement on any outcome measure in a 12-week placebo-controlled trial of clomipramine in OCD. However, a larger number of personality disorders was consistently related to poorer outcome, as was the presence of a DSM-III cluster, a personality diagnosis. A subsequent single-site study of the effect of a personality disorder on the response to fluoxetine failed to confirm that a cluster A diagnosis is a negative predictor of outcome (55).

The DSM-IV field trial of OCD established that a significant percentage of patients with OCD have poor insight (56). The validity of this new diagnostic category is still in question. Data pertaining to the effect of poor insight or overvalued ideation on behavioral treatment response have been inconsistent (57 ,58). Eisen and Rasmussen (59) retrospectively assessed the course of illness in four subgroups of OCD: OCD and schizophrenia, OCD and schizotypal personality disorder, OCD with poor insight, and OCD without psychotic features. A deteriorative course was noted in 82% of the patients with coexisting schizophrenia, 69% of those with

coexisting schizotypal personality disorder, 17% of those with poor insight, and only 8% of those without psychotic features. This study was hampered by the lack of a valid and reliable scale to measure poor insight and by the retrospective assessment of the course. We have recently published data on the reliability and validity of a new scale, the Brown Assessment of Beliefs Scale (BABS), that has demonstrated excellent sensitivity to change with short-term treatment (60). Eisen et al. (61) assessed change in BABS scores in patients with OCD who participated in the first phase of a double-blinded relapse study of sertraline in OCD. They found no significant correlation between degree of insight as measured by the BABS and outcome after 16 weeks of sertraline. The role of insight in remission and relapse deserves further scrutiny.

Obsessive-compulsive disorder has been linked to alterations in neurologic function involving the basal ganglia after head trauma, encephalitis, and birth events (62). Hollander et al. (63) described a subset of patients with OCD who had an increased number of neurologic soft signs and neuropsychological abnormalities in comparison with a control group matched for age, sex, and handedness. The examination of soft signs involved fine motor coordination, involuntary movements, sensory function, and visuospatial tasks. Soft signs were correlated with severity of obsessions. Receiver operating characteristic analysis found that a cutoff of three or more signs yielded the minimum number of combined errors of sensitivity and specificity in blindly distinguishing OCD subjects from controls. When these criteria were used, 25 of 40 subjects were considered to have soft neurologic signs. A second study of OCD adolescents found a high frequency of age-inappropriate synkinesias and lateralization of deficits to the left side of the body (64). In a nonblinded study in which a clinical neurologic examination was performed in childhood and adolescent OCD subjects, most of the patients had abnormal neurologic findings, including choreiform movements, nonspecific neurodevelopmental signs, and left hemisyndrome (65). Significantly more signs of central nervous system dysfunction were observed in the OCD group, manifested by abnormalities in fine motor coordination, involuntary movements, and abnormal sensory and visuospatial function. Some evidence was found of an increased number of left-sided signs that were suggestive of right-sided dysfunction. The hypothesis that this subgroup is etiologically distinct requires further validation in a study of predictive outcome.

Another important area of investigation is symptom subtype in OCD. Thus far, specific obsessions and compulsions have not predicted outcome in the vast majority of follow-up studies. In a preliminary analysis of 544 patients from a multicenter trial of acute clomipramine, the authors failed to find any significant correlation between symptom subtype, identified by the Y-BOC Symptom Checklist, and outcome (66). Recent work in which factor analysis was used to cluster groups of obsessions and compulsions suggests that certain symptom clusters may identify subtypes of OCD (67 ,68). Data obtained with positron emission tomography (PET) have shown that regional activation of the prefrontal cortex varies according to factor (69), and emerging genetic data suggest that familial loading varies according to factor (70). Symmetry and certain obsessions, such as aggressive and sexual obsessions, are more frequent in patients with OCD and chronic tics (71). One family study suggests that the rate of OCD is higher in first-degree relatives of probands with aggressive obsessions (49). The analytic technique used to identify factors from the Y-BOC Symptom Checklist may be fruitful in predicting the course of OCD. Evidence is increasing that patients in whom hoarding is a primary obsessive-compulsive symptom are resistant to traditional behavioral and pharmacologic interventions (72 ,73 and 74). In addition, hoarding was the only compulsion associated with a lower probability of remission in our pilot study.

The data regarding early onset and outcome of OCD are quite inconsistent. In a number of studies, an earlier age at onset of OCD was associated with a worse prognosis. In the study of Ravizza et al. (75), the age at onset of patients who failed to respond to a trial of an SSRI was earlier than that of responders. Thomsen (36) reported that attainment of puberty by the time of referral predicted a better prognosis than a prepubertal onset. In a reanalysis of the multicenter efficacy and safety data for clomipramine, Ackerman et al. (76), using stratification and logistic regression techniques to identify multiple prognostic factors and control for confounds, found a later age at onset to be a strong predictor of response. Skoog and Skoog (42) reported that onset of OCD before age 20 was related to a poorer outcome, especially in men. In other studies, age at onset did not predict severity of illness at follow-up. Adolescents in the study of Berg et al. (37) had an extremely variable course. In our sample, the onset of major obsessive-compulsive symptoms before age 14 predicted a higher likelihood of remission.

The severity of OCD symptoms at baseline was not predictive of long-term outcome in most studies (77 ,78 ,79 ,80 and 81), although the truncated pretreatment range of severity makes such a relationship difficult to demonstrate. It seems likely that more severe symptoms are associated with a greater degree of functional impairment and a greater number of comorbid conditions, although this hypothesis remains to be tested. However, in the only study we could locate that examined level of functioning in OCD, pretreatment functioning did not predict follow-up outcome (73). Duration of symptoms was not predictive in any study (78 ,79 ,81 ,82), although it is possible that chronicity accompanied by comorbidity may worsen prognosis. Type of ritual (washing vs. checking) was not predictive in two studies (11 ,79) and predicted erratically in others (77 ,83), a finding that argues against any consistent relationship of symptom type to outcome.

PHENOMENOLOGIC SUBTYPES AND THEIR STABILITY OVER TIME

The beginning clinician is often struck by the diversity of the clinical presentations of OCD. However, this initial impression is soon replaced by the realization that the obsessions and compulsions are remarkably limited in number and stereotypic. During the last 15 years, we have characterized the phenomenologic and clinical features of more than 1,000 patients with OCD. The basic types and frequencies of obsessive-compulsive symptoms have been found to be consistent across cultures and time (84). Why particular symptom patterns develop in given persons remains unknown. The most common obsessions include contamination, pathologic doubt, aggressive and sexual thoughts, somatic concerns, and the need for symmetry and precision. The most common rituals are checking, cleaning, and counting.

Aside from a relatively gross analysis of the course in terms of variation in overall intensity of symptoms, finer analyses of variations in symptom focus or symptom mix have not been attempted. Nevertheless, in their study of childhood OCD, Swedo and Leonard (22) reported that 90% of patients experienced some change in symptom pattern over time, often starting with a solitary ritual without associated obsessive thoughts (notably uncommon in adults), then later adding new symptoms that sometimes became predominant over earlier ones. More work is needed to delineate the frequency and magnitude of the cyclic variations in intensity and focus of obsessive-compulsive symptoms.

COMORBIDITY

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

Biological markers and neuropharmacologic challenge studies depend on the selection of homogeneous clinical populations that reduce the variance. In studying a disorder like OCD, the presence of other axis I disorders is a serious obstacle for researchers wishing to obtain homogeneous subgroups. The majority (57%) of OCD patients presenting to our clinic have at least one other DSM-III-R diagnosis. To complicate matters further, OCD is a chronic illness, and an even higher percentage of our patients have a lifetime history of another axis I disorder. Distinguishing primary from secondary diagnoses can often be difficult, if not impossible.

Studies examining the coexistence of OCD and other psychiatric disorders can be divided into two groups: (a) those examining the coexistence of other psychiatric disorders in a clinically defined population of patients with OCD and (b) those primarily focused on recording the incidence of obsessive-compulsive symptoms in other diagnostic groups. The coexistence of other anxiety states, depression, and psychotic symptoms with obsessive-compulsive symptoms was well documented in the early literature. However, few systematic clinical psychopathologic studies had been completed before 1985. Earlier studies were retrospective and failed to utilize standardized diagnostic criteria or reliable structured instruments.

The dispute about the relationship between OCD and schizophrenia has been of central interest. Controversy centers on whether a psychopathologic continuum exists for the two disorders. Some investigators have suggested that obsessions are a preliminary sign of schizophrenia, whereas others have claimed that obsessional thoughts are a neurotic defense against psychotic decompensation. Most current researchers feel that the two disorders are different entities without any true relationship. If OCD was closely related to schizophrenia, one would expect that schizophrenia would develop in a significant percentage of OCD patients. However, follow-up studies have shown that the incidence of progression to schizophrenia in primary OCD probands is low, between 1.0% and 3.3%. Rosen (85), in a retrospective chart review of 850 inpatients with schizophrenia, found that approximately 10% exhibited prominent obsessive-compulsive symptoms. This finding was replicated by Fenton and McGlashan (86), who found that 10% of schizophrenics in a Chestnut Lodge (Rockville, Maryland) follow-up study exhibited prominent obsessive-compulsive symptoms. These obsessive-compulsive schizophrenic patients tended to have a more chronic course and a greater frequency of social or occupational impairment in comparison with a matched sample of schizophrenics without obsessive-compulsive features. Recently, Eisen et al. (87) interviewed 77 patients who met SADS criteria for schizophrenia and found that 7.8% met DSM-III-R criteria for OCD. The average Y-BOCS score for those meeting the criteria for OCD was 22.3 ± 5.2.

The relationship between obsessions, compulsions, and depression was the subject of several early studies. These were primarily retrospective and failed to use diagnostic criteria or structured interviewing. Thus, many aspects of the association between depression and OCD remain unclear. One aspect that deserves further study is whether the affective episodes in OCD are primary or secondary. Dividing depressed obsessional patients into these two categories (i.e., primary and secondary) was originally advocated by Lewis (24). No systematic study of the frequency of obsessions and compulsions in a sample of depressed patients existed until recently. Although a great deal of interest has been shown in the question of whether compulsive personality increases the risk for development of a major depression, the results remain inconclusive, with wide variations in percentages across studies possibly caused by the lack of standardized diagnostic criteria.

The phenomenologic and biological evidence relating OCD to affective disorder has been reviewed by Insel (88). It has been noted that obsessive-compulsive features are rarely, if ever, seen in mania. We reported a case of OCD

in a patient with bipolar disorder whose obsessions and compulsions worsened in direct proportion to the severity of his depression and totally disappeared when he became manic (89). Although preliminary evidence suggests that OCD is rarely seen in mania, no systematic data on the frequency of obsessive-compulsive symptoms in a bipolar population were available until recently. Kruger et al. (90) and Chen and Dilsaver (91) reported on the frequency of OCD in bipolar and unipolar populations. Chen et al. found that 21% of patients with bipolar disorder, 12.2% of patients with unipolar depression, and 5.9% of patients with other disorders had OCD in the National Epidemiology Catchment Area Survey sample. Kruger et al. found that 35% of patients with both bipolar and unipolar depression had an obsessive-compulsive syndrome. Many of these depressed patients suffer from obsessions, which are at times difficult to differentiate from ruminations.

In our subsample of 250 patients who met DSM-III criteria for OCD, only 25% denied depression on admission (72). The majority admitted to feelings of inadequacy and hopelessness, and only one patient gave a history of euphoria. During the course of their illness, most reported that depression developed after the obsessive-compulsive symptoms; therefore, the patients were classified as having secondary depression. A minority (8%) of patients had a simultaneous onset of obsessive-compulsive symptoms and depressive episodes.

Kringlen (27) reported that more than 50% of 91 obsessional patients in his series had phobic symptoms. Among the 104 depressed obsessional patients of Videbach (92), 42 (40%) described phobic symptoms. In contrast, Welner et al. (93) found associated phobias in only 7 (5%) of 150 patients with severe OCD. Additional evidence supporting a shared vulnerability to OCD and other anxiety disorders is the high incidence of childhood phobias reported by obsessional patients. Lo (28) reported that 21 (35%) of his 59 obsessional patients had had significant phobias during childhood. Videbach (92) observed the same in 52 (50%) of his 104 depressed, ruminative patients. Similarly, Ingram (26) reported that 22 (25%) of 89 OCD patients had had significant phobias in childhood. During the last 5 years, several studies have examined the association of OCD with other anxiety disorders. In a study of 60 patients with panic disorder in which the SADS-LA and personal interviews were used, Breier et al. (94) found that 17% met the DSM-III criteria for OCD. Subsequent studies by Mellman and Uhde (95) and Barlow (96) confirmed these initial findings of the overlap between panic and OCD. Insel (88) pointed out the importance of the distinction between primary and secondary anxiety disorders. For example, it is often difficult to distinguish a primary social phobia with obsessive features from primary OCD that is centered on obsessing about having to complete a ritual in public. The finding of a high frequency of current and lifetime anxiety disorders suggests that OCD patients are vulnerable to virtually all types of anxiety. The high prevalence of anxiety states in these patients may be a consequence of common developmental/temperamental traits whose phenotypic expression is secondary to shared genotypic and psychosocial factors. Of particular interest in this regard is the high lifetime prevalence (12%) of separation anxiety in this group of patients (97), a finding that has also been well documented in panic disorder (20).

Table 111.4 summarizes the common axis I disorders associated with OCD in the Brown cohort. Two-thirds of obsessive-compulsive patients have a lifetime history of a major depression, and one-third have a major depression at the time of first evaluation. The majority (85%) have a mood disorder secondary to their OCD, and 15% appear to have a concurrent unipolar recurrent depression. A significant overlap is also seen with the other axis I anxiety disorders, including panic disorder, panic disorder with agoraphobia, social phobia, generalized anxiety disorder, and separation anxiety disorder. Other comorbid conditions

that appear more frequently than one would expect include eating disorders, Tourette syndrome, and schizophrenia. Comorbid axis I conditions can influence the course of illness and affect choice and order of treatment.

TABLE 111.4. COEXISTING AXIS I DIAGNOSES IN PRIMARY OBSESSIVE-COMPULSIVE DISORDER

Diagnosis	Current Semistructured (*n* = 100) (%)	Lifetime Semistructured (*n* = 100) (%)	From SADS (*n* = 60) (%)
Major depressive disorder	31	67	78
Simple phobia	7	22	28
Separation anxiety disorder	—	2	17
Social phobia	11	18	26
Eating disorder	8	17	8
Alcohol abuse (dependence)	8	14	16
Panic disorder	6	12	15
Tourette syndrome	5	7	6

SADS, Schedule for Affective Disorders and Schizophrenia.
Adapted from Jenike et al., with permission.

Special attention has been focused recently on patients with tics and OCD. Approximately 20% of patients with OCD have a lifetime history of multiple tics, and 5% to 10% have a lifetime history of Tourette syndrome (82). The age at onset in this subgroup is earlier, and they have family pedigrees loaded for both Tourette syndrome and OCD (49). Miguel et al. (98) studied similarities and differences in the clinical symptoms of 15 outpatients with OCD but not tics and 12 adult patients with Tourette syndrome but not OCD. All patients with OCD reported that some cognition preceded their compulsions, whereas only 2 of 12 patients with Tourette syndrome reported any cognition. In contrast, all patients with Tourette syndrome reported that sensory phenomena preceded their repetitive behaviors; no OCD patients reported such sensations (97).

Considerable interest has been shown in the overlap of OCD with eating disorders, particularly anorexia nervosa. In the study of Thiel et al. (99), 37% of 93 women who met criteria for anorexia or bulimia nervosa also met DSM-III-R criteria for OCD and had a score of 16 or higher on the Y-BOCS. Rastam et al. (100) also reported a high rate of OCD in a sample of 16-year-old girls in whom anorexia nervosa had been diagnosed.

Axis II conditions in OCD are covered extensively elsewhere in this volume. The most commonly encountered diagnoses are dependent, avoidant, passive–aggressive, and compulsive. Schizotypal, paranoid, and borderline personalities are found less commonly in OCD but appear to be associated with a poor outcome.

RELATIONSHIP OF HETEROGENEITY TO COMORBIDITY

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

We have become increasingly interested in developing a model for subtyping patients with OCD according to what we see as the three core features of the disorder: abnormal risk assessment, pathologic doubt, and incompleteness. These features cut across phenomenologic subtypes, such as checking, washing, or the need for symmetry, although some subtypes are more closely associated with one core feature than another.

Like most phobics, persons with OCD continually worry that if there’s a one in a million chance that something terrible will happen, it will happen to them. If there’s a one in a million chance that the elevator cable will snap, the phobic patient is certain that it will snap when he is in the elevator. In the same way, many of the thoughts of the patient with OCD are dominated by improbable events that most of us would not think twice about. Many checkers suffer from “what if?” What if I don’t unplug the coffee machine and there’s a fire? The patients with sexual or aggressive obsessions also worry. What if I do pick up the knife?

On the opposite side of the spectrum are the patients with OCD who experience little or no anxiety that something terrible will happen. Janet observed that many patients with OCD are tormented by an inner sense of imperfection. Their actions are never completely achieved to their satisfaction. Many of our patients describe an inner drive that is connected with a wish to have things perfect, absolutely certain, or completely under control. When they achieve such perfection, they describe a curious sensation that they can compare to no other feeling. Janet called it the “occasional brief appearance of sublime ecstasy.” This absolute feeling of certainty or perfection is rarely attained, and therefore the patients experience a feeling of incompleteness.

Feelings of going exactly through the middle of a door, of having both shoelaces tied to exactly the same tension, of having ones hands perfectly clean, of saying ones prayers exactly right, or of having one’s hair parted precisely down the middle are clinical examples. Most of us can relate to the feeling of wanting to have something just so or perfect and the feeling of accomplishment when we finally get it that way, and to feelings of frustration and incompleteness when it’s not that way. But for the obsessive, this feeling becomes attached to an action that would hold little significance for most of us, just as most of us do not think about the one in a million chance that something will go wrong. Patients with trichotillomania or Tourette syndrome also describe a feeling of incompleteness with continued tension until they have finished pulling out an entire patch of hair or completed a sequence of tics to their satisfaction. Both say that it is impossible to stop in the middle of a compulsive action despite the consequences.

The core features appear to relate both to the clinical features of OCD and to the comorbid disorders. In patients with abnormal risk assessment, high levels of anxiety are associated with symptoms. They are also likely to have comorbid axis I generalized anxiety disorder or social phobia, avoidant and dependent personality features, and a family history of an anxiety disorder. In contrast, patients with incompleteness are likely to manifest low levels of anxiety, comorbid multiple tics or habit disorders (e.g., trichotillomania, onychophagia), and compulsive personality features. Empiric validation of these subgroups may have important implications for diagnosis and treatment. Some evidence has already been found that patients with treatment-resistant OCD and tic spectrum disorder are particularly responsive to dopaminergic antagonists. These patients are also more likely to exhibit incompleteness.

Baer et al. (67) applied principal component analysis to 107 patients with OCD who completed the Y-BOC Symptom Checklist and examined the correlations between the factor scores and the presence of comorbid tic or personality disorders. Three factors, symmetry/hoarding, contamination/cleaning,

and pure obsessions, best explained the variance. Only the first factor was significantly related to OCPD (obsessive-compulsive personality disorder) or a lifetime history of Tourette syndrome.

COMMENT

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

During the past 15 years, significant advances have revolutionized the way we conceptualize and treat OCD. Epidemiologic studies have confirmed that OCD is an underrecognized common major psychiatric disorder with a lifetime prevalence of 2% to 3% in the general population, and they have been instrumental in focusing the attention of researchers, clinicians, and the media on OCD. Studies of the clinical features and course of the disorder and associated comorbid conditions have appeared in the literature since the turn of the twentieth century and have been the subject of numerous prospective and retrospective studies of its course, reviewed here.

Finally, future studies will continue to benefit from further refinement of our thinking about the heterogeneity and comorbidity of OCD and the search for homogeneous subtypes. The identification of an OCD–tic subtype has already led to important new genetic and biological studies and has been directly relevant to treatment. The recent effort to characterize pediatric autoimmune neuropsychiatric disorders and their relationship to genetic vulnerability to streptococcal infection offers a promising lead for furthering our understanding of the pathophysiology of OCD. It is possible that we will increase our understanding of predictions of remission and relapse related to possible homogeneous subtypes of illness. A review of these studies suggests that the course of OCD, long thought to be chronic, may be more episodic than previously believed, particularly in children and adolescents. It also appears that in some subjects, pharmacologic and behavioral treatments may alter the natural course of illness. However, a long-term prospective follow-up study of a large number of patients with OCD is needed to confirm these observations. In addition, the effectiveness of these treatments in routine practice are not known.

SUMMARY

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

The prevailing notion that the course of OCD is chronic and deteriorating has not been consistently borne out by the evidence, particularly in children followed prospectively. Furthermore, the natural course of this disorder appears to have been altered by the availability of effective pharmacologic and behavioral therapy. In their review of follow-up studies, Goodwin et al. (29) found that the course of OCD can be categorized as (a) unremitting and chronic, (b) phasic with periods of complete remission, or (c) episodic with incomplete remission that permits normal social functioning. Although the results of studies varied considerably in regard to the percentage of patients in each category, the majority of patients in each study were always in the last group, and the course of about 10% of patients was marked by progressive deterioration. These figures are consistent with our own study of patients meeting DSM-III criteria for OCD (Table 111.3). Although previous descriptive studies found a chronic waxing and waning course in 85% of patients, no attempt was made in previous studies to subdivide the waxing and waning course into predictable patterns or subtypes. More recent studies in which a prospective design and standardized criteria were used have shown that the episodic form of this disorder (clear periods of remission while the patient is off medication) is uncommon. The periodicity, duration, and severity of episodes in patients with OCD vary considerably. Once established, obsessions and compulsions usually persist, although the content, intensity, and frequency of the symptoms change over time.

The introduction of the SSRIs has led to a significantly improved prognosis for patients with OCD during the last decade. In a follow-up study by Orloff et al. (44) of a cohort of 83 OCD patients assessed 1 to 3 years after initial evaluation, 64% had a decrease of more than 50% in Y-BOCS score, and 33% had a decrease of more than 75% in Y-BOCS score at follow-up. These results are at odds with those of two other prospective longitudinal observational studies of the course of OCD that have recently been initiated at our site. Eisen et al. (38) examined 68 obsessive-compulsive outpatients evaluated at the Yale–Brown clinics and followed them prospectively during a 2-year period. Of the 51 patients who started the study meeting full criteria, 57% still met full criteria after 2 years. Survival analysis revealed a 47% probability of achieving at least partial remission during the 2-year study period. In another prospective study, by Steketee et al., 107 clinic patients with OCD were followed for up to 5 years after intake. The probability of partial remission for at least a 2-month period was 53%, and for full remission (no longer meeting criteria) at 5 years it was 22% (41).

DISCLOSURE

Part of "111 - The Course and Clinical Features of Obsessive-Compulsive Disorder "

Dr. Rasmussen receives research support from Solvay Pharmaceuticals and Pfizer.

REFERENCES

1. Robins LN, Helzer JE, Weissman MM, et al. Lifetime prevalence of specific psychiatric disorders in three sites. *Arch Gen Psychiatry* 1984;41:958–967.

2. Weissman MM, Bland RC, Canino GJ et al, The cross-national epidemiology of obsessive-compulsive disorder *J Clin Psychiatry* 1994;55:5–10.

3. Eisen JL, Steketee G. Course of illness in obsessive-compulsive disorder. In: Dickstein JJ, Riba MB, Oldham JM, eds. *1997 Review of psychiatry*, vol 16.

4. Karno M, Goldin JM, Sorenson SB, et al. The epidemiology of obsessive-compulsive disorder in five U.S. communities. *Arch Gen Psychiatry* 1988;45:1094–1099.

5. Murray CJ, Lopez AD. The global burden of disease. *WHO Report 1996*. Cambridge, MA: Harvard University (1997).

6. Dupont R, Rice D, et al. Economic costs of obsessive-compulsive disorder. *Med Interface* 1995;8:102–109.

7. Rapoport JL, ed. *Obsessive-compulsive disorder in children and adolescents*. Washington, DC: American Psychiatric Association, 1989.

8. Kohn R, Westlake RJ, Rasmussen SA, et al. Clinical features of obsessive-compulsive disorder in elderly patients. *Am J Geriatr Psychiatry* 1997;5:211–215.

9. Jenike MA. Theories of etiology. In: Jenike MA, Baer L, Minichiello WE, eds. *Obsessive-compulsive disorders practical management*, third ed. St. Louis: Mosby, 1998:203–221.

10. Rauch SL, Whalen PJ, Dougherty D, et al. Neurobiologic models of obsessive-compulsive disorder. In: Jenike MA, Baer l, Minichiello WE, eds. *Obsessive-compulsive disorders practical management*, third ed. St. Louis: Mosby, 1998:222–253.

11. First MB, Gibbon M, Spitzer RL, et al. *Structured clinical interview for DSM-IV axis II personality disorders (SCID-II)*. Washington, DC: American Psychiatric Press, 1997.

12. Rachman S, Hodgson RL. *Obsessions and compulsions*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

13. Apter A, Fallon TJ, King RA, et al. Obsessive-compulsive characteristics: from symptoms to syndrome. *J Am Acad Child Adolesc Psychiatry* 1996;35:907–912.

14. Degonda M, Wyss M, Angst J. The Zurich Study. XVIII. Obsessive-compulsive disorders and syndromes in the general population. *Eur Arch Psychiatry Clin Neurosci* 1993;243:16–22.

15. Goodman W *(personal communication, 1997)*.

16. Leonard HL, Goldberger EL, Rapoport JL, et al. Childhood rituals: normal development or obsessive-compulsive symptoms? *J Am Acad Child Adolesc Psychiatry* 1990;29:17–22.

17. Greist JH, Jefferson JW, Kobak KA, et al. Efficacy and tolerability of serotonin transport inhibitors in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1995;52:53–60.

18. Alsobrook JP II, Pauls DL. The genetics of obsessive-compulsive disorder. In: Jenike MA, Baer L, Minichiello WE, eds. *Obsessive-compulsive disorders practical management*, third ed. St. Louis: Mosby, 1998:276–288.

19. Janet P. *Les obsessions et psychasthénie*, second ed. Paris: Bailliere, 1904.

20. Kagan J, Reznick JS, Snidman N. The physiology and psychology of behavioral inhibition in children. *Child Dev* 1987;58:459.

21. Jenike MA, Baer L, Minichiello WE. An overview of obsessive-compulsive disorder. In: Jenike MA, Baer L, Minichiello WE, eds. *Obsessive-compulsive disorders practical management*, third ed. St. Louis: Mosby, 1998:3–11.

22. Swedo SE, Leonard HL. Childhood movement disorders and obsessive-compulsive disorder. *J Clin Psychiatry* 1994;55:32–37.

23. Swedo SE, Leonard HL, Garvey M, et al. Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections: clinical description of the first 50 cases. *Am J Psychiatry* 1998;155:264–271.

24. Lewis AJ. Problems of obsessional illness. *Proc R Soc Med* 1936;29:325–336.

25. Pollitt J. Natural history of obsessional states: a study of 150 cases. *Br Med J* 1957;1:194–198.

26. Ingram IM. Obsessional illness in mental hospital patients. *J Ment Sci* 1961;107:382–402.

27. Kringlen E. Obsessional neurotics: a long-term follow-up. *Br J Psychiatry* 1965;111:709–722.

28. Lo WH. A follow-up study of obsessional neurotics in Hong Kong Chinese. *Br J Psychiatry* 1967;113:823–832.

29. Goodwin DW, Guze SB, Robins E. Follow-up studies in obsessional neurosis. *Arch Gen Psychiatry* 1969;20:182–187.

30. Rasmussen SA, Tsuang MT. DSM-III obsessive compulsive disorder: clinical characteristics and family history. *Am J Psychiatry* 1986;143:317–322.

31. Coryell W. Obsessive compulsive disorder and primary unipolar depression: comparisons of background, family history, course and mortality. *J Nerv Ment Dis* 1981;169:220–224.

32. Berman L. The obsessive-compulsive neurosis in children. *J Nerv Ment Dis* 1942;95:26–39.

33. Flament MF, Whitaker A, Rapoport JL, et al. Obsessive-compulsive disorder in adolescence: an epidemiologic study. *J Am Acad Child Adolesc Psychiatry* 1988;27:764–771.

34. Flament MF, Koby E, Rapaport JL, et al. Childhood obsessive-compulsive disorder: a prospective follow-up study. *J Child Psychol Psychiatry* 1990;31:363–380.

35. Valleni-Basile LA, Garrison CZ, Jackson KL, et al. Frequency of obsessive-compulsive disorder in a community sample of young adolescents. *J Am Acad Child Adolesc Psychiatry* 1995;34:782–791.

36. Thomsen PH. Obsessive-compulsive disorder in children and adolescents in Denmark. *Acta Psychiatr Scand* 1993;88:212–217.

37. Berg CL, Rapoport JL, Whitaker A, et al. Childhood-obsessive compulsive disorder. *Am Acad Child Adolesc Psychiatry* 1989;28:528–533.

38. Eisen JL, Goodman W, Keller MB, et al. Patterns of remission and relapse in OCD: a 2-year prospective study. *J Clin Psychiatry* 1999;60:346–351.

39. Goodman WG, Rasmussen SA, Price LA, et al. The Yale Brown Obsessive-Compulsive Rating Scale (Y-BOCS) reliability and validity. *Arch Gen. Psychiatry* 1989;46:1006–1011.

40. Keller MB, Lavori PW, Friedman B, et al. The longitudinal interval follow-up evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Arch Gen Psychiatry* 1987;44:540–548.

41. Steketee G, Eisen J, Dyck I, et al. Related articles: predictors of course in obsessive-compulsive disorder. *Psychiatry Res* 1999;89:229–238.

42. Skoog G, Skoog I. A 40-year follow-up of patients with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1999;56:121–127.

43. Leonard HL, Swedo SE, Lenane MC, et al. A two- to seven-year follow-up study of 54 obsessive-compulsive children and adolescents. *Arch Gen Psychiatry* 1993;50:429–439.

44. Orloff LM, Battle MA, Baer L, et al. Long-term follow-up of 85 patients with obsessive-compulsive disorder. *Am J Psychiatry* 1994;151:441–442.

45. Koran LM, Thienemann ML, Davenport R. Quality of life for patients with obsessive-compulsive disorder. *Am J Psychiatry* 1996;153:783–788.

46. Judd LL, Akiskal HS, Zeller PH, et al. Psychosocial disability during the long-term course of unipolar major depressive disorder. *Arch Gen Psychiatry* 2000;57:375–380.

47. Leonard HL, Lenane MC, Swedo SE, et al. Tics and Tourette's disorder: a two- to seven-year follow-up of 54 obsessive compulsive children. *Am J Psychiatry* 1992;149:1244–1251.

48. Pitman RK, Green RC, Jenike MA, et al. Clinical comparison of Tourette's disorder and obsessive-compulsive disorder. *Am J Psychiatry* 1987;144:1166–1171.

49. Pauls DL, Alsobrook MP, Goodman W, et al. A family study of obsessive-compulsive disorder. *Am J Psychiatry* 1995;152:76–84.

50. Nestadt G, Samuels J, Riddle M, et al. A family study of obsessive-compulsive disorder. *Arch Gen Psychiatry* 2000;57:358–363.

51. McDougle CJ, Goodman WK, Leckman JF, et al. Haloperidol addition in fluvoxamine-refractory obsessive-compulsive disorder: a double-blind placebo-controlled study in patients with and without tics. *Arch Gen Psychiatry* 1994;51:302–308.

52. Leckman JL, Walker DE, Goodman WK, et al. "Just right" perceptions associated with compulsive behavior in Tourette's syndrome. *Am J Psychiatry* 1994;151:675–680.

53. Jenike MA, Baer L, Minichiello WE, et al. Concomitant obsessive-compulsive disorder and schizotypal personality disorder. *Am J Psychiatry* 1986;143:530–533.

54. Baer L, Jenike MA, Black DW, et al. Effect of axis II diagnoses on treatment outcome with clomipramine in 54 patients with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:862–866.

55. Tollefson GD, Birkett M, Koran L, et al. Continuation treatment of OCD: double-blind and open-label experience with fluoxetine. *J Clin Psychiatry* 1994;55[10 Suppl]:69–76.

56. Foa EB, Kozak MJ. DSM-IV field trial: obsessive-compulsive disorder. *Am J Psychiatry* 1995;152:90–96.

57. Foa EB. Failure in treating obsessive-compulsives. *Behav Res Ther* 1979;17:169–176.

58. Lelliott PT, Noshirvani HF, Basoglu M, et al. Obsessive-compulsive beliefs and treatment outcome. *Psychol Med* 1988;18:697–702.

59. Eisen JL, Rasmussen SA. Obsessive-compulsive disorder with psychotic features. *J Clin Psychiatry* 1993;54:373–379.

60. Eisen JL, Phillips KA, Baer L, et al. The Brown Assessment of Beliefs Scale: reliability and validity. *Am J Psychiatry* 1998;155:102–108.

61. Eisen JL, Rasmussen SA, Phillips KA, et al. Insight and treatment outcome in obsessive-compulsive disorder. *Compr Psychiatry* 2001;42:494–497.

62. Insel T, Donnelly E, Lalakea M, et al. Neurological and neuropsychological studies of patients with obsessive-compulsive disorder. *Biol Psychiatry* 1983;18:741–751.

63. Hollander E, Schiffmann E, Cohen B, et al. Signs of central nervous system dysfunction in obsessive compulsive disorder. *Arch Gen Psychiatry* 1990;47:27–32.

64. Behhar et al. Computerized tomography and neuropsychological test measures in adolescents with obsessive-compulsive disorder. *Am J Psychiatry* 1984;141:363–369.

65. Denckla MB. Neurologic examination. In: Rapoport JL, ed. *Obsessive-compulsive disorder in children and adolescents*. Washington, DC: American Psychiatric Association, 1988:107–118.

66. DeVeaugh-Geiss J, Landau P, Katz R. Treatment of obsessive-compulsive disorder with clomipramine. *Psych Ann* 1989;19:97–101.

67. Baer L. Factor analysis of symptom subtypes of obsessive-compulsive disorder and their relationship to personality and tic disorders. *J Clin Psychiatry* 1994;55[Suppl 3]:18–23.

68. Leckman JF, Boardman J, Grice DE, et al. Symptoms of obsessive-compulsive disorder. *Am J Psychiatry* 1997;154:911–917.

69. Rausch SL, Dougherty D, Shin L, et al. Neuroimaging in OCD: clinical implications. *CNS Spectrums* 1998;5:26–29.

70. Pauls DL. The genetics of obsessive-compulsive disorder and Gilles de la Tourette's syndrome. *Psychiatr Clin North Am* 1992;15:759–766.

71. Holzer JC, GoodmanWK, McDougle CJ, et al. Obsessive compulsive disorder with and without a chronic tic disorder. *Br J Psychiatry* 1994;164:469–473.

72. Rasmussen SA, Eisen JL. Clinical findings of significance to neuropharmacologic trials in OCD. *Psychopharmacol Bull* 1988;24:466–470.

73. Steketee G. *Treatment of obsessive-compulsive disorder*. New York: Guilford Press, 1993.

74. Ball SG, Baer L, Otto MW. Symptom subtypes of obsessive-compulsive disorder in behavioral treatment studies: a quantitative review. *Behav Res Ther* 1996;34:47–51 (revised May 21, 1997).

75. Ravizza L, et al. Predictors of drug treatment response in obsessive-compulsive disorder. *J Clin Psychiatry* 1995;56:368–373.

76. Ackerman DL, Greenland S, et al. Predictors of treatment response in obsessive-compulsive disorder: multivariate analyses from a multicenter trial of clomipramine. *J Clin Psychiatry* 1994;14:247–254.

77. Basoglu M, Lax T, Kasvikis Y, et al. Predictors of improvement in obsessive-compulsive disorder. *J Anxiety Disord* 1988;2:229–317.

78. Foa EB, Grayson JB, Steketee GS, et al. Success and failure in the behavioral treatment of obsessive-compulsives. *J Consult Clin Psychol* 1983;51:287–297.

79. Hoogduin CAL, Duivenvoorden HJ. A decision model in the treatment of obsessive-compulsive neurosis. *Br J Psychiatry* 1998;144:516–521.

80. Marks IM, Stern RS, Mawson D, et al. Clomipramine and exposure for obsessive-compulsive rituals. I. *Br J Psychiatry* 1980;136:1–25.

81. O'Sullivan G., Noshirvani H, Marks I, et al. Six-year follow-up after exposure and clomipramine therapy for obsessive-compulsive disorder. *J Clin Psychiatry* 1991;52:150–155.

82. Cottraux J, Messy P, Marks IM, et al. Predictive factors in the treatment of obsessive-compulsive disorders with fluvoxamine and/or behavior therapy. *Behav Psychol* 1993;21:45–50.

83. Drummond LM. The treatment of severe, chronic, resistant obsessive-compulsive disorder. *Br J Psychiatry* 1993;163:223–229.

84. Rasmussen SA, Eisen, JL. The epidemiology and clinical features of obsessive-compulsive disorder. Psychiatr Clin North Am 1992;15:743–758.

85. Rosen I. The clinical significance of obsessions in schizophrenia. *J Ment Sci* 1957;103:773–785.

86. Fenton WS, McGlashan TH. The prognostic significance of obsessive-compulsive symptoms in schizophrenia. *Am J Psychiatry* 1986;143:437–441.

87. Eisen JL, Beer D, Pato MT, et al. Obsessive-compulsive disorder in patients with schizophrenia or schizoaffective disorder. *Am J Psychiatry* 1997;154:271–273.

88. Insel TR. Obsessive compulsive disorder—five clinical questions and a suggested approach, *Compr Psychiatry* 1982;23:241–251.

89. Gordon A, Rasmussen SA. Mood-related obsessive-compulsive symptoms in a patient with bipolar affective disorder. *J Clin Psychiatry* 1988;49:27–28.

90. Kruger S, Cooke RG, Hasey GM, et al. Comorbidity of obsessive-compulsive disorder in bipolar disorder. *J Affect Disord* 1995;34:117–120.

91. Chen YW, Dilsaver C. Comorbidity for obsessive-compulsive disorder in bipolar and unipolar disorders. *Psychiatry Res* 1995;59:57–64.

92. Videbach T. The psychopathology of anacastic endogenous depression *Acta Psychiatr Scand* 1975;52:336–373.

93. Welner A, Reich T, Robins E, et al. Obsessive-compulsive neurosis: record, follow-up, and family studies. I. Inpatient record study. *Compr Psychiatry* 1976;17:527–539.

94. Breier A, Charney DS, Heninger GR. Agoraphobia and panic disorder: development, diagnostic stability ad course of illness. *Arch Gen Psychiatry* 1986; 43:1029–1036.

95. Mellman TA, Uhde TW. Obsessive-compulsive symptoms in panic disorder. *Am J Psychiatry* 1986;144:1573–1576.

96. Barlow DH. *Anxiety and its disorders*. New York: Guilford Press, 1988, 319–355.

97. Lipsitz JD, Martin LY, Mannuzza S, et al. Childhood separation anxiety disorder in patients with adult anxiety disorders. *Am J Psychiatry* 1994;151:927–929.

98. Miguel EC, Coffey BJ, Baer L, et al. Phenomenology of intentional repetitive behaviors in obsessive-compulsive disorder and Tourette's disorder. *J Clin Psychiatry* 1995;56:246–255.

99. Thiel A, Broocks A, Ohlmeier M, et al. Obsessive-compulsive disorder among patients with anorexia nervosa and bulimia nervosa. *Am J Psychiatry* 1995;152:72–75.

100. Rastam M, Gillberg IC, Gillberg C. Anorexia nervosa 6 years after onset: part II. Comorbid psychiatric problems. *Compr Psychiatry* 1995;36:70–76.

112

The Pathophysiology and Genetics of Obsessive-Compulsive Disorder

David L. Pauls

Emanuela Mundo

James L. Kennedy

David L. Pauls: Child Study Center, Yale University School of Medicine, New Haven, Connecticut.

Emanuela Mundo and James L. Kennedy: Department of Psychiatry, University of Toronto, Toronto, Canada.

Multiple converging lines of evidence suggest that neurobiology plays a significant role in the etiology of obsessive-compulsive disorder (OCD). During the past decade, there has been considerable progress in the identification of neuroanatomic substrates involved in the expression of OCD. The brain areas most frequently identified by *in vivo* neuroimaging studies as potentially involved in the manifestation of OCD are the orbitofrontal cortex (OFC), the anterior cingulate area (ACA), and the head of the caudate nucleus (1). Furthermore, pharmacologic and neurobiological studies have implicated several central neurotransmitter systems in the pathophysiology of OCD and related conditions. The strongest pharmacologic evidence concerns the serotonergic system and the well-established efficacy of potent serotonin reuptake inhibitors in the treatment of OCD (2 ,3); however, other systems have also been implicated. A growing body of evidence suggests that the pathophysiology OCD is complex and that, despite the fundamental role played by serotonin (5-HT) in the pathogenesis of obsessions and compulsions, a serotonergic dysfunction may explain no more than 50% of the variability of the disease. The most widely accepted alternative neurochemical theory for OCD suggests that the dopamine (DA) neurotransmission system also may be important in the pathophysiology of some cases of OCD (3 ,4 ,5 and 6). Specifically, the DA hypothesis has been proposed for those cases of OCD that appear to be related to Gilles de la Tourette syndrome (GTS) or other tics disorders, and/or those that occur with schizotypal personality disorder and/or poor insight. There is also a new etiologic hypothesis for OCD involving an autoimmune mechanism, particularly relevant for early-onset cases.

At the same time, there has been considerable research that has documented the familial nature of OCD (7). These family data, when taken together with twin studies, suggest that genetic factors are important in the manifestation of this illness. Segregation, linkage, and association studies have begun and the results are similar to those observed for other major psychiatric disorders: The mode of transmission within families is complex and the precise genetic mechanism is not known.

In this chapter the main pathophysiologic findings for OCD are reviewed as well as the evidence that genetic factors are of etiologic importance. Finally, the findings from studies examining candidate genes proposed as the result of several lines of investigation that implicate both the serotonergic and dopaminergic neurotransmitter systems are summarized.

THE PATHOPHYSIOLOGY OF OCD

Part of "112 - The Pathophysiology and Genetics of Obsessive-Compulsive Disorder "

The Serotonin Hypothesis

Historically, the serotonin (5-HT) hypothesis has its basis in the pharmacology of OCD. In the late 1960s it was observed that clomipramine, the only tricyclic antidepressant with potent 5-HT reuptake blocking properties, had antiobsessional activity (8 ,9). Subsequently, several studies have shown that clomipramine and several other selective serotonin reuptake inhibitors (SSRIs) are effective antiobsessional agents. In fact, results were taken as evidence that serotonin plays a fundamental role in the pathogenesis of OCD (10 ,11 ,12 ,13 ,14 ,15 ,16 ,17 and 18). These observations have led to the examination of the serotonin system and its function in OCD patients. Peripheral markers for the 5-HT system and a number of parameters of the 5-HT function have been investigated. These include CSF 5-hydroxyindoleacetic acid (the major metabolite of serotonin) (19 ,20 ,21 and 22), whole blood levels of 5-HT (23 ,24 and 25), platelet 5-HT concentrations (26),

and platelet imipramine binding (thought to be reflective of 5-HT uptake) (27 ,28 ,29 and 30). The results of these studies, although not definitive, suggest that a 5-HT dysfunction is present in OCD. More detailed information has come from pharmacologic challenge studies in which compounds were administered that, acting presynaptically or postsynaptically, stimulate 5-HT transmission. In these studies behavioral and neuroendocrine responses in OCD patients were assessed after challenges with meta-chloro-phenyl-piperazine (mCPP) (20 ,31 ,32 ,33 ,34 ,35 and 36), intravenous clomipramine (37 ,38), the 5-HT precursor tryptophan (39 ,40), the 5-HT releasing agent fenfluramine (41 ,42 ,43 ,44 and 45), ipsapirone (46), buspirone (47), and sumatriptan (48 ,49). These studies have also yielded conflicting results, similar to those derived from the challenge studies employing the 5-HT antagonist metergoline (50 ,51) and tryptophan depletion (53). Overall, about 50% of the OCD patients challenged acutely with proserotonergic compounds experienced a transient worsening of obsessive symptoms. These results suggest that for some OCD patients there would be a basal hyperactivity of the 5-HT neurotransmission system, owing either to a hypersensitivity of the postsynaptic receptors or to a hypoactivity of the presynaptic ones, which usually provide self-regulation (54). This could explain both the worsening of OCD symptoms after acute 5-HT stimulation and the clinical efficacy (i.e., improvement of OCD symptoms) after chronic administration of proserotonergic compounds (37 ,38).

The chronic administration of clomipramine or SSRIs induces an enhanced 5-HT release in the orbitofrontal cortex, probably as a consequence of the desensitization of the terminal 5-HT autoreceptors, and this has been hypothesized to be the neurobiological substratum for the effects of SSRIs in the treatment of OCD. The involvement of the presynaptic desensitization as a key step for the neurobiological mechanism of the antiobsessional response to proserotonergic compounds is also suggested by both the long latency of clinical efficacy (6 to 8 weeks, longer than the latency for the antidepressant response induced by the same compounds) and the high doses required (54).

Nevertheless, the fact that not all OCD patients respond to clomipramine or SSRIs and approximately 40% of them have no clinical improvement (55 ,56) may reflect the biological heterogeneity of the OCD phenotype already suggested by the variability of the response to acute 5-HT challenges. Thus, consideration of more homogeneous subgroups of OCD patients defined by response to biological challenges or different symptom subtypes could lead to clarification of the pathogenesis of the disease and the role of alternative hypotheses to the serotonergic one.

The Dopamine Hypothesis

There is now considerable evidence that some forms of OCD are etiologic related to GTS (57). GTS appears to be predominantly dopaminergically mediated, as evidenced by the well-documented clinical response to haloperidol and other dopamine antagonists (58), by the exacerbation with L-dopa and central nervous system stimulants (such as amphetamines) (59 ,60), and reports of lower CSF levels of the dopamine metabolite homovanillic acid (HVA) (61). Moreover, OCD patients with comorbid tic disorder or GTS are usually resistant to conventional pharmacotherapy with proserotonergic compounds, and may benefit from adjuvant treatment with dopamine (DA) or DA/5-HT blockers (6 ,55 ,56 ,62 ,63). This body of evidence suggests that there is an involvement of DA in at least some OCD patients.

With respect to the peripheral markers of the DA transmission, normal CSF levels of HVA have been reported (19 ,64), whereas the administration of fenfluramine produced increased inhibition of HVA secretion (33 ,65). The DA involvement has been assessed by measures of growth hormone response to apomorphine (66 ,67), and challenge with *d*-amphetamine (68) and methylphenidate (69), with conflicting results.

The serotonin and dopamine systems interact extensively, particularly in the basal ganglia (31), an area that has been implicated in the pathogenesis of obsessive-compulsive phenomenology by several studies (1 ,70 ,71 and 72). Indirect support for the involvement of both transmitter systems includes the observation of the emergence of *de novo* OC symptoms in patients on clozapine or risperidone, atypical antipsychotics with both D2 and 5-HT2 blocking properties (73 ,74 and 75), together with the demonstrated antidopaminergic activity of two antiobsessional agents, clomipramine and fluoxetine (76 ,77).

Other Neurobiological Hypotheses

Alternative neurobiological mechanisms have been proposed for OCD but they are in need of further confirmation. As already reported, functional neuroimaging studies have demonstrated dysfunction in the orbitofrontal cortex, basal ganglia and striatum, which normalize with successful treatment (78 ,79). Neuroendocrine mechanisms were implicated in the pathogenesis of obsessions and compulsions, based on studies employing oxytocin, vasopressin, and somatostatin (64 ,80 ,81 and 82). These studies also need further replication.

The Autoimmune Hypothesis

Allen, Leonard, and Swedo first proposed the intriguing autoimmune hypothesis of OCD (83) after a thoughtful review of the literature. An association was drawn between infection with group A B-hemolytic Streptococcus (as well as other agents, including viruses), and the onset or the exacerbation of OCD in some children. The observation of an association between Sydenham’s chorea (an involuntary

movement disorder related to group A B-hemolytic Streptococcus-induced autoantibodies reacting with the basal ganglia) and OCD led to the characterization of the “pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections” (PANDAS) by Swedo and colleagues (84 ,85), including OCD. Similar links between group A B-hemolytic Streptococcus and GTS have been observed (86 ,87). A monoclonal antibody that identifies B-lymphocyte antigen (D8/17) has been shown to be a marker for susceptibility to rheumatic fever, PANDAS, and early-onset OCD in general (88 ,89). Moreover, D8/17 levels have been found to follow a segregation pattern most consistent with autosomal recessive inheritance in rheumatic fever families (89 ,90). There is evidence that PANDAS is familial, with dramatically increased rates of clinical and subclinical OCD observed in parents of children with PANDAS (91).

The importance of this hypothesis cannot yet be determined, however, because it is not known how many childhood-onset cases of OCD are associated with this autoimmune process. It is possible that inherited genetic factors interact with the autoimmune mechanisms, making a subject more or less susceptible to the cross reaction created by the infection. Genetic variants in the human leukocyte antigen (HLA) system may be interesting candidates to examine in this group of OCD subjects.

THE GENETICS OF OCD

Part of "112 - The Pathophysiology and Genetics of Obsessive-Compulsive Disorder "

There has been considerable controversy regarding the inheritance of OCD. This is surprising since the familial nature of OCD has been documented since the 1930s and twin studies have provided evidence for the importance of genetic factors in the manifestation of OCD.

Twin Studies

In 1986, Rasmussen and Tsuang (92) reviewed the literature on OCD twins and found 32 of 51 (63%) MZ twins were concordant for OCD. Furthermore, when those twins where zygosity was in doubt were eliminated from the sample, 13 of 20 (65%) were concordant for OCD. These MZ concordance rates are similar to those reported for affective and anxiety disorders. However, the results need to be interpreted with caution because no data from DZ twins were available for comparisons.

This shortcoming was addressed by Carey and Gottesman (93) who studied a consecutive series of 15 MZ and 15 DZ twins ascertained from the Maudsley Twin Register. The index twin in each pair had received a psychiatric diagnosis of obsessional neurosis, obsessional personality, or phobic neurosis at local hospitals during a 32-year interval (1948 to 1979). Each twin pair was followed-up by personal interview and assessment of psychiatric status. The hospitals’ notes on the index cases and family members were also used to assign diagnoses. When a diagnosis of obsessive symptoms was used, these investigators found that 87% of the MZ twins were pair wise concordant compared to 47% of DZ twins, yielding a heritability estimate of approximately 80%.

In a more recent study, Torgersen (94) investigated the concordance of anxiety disorders (including obsessive-compulsive disorder) in the co-twins of 32 MZ and 53 DZ same-sex Norwegian twins. The sample consisted of all twins born between 1910 and 1955 who were admitted for treatment of neurotic or borderline psychotic disorders at any time before 1977. After ascertainment each twin was interviewed using a structured psychiatric interview that recorded lifetime occurrence of psychiatric symptoms; this information was then combined with the hospital records to make DSM-III lifetime diagnoses. A group of six DSM-III anxiety disorders was examined: panic disorder, agoraphobia with and without panic, social phobia, OCD, and generalized anxiety disorder (GAD). No twins were found to be concordant for the same anxiety disorder. Thus, the author examined concordances in the larger context of an “anxiety spectrum.” When the category that included only panic disorder, agoraphobia, social phobia, and OCD was, a statistically significant difference in concordance rates was seen: 45% in MZ pairs compared to 15% in DZ pairs (P <.02). This difference was not seen when considering GAD alone, nor when a combined proband diagnostic category including GAD was used.

Two important aspects of these studies critically limit their usefulness. The first limitation is the lack of standardized diagnostic criteria across studies. It is difficult to interpret results when different diagnostic criteria are used in the different studies being compared. The second limitation is the lack of blindness in evaluating the twins. The investigators doing the evaluations of the co-twin, knew the diagnosis of the index case. The lack of any procedural blind for obtaining diagnostic information or for making the actual diagnoses of a co-twin is a serious source of bias that could lead to spurious results.

Two studies were completed that used twins ascertained through twin registries. Furthermore, the evaluations of the twins were done blindly. Clifford (95) and Clifford and associates (96) analyzed data collected from 419 pairs of unselected twins who had been given the Eysenck Personality Questionnaire (EPQ) and the 42-item version of the Leyton Obsessional Inventory (LOI). Multivariate analyses provided separate heritability estimates of 44% for obsessional traits (as defined by the 10-item “Trait Scale” of the LOI) and 47% for obsessional symptoms necessary for a diagnosis of OCD (as defined by the 32-item “Symptom Scale” of the LOI). In a separate study using twins from the Australian Twin Registry, Andrews and associates (97) administered structured psychiatric interviews to 186 MZ and 260 DZ twin pairs. Ascertainment was not based on psychiatric caseness. Lifetime data for OCD, GAD, panic

disorder, social phobia, and major depressive disorder were obtained. The findings were similar to those of Torgersen in that, no differences in MZ/DZ concordances were observed when individual disorders were examined. However, when the diagnoses were combined into a single category of “neuroticism,” the MZ correlations were significantly higher than for the DZ twins; (0.58 versus 0.31) for female twins and (0.44 versus 0.27) for male twins.

In summary, all twin studies to date are consistent with the hypothesis that genetic factors are important for the expression of OCD and the specific symptoms necessary for a diagnosis of OCD. Furthermore, the two most recent studies (94 ,97) suggest that some of the same genetic factors may be important for the manifestation of some other anxiety symptoms.

Family Studies

Data from the majority of family studies completed over the past 60 years suggest that OCD is familial (7); however, rates of illness among relatives vary from study to study. Many of the studies completed prior to 1990 are difficult to interpret because of differences in diagnostic criteria and assessment methodologies and the lack of control samples or reliable estimates of population prevalence. Some of the shortcomings of this early research were addressed in six recent studies (57 ,98 ,99 ,100 ,101 and 102). All of these studies demonstrate that OCD and related conditions are familial.

Lenane and colleagues (98) studied families of 46 children and adolescents and found that 17% of the parents had OCD (25% of fathers and 9% of mothers). In a second study of the families of children and adolescents, Leonard and co-workers (99) found that 13% of first-degree first-degree relatives of OCD probands met DSM criteria for OCD. Bellodi and associates (100) reported that 3.4% of the relatives in 92 families had OCD. Although this rate is somewhat lower than other studies, it nevertheless represents a twofold increase over available population prevalence estimates. Of note is that when probands were separated on the basis of age at onset, the morbid risk for OCD among relatives of early onset (before age 14) probands was 8.8% compared to 3.2% among the relatives of later onset probands. A shortcoming of these three studies is that none included a control sample. However, assessments in all of them were done using structured interviews that were used in epidemiologic surveys. The best estimate of the population prevalence for OCD from the most recent epidemiologic study (103) is 2.3 ±0.3% in Baltimore (both of the Lenane and Leonard studies were done at the NIH in Washington, DC). Using this estimate of prevalence, the relative risk (γ) (the ratio of illness among relatives to the population prevalence) (103) for these two studies ranged between 4.4 and 10.2.

The three remaining family studies included control samples. Black and colleagues (101) studied families of 32 adult OCD probands. First-degree relatives over the age of 18 of 32 probands with OCD and 33 controls were systematically interviewed using the Diagnostic Interview Schedule (DIS). Altogether, 249 relatives (nearly 60%) were directly interviewed. Although the prevalence of OCD was not significantly increased among relatives of OCD probands, the results support the concept of an OCD spectrum. In particular, there was an increase in the rate of “subclinical” OCD among parents of OCD probands when compared to parents of controls. The following caveat should be noted when interpreting these findings. In most studies of psychiatric disorders, direct interviews, family history data, and medical record data are used to make “best estimate” diagnoses (104). In this study, Black and colleagues used only data collected from direct interviews of the relatives about themselves to assign diagnoses. Given the secrecy of many OCD patients and their tendency to hide their illness, it is possible that some family members denied symptomatology on direct interview. As noted in the report, family history data were collected by Black and colleagues but were not used in the diagnostic assignment reported. When those family history data are included, the recurrence risk among first-degree relatives is 9.8%; a rate not significantly different from recurrence risks reported in other studies. Using this estimate for the risk to relatives, the γ ranges between 1.8 and 12.1.

In 1995, Pauls and colleagues (57) reported the results of a study of 466 first-degree relatives of 100 OCD probands and 113 control subjects. The age corrected rate of OCD (10.3%) was significantly higher in the relatives of OCD probands when compared to controls (1.9%). A familial relationship has been reported between OCD and GTS and chronic tics (CT) (105) and it has been speculated that familial OCD is that type that is related to GTS. Although the rates of GTS and chronic tics were also significantly higher among relatives of OCD probands than among controls in this study, the patterns within the families suggested that much of OCD is not related to GTS; the majority of OCD individuals did not have a personal or family history of GTS or tics; however, many did have a positive family history of OCD. Thus, some forms of OCD that appear to be unrelated to tics or GTS are familial.

Finally, in the most recent and methodologically sound family study of OCD, Nestadt and colleagues (102) reported that 11.7% of 326 first-degree relatives of 85 OCD probands met DSM-III-R criteria for OCD compared to only 2.7% of controls. The methods used in this study were essentially identical to those used in the Pauls and associates (57) study. That is, the investigators directly interviewed all available first-degree relatives and obtained family history data for all first-degree relatives. Best estimate procedures were used in assigning diagnoses. The difference between the Pauls and Nestadt study is that, the latter was able to interview many more first-degree relatives than the former. Nevertheless, it is remarkable that the estimates of recurrence

risk obtained in the two studies were not significantly different. Using available population prevalence estimates from each site and the recurrence risks reported for first-degree relatives, the estimated γ ranges between 3.4 and 27.8.

Clinical Heterogeneity and Its Relationship to Familiality

Although the assessment and diagnosis of OCD is highly reliable and valid, it has also become clear over the last decade that there is considerable variability of symptomatology across individuals who have a diagnosis of OCD. Given this variability, a number of investigators have begun research to explore the possibility that subtypes/components of OCD might be distinguished on the basis of some features of the disorder. Several analyses have been completed to determine whether more homogeneous groups of OCD patients could be identified that were also more likely to be familial. One way of grouping individuals that has helped identify heritable subtypes in other conditions has been to examine age at onset. Analyses of age at onset of OCD indicate that early-onset OCD is more likely to be familial (57 ,102). However, there is still considerable familial heterogeneity within this group because a substantial proportion of early-onset OCD cases are not familial (7).

Another approach that has proven useful in identifying components of the phenotype rather than subtypes of patients is factor analysis. A number of investigators have completed factor analyses on at least four independent samples of individuals with OCD (106 ,107 ,108 and 109). In all of these analyses similar factors emerged that accounted for a substantial amount of the variance in each data set. One factor was best characterized by aggressive, sexual, religious, and somatic obsessions and related checking behavior (in the most recent set of analyses, this factor appeared to split into two separate factors). Another factor was characterized by the need for symmetry or exactness, repeating rituals, counting compulsions, and ordering/arranging compulsions. Another factor was characterized by contamination obsessions and washing/cleaning compulsions. And a final factor was characterized by hoarding obsessions and compulsions.

Preliminary analyses have been undertaken to examine whether any of these factors are related to family risk patterns (110). One of the samples included in the factor analyses reported by Leckman and associates (107) consisted of the probands for the family study of OCD reported by Pauls and co-workers (57). Additional factor analyses of the family data, which included all first-degree relatives, demonstrated that the relatives showed the same factor structure as the probands. Further analyses suggested that there were different recurrence risks for OCD among relatives of probands with different combinations of factors scores. The age-corrected recurrence risk among relatives of probands who had positive scores on the factor characterized by aggressive, sexual, religious and somatic obsessions and related checking behavior was 23.6% compared to only 13.9% among relatives of probands who had negative scores on that factor ($x^2 = 7.57$, $P < .006$). For the factor characterized by the need for symmetry or exactness, repeating rituals, counting compulsions and ordering/arranging compulsions, the age-corrected recurrence risk among relatives of probands who had positive scores was 22.7% compared to 13.5% among relatives of probands who had negative scores on that factor ($x^2 = 7.57$, $P < .019$). There was no relationship between risk to relatives and proband factor scores for the other factors. As discussed, results of complex segregation analyses that incorporated these factors scores suggested that there were different patterns of transmission within families that were related to the factor scores of the probands. Unfortunately, the number of affected relatives for whom it was possible to generate factor scores was too small to allow meaningful analyses designed to determine whether the factor scores of affected relatives were correlated with the factor scores of the probands.

Family patterns demonstrate that OCD is a complex disorder with different familial patterns being associated with different clinical characteristics of OCD. Given these familial patterns, it is likely that several genes contribute to the manifestation of the disorder. It is quite plausible that unique genes could be involved in the expression of different domains of symptomatology. Separately examining these component parts of the phenotypic spectrum with regard to their transmission within families and the possible role of genetic factors could facilitate the identification of the genes involved in the manifestation of OCD.

Segregation Analyses

Together, the family and twin study data provide compelling evidence that some forms of OCD are familial and genetic. Furthermore, segregation analyses reveal that the patterns within families are consistent with genetic models that include genes of major effect. Nicolini and colleagues (111) performed segregation analyses on data collected from 24 OCD families and Cavallini and co-workers (112) completed complex segregation analyses with data from a sample 107 families ascertained through an OCD proband. In both studies, the most parsimonious result suggested that the mode of transmission within families was most similar to an autosomal dominant pattern; however, other major gene solutions could also adequately explain the observed familial patterns.

More recently, complex segregation analyses were completed (110) on the family study data reported on by Pauls and associates (57). Analyses were done using the computer program POINTER (113). One hundred families with 466 first-degree relatives composed of 191 parents (95 fathers and 96 mothers), 217 siblings (105 brothers and 112 sisters),

and 58 offspring (39 sons and 19 daughters) were included in the analysis.

Analyses were performed on: (a) the complete data set; (b) a subset of clearly familial families; and (c) family subsets based on the previously described symptom factor structure in order to explore the possibility that these factors represent homogeneous heritable components of OCD.

Using the entire data set, only the model of no transmission could be rejected ($x^2_{(4)} = 135.49$, $P < 10^{-6}$). No specific genetic models could be ruled out. The lack of definitive results with the total sample could be due to the fact that approximately half of the families in this sample did not have any relatives affected with OCD. This pattern of familiality has also been observed in at least three other studies: Nicolini and associates (111) reported that 11 of 24 probands in their study did not have a family history of OCD; Eapen and co-workers (114) observed that approximately half of the probands in their study were isolated cases; and Cavallini and colleagues (112) reported that 54 of their 107 probands were sporadic cases.

Given these findings, the next analyses included only the subset of families in which there were at least two individuals (the proband and one other first-degree relative) affected with OCD. A total of 52 families were included in these analyses. Because OCD in another family member was a criterion for inclusion of a family, one of the affected first-degree relatives in each family was randomly assigned as a secondary proband and the analyses were appropriately corrected for this ascertainment. In these analyses, all transmission models could be rejected except a multigenic model that included both a gene of major effect and polygenic background.

Finally, these investigators divided all families on the basis of four dichotomous classification schemes that were derived from the probands' factor scores obtained from the factor analyses described in the preceding (107). For the analyses performed on families whose probands had positive scores for hoarding obsessions and compulsions, contamination obsessions and compulsions, or aggression/checking obsessions and compulsions, only the hypothesis of no transmission could be rejected. That is, there was no genetic model that could be identified as the most parsimonious. On the other hand, analyses of families whose probands had positive scores on the factor characterized by symmetry (counting obsessions and compulsions) yielded results consistent with a model that included genes of major effect.

The results of these segregation analyses demonstrate that the transmission of OCD is generally difficult to model (at least within the confines of current methods for complex segregation analysis). In most cases, no single genetic model provided the most parsimonious solution to the patterns of transmission. This could be owing to the fact that OCD is etiologically heterogeneous with only half of the cases representing familial forms. It is noteworthy that when analyses were limited to those families in which the disorder was clearly familial (i.e., families in which there are at least two individuals with OCD), the most parsimonious explanation of transmission was that it was multigenic, with at least one gene of major effect on a polygenic background.

OCD is clinically heterogeneous and most likely is also genetically heterogeneous. Although it is possible that there is at least one locus that has an appreciable impact on the manifestation of OCD, it is highly likely that it is not a single-gene disease. The familial transmission patterns are not consistent with single-gene inheritance. As reviewed in the section on pathophysiology, it appears that there are several different neurochemical pathways that are involved in the expression of OCD. Thus, it is likely that there are a number of different molecular paths to the behavioral outcome; each influenced by a different gene or genes.

In summary, there is compelling evidence from twin, family, neuroanatomic, and neurobiological studies that biological/genetic factors are important in the expression of OCD. A more complete understanding of the genetic basis and of the interactions between relevant genotypes and relevant environmental factors will be important for eventual clarification of the etiology and pathogenesis of this complex disorder. Results from all segregation analyses suggest that the underlying genetic mechanisms for OCD involve genes of major effect. The next necessary step in our goal of understanding the genetics of OCD is to localize and characterize the genes that confer susceptibility.

MOLECULAR GENETIC STUDIES

Part of "112 - The Pathophysiology and Genetics of Obsessive-Compulsive Disorder "

Genetic linkage has long been recognized as one of the methods useful in clarifying the role of genetic and environmental factors in the expression of complex disorders like OCD. Historically, the method has had limited applicability because of the small number of sufficiently polymorphic genetic markers available for study in humans. With the sequencing of the human genome, this situation has changed dramatically. Theoretical and empirical work suggests linkage studies can identify the location and thereby verify the existence of genetic loci important in the expression of these disorders; however, multiple strategies need to be employed in the study of complex non-mendelian disorders (115 ,116). Although the sib-pair approach has been available for some time, only recently has it become evident that its application is increasingly important in the genetic study of disorders where there may be genetic heterogeneity and where the mode of inheritance is complex (116). An important advantage in the use of sib-pairs is that no prior assumptions regarding specific genetic model parameters are required. Furthermore, with the increasing number of polymorphic markers available, association studies may prove feasible in the search for susceptibility genes (117).

Given the family patterns observed and the findings from

segregation analyses regarding the mode of transmission of OCD, a methodologic approach that does not require specification of a particular genetic model should be more efficacious in identifying some of the OCD susceptibility genes. Furthermore, as discussed by Pauls (118) exclusive reliance on large multigenerational families for the detection of linkage is not indicated when the disorder is common and the most likely mode of transmission is multigenic.

At the present time, there are no published linkage studies of OCD. Thus, the remainder of this chapter focuses on association studies of candidate genes.

Association Studies

Genes of the Serotonergic System

Gene testing in OCD has begun, focusing on candidates derived from the hypothesized etiologic importance of the serotonin and dopamine systems. With respect to the serotonin system, the serotonin transporter gene (SLC6A4) has been implicated in OCD as the site at which SSRIs initially exert their effects. Lesch and co-workers (119) evoked considerable interest in the transporter gene by demonstrating an association between the short allele of the 44bp insertion/deletion polymorphism in the promoter region and the anxiety-related personality traits of Neuroticism and Harm Avoidance in 505 individuals. This polymorphism has been shown to affect gene function *in vitro;* the longer allele (*l*) is associated with threefold increases in gene expression (120). Furthermore, the *l* allele of this polymorphism has been associated with elevated blood SLC6A4 levels in a sample of 70 OCD subjects (121). The first study of the promoter polymorphism revealed a trend toward increased homozygosity (*l*/*l* and *s*/*s*) in OCD patients, but the overall results were indeterminate (122). Subsequent investigation of OCD trios revealed significantly increased transmission of the *l* allele to OCD probands (123).

The 5-HT2A-receptor gene has been investigated in an association study, including 67 OCD patients (124) with inconclusive results. Kim and associates sequenced the 5-HT2B-receptor (125). One single nucleotide polymorphism was found in intron 1 of the gene, but no evidence for a functional mutation was found. Finally, another association study by Cavallini and colleagues (126) of a 5-HT2C polymorphism in 109 OCD subjects and matched controls also gave negative results.

Very recently, an association between OCD and a polymorphism of the 5-HT1DB receptor gene has been reported (127). This result appears to be particularly interesting with respect to the pathogenesis of OCD and surely deserves further investigation. The 5-HT1DB receptor is a terminal autoreceptor involved in the regulation of 5-HT transmission, and challenge studies with selective ligands (i.e., sumatriptan) (48) showed an acute worsening of OC symptoms, whereas chronic treatment with the same compound can induce improvement in OCD patients resistant to conventional pharmacotherapy (128).

Genes of the Dopaminergic System

Turning to dopamine system genes, two early case–control studies showed a lack of association with polymorphic sites in the dopamine D2 and D3 receptor genes (129 ,130). These negative results also have been replicated (124). The most promising gene in the dopamine system appears to be the D4 receptor gene. Preliminary investigations have shown a positive association between a 48-bp VNTR polymorphism and OCD (131); Cruz and co-workers (132) reported an association between the seven-repeat variant of the same polymorphism and the subtype of OCD with comorbid tic disorders.

Additional Genes

There are preliminary investigations of two enzymes involved in the metabolism of biogenic amines, catechol-*O*-methyltransferase (COMT) and monoamine oxidase A (MAO-A). Camarena and associates (133) found increased frequency of the low activity MAO-A allele in females, among a total sample of 41 OCD patients and controls. COMT modulates dopaminergic and noradrenergic neurotransmission via inactivation of catecholamines. Support for involvement of this gene derives from observations of an association between velo-cardio-facial syndrome, a rare congenital malformation associated with a microdeletion on chromosome 22q11, and obsessive-compulsive symptoms (134 ,135). Of significant interest, the COMT gene maps to this region. Two studies by Karayiorgou and colleagues (136 ,137) have shown an association between susceptibility to OCD in males and a common functional COMT allele that leads to a reduced enzyme activity. In an attempt to replicate this finding, Alsobrook and associates (138) completed family based association studies in 50 OCD trios and found no significant association in the total sample or in male probands; however, these investigators found an association in female probands (P = 0.051). Schindler and colleagues (139) have found evidence for association between homozygosity of either COMT allele and OCD. In sum, these COMT results are difficult to understand at this point. It is possible that genetic heterogeneity and population stratification may be contributing to the complexity of the findings.

CONCLUSION

Part of "112 - The Pathophysiology and Genetics of Obsessive-Compulsive Disorder "

The identification and characterization of genes important in the expression of OCD will be a major step forward in understanding the genetic and biological risk factors important for the expression of this disorder. In addition, this

work will allow the potential identification of other nongenetic factors associated with the manifestation or amelioration of the symptoms of the disorders. On the one hand, the identification of a linked marker will permit the design of much more incisive studies to illuminate the physiologic/biochemical etiology of OCD by examination of the gene product and its impact on the development of the disorders. On the other hand, by controlling for genetic factors, through the genetic case-control research paradigm, it will be possible to document more carefully the environmental and nongenetic factors important for the expression of OCD and other possibly related conditions. Studying genetic marker data together with data characterizing phenotypic expression in the context of specific environments should allow a more complete examination of the cocontribution of genetic and nongenetic factors.

ACKNOWLEDGMENTS

Part of "112 - The Pathophysiology and Genetics of Obsessive-Compulsive Disorder "

The work was supported in part by NIH grants NS 16648, HD 21887, HD 03008, HD 35482, and NS 40024 to DLP; an Ontario Mental Health Foundation grant to JLK; and Medical Research Council of Canada grant MOP-38077 to JLK and EM.

REFERENCES

1. Insel TR. Toward a neuroanatomy of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:739–744.

2. Zohar J, Insel TR. Obsessive-compulsive disorder: psychobiological approaches to diagnosis, treatment and pathophysiology. *Biol Psychiatry* 1987;2:667–687.

3. Goodman WK, McDougle CJ, Price LH, et al. Beyond the serotonin hypothesis: a role for dopamine in some forms of obsessive compulsive disorder? *J Clin Psychiatry* 1990;51(Suppl):36–43.

4. Hanna GL, McCracken JT, Cantwell DP. Prolactin in childhood obsessive-compulsive disorder: clinical correlates and response to clomipramine. *J Amer Acad Child Adolesc Psychiatry* 1991;30:173–178.

5. McDougle CJ, Goodman WK, Leckman JF, et al. The efficacy of fluvoxamine in obsessive compulsive disorder: effects of comorbid chronic tic disorder. *J Clin Psychopharmacol* 1993;13:354–358.

6. McDougle CJ, Goodman WK, Leckman J, et al. Haloperidol addition in fluvoxamine-refractory obsessive-compulsive disorder. *Arch Gen Psychiatry* 1994;51:302–308.

7. Alsobrook JP II, Pauls DL. The genetics of obsessive compulsive disorder. In: Jenike MA, Baer L, Minichiello WE, eds. *Obsessive-compulsive disorders: theory and management: a guide for clinicians, patients and families*, 3rd ed. Littleton, MA: Year Book, 1998.

8. Fernandez CE, Lopez-Ibor JJ. Monochlorimipramine in the treatment of psychiatric patients resistant to other therapies. *Actas Luso Esp Neurol Psiquiatr Cienc Afines* 1967;26:119–147.

9. Lopez-Ibor JJ. Intravenous perfusion of monochlorimipramine: technique and results. In Cerletti A, Bove FJ, eds. *The present status of psychotropic drugs*. Amsterdam: Excerpta Medica, 1969.

10. Goodman WK, Price LH, Rasmussen SA, et al. Efficacy of fluvoxamine in obsessive-compulsive disorder. A double-blind comparison with placebo. *Arch Gen Psychiatry* 1989;46:36–44.

11. Jenike MA, Buttolph L, Baer L, et al. Open trial of fluoxetine in obsessive-compulsive disorder. *Am J Psychiatry* 1989;146:909–911.

12. Clomipramine Collaborative Study Group. Clomipramine in the treatment of patients with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1991;48:730–738.

13. Chouinard G. Sertraline in the treatment of obsessive compulsive disorder: two double-blind, placebo-controlled studies. *Int Cl Psychopharmacol* 1992;7(Suppl 2):37–41.

14. Greist JH, Jefferson JW, Kobak KA, et al. Efficacy and tolerability of serotonin transport inhibitors in obsessive-compulsive disorder. A meta-analysis. *Arch Gen Psychiatry* 1995;52(1):53–60.

15. Tollefson GD, Rampey AH, Potvin JH, et al. A multicenter investigation of fixed-dose fluoxetine in the treatment of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1995;51:559–567.

16. Mundo E, Bareggi SR, Pirola R, et al. Long-term pharmacotherapy of obsessive-compulsive disorder: a double-blind controlled study. *J Clin Psychopharmacol* 1997;17(1):4–10.

17. Mundo E, Bianchi L, Bellodi L. Efficacy of fluvoxamine, paroxetine, and citalopram in the treatment of obsessive-compulsive disorder: a single-blind study. *J Clin Psychopharmacol* 1997;17(4):267–271.

18. Zohar J, Judge R, OCD Paroxetine Study Investigators. Paroxetine versus clomipramine in the treatment of obsessive-compulsive disorder. *Br J Psychiatry* 1996;169:468–474.

19. Thoren P, Asberg M, Bertillson L, et al. Clomipramine treatment of obsessive-compulsive disorder. II. Biochemical aspects. *Arch Gen Psychiatry* 1980;37:1289–1294.

20. Zohar J, Mueller E, Insel T, et al. Serotonergic responsivity in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1987;44:946–951.

21. Kruesi MJ, Rapoport JL, Hamburger S, et al. Cerebrospinal fluid monoamine metabolites, aggression, and impulsivity in disruptive behavior disorders of children and adolescents. *Arch Gen Psychiatry* 1990;47:419–426.

22. Lydiard FB, Ballenger JC, Ellinwood E, et al. CSF monoamine metabolites in obsessive-compulsive disorder. Presented at the 143rd Annual Meeting of the American Psychiatric Association, New York, May, 1990.

23. Yaryura-Tobias J, Bhagavan H. L-tryptophan in obsessive-compulsive disorders. *Am J Psychiatry* 1977;134:1298–1299.

24. Hanna GL, Yuwiler A, Cantwell DP. Whole blood serotonin in juvenile obsessive-compulsive disorder. *Biol Psychiatry* 1991;29:738–744.

25. Hanna GL, Yuwiler A, Coates JK. Whole blood serotonin and disruptive behaviors in juvenile obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry* 1995;34:28–35.

26. Flament J, Rapoport J, Murphy D, et al. Biochemical changes during clomipramine treatment of childhood obsessive-compulsive disorder. *Arch Gen Psychiatry* 1987;44:219–225.

27. Insel T, Mueller E, Alterman I, et al. Obsessive-compulsive disorder and serotonin: is there a connection? *Biol Psychiatry* 1985;20:1174–1188.

28. Weizman A, Carmi M, Hermesh H, et al. High-affinity imipramine binding and serotonin uptake in platelets of eight adolescent and ten adult obsessive-compulsive patients. *Am J Psychiatry* 1986;143:335–339.

29. Bastani B, Arora RC, Meltzer HY. Serotonin uptake and imipramine binding in the blood platelets of obsessive-compulsive disorder patients. *Biol Psychiatry* 1991;30:131–139.

30. Marazziti D, Hollander E, Lensi P, et al. Peripheral markers of serotonin and dopamine function in obsessive-compulsive disorder. *Psychiatry Res* 1992;42(1):41–51.

31. Goodman WK, Price LH, Woods SW, et al. Pharmacologic challenges in obsessive-compulsive disorder. In: Zohar J, Insel T, Rasmussen SA, eds. *The psychobiology of obsessive-compulsive disorder*. New York: Springer, 1991:162–186.

32. Pigott TA, Hill JL, Grady TA, et al. A comparison of the behavioral effects of oral versus intravenous *m*CPP administration in OCD patients and effect of metergoline prior to IV *m*CPP. *Biol Psychiatry* 1993;33:3–14.

33. Hollander E, DeCaria CM, Nitescu A, et al. Serotonergic function in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:21–28.

34. Barr LC, Goodman WK, Price LH, et al. The serotonin hypothesis of obsessive compulsive disorder: implications of pharmacologic challenge studies. *J Clin Psychol* 1992;53:17–28.

35. Goodman WK, McDougle CJ, Price LH, et al. *m*-Chlorophenylpiperazine in patients with obsessive-compulsive disorder: absence of symptom exacerbation. *Biol Psychiatry* 1995;38:138–149.

36. Brooks A, Pigott TA, Hill JL, et al. Acute intravenous administration of ondansetron and m-CPP, alone and in combination, in patients with OCD: behavioral and biological results. *Psychiatry Res* 1998;79:11–20.

37. Mundo E, Bellodi L, Smeraldi E. Effect of acute intravenous clomipramine on obsessive-compulsive and response to chronic treatment. *Biol Psychiatry* 1995;38:525–531.

38. Mundo E, Bareggi SR, Pirola R, et al. Effect of acute intravenous clomipramine and antiobsessional response to proserotonergic drugs: Is gender a predictive variable? *Biol Psychiatry*,1999;45(3):290–294.

39. Charney D, Goodman WK, Price LH, et al. Serotonin function in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1988;45:177–185.

40. Fineberg NA, Cowen PJ, Kirk JW, et al. Neuroendocrine responses to intravenous L-tryptophan in obsessive compulsive disorder. *J Affective Dis* 1994;32:97–104.

41. Hollander E, Fay M, Liebowitz M. Serotonergic and noradrenergic function in obsessive-compulsive disorder. *Am J Psychiatry* 1988;1145:1015–1017.

42. Hollander E, Cohen LJ, DeCaria C, et al. Timing of neuroendocrine responses and effect of m-CPP and fenfluramine plasma levels in OCD. *Biol Psychiatry* 1993;34:407–413.

43. Lucey JV, O'Keane V, Butcher G, et al. Cortisol and prolactin responses to d-fenfluramine in non-depressed patients with obsessive-compulsive disorder: a comparison with depressed and healthy controls. *Br J Psychiatry* 1992;161:517–521.

44. Monteleone P, Catapano F, Di Martino S, et al. Prolactin response to d-fenfluramine in obsessive-compulsive patients, and outcome of fluvoxamine treatment. *Br J Psychiatry* 1997;170:554–557.

45. Fineberg NA, Roberts A, Montgomery SA, et al. Brain 5-HT function in obsessive-compulsive disorder. *Br J Psychiatry* 1997;171:280–282.

46. Lesch KP, Hoh A, Disselkamp-Tietze J, et al. 5-Hydroxytryptamine-1A receptor responsivity in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1991;48:540–547.

47. Norman TR, Apostolopoulos M, Burrows GD, et al. Neuroendocrine responses to single doses of buspirone in obsessive-compulsive disorder. *Int Clin Psychopharmacol* 1994;9:89–94.

48. Zohar, J. Is 5-HT1D involved in obsessive compulsive disorder? *Eur Neuropsychopharmacol* 1996;6(Suppl):4S4–54.

49. Pian KLH, Westenberg HGM, van Megen HJGM, et al. Sumatriptan (5-HT1D receptor agonist) does not exacerbate symptoms in obsessive compulsive disorder. *Psychopharmacology* 1998;140:365–370.

50. Benkelfat C, Murphy DL, Zohar J, et al. Clomipramine in obsessive-compulsive disorder: further evidence for a serotonergic mechanism of action. *Arch Gen Psychiatry* 1989;46:23–28.

51. Greenberg BD, Benjamin J, Martin JD, et al. Delayed obsessive-compulsive disorder symptom exacerbation after a single dose of a serotonin antagonist in fluoxetine-treated but not untreated patients. *Psychopharmacology* 1998;140:434–444.

52. Deleted in press.

53. Barr LC, Goodman WK, McDougle CJ, et al. Tryptophan depletion in patients with obsessive-compulsive disorder who respond to serotonin reuptake inhibitors. *Arch Gen Psychiatry* 1994;51:309–317.

54. Blier P, de Montigny C. Possible serotonergic mechanisms underlying the antidepressant and anti-obsessive-compulsive disorder responses. *Biol Psychiatry* 1998;44:313–323.

55. Ravizza L, Barzega G, Bellino S, et al. Predictors of drug treatment response in obsessive-compulsive disorder. *J Clin Psychiatry* 1995;56(8):368–373.

56. Stein DJ, Bouwer C, Hawkridge S, et al. Risperidone augmentation of serotonin reuptake inhibitors in obsessive-compulsive and related disorders. *J Clin Psychiatry* 1997;58(3):119–122.

57. Pauls DL, Alsobrook JP II, Goodman W, et al. A family study of obsessive compulsive disorder. *Am J Psychiatry* 1995;152:76–84.

58. Shapiro AK, Shapiro E, Eisenkraft GJ. Treatment of Gilles de la Tourette syndrome with pimozide. *Am J Psychiatry* 1983;140:1183–1186.

59. Golden GS. Gilles de la Tourette syndrome following methylphenidate administration. *Dev Med Child Neurol* 1974;16:76–78.

60. Lowe TL, Cohen DJ, Detlor J, et al. Stimulant medications precipitate Tourette's syndrome. *JAMA* 1982;247:1729–1731.

61. Cohen DJ, Shaywitz BA, Caparulo B, et al. Chronic, multiple tics of Gilles de la Tourette's disease: CSF acid monoamine metabolism after probenecid administration. *Arch Gen Psychiatry* 1978;35:245–250.

62. McDougle CJ. Update on pharmacologic management of OCD: agents and augmentation. *J Clin Psychiatry* 1997;58(Suppl 12):11–17.

63. Jacobsen FM. Risperidone in the treatment of affective illness and obsessive-compulsive disorder. *J Clin Psychiatry* 1995;56:423–429.

64. Swedo SE, Leonard HL, Kruesi MJP, et al. Cerebrospinal fluid neurochemistry in children and adolescents with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:29–36.

65. Benkelfat C, Mefford IN, Masters CF, et al. Plasma catecholamines and their metabolites in obsessive-compulsive disorder. *Psychiatry Res* 1991;37:321–332.

66. Pitchot W, Hansenne M, Moreno AG, et al. Growth hormone response to apomorphine in obsessive-compulsive disorder. *J Psychiatry Neurosci* 1996;21(5):343–345.

67. Brambilla F, Bellodi L, Perna G, et al. Dopamine function in obsessive-compulsive disorder: growth hormone response to apomorphine stimulation. *Biol Psychiatry* 1997;42:889–897.

68. Insel TR, Hamilton JA, Guttmacher LB, et al. D-Amphetamine in obsessive-compulsive disorder. *Psychopharmacology* 1983;80:231–235.

69. Swinson RP, Joffe RT. Biological challenges in obsessive compulsive disorder. *Progr Neuro-Psychopharmacol Biol Psychiatry* 1988;12:269–275.

70. Swedo SE, Rapoport JL, Cheslow DL, et al. High prevalence of obsessive-compulsive symptoms in patients with Sydenham's chorea. *Am J Psychiatry* 1989;146:246–9249.

71. Rapoport JL. Recent advances in obsessive-compulsive disorder. *Neuropsychopharmacology* 1991;5:1–10.

72. Baxter LR, Schwartz JM, Guze BH. Brain imaging: toward a neuroanatomy of OCD. In: Zohar J, Einsel T, Rasmussen S, eds. *The psychobiology of obsessive-compulsive disorder*. New York: Springer, 1991.

73. Baker RW, Chengappa KN, Baird JW, et al. Emergence of obsessive compulsive symptoms during treatment with clozapine. *J Clin Psychiatry* 1992;53(12):439–442.

74. Allen L, Tejera C. Treatment of clozapine-induced obsessive-compulsive symptoms with sertraline. *Am J Psychiatry* 1994;151:1096–1097.

75. Remington G, Adams M. Risperidone and obsessive-compulsive symptoms. *J Clin Psychopharmacol* 1994;14(5):358–359.

76. Lipinski JF, Mallya G, Zimmerman P, et al. Fluoxetine-induced akathisia: clinical and theoretical implications. *J Clin Psychiatry* 1989;50:339–342.

77. Austin LS, Lydiard B, Ballenger JC, et al. Dopamine blocking activity of clomipramine in patients with obsessive-compulsive disorder. *Biol Psychiatry* 1991;30:225–232.

78. Baxter LR, Schwartz JM, Bergman KS, et al. Caudate glucose metabolic rate changes with both drug and behaviour therapy for obsessive compulsive disorder. *Arch Gen Psychiatry* 1992;49:681–689.

79. Baxter LR, Saxena S, Brody AL, et al. Brain meditation of obsessive-compulsive disorder symptoms: evidence from functional brain imaging studies in the human and nonhuman primate. *Semin Clin Neuropsychiatry* 1996;1:32–47.

80. Leckman JF, Goodman WK, North WG, et al. The role of central oxytocin in obsessive compulsive disorder and related normal behavior. *Psychoneuroendocrinology* 1994;19:723–749.

81. Altemus M, Swedo SE, Leonard HL, et al. Changes in cerebrospinal fluid neurochemistry during treatment of obsessive compulsive disorder with clomipramine. *Arch Gen Psychiatry* 1994;5:794–803.

82. Kruesi MJ, Swedo S, Leonard H, et al. CSF somatostatin in childhood psychiatric disorders: a preliminary investigation. *Psychiatry Res* 1990;33:277–284.

83. Allen AJ, Leonard HL, Swedo SE. Case study: a new infection-triggered, autoimmune subtype of pediatric OCD and Tourette's syndrome. *J Am Acad Child Adolesc Psychiatry* 1995;34(3):307–311.

84. Swedo SE. Sydenham's chorea: a model for childhood autoimmune neuropsychiatric disorders. *JAMA* 1994;272(2):1788–1791.

85. Swedo SE, Leonard HL, Garvey M, et al. Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections: clinical description of the first 50 cases. *Am J Psychiatry* 1998;155(2):264–271.

86. Swedo SE, Leonard HL, Kiessling LS. Speculations on antineuronal antibody-mediated neuropsychiatric disorders of childhood. *Pediatrics* 1994;93(2):323–326.

87. Kurlan R. Tourette's syndrome and 'PANDAS': will the relation bear out? *Am Acad Neurol* 1998;50:1530–1534.

88. Swedo SE, Leonard HL, Mittleman BB, et al. Identification of children with pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections by a marker associated with rheumatic fever. *Am J Psychiatry* 1997;154:110–112.

89. Khanna AK, Buskirk DR, Williams RC, et al. Presence of a non-HLA B cell antigen in rheumatic fever patients and their families as defined by a monoclonal antibody. *J Clin Invest* 1989;83:1710–1716.

89. Murphy TK, Goodman WK, Fudge MW, et al. B Lymphocyte antigen D8/17: a peripheral marker for childhood-onset obsessive-compulsive disorder and Tourette's syndrome? *Am J Psychiatry* 1997;154(3):402–407.

90. Gibofsky A, Khanna A, Suh E, et al. The genetics of rheumatic fever: relationship to streptococcal infection and autoimmune disease. *J Rheumatol* 1991;18(Suppl 30):1–5.

91. Lougee L, Perlmutter SJ, Garvey M, et al. OCD and Tic Disorders in parents of children with pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections. Presented at the Annual Meeting of the American Psychiatric Association, Toronto, Canada, June 1998.

92. Rasmussen SA, Tsuang MT. Clinical characteristics and family history in DSM-III obsessive-compulsive disorder. *Am J Psychiatry* 1986;143:317–322.

93. Carey G, Gottesman II. Twin and family studies of anxiety, phobic and obsessive disorders. In: Klien DF, Rabkin J, eds. *Anxiety: new research and changing concepts*. New York: Raven, 1981:117–136.

94. Torgersen S. Genetic factors in anxiety disorder. *Arch Gen Psychiatry* 1983;40:1085–1089.

95. Clifford CA. *Twin studies of drinking behavior and obsessionality*. Doctoral dissertation, Institute of Psychiatry, London, 1983.

96. Clifford CA, Murray RM, Fulker DW. Genetic and environmental influences on obsessional traits and symptoms. *Psychol Med* 1984;14:791–800.

97. Andrews G, Stewart G, Allen R, et al. The genetics of six neurotic disorders: a twin study. *J Affective Disord* 1990;19:23–29.

98. Lenane MC, Swedo SE, Leonard H, et al. Psychiatric disorders in first-degree relatives of children and adolescents with obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry* 1990;29:407–412.

99. Leonard HL, Lenane MC, Swedo SE, et al. Tics and Tourette's disorder: a 2- to 7-year follow-up of 54 obsessive-compulsive children. *Am J Psychiatry* 1992;149:1244–1251.

100. Bellodi L, Sciuto G, Diaferia G, et al. Psychiatric disorders in the families of patients with obsessive-compulsive disorder. *Psychiatry Res* 1992;42:111–120.

101. Black DW, Noyes R Jr, Goldstein RB, et al. A family study of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:362–368.

102. Nestadt G, Samuels J, Riddle M, et al. A family study of obsessive-compulsive disorder. *Arch Gen Psychiatry* 2000;57:358–63.

103. Karno M, Golding JM, Sorenson SB, et al. The epidemiology of obsessive-compulsive disorder in five U.S. communities. *Arch Gen Psychiatry* 1988;45:1084–1099.

103. Risch N. Linkage strategies for genetically complex traits. I multilocus models. *Am J Hum Genet* 1990;46:222–228.

104. Leckman JF, Sholomskas D, Thompson WD, et al. Best estimate of lifetime psychiatric diagnoses: a methodological study. *Arch Gen Psychiatry* 1982;39:879–883.

105. Pauls DL, Towbin KE, Leckman JF, et al. Gilles de la Tourette syndrome and obsessive compulsive disorder: evidence supporting a genetic relationship. *Arch Gen Psychiatry* 1986;43:1180–1182.

106. Baer L. Factor analysis of symptom subtypes of obsessive compulsive disorder and their relation to personality and tics. *J Clin Psychiatry* 1993;55:18–23.

107. Leckman JF, Grice DE, Boardman J, et al. Symptoms of obsessive compulsive disorder. *Am J Psychiatry* 1997;154:911–917.

108. Summerfeldt LJ, Richter MA, Antony MM, et al. Symptom structure in obsessive-compulsive disorder: a confirmatory factor-analytic study. *Behav Res Ther* 1999;37:297–311.

109. Mataix-Cols D, Rauch SL, Manzo PA, et al. Use of factor-analyzed symptom dimensions to predict outcome with serotonin reuptake inhibitors and placebo in the treatment of obsessive-compulsive disorder. *Am J Psychiat* 1999;156:1409–1416.

110. Alsobrook II JP, Leckman JF, Goodman WK, et al. Segregation analysis of obsessive-compulsive disorder using symptom-based factors. *Am J Med Genet* 1999;88:669–675.

111. Nicolini H, Hanna G, Baxter L, et al. Segregation analysis of obsessive compulsive and associated disorders. Preliminary results. *Ursus Med* 1991;1:25–28.

112. Cavallini MC, Pasquale L, Bellodi L, et al. Complex segregation analysis for obsessive compulsive disorder and related disorders. *Am J Med Gen* 1999;88(1):38–43.

113. Lalouel JM, Rao DC, Morton NE, et al. A unified model for complex segregation analysis. *Am J Hum Genet* 1983;35:816–826.

114. Eapen V, Robertson MM, Alsobrook JP II, et al. Obsessive compulsive symptoms in Gilles de la Tourette's syndrome and obsessive compulsive disorder: differences by diagnosis and family history. *Am J Med Genet (Neuropsychiatr Genet)* 1997;74:432–438.

115. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994;265:2037–2048.

116. Kruglyak L, Lander ES. Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet* 1995;57:439–454.

117. Risch N, Merikangas K. The future of genetic studies of complex human disorders. *Science* 1996;273:1516–1517.

118. Pauls DL. Behavioural disorders: lessons in linkage. *Nat Genet* 1993;3:4–5.

119. Lesch KP, Bengel D, Heils A, et al. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 1996;274(5292):1527–1531.

120. Heils A, Teufel A, Petri S, et al. Allelic variation of human serotonin transporter gene expression. *J Neurochem* 1996;66:2621–2624.

121. Hanna GL, Himle JA, Curtis GC, et al. Serotonin transporter and seasonal variation in blood serotonin in families with obsessive-compulsive disorder. *Neuropsychopharmacology* 1998;18(2):102–111.

122. Billet EA, Richter MA, King N, et al. Obsessive compulsive disorder, response to serotonin reuptake inhibitors and the serotonin transporter gene. *Mol Psychiatry* 1997;2:403–406.

123. McDougle CJ, Epperson CN, Price LH, et al. Evidence for linkage disequilibrium between serotonin transporter protein gene (SLC6A4) and obsessive-compulsive disorder. *Mol Psychiatry* 1998;3:270–273.

125. Kim SJ, Veenstra-Vander Weele J, Hanna GL, et al. Mutation screening of human 5-HT(2B) receptor gene in early-onset obsessive-compulsive disorder. *Mol Cell Probes* 2000;14:47–52.

128. Stern L, Zohar J, Cohen R, et al. Treatment of severe, drug resistant obsessive-compulsive disorder with the 5-HT1D agonist sumatriptan. *Eur Neuropsychopharmacol* 1998;8(4):325–328.

124. Nicolini H, Cruz C, Camarena B, et al. DRD2, DRD3 and 5-HT2A receptor genes polymorphisms in obsessive-compulsive disorder. *Mol Psychiatry* 1996;1(6):461–465.

129. Catalano M, Sciuto G, Di Bella D, et al. Lack of association between obsessive-compulsive disorder and the dopamine D3 receptor gene: some preliminary considerations. *Am J Med Genet* 1994;54:253–255.

130. Novelli E, Nobile M, Diaferia G, et al. A molecular investigation suggests no relationship between obsessive-compulsive disorder and the dopamine D2 receptor. *Neuropsychobiology* 1994;29:61–63.

126. Cavallini MC, Di Bella D, Pasquale L, et al. 5-HT2C CYS23/SER23 polymorphism is not associated with obsessive-compulsive disorder. *Psychiatry Res* 1998;77:97–104.

131. Billett EA, Richter MA, Swinson RP, et al. Investigation of dopamine system genes in obsessive-compulsive disorder. *Psychiatric Genet* 1998;8(3):163–169.

127. Mundo E, Richter MA, Sam F, et al. Is the 5-HT1DB receptor implicated in the pathogenesis of obsessive-compulsive disorder? *Am J Psychiatry* 2000;157:1160–1661.

132. Cruz C, Camarena B, King N, et al. Increased prevalence of the seven-repeat variant of the dopamine D4 receptor gene in patients' obsessive-compulsive disorder with tics. *Neurosci Lett* 1997;231:1–4.

133. Camarena B, Cruz C, de la Fuente JR, et al. A higher frequency of a low activity related allele of the MAO-a gene in females with obsessive-compulsive disorder. *Psychiatric Genet* 1998;8:255–257.

134. Pulver AE, Nestadt G, Goldberg R, et al. Psychotic illness in patients diagnosed with velo-cardio-facial syndrome and their relatives. *J Ner Ment Dis* 1994;182(8):476–478.

135. Papolos DF, Faedda GL, Veit S, et al. Bipolar spectrum disorders in patients diagnosed with velo-cardio-facial syndrome: does a hemizygous deletion of chromosome 22q11 result in bipolar affective disorder? *Am J Psychiatry* 1996;153(12):1541–1547.

136. Karayiorgou M, Altemus M, Galke B, et al. Genotype determining low catechol-o-methyltransferase activity as a risk factor for obsessive-compulsive disorder. *Nat Acad Sci* 1997;94:4572–4575.

137. Karayiorgou M, Sobin C, Blundell JL, et al. Family-based association studies support a sexually dimorphic effect of *COMT* and *MAOA* on genetic susceptibility to obsessive-compulsive disorder. *Biol Psychiatry* 1999;45:1178–1189.

138. Alsobrook JP II, Zohar AH, Leboyer M, et al. Association between the COMT locus and Obsessive Compulsive Disorder in females but not males. *Am J Med Genet* in press.

139. Schindler KM, Richter MA, Kennedy JL, et al. Association between homozygosity at the COMT gene locus and obsessive compulsive disorder. *Am J Med Genet* 2000;96(6):721–724.

113

Imaging and Neurocircuitry of OCD

David R. Rosenberg

Shauna N. MacMillan

David R. Rosenberg and Shauna N. MacMillan: Wayne State University School of Medicine, Detroit, Michigan.

> The history of the living world can be summarized as an elaboration of ever more perfect eyes, within a cosmos in which there is always something more to be seen.
>
> --(Pierre Teihlard de Chardin, *The Phenomenon of Man*)

Once considered a rare condition, obsessive-compulsive disorder (OCD) is now recognized as a severe and often chronically disabling illness that affects 1% to 3% of the world's population (1 ,2 ,3 ,4 and 5). The condition is characterized by intrusive ritualistic thoughts, ideas, and behaviors (obsessions and compulsions, respectively) over which the person has little if any control. Recent investigation further suggests that the illness has its onset in childhood and adolescence in at least 80% of all cases, although it often goes undiagnosed until adulthood (6).

In recent years, neuroimaging studies have begun to bridge the gap between our understanding of the neurobiologic underpinnings of OCD and the development of effective clinical assessment measures and treatments for the illness. This is critical because even as use of the selective serotonin reuptake inhibitors (SSRIs) and cognitive behavioral therapy (CBT) has become routine in the treatment of OCD, at least one-third of OCD patients do not respond at all to currently available treatments (7). Many patients who respond to treatment do so only partially; continued functional impairment is the rule. It is the rare patient who experiences complete remission of his or her symptoms. A clearer understanding of the neuropathophysiologic mediation and expression of OCD may result in the identification of new and more effective treatments for this chronic and often crippling illness.

Indeed, this view is supported by many other neurologic and medical conditions for which treatment includes psychosocial and medication interventions acting on specific somatic substrates. Up until the 1930s, epilepsy was considered more a psychiatric than neurologic condition (8 ,9). Indeed, in the Middle Ages, persons with epilepsy were often believed to be possessed by the devil. Epilepsy carried considerable stigmata and was often a cause for shame; its victims were often considered to be "crazy" and in need of psychiatric treatment. This reflected, in part, a lack of knowledge about the neurobiologic substrates underlying epilepsy. With the advent of the EEG, it became evident that electrical abnormalities in the brain underlie various epileptiform conditions. The EEG also helps guide choice of treatment intervention (10); for example, different medications are used for persons suffering from temporal lobe epilepsy (e.g., carbamazepine) versus those suffering from grand mal seizures (e.g., phenobarbital). Treatment for other medical conditions such as diabetes, rheumatoid arthritis, and asthma also includes psychosocial and medication interventions acting on specific somatic substrates. Although we have not achieved comparable understanding of neuropsychiatric disorders such as OCD, 10 years from now we may have identified 10 different subtypes of OCD characterized by specific neurobiologic abnormalities that may, in turn, necessitate individualized treatment interventions (11). Neuroimaging allows for precise measurement of brain structure, chemistry and function, which can be correlated with both baseline and clinical outcome measures (e.g., Yale-Brown-Obsessive-Compulsive Scale, SSRI, CBT treatment) (12). This very powerful approach is answering hypothesis-driven investigations regarding pathophysiology, psychobiology, and treatment response in psychiatric disorders such as OCD. Brain imaging techniques in conjunction with advances in neuroscience and neuropsychopharmacology can contribute a great deal to the outstanding questions in child psychiatry relevant to enhanced diagnostic rigor and treatment development (13).

In this chapter, the authors review: (a) current neurocircuitry models of OCD that have converged to implicate abnormalities in cortical-striatal-thalamic-cortical circuits (14 ,15 ,16 ,17 ,18 ,19 and 20); (b) the relevant neuroimaging techniques utilized in investigating the neurocircuitry of OCD; and (c) neuroimaging studies conducted in OCD patients. The developmental

implications of these investigations are then discussed and a new clinical neurodevelopmental model of OCD is described as it may relate to treatment development.

NEUROCIRCUITRY OF OCD: CORTICOSTRIATAL CIRCUIT THEORIES

Part of "113 - Imaging and Neurocircuitry of OCD "

Basal Ganglia

As early as 1931, von Economo (21) described postmortem globus pallidus abnormality in OCD associated with postencephalitic parkinsonism. OCD behaviors are also increased in other basal ganglia disorders, including Huntington disease, Tourette syndrome, pediatric autoimmune neuropsychiatric disorders associated with group A ß-hemolytic Streptococcal infections (including Sydenham's chorea), neuroacanthocytosis, and progressive supranuclear palsy (17 ,22 ,23 ,24 ,25 ,26 ,27 ,28 and 29). Aberrations in basal ganglia–frontal cortical networks may play an especially critical role in the emergence of OCD symptoms (14 ,16).

Frontal Cortex

Fronto-striatal abnormalities have been hypothesized to represent the core pathology in OCD (15 ,16 ,30 ,31 ,32 ,33 and 34). Ventral prefrontal cortical regions, particularly anterior cingulate and medial orbital frontal cortex and their striatal target fields, have been most consistently implicated in the pathogenesis and maintenance of OCD (16 ,31 ,35 ,36 ,37 ,38 ,39 ,40 ,41 and 42). This may, in part, reflect the critical role of anterior cingulate and medial orbital frontal cortex in regulation of affect and motivation (31 ,35 ,36 and 37). Lesions to these brain regions results in the inability to inhibit context inappropriate responses and inappropriate impulse modulation and behavior. Indeed, neuropsychological studies suggest that a deficit in response inhibition abilities may represent a core deficit in OCD (43 ,44 ,45 ,46 ,47 and 48). Neurosurgical lesions (e.g., cingulotomy) have also been demonstrated to be effective in reducing OCD symptoms in treatment-refractory patients (49). In contrast, evidence for abnormalities in other frontal lobe regions (e.g., dorsolateral prefrontal cortex) are much less compelling.

The anterior cingulate and medial orbital frontal cortex and their striatal target fields in ventral striatum receive dense projections from mesiotemporal lobe, particularly the amygdala and hippocampus (31 ,50). Mesiotemporal structures including the amygdala and hippocampus play an especially important role in processing and responding to the emotional valence of affective stimuli (30 ,51 ,52 ,53 ,54 and 55). In fact, animal studies have demonstrated that serotonin reuptake inhibitors known to be effective in the treatment of OCD have particularly potent effects on receptors in the amygdala (56 ,57 ,58 and 59). Wise and Rapoport (60) have hypothesized that excess activity in mesial temporal-orbitofrontal and anterior cingulate regions and their striatal target fields could disinhibit particular regions of the thalamus leading to OCD behaviors.

Thalamus

The thalamus is a highly evolved sensory and motor gateway to the cortex that serves as the final subcortical input to frontal cortex and plays a critical role in consciousness, perception, and integration of information (32 ,61). The thalamus is an important component of frontal and limbic circuits; thalamic lesions result in neuropsychological and behavioral disturbances similar to deficits observed in OCD patients (17 ,31 ,62). In fact, "frontal lobe"-type syndromes frequently appear indistinguishable from vascular and degenerative disorders of the thalamus (17). The thalamus serves as a filter in integrating information before it reaches the cortex—a task facilitated by its many cortical and subcortical connections.

When released from the inhibitory influence of the striatum, the thalamus stimulates cortical output. This has led to descriptions of "direct" and "indirect" pathways modulating frontal cortical output to ensure context-appropriate responses (31 ,32 and 33). Although the "direct" pathway releases the inhibitory tonic influence of the striatum, thereby stimulating thalamic stimulation of the cortex so that instinctual and protective hard-wired behaviors are enhanced, the "indirect" pathway facilitates the cortex in shifting sets and responding to new situations on the basis of the particular circumstance and prior stored information by inhibiting the thalamus. Baxter and associates (32) have hypothesized neural hyperactivity in the direct versus indirect pathways and that this imbalance may result in obsessive and compulsive behaviors characteristic of the illness.

Preclinical investigation has demonstrated that compulsive behaviors can be provoked by altering thalamic function (63 ,64), whereas thalamic stimulation can result in compulsive behaviors in humans (65). This is particularly intriguing because neurosurgical lesion of the thalamus (e.g., partial thalamotomy) has actually been reported to reduce OCD symptoms in treatment-refractory OCD patients (49).

In summary, neurobiologic studies from preclinical and clinical laboratories have consistently implicated a cortico-striato-thalamo-cortical network in the pathogenesis of OCD (15 ,16 ,32 ,60 ,66). Advances in neuroimaging (discussed in the following) provide an unprecedented opportunity for developing a mechanistic understanding of the developmental neurobiology of OCD as it relates to the behaviors that characterize the illness. These methods allow for the direct, *in vivo* and noninvasive "biopsy" of the brain at levels of spatial and temporal resolution heretofore possible only in animal or postmortem human studies. Neuroimaging studies also facilitate our taking advantage of advances in neuroscience and applying them directly to clinical neuropsychiatric conditions such as OCD. Such an approach

may ultimately result in new diagnostic and therapeutic approaches with the identification of surrogate neurobiological markers predicting treatment response (or lack, thereof) (66 ,67 and 68). Although a detailed review of brain imaging methodology is beyond the scope of this chapter, a brief review is presented in the following. For more detailed descriptions, the reviewer is referred to relevant textbooks on brain imaging (69 ,70).

BRAIN IMAGING TECHNIQUES

Part of "113 - Imaging and Neurocircuitry of OCD "

Neuromorphometry

Newer noninvasive neuroimaging procedures permit measurement of changes in regional brain anatomy, chemistry, and function. Pioneering work in psychiatric neuroimaging has already resulted in findings of critical relevance to psychiatric disorders despite the brief period during which these techniques have been in use (71). For example, less than 15 years ago, measurement of structural abnormalities in neuropsychiatric disorders was limited to postmortem brain studies (70). Understandably, many were (and some remain) skeptical about applying brain imaging to neuropsychiatric disorders such as OCD. This may, in part, reflect the lack of sensitivity of earlier neuroimaging techniques (e.g., skull x-ray and ventriculogram) (70). The emergence of computerized tomography (CT) ushered in a new era of brain research during which evaluation of living tissue would become routine.

Moreover, although many discounted Johnstone and colleagues' (72) CT finding of cerebral ventricular abnormalities in schizophrenia, these currently represent some of the most replicated and best-established findings in psychiatry (69).

Quantitative CT and the more recent emergence of magnetic resonance imaging (MRI), which allows for enhanced three-dimensional (3D) acquisitions, tissue differentiation without putative ionizing radiation risks have revolutionized our ability to conduct precise, *in vivo* and noninvasive quantitative neuromorphometric studies of regional brain anatomy. These studies are critical because the volume of a brain region of interest in neuropsychiatric disorders has been demonstrated to reflect that region's function (73 ,74). Volumetric neuroimaging studies are also critical to guide functional, metabolic, and neurochemical studies and can also help characterize neurodevelopmental and degenerative effects of neuropsychiatric disorders (75). Interpretation of functional neuroimaging studies is, therefore, predicated on controlling for volumetric differences between patients and controls as well as psychotropic medication-induced volumetric changes (76 ,77 and 78).

Functional Neuroimaging

Although quantitative measures of brain volume are critical, newer and more sophisticated functional neuroimaging techniques may be more sensitive in identifying subtle brain abnormalities in neuropsychiatric disorders (79). Positron emission tomography (PET) and single photon emission computerized tomography (SPECT) permit measurement of cerebral blood flow, metabolism, neurochemistry, and receptor function. These techniques have been instrumental in elucidating the pathophysiology of OCD. Both SPECT and PET techniques use ligands labeled with radiation. These putative radiation risks limit their viability in pediatric populations and in longitudinal repeated studies assessing neurodevelopment and neurodegenerative effects. For example, many institutional review boards preclude PET studies in healthy control children. This may, however, become more feasible with advances in methodology including 3D PET, which may reduce radiation exposure significantly. Although radiation risks associated with SPECT are considerably lower than PET, the decreased resolution of SPECT limits its usefulness.

The recent advent of magnetic resonance spectroscopy (MRS) and functional MRI (fMRI) represent powerful approaches for the direct, *in vivo* and noninvasive measurement of brain chemistry and function without ionizing radiation risks. Recent investigation in OCD suggests that these techniques are very sensitive in identifying brain abnormalities not evident even with sophisticated morphometric MRI (79). In contrast to PET, fMRI allows for enhanced temporal resolution of brain function rather than focusing on measurement of brain activity during a single brief time interval (67). Event-related fMRI allows for second-to-second temporal resolution (80). Newer techniques including magnetoelectroencephalography (MEG) or electroencephalography (EEG) with fMRI studies allow acquisition and analysis brain activity in real-time with a spatial resolution in millimeters and a temporal resolution of milliseconds (81). Although not currently available, this is an active area of investigation and expected to be available in the very near future.

In summary, previous *in vivo* investigation of neuropsychiatric disorders such as OCD was limited by measures that provided only a peripheral index of brain function. Direct assessment of brain was limited to postmortem neuropathologic study or *in vivo* study in animal models. The aforementioned neuroimaging techniques provide an unprecedented opportunity to measure brain anatomy, chemistry, and function to elucidate the neurobiologic underpinnings of neuropsychiatric disorders such as OCD, which may translate into the development of new assessment procedures and treatments. In the following section, we discuss a series of neuroimaging studies that have consistently implicated regional changes in cortico-striatal-thalamic-cortical neural circuitry (66) being involved in the pathogenesis and maintenance of OCD.

NEUROMORPHOMETRY OF OCD (TABLE 113.1)

Part of "113 - Imaging and Neurocircuitry of OCD "

Structure	Volumetric Findings		
	Decreased	Increased	No Difference
Total cerebral volume/ intracranial volume			*Giedd et al., 2000*[a] *Rosenberg et al., 1997* *Gilbert et al., 2000* Aylward et al., 1991 Breiter et al., 1994 Jenike et al., 1996 Robinson et al., 1995 Stein et al., 1997
Ventricular brain ratios		*Behar et al., 1984*	Insel et al., 1983
Lateral ventricles		Stein et al., 1993	*Rosenberg et al., 1997* *Luxenberg et al., 1998* Insel et al., 1983 Kellner et al., 1991 Robinson et al., 1995 Jenike et al., 1996 Aylward et al., 1996 Stein et al., 1997
Third ventricle		*Rosenberg et al., 1997*	
Total prefrontal cortex		Grachev et al., 1998 (in 10 PU subregions)	*Rosenberg et al., 1997* Robinson et al., 1995 Jenike et al., 1996
Orbital frontal cortex	Szeszko et al., 1999		Grachev et al., 1998
Dorsolateral prefrontal cortex			*Rosenberg and Keshavan, 1998*
Anterior cingulate cortex		*Rosenberg and Keshavan, 1998*	Szeszko et al., 1999 Grachev et al., 1998
Posterior cingulate cortex			*Rosenberg and Keshavan, 1998*
Basal ganglia caudate	*Luxenberg et al., 1988* Robinson et al., 1995	Scarone et al., 1992 *Giedd et al., 2000*[a]	*Rosenberg et al., 1997* Kellner et al., 1991 Stein et al., 1993, 1997 Aylward et al., 1991, 1996 Jenike et al., 1996 *Peterson et al., 2000*[b]
Putamen	*Rosenberg et al., 1997*	*Giedd et al., 1995, 2000*[a] *Peterson et al., 2000*[b]	Jenike et al., 1996 Aylward et al., 1996
Globus pallidus		*Giedd et al., 2000*[a] *Peterson et al., 2000*[b]	Jenike et al., 1996
Thalamus		*Gilbert et al., 2000*	Jenike et al., 1996 *Giedd et al., 2000*[a]
Amygdala	Szeszko et al., 1999		*Rosenberg and Keshavan, 1998* Jenike et al., 1996
Hippocampus			*Rosenberg and Keshavan, 1998* Szeszko et al., 1999 Jenike et al., 1996
Superior temporal gyrus			*Rosenberg and Keshavan, 1998*

Italics indicate studies in the pediatric age-range.
[a]Studied patients with pediatric autoimmune neuropsychiatric disorders associated with Group A β-hemolytic streptococcal infection.
[b]Study by Peterson et al. included both pediatric and adult subjects. Putamen and globus pallidus were increased in patients with obsessive-compulsive disorder and attention deficit hyperactivity disorder who had increased antistreptolysin O and antideoxyribonuclease B antibody titers.

TABLE 113.1. STRUCTURAL NEUROIMAGING STUDIES IN OBSESSIVE-COMPULSIVE DISORDER

Global Changes of Ventricles, Cerebral Volume, and Atrophy

Although total brain volume has not been found to differ between OCD patients and controls (76 ,82 ,83 ,84 ,85 ,86 ,87 and 88), Behar and associates (89) reported significantly increased ventricular brain ratio (VBR) in adolescent OCD patients compared to controls. Rosenberg and co-workers (87) observed significantly increased third ventricular volumes in 19 treatment-naive, pediatric OCD patients compared to 19 age- and sex-matched controls but no differences in lateral ventricular volume as measured by volumetric MRI. In contrast, Insel

and colleagues (90) observed no significant differences in VBR between adult OCD patients and healthy comparison subjects. Examination of cortical gray and white matter has, however, demonstrated increased gray-white matter ratios in adult OCD patients (83 ,84). Such abnormalities could be owing to aberrations in prenatal programmed cell death or postnatal reductions or delays in myelination (84). Recent investigation has suggested abnormalities of postnatal myelination in pediatric OCD patients (91 ,92).

Striatum

Despite the striatum being posited as a primary site of pathology in OCD (30), structural neuroimaging studies of the caudate nucleus in adult OCD patients have revealed contradictory findings. Scarone and colleagues (93) reported increased right caudate size in OCD patients compared to controls, whereas Robinson and co-workers (85) found bilateral reductions in caudate volume in OCD patients. Four MRI studies have reported no significant differences in caudate size between OCD patients and controls (79 ,82 ,94 ,95). Investigation of the other components of the basal ganglia, including the putamen and globus pallidus, has not demonstrated volumetric differences between adult OCD patients and controls (84 ,94 ,96 ,97). These studies were potentially confounded by several factors including illness chronicity, past treatments, heterogeneity of OCD, and differences in imaging methodology used. Structural imaging studies in children may prove especially instructive because they allow for examination of neurodevelopmental factors and repeated studies for longitudinal assessment.

Nonetheless, structural neuroimaging studies in pediatric OCD patients have not been entirely consistent. Behar and associates' (89) observation of increased VBR in adolescent OCD patients is consistent with reductions in striatal volume (information on striatal volumes was not provided). More recently, Luxenberg and co-workers (96) reported bilateral reductions in caudate volume in adolescent men with OCD compared to controls using quantitative CT. Using volumetric MRI, two recent investigations (87 ,98) found no significant differences in caudate volume between treatment-naive pediatric OCD patients and age- and sex-case matched controls. Localized reductions in putamen volume associated with OCD symptom severity but not illness duration, however, were observed in pediatric OCD patients compared to controls (Fig. 113.1). Reduced putamen volumes have been reported in Tourette syndrome (99), a condition frequently associated with OCD symptoms. Putaminal lesions associated with OCD also have been reported in isolated case reports (100 ,101) and pediatric OCD patients have antibodies directed at the putamen at rates significantly greater than in healthy pediatric comparison subjects (102).

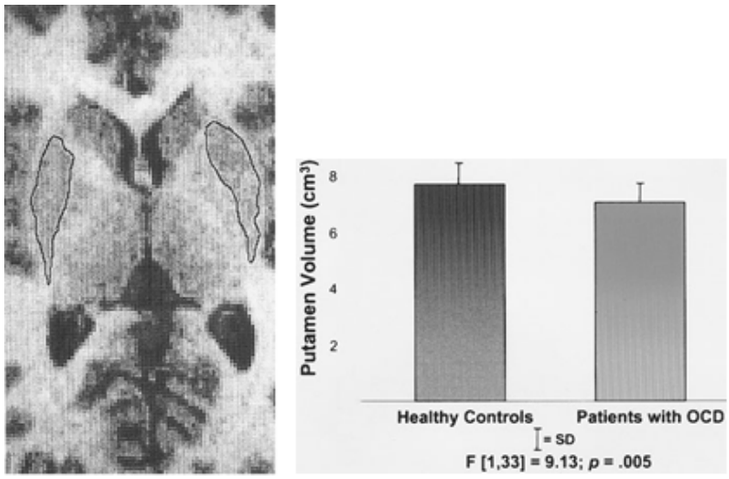


FIGURE 113.1. Measurement of putamen in the axial plane using a manual tracing technique *(left)*. Putamen volume by group *(right)*. OCD, obsessive-compulsive disorder. Adapted from Rosenberg DR, Keshavan MS, O'Hearn KM, et al. Frontostriatal measurement in treatment-naive children with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1997;54:824–830.

In this regard, it is important to point out that Giedd and associates (24 ,26) have reported increased volumes in caudate, putamen, and globus pallidus in children with OCD or tics associated with group A ß hemolytic streptococcal (GABHS) infections and pediatric patients with Sydenham chorea and associated OCD and tic behaviors compared to healthy children. These conditions are now referred to as pediatric autoimmune neuropsychiatric disorders associated with streptococcal infection (PANDAS) (103 ,104) and may represent discrete subtypes of OCD and tic disorders. Increased basal ganglia volumes may be consistent with hypothesized antibody-mediated inflammation of the basal ganglia in poststreptococcal or OCD or tic disorders (103 ,104). Giedd and colleagues (24 ,26), however, did not observe an association between basal ganglia size and symptom severity of OCD or tics, suggesting an indirect relationship between basal ganglia size and the pathophysiology of the condition. Allen and associates (104) observed plasmapheresis to be dramatically effective in PANDAS patients in reducing OCD and tic symptom severity.

In fact, Giedd and colleagues (25) conducted serial MRI scans and observed a striking relationship among basal ganglia volume, OCD symptom severity, and treatment with plasmapheresis in an adolescent with autoimmune OCD (PANDAS) (Fig. 113.2). Recently, Peterson and colleagues (27) reported that higher antistreptolysin O antibody titers were associated with larger basal ganglia volumes in OCD patients with chronic or recurrent streptococcal infections. This finding was not specific to OCD; however, as higher antibody titers were also associated with enlarged basal ganglia volumes in attention deficit hyperactivity disorder (ADHD) patients with chronic or recurrent streptococcal infections (Fig. 113.3 ; Table 113.2). In fact, Peterson and colleagues (27) found robust associations between diagnosis of ADHD and titers of antistreptolysin O and antideoxyribonuclease B titers, whereas no such association was seen between antibody titers and a diagnosis of OCD or chronic tic disorders. We also do not know the impact of chronic or recurrent streptococcal infections on basal ganglia volume in children who do not develop OCD, tic disorders, or ADHD. Further study is clearly warranted.

TABLE 113.2. ANTIBODY AND DIAGNOSIS ASSOCIATIONS WITH BASAL GANGLIA VOLUMES

Independent Variable	Basal Ganglia		Type 3 Sum of Squares	F[a]	p[b]	Independent Varible	Basal Ganglia		Type 3 Sm of Squares	F[a]	p[b]
Age	Caudate	R	185 622.2	0.70	.41	ASO titer	Caudate	R	77 976.8	0.29	.59
		L	329 784.1	1.16	.28			L	163 312.5	0.57	.45
	Putamen	R	403 527.5	1.82	.18		Putamen	R	719 418.1	3.25	.07
		L	1 949 333.2	9.24	.003			L	162 392.0	0.77	.38
	GP	R	68 399.0	1.80	.18		GP	R	124 003.6	3.27	.07
		L	1426.6	0.04	.85			L	311 963.2	7.78	.006
Sex	Caudate	R	26 436.7	0.10	.75	ASO × CTD	Caudate	R	14 529.1	0.05	.81
		L	20 262.5	0.07	.79			L	888.4	0.003	.96
	Putamen	R	41 336.3	0.19	.67		Putamen	R	479 307.9	2.16	.14
		L	3410.7	0.02	.90			L	564 780.3	2.68	.10
	GP	R	75 676.4	2.0	.16		GP	R	2594.2	0.07	.79
		L	3287.1	0.08	.77			L	6767.6	0.16	.69
Cerebral volume	Caudate	R	9 399 861.6	35.47	.0001	ASO × OCD	Caudate	R	17 949.4	0.07	.79
		L	9 539 401.1	33.61	.0001			L	31 020.8	0.11	.74
	Putamen	R	14 425 018.5	65.15	.0001		Putamen	R	554 597.5	2.51	.12
		L	11 771 290.4	55.82	.0001			L	180 705.5	0.86	.36
	GP	R	2 509 899.4	66.13	.0001		GP	R	47 453.5	1.25	.27
		L	2 017 826.0	50.34	.0001			L	271 584.3	6.78	.01
CTD	Caudate	R	447 291.1	1.69	.20	ASO × ADHD	Caudate	R	19 359.1	0.07	.79
		L	740 785.4	2.61	.11			L	44 599.2	0.16	.69
	Putamen	R	2 122 069.4	9.58	.003[c]		Putamen	R	382 533.3	1.73	.19
		L	795 020.5	3.78	.05[c]			L	1 284 129.5	6.10	.01
	GP	R	16 303.4	0.43	.51		GP	R	320 157.8	8.44	.005
		L	52 985.6	1.32	.25			L	121 207.3	3.02	.08
OCD	Caudate	R	2751.9	0.01	.92	ASO × ADHD × OCD	Caudate	R	24 437.7	0.09	.76
		L	7.9	0.00	.99			L	7602.8	0.03	.87
	Putamen	R	154 842.7	0.70	.41		Putamen	R	713 789.6	3.22	.07
		L	34 663.3	0.16	.69			L	1 223 781.8	5.80	.01
	GP	R	43 768.0	1.15	.28		GP	R	272 664.1	7.18	.009
		L	73 499.1	1.83	.18			L	98 375.9	2.45	.12
ADHD	Caudate	R	2851.2	0.01	.92	Error	Caudate	R	26 502 840.2	—	—
		L	125.9	.000	.98			L	28 386 288.6	—	—
	Putamen	R	37 475.8	0.17	.68		Putamen	R	22 140 347.0	—	—
		L	56 711.5	0.27	.61			L	21 087 448.1	—	—
	GP	R	179 236.2	4.72	.03[d]		GP	R	3 795 193.9	—	—
		L	103 525.9	2.58	.11			L	4 008 318.0	—	—

[a]df = 1 for all except error, where df = 100.
[b]P values are Bonferroni, adjusted.
[c]The estimated marginal mean of the right and left putamen is smaller in subjects with CTD.
[d]The estimated marginal mean of the right globus pallidus in smaller in subjects with ADHD.
ADHD, attention-deficit/hyperactivity disorder;
ASO, antistreptolysin O; CTD, chronic tic disorder; ellipses, error terms; GP, globus pallidus; L, left; OCD, obsessive-compulsive disorder; R, right; x, statistical interaction of adjacent terms; lighter shading indicates $P < .05$; Overall, $R^2 = .34$.
Reprinted from Peterson BS, Leckman JF, Tucker D, et al. Preliminary findings of antistreptococcal antibody titers and basal ganglia volumes in tic, obsessive-compulsive, and attention deficit/hyperactivity disorders. *Arch Gen Psychiatry* 2000;57:364–372.
Multivariate analysis of variance assessing the strength of the association of diagnosis-by-antibody interactions with basal ganglia volumes.

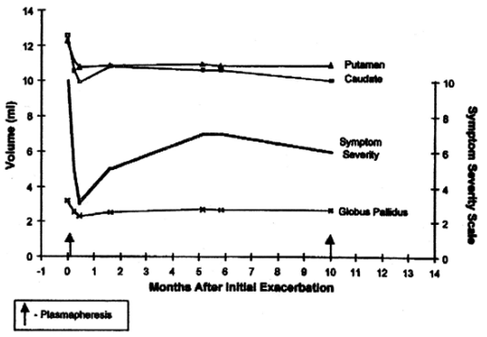


FIGURE 113.2. Sequential magnetic resonance imaging of basal ganglia volumes in a male adolescent undergoing plasma exchange for infection-related obsessive-compulsive disorder. Reprinted from Giedd JN, Rapoport JL, Leonard HL, et al. Case study: acute basal ganglia enlargement and obsessive-compulsive symptoms in an adolescent boy. *J Am Acad Child Adolesc Psychiatry* 1996;3S(7):913–915.

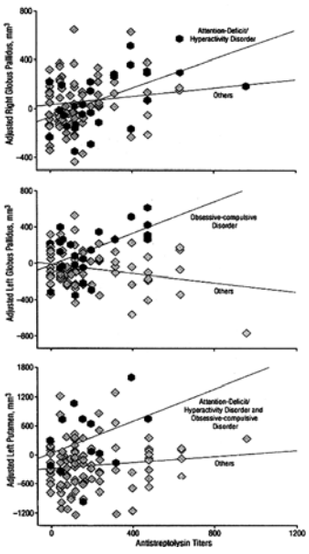


FIGURE 113.3. Association of titers, diagnoses, and basal ganglia volumes. The interaction of antistreptolysin 0 titers with attention-deficit/hyperactivity disorder and obsessive-compulsive disorder diagnoses are presented graphically. Basal ganglia volumes are adjusted for the effects of all independent variables in the multivariate analysis of covariance (Table 113.2); hence, the residuals of the volumes after the adjustment can be positive or negative. These volume residuals are plotted against the raw antistreptolysin 0 values for each of the relevant diagnostic groups. Titers are plotted in dark circles for the noted diagnostic group and in lighter diamonds for all other subjects. Reprinted from Peterson BS, Leckman JF, Tucker D, et al. Preliminary findings of antistreptococcal antibody titers and basal ganglia volumes in tic, obsessive-compulsive, and attention deficit/hyperactivity disorders. *Arch Gen Psychiatry* 2000;57:364–372.

Taken together, these findings underscore the need for standardization of research studies in patients with OCD and controlling for potential confounds of comorbidity, treatment effects, and illness duration. It also illustrates how brain imaging is exploiting advances in developmental neurobiology with important implications for neurodiagnostic assessment and treatment development. A neurodevelopmental perspective is equally critical as illustrated in the following.

Prefrontal Cortex

Morphometric MRI measurement of the prefrontal cortex has also yielded conflicting findings. Total prefrontal cortical volumes have not been found to differ between adult OCD patients and controls (84 ,85 ,105). Jenike and associates (84) did observe increased opercular volumes in OCD patients. Grachev and co-workers (105) reanalyzed the 10 adult female OCD patients and matched controls studied by Jenike and associates (84) using a sophisticated topographic parcellation method (106) and found an increase in six right frontal and four left parcellation units in OCD patients. Anterior cingulate, orbitofrontal, and opercular cortical volumes did not differ significantly between OCD patients and controls. Grachev and associates (105) also noted a significant correlation between increased volume of right inferior frontal pars triangularis and right midfrontal cortical volumes and poor cognitive performance on nonverbal immediate recall testing. More recent investigation (107) found localized reduced bilateral orbital frontal volumes in OCD patients versus healthy comparison subjects. Superior frontal gyrus and anterior cingulate volumes did not differ between OCD patients and controls.

Consistent with findings in adults, Rosenberg and colleagues (87) reported no significant differences between treatment-naive pediatric OCD patients and controls in total prefrontal cortical volume; however, a neurocognitive study of a similar sample of pediatric OCD patients (44) revealed a selective deficit in the core prefrontal cognitive function, neurobehavioral response inhibition, with no abnormalities in working memory (delayed response) or preparatory set (Fig. 113.4). Monkey studies and human clinical studies suggest that ventral prefrontal cortex plays a critical role in mediating the suppression of context inappropriate responses (17 ,108 ,109 ,110 ,111 ,112 ,113 ,114 ,115 and 116), whereas dorsal prefrontal cortex may play a more specific role in mediating delayed response and preparatory set (117 ,118). Subsequent investigation demonstrated increased corpus callosal area in treatment-naive pediatric OCD patients compared to controls,

particularly in the regions of the genu and splenium (91). The corpus callosum connects the cerebral hemispheres so that the genu connects ventral prefrontal cortex and the striatum, whereas the splenium connects temporal lobe regions (119 ,120).

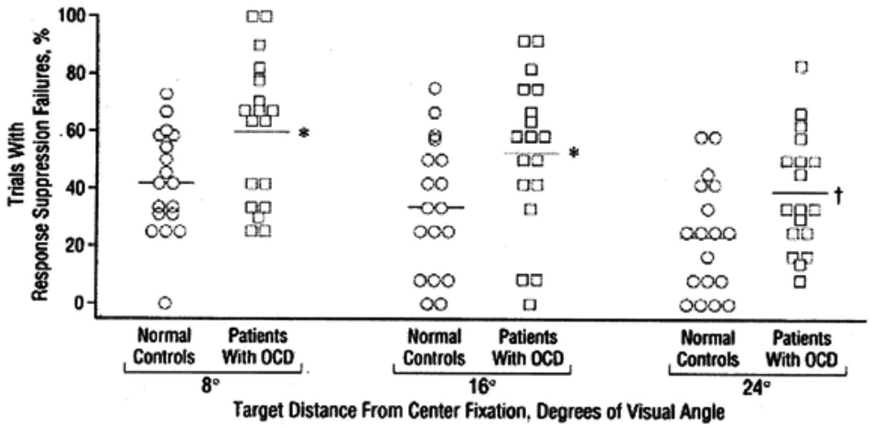


FIGURE 113.4. Mean response suppression failures for nondepressed, psychotropic medication-naive pediatric patients with obsessive-compulsive disorder (OCD) and normal controls performing the antisaccadic task. Lines through distributions represent the mean value. Asterisks indicate *P* = .01; dagger, *P* = .02. Reprinted from Rosenberg DR, Averbach DH, O'Hearn KM, et al. Oculomotor response inhibition abnormalities in pediatric obsessive-compulsive disorder. *Arch Gen Psychiatry* 1997;54:831–838.

Rosenberg and associates (91) also noted that the age-related increase in corpus callosal area in healthy children and adolescents was absent in OCD patients (Fig. 113.5A). Controls achieved comparable corpus callosal areas to their age-matched OCD counterparts between 16 and 18 years of age, which is consistent with prior findings of no significant differences in corpus callosal area between adult OCD patients and controls (83 ,84). Postnatal reduction or delay in myelination in OCD has been hypothesized to be involved in the pathogenesis of OCD (84). In support of this hypothesis, MacMaster and colleagues (92) reported increased signal intensity localized to the genu region of the corpus callosum in pediatric OCD patients

compared to controls. Increased genu area in pediatric OCD patients could be related to excess myelin sheath thickness (92). An alternative explanation is abnormal pruning or reduction of neural elements within the corpus callosum. This may be less likely because neuronal apoptosis occurs very early in development (121), whereas myelinization takes place during the peak periods of onset of pediatric OCD (122).

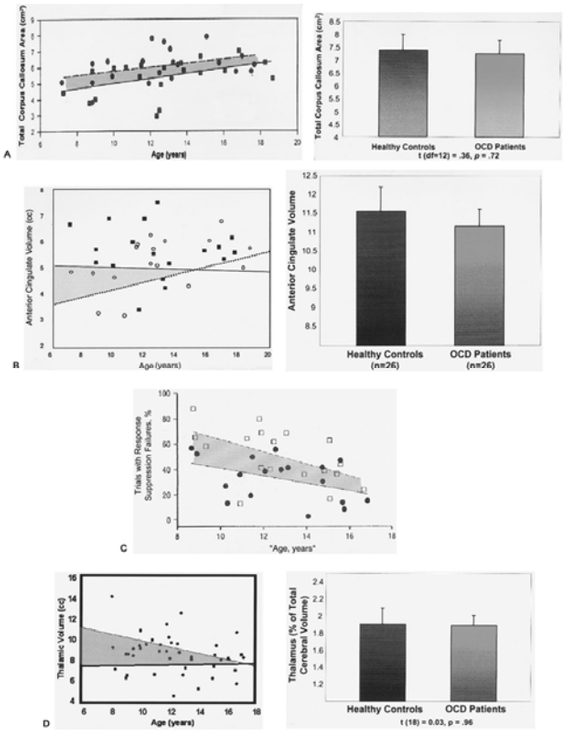


FIGURE 113.5. A: Corpus callosum area (A) in pediatric *(left* [•, —, normal controls; ▪, - - - - - -, OCD patients.]) and adult patients with obsessive-compulsive disorder *(right)*. Reprinted from Rosenberg DR, Keshavan MS, Dick EL, et al. Corpus callosal morphology in treatment naive pediatric obsessive-compulsive disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 1997;21:1269–1283. Breiter HC. Retrocallosal white matter abnormalities in patients with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1994;51:663–664. **B:** Anterior cingulate volume in pediatric *(left* [ô, ⋯⋯⋯, controls (r = .45, p = .055); ▪, —, OCD patients (r = -.12, p = .62)]) and adult patients with obsessive-compulsive disorder *(right)*. Reprinted from Rosenberg DR, Keshavan MS. Toward a neurodevelopmental model of obsessive-compulsive disorder. *Biol Psychiatry* 1998;43:623–640. Szeszko PR, Robinson D, Alvir JM et al. Orbital frontal and amygdala volume reductions in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1999;56:913–919. **C:** Response suppression failures as a function of age for nondepressed, psychotropic medication-naive pediatric patients with OCD and normal controls performing the antisaccadic task. Note the marked inverse correlation between age and total number of response suppression errors in patients with OCD and a trend for such an effect in controls. (•, —, normal controls [r = -.44, p = .07; □, - - - - - -, OCD patients [r = -.54, p = .02]). Reprinted from Rosenberg DR, Averbach DH, O'Hearn KM, et al. Oculomotor response inhibition abnormalities in pediatric obsessive-compulsive disorder. *Arch Gen Psychiatry* 1997;54:831–838. **D:** Thalamic volume in pediatric *(left* [•, —, healthy control (r = .012, p = .959); •, ⋯⋯⋯, pediatric OCD (r = .425, p = .055)]) and adult patients with obsessive-compulsive disorder *(right)*. Adapted from Gilbert AR, Moore GJ, Keshavan MS, et al. Decrease in thalamic volumes of pediatric obsessive-compulsive disorder patients taking paroxetine. *Arch Gen Psychiatry* 2000;57(5):449–456. Jenike MA, Breiter HC, Baer L, et al. Cerebral structural abnormalities in obsessive-compulsive disorder. A quantitative morphometric magnetic resonance imaging study. *Arch Gen Psychiatry* 1996;53:625–632.

Subsequent investigation has revealed significantly increased ventral prefrontal cortical volumes in anterior cingulate cortex in 21 treatment-naive pediatric OCD patients compared to age and sex case-matched controls (88) (Fig. 113.5B). Increased anterior cingulate volumes were inversely correlated with reduced striatal volumes in OCD patients. Oculomotor response inhibition abnormalities also correlated with increased anterior cingulate volumes and reduced striatal volumes in pediatric OCD patients (Fig. 113.5C). No significant differences were observed in posterior cingulate or dorsolateral prefrontal cortical volumes between pediatric OCD patients and controls (88). Thus, prefrontal cortical abnormalities in pediatric OCD may be localized to ventral prefrontal anterior cingulate circuits, particularly in younger patients.

Temporal Cortex

The temporal limbic structures, comprising the amygdala and hippocampi are critically involved in regulating emotion in both health and disease (123 ,124 and 125). This function undergoes striking changes throughout childhood, adolescence, and early adulthood (126 ,127). Medial orbital frontal cortex, anterior cingulate cortex, and ventral striatum receive dense afferent projections from limbic regions, including the amygdala and hippocampus (31 ,128).

Initial MRI investigation in adult OCD patients and healthy controls revealed no significant differences in mesiotemporal lobe brain structures (84 ,95 ,105). More recent investigation by Szesko and colleagues (107) using criteria from postmortem histologic analysis (129) with a semiautomated computerized system (130) demonstrated bilateral reductions in amygdala volume in OCD patients as compared to healthy controls. No significant differences between OCD patients and controls were observed in the hippocampus.

Recent investigation in pediatric OCD patients has also implicated the amygdala (87 ,91). Specifically, Rosenberg and associates (87) reported reduced putamen but not caudate volumes in treatment-naive pediatric OCD patients. The putamen receives more projections from the amygdala than the caudate and reduced putamen volumes are observed after temporal lobe lesions (131). In pediatric OCD patients, a positive correlation was observed between putamen and amygdala volumes but not amygdala and caudate volumes (87). Subsequent investigation also revealed significant differences between pediatric OCD patients and controls in the size of the splenium, the region of the corpus callosum that connects temporal-limbic regions (91). However, direct measurement of whole temporal lobe, amygdala, and hippocampal and superior temporal gyrus volumes failed to reveal any significant differences between pediatric OCD patients and age- and sex-matched controls (88). Perhaps, volumetric abnormalities of the amygdala only become apparent later in development. Alternatively, our methods may not have been sensitive enough to distinguish subtle abnormalities in this circuitry. It should be noted that it is often difficult to distinguish the amygdala and hippocampus even at the histologic level (132).

Thalamus

Volumetric abnormalities in ventral prefrontal cortex and the striatum in pediatric OCD patients led to our studying the thalamus, the final subcortical input to frontal cortex. Jenike and co-workers (84) reported no significant differences in thalamic volume in adult OCD patients, many of whom had been treated with psychotropic medication and had long-term illness duration. In contrast, Gilbert and co-workers (76) demonstrated significantly increased thalamic volume as measured by volumetric MRI in 21 treatment-naive, pediatric OCD patients compared to 21 age- and sex-matched healthy controls (Fig. 113.6). As in the corpus callosum and anterior cingulate cortex, volumetric abnormalities in the thalamus were particularly pronounced in younger patients with OCD (Fig. 113.5D). After monodrug therapy with the SSRI, paroxetine, thalamic volumes decreased to levels comparable to those observed in healthy children. Reduction in thalamic volume was positively correlated with reduction in OCD symptom severity with increased pretreatment thalamic volume predicting better response to paroxetine treatment (Fig. 113.7). In contrast, thalamic volume did not decrease after 12 weeks of CBT in 11 treatment-naive pediatric OCD patients who received no adjunctive medication treatment (133). No significant changes were observed in total brain volume, the striatum, or anterior cingulate cortex with either CBT or paroxetine treatment. In view of serotonin's critical role in thalamocortical development and activity (134), thalamic volumetric reductions in pediatric OCD patients may be specific to SSRI treatment as opposed to a generalized treatment response or spontaneous resolution of symptoms.

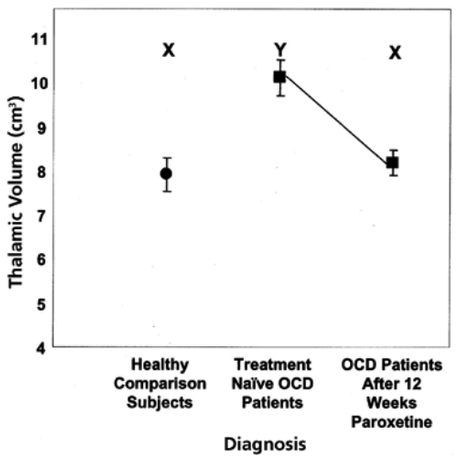


FIGURE 113.6. Thalamic volume by diagnostic and treatment condition. Groups not sharing the same letter are significantly different at $P < .05$. Adapted from Gilbert AR, Moore GJ, Keshavan MS, et al. Decrease in thalamic volumes of pediatric obsessive-compulsive disorder patients taking paroxetine. *Arch Gen Psychiatry* 2000;57(5):449–456.

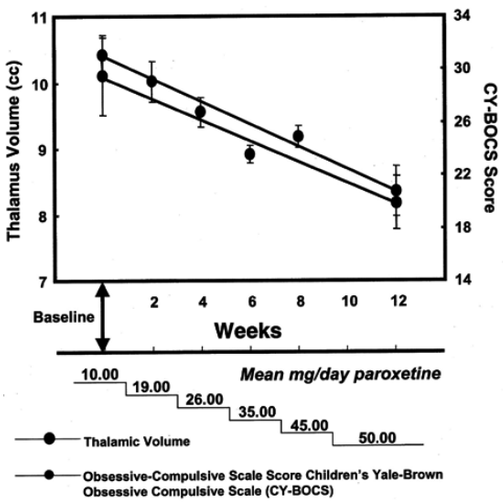


FIGURE 113.7. Decrease in thalamic volume associated with reduction in Obsessive-Compulsive Score of the Children's Yale-Brown Obsessive-Compulsive Scales. Reprinted from Gilbert AR, Moore GJ, Keshavan MS, et al. Decrease in thalamic volumes of pediatric obsessive-compulsive disorder patients taking paroxetine. *Arch Gen Psychiatry* 2000;57(5):449–456.

It should be noted that the techniques utilized in these investigations were unable to discriminate specific subdivisions of the thalamus. Thus, we were unable to localize changes to specific thalamic target fields. The dorsomedial

nucleus of the thalamus has been most implicated in the pathogenesis of OCD (15 ,31). Morphometric analysis of regional subdivisions of the thalamus is currently an active area of investigation in our laboratory.

In summary, although not entirely consistent, structural neuroimaging studies implicate abnormalities in cortico-striato-thalamo-cortical circuits. The neural network dysplasia of reduced striatal volumes and increased ventral prefrontal and thalamic volumes in treatment-naive pediatric OCD patients is especially intriguing. Rosenberg and Keshavan (88) have hypothesized that a dysplasia in postnatal synaptic pruning may be involved resulting in excess pruning in the striatum and reduced or delayed pruning in ventral prefrontal cortex and the thalamus in pediatric OCD,

which may be owing to abnormalities in glutamatergic–serotonin neurotransmission. Although morphometric brain imaging studies are instructive, functional neuroimaging studies that actively drive the system and can measure brain chemistry and receptor function may be more sensitive in their ability to detect more subtle and localized abnormalities in brain circuitry (79).

FUNCTIONAL NEUROIMAGING STUDIES OF OCD

Part of "113 - Imaging and Neurocircuitry of OCD "

Functional Neurocircuitry of OCD

Although structural neuroimaging studies measure the brain in the resting or neutral state, functional neuroimaging procedures including PET, SPECT, and fMRI allow for localized measurement of dynamic rather than static brain function by measurement of regional cerebral blood flow, glucose metabolism, and brain activation (135). To date, most studies have assessed function over a period of seconds to minutes during single session studies so that the temporal dimension has been relatively unexplored. Recent advances are beginning to permit “real-time” analysis of brain function in response to differential stimulation and diagnostic and treatment conditions. This allows for a connectionist approach to brain circuitry and not simply monitoring whether a certain region of interest activates but when and for how long. Such an approach is especially promising for repeated longitudinal assessment of OCD patients before, during, and after treatment intervention.

PET and fMRI studies suggest excess activity in the caudate nucleus, orbitofrontal cortex, thalamus, and amygdala in OCD (38 ,42 ,66 ,136 ,137 ,138 ,139 ,140 and 141). Symptom provocation of OCD symptoms using individually tailored noxious stimuli in adult OCD patients results in an increase in regional cerebral blood flow and activation in these regions (42 ,140 ,141) (Fig. 113.8). It is not entirely clear, however, whether increased activity in these circuits is specific to OCD or

represents a nonspecific finding generalized to all anxiety states (142). For example, increased regional cerebral blood flow in frontal cortex has been found to be associated with anxiety in OCD patients (143 ,144). In healthy adult volunteers, cholecystokinin induces anxiety with increased regional cerebral blood flow in anterior cingulate and the amygdala, whereas anticipation of an anxiety-provoking stimulus has been associated with increased activation of orbital frontal cortex (145). Therefore, Rauch and associates (146) reviewed pooled data from their PET symptom provocation studies of patients with OCD, simple phobia and posttraumatic stress disorder in an effort to determine which patterns of activation were specific to OCD and which were generalized across different anxiety disorders. They found that activation of paralimbic circuits, including posterior medial orbital frontal cortex, anterior cingulate, and temporal limbic regions were associated with all anxiety conditions and not specific to OCD. In contrast, activation of anterior orbital frontal cortex and the caudate nucleus were specific to OCD, suggesting that these regions may be primary loci of abnormality in the illness (Fig. 113.9).

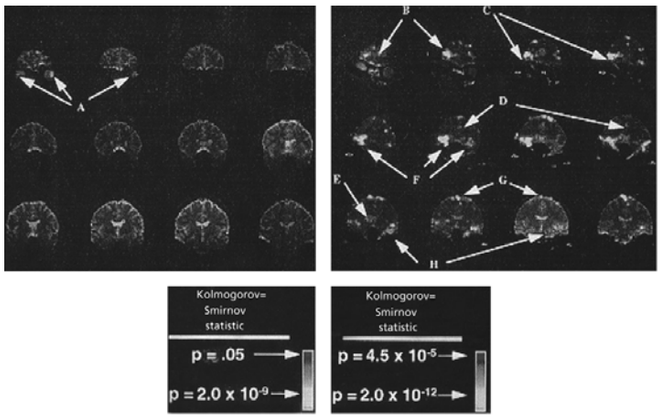


FIGURE 113.8. Results for one normal subject (left) and one patient with obsessive-compulsive disorder (right) (normal subject 2 and patient 9 [trial B]), juxtaposed for comparison. The gradient echo functional data are shown as a - log (p) map (Kolmogorov-Smirnov statistic) in color, superimposed over a T_2-weighted high-resolution instascan image in gray tone, for anatomic reference. Twelve contiguous slices are shown for each subject. The threshold for the control subject is at a lower level to emphasize the absence of activation, while the patient's threshold is at a more stringent level ($P < 10_{-7}$, approximating Bonferroni-corrected $P < .01$). A, eye movement; B, middle frontal cortex; C, inferior frontal cortex; D, cingulate cortex; E, caudate nucleus; F, orbital frontal cortex; G, superior frontal cortex; H, temporal cortex. Reprinted from Breiter HC, Rauch SL, Kwong KK, et al. Functional magnetic resonance imaging of symptom provocation in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1996;53:595–606. See color version of figure.

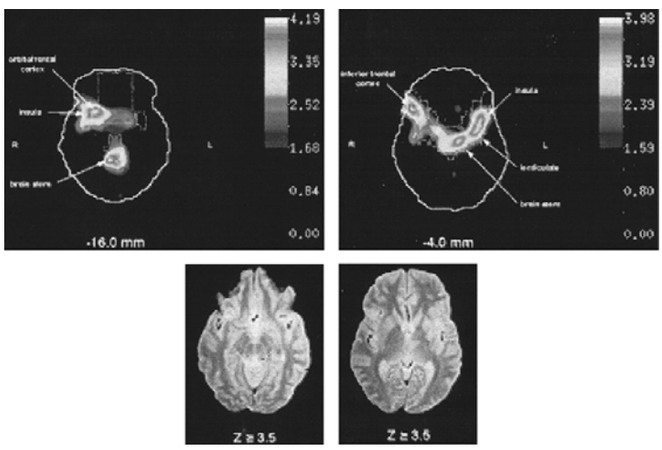


FIGURE 113.9. Positron emission tomography (PET) statistical parametric map of provoked minus control conditions for pooled data from three anxiety disorders. Horizontal slices from the statistical parametric map, at 16 and 4 mm inferior to the anterior commissure-posterior commissure plane respectively, are oriented according to neuroimaging convention (*top*, anterior; *bottom*, posterior; *right*, left; *left*, right), and each is displayed in two ways. In the upper panels, Z-score values are illustrated via a Sokoloff color scale. White dashed outlines reflecting the boundaries of brain regions of interest, as defined via a digitized version of the Talairach atlas, are superimposed for anatomic reference; solid lines demarcate the boundaries of the whole brain slice. In the lower panels, contiguous pixels exceeding a Z-score threshold of 3.5 (corresponding to approximately $P < .05$, corrected for multiple comparisons) are shown in red. The findings are superimposed over a structural (T2-weighted) magnetic resonance image transformed to Talairach space for approximate anatomic reference. The magnetic resonance reference image is from a nominally normal subject, who did not participate in these PET studies. Reprinted from Rauch SL, Savage CR, Alpert NM, et al. The functional neuroanatomy of anxiety: a study of three disorders using positron emission tomography and symptom provocation. *Biol Psychiatry* 1997;42(6):446–452. See color version of figure.

Functional Neuroimaging Studies in OCD: Implications for Treatment Development

Recent investigation suggests that differential baseline patterns of brain activity in OCD patients may predict differential response to specific treatment interventions (e.g., CBT,

SSRI) (147). For example, PET scans performed before and after 10 weeks of treatment with either CBT or the SSRI, fluoxetine have identified significant and comparable reductions in right caudate glucose metabolism associated with reduction in OCD symptom severity (38 ,148) (Fig. 113.10). Pathological correlations among orbital frontal cortex, the caudate nucleus and the thalamus were observed in pretreatment OCD patients but not in healthy volunteers. These pathologic correlations were eliminated after effective treatment with either SSRI or CBT. Subsequent analysis of this data (38 ,148) has demonstrated that specific patterns of metabolic activity in left orbital frontal cortex predicted response to CBT and the SSRI, fluoxetine (68). Specifically, decreased left orbital frontal-to-hemisphere metabolic ratios at baseline predicted better response to fluoxetine, whereas increased left orbital frontal-hemisphere metabolic ratios at baseline predicted better response to CBT. Saxena and colleagues (147) has extended this finding in OCD patients treated with paroxetine noting significant decreases in glucose metabolism in right anterior orbital frontal cortex and the right caudate nucleus in treatment responders but not in nonresponders (Fig. 113.11). Decreased metabolic activity in the left and right orbital frontal cortex predicted better response to paroxetine with greater reduction in OCD symptom severity (Fig. 113.12).

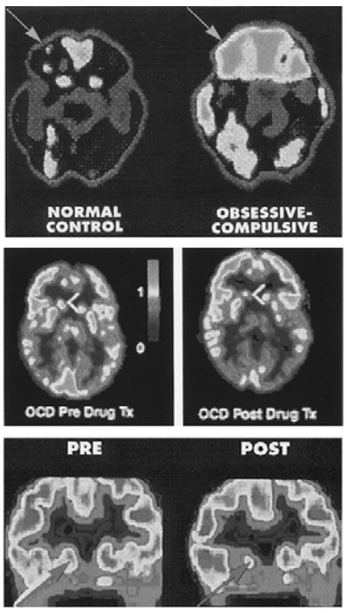


FIGURE 113.10. Positron emission tomography studies demonstrating increased right caudate glucose metabolism in a patient with obsessive-compulsive disorder compared to a healthy volunteer *(top)*. Note the significant decrease in right caudate glucose metabolism after pharmacotherapy *(middle)* and cognitive behavioral therapy *(bottom)*. Reprinted from Schwartz JM. *Brain lock*. New York: HarperCollins, 1996. Baxter LR, Schwartz JM, et al. Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:681–689. See color version of figure.

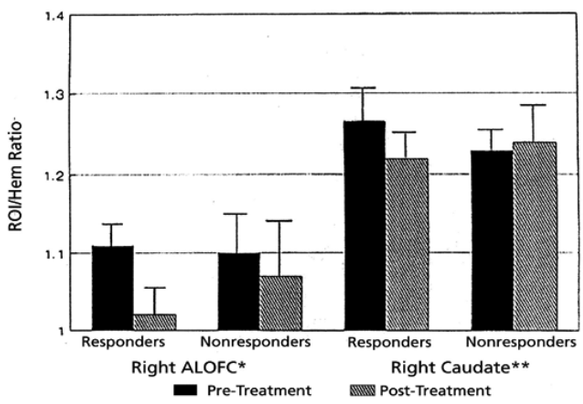


FIGURE 113.11. Mean pretreatment and posttreatment metabolic ratios (± SD) in right anterior lateral orbital frontal cortex (ALOFC) and right caudate, normalized to ipsilateral hemisphere (ROI/Hem), in responders versus nonresponders to paroxetine. There was a significant difference in the magnitude of change in right ALOFC/Hem between responders (1.11 ± .05 pretreatment to 1.02 ± 0.05 posttreatment) and nonresponders (1.10 ± .06 pretreatment to 1.07 ± .08 posttreatment). Mean right Cd/Hem decreased significantly in treatment responders (1.27 ± .06 to 1.22 ± .05) but not in nonresponders (1.23 ± .04 to 1.24 ± .05). *P = .04; **P = .01). Reprinted from Saxena S, Brody AL, Maidment KM, et al. Localized orbitofrontal and subcortical metabolic changes and predictors of response to paroxetine treatment in obsessive-compulsive disorder. *Neuropsychopharmacology* 1999;21:683–693.

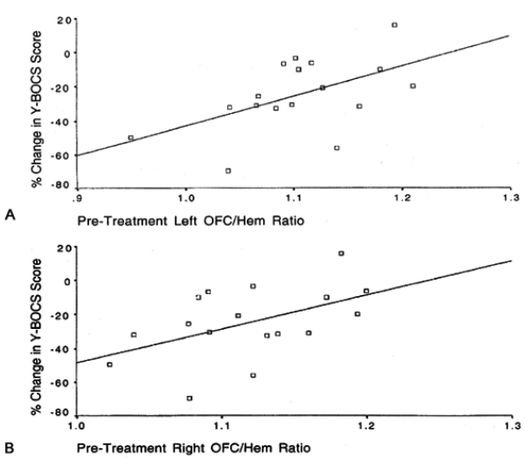


FIGURE 113.12. A: Scatter plot of pretreatment glucose metabolic rate in left orbitofrontal cortex, normalized to ipsilateral hemisphere (left OFC/Hem), and change in Y-BOCS score after paroxetine treatment (Kendall's tau = -.39, P = .01); **B:** Scatter plot of pretreatment glucose metabolic rate in right orbitofrontal cortex, normalized to ipsilateral hemisphere (right OFC/Hem), and change in YBOCS score after paroxetine treatment (Kendall's tau = -.35, P = .02). Reprinted from Saxena S, Brody AL, Maidment KM, et al. Localized orbitofrontal and subcortical metabolic changes and predictors of response to paroxetine treatment in obsessive-compulsive disorder. *Neuropsychopharmacology* 1999;21:683–693.

It should be noted, however, that functional imaging data sets have not been entirely consistent in OCD patients studied before and after treatment intervention. Benkelfat and colleagues (149), for example, observed a significant decrease in caudate and anterior orbital frontal glucose metabolism but only the decrease in caudate metabolism was associated with reduction in OCD symptom severity. In contrast, Swedo and co-workers (150) reported no change in caudate metabolic activity after 2 months of clomipramine treatment in OCD patients with childhood onset of

illness. Decreased baseline right orbitofrontal and anterior cingulate metabolic rates, however, did predict better response to clomipramine treatment. Rubin and associates (151) also observed no caudate metabolic changes before and after SSRI treatment but found that decreased metabolism in orbitofrontal cortex before treatment predicted greater reduction in OCD symptom severity.

Strikingly, there have been no published functional neuroimaging studies in children or adolescents with OCD. Techniques such as fMRI may be especially relevant to the study of childhood populations because there are no putative ionizing radiation risks facilitating repeated study for longitudinal follow-up with a neurodevelopmental perspective. This is an active area of investigation in our laboratory where preliminary symptom provocation modeling investigation of adult OCD patients (42 ,141) has demonstrated increased activation in ventral prefrontal–striatal circuitry (unpublished data). This is especially relevant in view of recent clinical neurodevelopmental models of OCD (88).

Taken together, functional neuroimaging studies are already beginning to integrate and translate advances in neuroscience into treatment development so that specific neural network activations might help predict patients more (or less) likely to respond to a particular treatment (e.g., CBT or SSRI) (147). Recent advances allowing for the noninvasive real-time measurement of brain activity provide an unprecedented window of opportunity for unlocking the mechanisms underlying the pathogenesis of OCD with important implications for treatment development.

NEUROCHEMICAL STUDIES IN OCD

Part of "113 - Imaging and Neurocircuitry of OCD "

Neuronal Viability

To our knowledge, there have only been four neuroimaging studies directly measuring brain chemistry in OCD. Proton magnetic resonance spectroscopy (1-H MRS), which can measure compounds including the neuronal marker, *N*-acetyl-aspartate (NAA) (152), cytosolic choline containing compounds (Cho), glutamatergic compounds including glutamate, glutamine and GABA (Glx), creatine/phosphocreatine (Cr), and myoinositol (mI). Like fMRI, there are no ionizing radiation risks, making it a particularly child-friendly technique facilitating longitudinal monitoring of patients before and after treatment intervention.

Prior investigation in adult OCD patients and those with epilepsy has found that 1-H MRS NAA measurement may be a more sensitive method for identifying neuronal dysfunction than morphometric MRI assessment (79 ,153). Reduced striatal NAA levels without striatal volumetric differences were observed in OCD patients compared to controls, suggesting that 1-H MRS NAA measurement can detect neuronal loss at a magnitude undetectable by morphometric MRI (79). Ebert and colleagues (154) also reported reduced NAA/Cr levels in the striatum and anterior cingulate cortex but not parietal white matter of adult OCD patients. Volumetric data for the regions of interest were not provided.

Fitzgerald and co-workers (155) compared 11 treatment-naive pediatric OCD patients and 11 age- and sex-case-matched healthy comparison subjects and found localized

functional neurochemical marker abnormalities in right and left medial but not lateral thalamus. Reduced NAA/Cho +Cr levels associated with OCD symptom severity were observed in OCD patients compared to controls (Fig. 113.13). This is intriguing because the dorsomedial nucleus of the thalamus has been especially implicated in the pathogenesis of OCD (15).

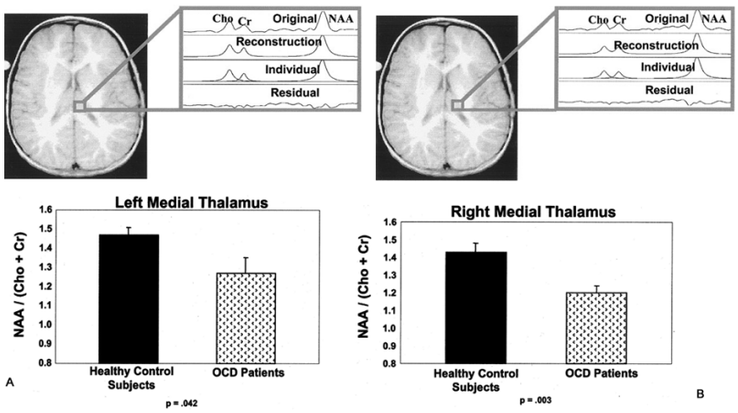


FIGURE 113.13. Sample spectra for voxels *(top)* placed in the left medial thalamus *(A)* and left lateral thalamus *(B)*. Individual peaks for choline compounds (Cho), creatine/phosphocreatine (Cr), and *N*-acetylaspartate (NAA) were resolved from the original spectrum, leaving a residual. NAA/(Cr + Cho) metabolite ratios by group *(bottom)* for left *(A)* and right *(B)* medial thalamus. OCD, obsessive- compulsive disorder. Adapted from Fitzgerald KD, Moore GJ, Paulson LD, et al. Proton spectroscopic imaging of the thalamus in treatment-naive pediatric obsessive-compulsive disorder. *Biol Psychiatry* 2000;47:174–182.

Neuronal dysfunction or loss in specific fronto-striatal-thalamic circuits in OCD could be related to excess brain activity measured in this circuitry with functional neuroimaging (154). Moreover, reduced NAA levels in fronto-striatal-thalamic regions may result from increased glutamatergic afferent projections to the thalamus and striatum (79 ,156). Excess glutamatergic afferent input could be neurotoxic to fronto-striatal neurons and, thereby, result in reduced NAA levels.

Glutamatergic Dysfunction in OCD

Rosenberg and Keshavan (88) hypothesized that anatomic and functional abnormalities in cortico-striatao-thalamo-cortical networks may result from disruptions in glutamatergic modulation of serotonin neurotransmission. The majority of axon terminals in the basal ganglia are glutamatergic afferents (31 ,157 ,158), with the caudate nucleus receiving an especially massive glutamatergic innervation from ventral prefrontal cortex (15 ,159 ,160). Ablation of frontal cortex results in a dramatic reduction in caudate glutamate concentrations (159 ,161). Becquet and associates (160) have shown that glutamate exerts a potent inhibitory effect on serotonin release in the caudate nucleus. Conversely, serotonergic neurons can modulate glutamate release (160) with stimulatory 5-HT2a receptors on GABAergic interneurons inhibiting glutamatergic projections from ventral prefrontal cortex to the striatum and thalamus (158). Increased glutamatergic afferent input from ventral prefrontal cortex to the striatum and thalamus, therefore, may be consistent with increased activation in this circuitry as measured by functional neuroimaging (141) and be related to the pathogenesis and maintenance of OCD. Chronic treatment with SSRIs results in a marked increase in 5-HT release in orbital frontal cortex of guinea pigs (162). SSRI treatment may, therefore, stimulate 5-HT2a receptors with consequent alterations in serotonergic release from cell bodies

in frontal cortex, which would alter cortico-striato-thalamic glutamate projections (Fig. 113.14).

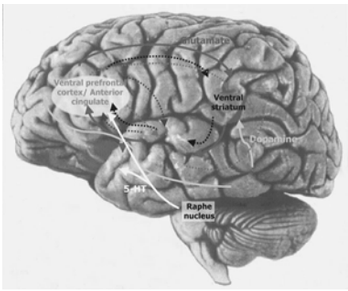


FIGURE 113.14. Illustration of cortico-striato-thalamo-cortico network in obsessive-compulsive disorder. Neurotransmitters: Glu, Glutamate; GABA-gamma-aminobutyric acid; DA, dopamine; 5-HT, serotonin. Brain regions: AC, anterior cingulate; RN, raphe nucleus; SN, substantia nigra; Thal, thalamus; TLC, temporal lobe cortex; VPFC, ventral prefrontal cortex; VS, ventral striatum; VT, ventral tegmentum. Adapted from Rosenberg DR, Keshavan MS. Toward a neurodevelopmental model of obsessive-compulsive disorder. *Biol Psychiatry* 1998;43:623–640.

Rosenberg and co-workers (98) recently reported significantly increased glutamatergic concentrations in the caudate nucleus but not occipital cortex of 11 treatment-naive pediatric OCD patients compared to 11 age- and sex-matched healthy controls (Fig. 113.15). After 12 weeks of monodrug therapy with paroxetine, a significant decrease in caudate but not occipital glutamatergic concentrations was observed (Fig. 113.16). Decrease in caudate glutamatergic concentrations was associated with reduction in OCD symptom severity so that higher pretreatment caudate glutamatergic concentrations predicted better response to paroxetine (Fig. 113.17). An active area of investigation in our laboratory involves comparisons of the impact of CBT, SSRI, and combination therapy on glutamatergic concentrations in the caudate and other brain regions as well as measuring the long-term stability of these changes (e.g., do changes persist after medication is discontinued?).

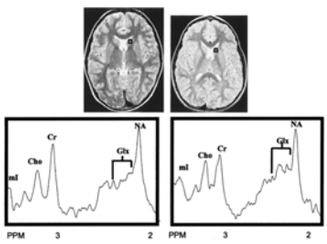


FIGURE 113.15. Spectra of a healthy control *(left)* compared to that of a patient with obsessive-compulsive disorder *(right)*. MI, myo-inositol; Cho, choline compounds; Cr, creatine/phosphocreatine; Glx, glutamate/glutamate/GABA; NA, N-acetylasparate. Adapted from Rosenberg DR, MacMillan SN, Moore GJ. Brain anatomy and chemistry may predict treatment response in pediatric obsessive-compulsive disorder. *Int J Neuropsychopharmacol* 2001;4:179–190.

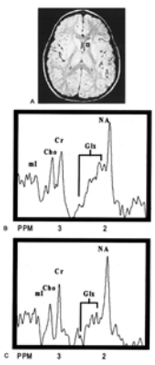


FIGURE 113.16. Magnetic resonance studies on a 9-year-old boy with obsessive-compulsive disorder. **A:** ^{1}H MR spectra were obtained from a 0.7-cc volume of interest centered in the left caudate as shown by the box on the T_1-weighted MR image. **B:** Spectrum obtained at baseline in the treatment-naive state. **C:** Spectrum obtained after a 12-week trial of the selective serotonin reuptake inhibitor, paroxetine. Cho, choline compounds; Cr, creatine/phosphocreatine; Glx, glutamate/glutamine; ml, myoinositol; MR, magnetic resonance; NA, *N*-acetylaspartate. Reprinted from Moore GJ, MacMaster FP, Stewart C, et al. Case study: caudate glutamatergic changes with paroxetine therapy for pediatric obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry* 1998;37(6):663–667.

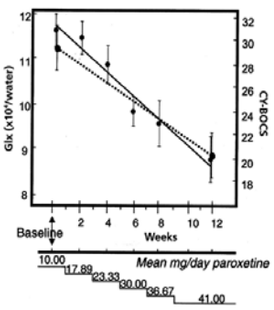


FIGURE 113.17. Decrease in left caudate glutamatergic concentrations associated with reduction in obsessive-compulsive score of the Children's Yale-Brown Obsessive-Compulsive Scales (CY-BOCS) (CY-BOCS, —; Glx [glutamate/glutamine/GABA], ·····). Reprinted from Rosenberg DR, Hanna GL. Review of genetic and imaging strategies in obsessive-compulsive disorder. Potential implications for treatment development. *Biol Psychiatry* 2000;48:1210–1222.

Findings of increased caudate glutamatergic concentrations that decreased after paroxetine treatment are consistent with functional neuroimaging studies that have reported increased metabolic rates, regional cerebral blood flow, and brain activation that decreased after SSRI treatment (38 ,147). Glutamatergic afferent terminals influence brain glucose metabolism so that regional brain glucose metabolism parallels glutamatergic activity (163). Serotonin agonists have also been shown to reduce glucose metabolism (164). Taken together, these data suggest a reversible glutamatergically mediated dysfunction in cortico-striato-thalamo-cortico circuitry.

Serotonergic Role in OCD

Perhaps the most compelling evidence for a serotonergic role in OCD comes from investigation that has consistently found serotonin reuptake inhibitors to be effective in reducing OCD symptoms, whereas medications affecting norepinephrine and dopamine appear to be less effective (165). Indirect support for the serotonin hypothesis of OCD comes from a plethora of platelet, cerebrospinal fluid (CSF), and pharmacologic challenge studies suggesting a serotonergic role in OCD. Platelet 3-H imipramine binding sites are considered to be putative markers of 5-HT function and are quite similar to those on presynaptic 5-HT neurons (166). Although a number of studies have reported decreased 5-HTPR in platelets of medication-free OCD patients (166 ,167 ,168 ,169 and 170), there are contradictory reports (171 ,172 ,173 and 174).

Studies of the serotonin metabolite, 5-hydroxy-indole acetic acid (5HIAA) have also found increased 5HIAA levels in OCD patients compared to controls (172 ,175), although contradictory reports exist (176). Swedo and colleagues (177) reported that higher pretreatment CSF 5HIAA levels were correlated with increased OCD symptom severity and predicted better response to clomipramine. In contrast, Asberg and co-workers (178) noted that OCD symptom severity was associated with lower pretreatment CSF 5HIAA levels, although higher pretreatment 5HIAA levels did predict better response to clomipramine. Blood and CSF serotonin measures can be influenced by a person's height, diet, season of the year, activity level, and menses (179), thereby limiting the reliability of these markers as indices of brain serotonin function.

5-HT receptor subtypes are localized to different areas of the brain (180). In order to better elucidate regional abnormalities of serotonin function that might contribute to OCD, Hollander and co-workers (181) combined pharmacologic challenge with functional neuroimaging. They used SPECT to measure the impact of the mixed 5-HT agonist–antagonist, meta-chlorophenylpiperazine (mCPP) in seven OCD patients. mCPP is a metabolite of trazodone, which acts most strongly at the postsynaptic 5-HT-2C receptor. Its administration results in exacerbation of OCD symptoms (181 ,182 ,183 ,184 ,185 ,186 ,187 and 188), although contradictory findings have been reported (189 ,190). Hollander and associates (181) noted a significant global increase in cortical blood flow in OCD patients that was associated with mCPP symptom exacerbation. Ho Pian and co-workers (190), also using SPECT, were unable to confirm this finding and, instead, reported decreased blood flow in the frontal cortex, striatum, thalamus, and cerebellum as well as global decreased blood flow in seven OCD patients with no mCPP response. The advent of more specific probes of serotonin synthesis and receptor function may help clarify putative serotonin dysregulation in OCD.

Recent PET investigation using the probe, α-C-11-methyl-L-tryptophan (AMT), an analogue of tryptophan may be especially instructive. Tryptophan is the amino acid precursor to serotonin so that AMT is converted to α-C-11—methyl serotonin. α-C-11-methylserotonin is a 5-HT variant that is not degraded by monoamine oxidase; therefore, AMT-derived serotonin is effectively “kept” in serotonin neurons making it measurable by PET (191). In healthy volunteers, Chugani and associates (191) used AMT PET studies and found high regional AMT uptake in frontal cortex, the striatum, thalamus, and temporal lobe, regions that receive dense serotonergic projections from the raphe nuclei in the midbrain. More recently, Chugani and colleagues (192) reported reduced AMT uptake, suggestive of decreased serotonin synthesis in the frontal cortex and thalamus in autistic boys 4 to 11 years old. The repetitive, ritualistic thoughts and behaviors of autism can be similar to those in OCD. SSRIs sometimes can be beneficial in treating compulsive behaviors observed in autistic patients such as head banging and other rituals (193).

Pilot AMT PET studies in our laboratory have compared AMT uptake in pediatric OCD patients, 8 to 17 years old

as compared to their unaffected siblings (194). Decreased uptake of AMT in the caudate nucleus and a trend for decreased uptake of AMT in anterior cingulate cortex was observed in the OCD patients. Because of the putative radiation risks associated with PET, we are not able to study healthy children as a comparison group. The radiation risks also make repeated studies in the same subjects less viable. Recent study has also suggested that the AMT tracer may be more reflective of free tryptophan than of serotonin synthesis (195). This is currently an active area of investigation in our center.

Other ligands such as 18F-altanserin, a 5-HT2a postsynaptic receptor antagonist may be relevant to the study of OCD. Goldman and associates (196) reported consistent 5-HT2a associations in two independent populations in two countries, implicating 5-HT2a in anorexia nervosa and OCD. Kaye and co-workers (197) have studied anorexic and bulimic women after long-term recovery with regular menstrual cycles and normalized eating and weight and found significantly reduced 18F altanserin binding in bilateral orbital frontal regions but not in other brain regions as compared to healthy female controls. The authors hypothesized that increased extracellular 5-HT could compete with 18F altanserin binding at 5-HT2a receptors and, thereby down-regulate 5-HT2a postsynaptic receptors. Anorexia nervosa and bulimia share certain characteristics with OCD and both conditions can also benefit from SSRI treatment (198).

This ligand as well as others being developed and in the pipeline may ultimately clarify more precisely the role of serotonin in OCD.

CONCLUSION

Part of "113 - Imaging and Neurocircuitry of OCD "

Advances in brain imaging and neuroscience are making the brain mechanisms involved in the pathogenesis and maintenance of OCD accessible as never before. Indeed, the time is now ripe to use brain imaging to exploit advances across several disciplines. For example, sophisticated brain imaging studies may facilitate more informed genetic studies and vice versa. Rauch (67) has argued persuasively that brain-imaging studies may help delineate specific endophenotypes for genetic studies in OCD. Such an approach might help discriminate between familial and sporadic OCD, thereby clarifying some of the genetic heterogeneity of OCD (11). To our knowledge, no prior study has employed both genetic and neuroimaging approaches in the same population, although there is precedent for such an approach in the study of ADHD (199). Obviously, costs of such an endeavor prohibit routine use of both techniques. It will be critical to delineate areas of maximum utility where these techniques are likely to provide the best yield and, therefore be most cost effective (67). For example, identified candidate serotonin or glutamatergic genes might be integrated into specific neuroimaging paradigms (11). Neuroimaging studies may also help clarify the role of established susceptibility genes as well as facilitating an enhanced understanding of the developmental neurobiologic underpinnings of OCD. Combination of these approaches may be especially effective in the identification of meaningful surrogate neurobiologic markers predictive of treatment response (or lack thereof). Ample evidence exists across various medical disciplines that increased understanding of the biologic mechanisms underlying an illness inevitably translates into critical advances in diagnosis and treatment (200).

ACKNOWLEDGMENTS

Part of "113 - Imaging and Neurocircuitry of OCD "

This work was supported in part by the State of Michigan Joe Young Sr. Psychiatric Research and Training Program, and grants MH-01372 and MH-59299 from the National Institute of Health, Bethesda, Maryland, and the National Obsessive-Compulsive Disorder Foundation, Milford, Connecticut to Dr. Rosenberg.

REFERENCES

1. Rasmussen SA, Eisen JL. The epidemiology and differential diagnosis of obsessive-compulsive disorder. *J Clin Psychiatry* 1994;55(Suppl):5–14.

2. Robins LN, Helzer JI, Weissman MM, et al. Lifetime prevalence of specific psychiatric disorders in three sites. *Arch Gen Psychiatry* 1984;41:949–958.

3. Hanna GL. Demographic and clinical features of obsessive-compulsive disorder in children and adolescents. *J Am Acad Child Adolesc Psychiatry* 1995;34:19–27.

4. Valleni-Basile LA, Garrison CZ, Jackson KL, et al. Frequency of obsessive-compulsive disorder in a community sample of young adolescents. *J Am Acad Child Adolesc Psychiatry* 1994;33:782–791.

5. Flament MF, Whitaker A, Rapoport JL. Obsessive-compulsive disorder in adolescence: an epidemiological study. *J Am Acad Child Adolesc Psychiatry* 1988;27:764–771.

6. Pauls DL, Alsobrook JP II, Phil M, et al. A family study of obsessive-compulsive disorder. *Am J Psychiatry* 1995;152:76–84.

7. Hollander E, Schiffman E, Cohen B, et al. Signs of central nervous system dysfunction in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1990;47:27–32.

8. Vanzan Paladin A. Epilepsy in twentieth century literature. *Epilepsia* 1995;36(10):1058–1060.

9. Karbowski K. Electroencephalography and epileptology in the 20th century. *Schweiz Rundsch Med Prax* 1995;84(49):1465–1473.

10. Menkes JH. Paroxysmal disorders. In Menkes JH, ed. *Textbook of child neurology*, fourth ed. Philadelphia: Lea and Febiger, 1990:602–674.

11. Rosenberg DR, Hanna GL. Review of genetic and imaging strategies in obsessive-compulsive disorder: potential implications for treatment development. *Biol Psychiatry* 2000;48:179–190.

12. Nemeroff CB, Kilts CD, Berns GS. Functional brain imaging: twenty-first century phrenology or psychobiological advances for the millenium? *Am J Psychiatry* 1999;156(5):671–673.

13. Keshavan MS, Rosenberg DR. New trends in developmental neuroimaging in psychiatry. *Prog Neuropsychopharmacol Biol Psychiatry* 1999;23(4):557–560.

14. Rapoport JL, Wise SP. Obsessive-compulsive disorder: evidence for basal ganglia dysfunction. *Psychopharmacol Bull* 1988;24:380–384.

15. Modell JG, Mountz JM, Curtis GC, et al. Neurophysiologic dysfunction in basal ganglia/limbic striatal and thalamocortical circuits as a pathogenetic mechanism of obsessive-compulsive disorder. *J Neuropsychiatry* 1989;1:27–36.

16. Insel TR. Toward a neuroanatomy of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:739–744.

17. Cummings JL. Frontal-subcortical circuits and human behavior. *Arch Neurol* 1993;50:873–880.

18. Ames D, Cummings JL, Wirshing WC, et al. Repetitive and compulsive behavior in frontal lobe degenerations. *J Neuropsychiatry Clin Neurosci* 1994;6:100–113.

19. Berthier ML, Kulisevsky J, Gironell A, et al. Obsessive-compulsive disorder associated with brain lesions: clinical phenomenology, cognitive function, and anatomic correlates. *Neurology* 1996;47:353.

20. Laplane D, Levasseur M, Pillon B, et al. Obsessive-compulsive and other behavioral changes with bilateral basal ganglia lesions. A neuropsychological, magnetic resonance imaging and positron tomography study. *Brain* 1989;112:699.

21. von Economo C. *Encephalitis lethargica: its sequelae and treatment*. London: Oxford University Press, 1931.

22. Pitman RE, Green RC, Jenike MA, et al. Clinical comparison of Tourette's disorder and obsessive-compulsive disorder. *Am J Psychiatry* 1987;144:1166–1171.

23. Swedo SE, Rapoport JL, Cheslow DL. High prevalence of obsessive-compulsive symptoms in patients with Sydenham's chorea. *Am J Psychiatry* 1989;146:246–249.

24. Giedd JN, Rapoport JL, Kruesi MJP, et al. Sydenham's chorea: magnetic resonance imaging of the basal ganglia. *Neurology* 1995;45:2199–2202.

25. Giedd JN, Rapoport JL, Leonard HL, et al. Case study: acute basal ganglia enlargement and obsessive-compulsive symptoms in an adolescent boy. *J Am Acad Child Adolesc Psychiatry* 1996;35(7):913–915.

26. Giedd JN, Rapoport JL, Garvey MA, et al. MRI assessment of children with obsessive-compulsive disorder or tics associated with streptococcal infection. *Am J Psychiatry* 2000;157(2):281–283.

27. Peterson BS, Leckman JF, Tucker D, et al. Preliminary findings of antistreptococcal antibody titers and basal ganglia volumes in tic, obsessive-compulsive, and attention deficit/hyperactivity disorders. *Arch Gen Psychiatry* 2000;57:364–372.

28. Husby G, van de Rijn I, Zabriskie JB, et al. Antibodies reacting with cytoplasm of subthalamic and caudate nuclei neurons in chorea and acute rheumatic fever. *J Exp Med* 1976;144:1094–1110.

29. Swedo SE. Sydenham's chorea: a model for childhood autoimmune neuropsychiatric disorders. *JAMA* 1994;272:1788–1791.

30. Rauch SL, Whalen PJ, Dougherty DD, et al. Neurobiological models of obsessive-compulsive disorders. *Obsessive-compulsive disorders: practical management*. Boston: Mosby, 1998:222–253.

31. Alexander GE, Crutcher MD, DeLong MR. Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. In: Uylings HBM, VanEden CG, DeBruin JPC, et al, eds. *Progress in Brain Research*. New York: Elsevier, 1990:119–146.

32. Baxter LR, Saxena S, Brody AL, et al. Brain mediation of obsessive-compulsive disorder symptoms: evidence from functional brain imaging studies in the human and nonhuman primate. *Semin Clin Neuropsychiatry* 1996;1:32–47.

33. Cummings JL. Frontal-subcortical circuits and human brain behavior. *Arch Neurol* 1993;50:873–880.

34. Mega MS, Cummings JL. Frontal-subcortical circuits and neuropsychiatric disorders. *J Neuropsychiatry Clin Neurosci* 1994;6:358–370.

35. Butter CM, Snyder DR. Alterations in aversive and aggressive behaviors following orbital frontal lesions in rhesus monkeys. *Acta Neurobiol Exp* 1972;32:525.

36. Butters N, Butter C, Rosen J, et al. Behavioral effects of sequential and one-stage ablations of orbital prefrontal cortex in the monkey. *Exp Neurol* 1973;39:204.

37. Fuster JM. *The prefrontal cortex: anatomy, physiology and neuropsychology of the frontal lobe*, second ed. New York: Raven, 1989.

38. Baxter LR, Schwartz JM, Bergman KS, et al. Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:681–689.

39. Baxter LR. Positron emission tomography studies of cerebral glucose metabolism in obsessive-compulsive disorder. *J Clin Psychiatry* 1994;55:54S–59S.

40. Zald DH, Kim SW. Anatomy and function of the orbital frontal cortex. I: Anatomy, neurocircuitry, and obsessive-compulsive disorder. *J Neuropsychiatry Clin Neurosci* 1996;8:125–138.

41. Zald DH, Kim SW. Anatomy and function of the orbital frontal cortex. II: Function and relevance to obsessive-compulsive disorder. *J Neuropsychiatry Clin Neurosci* 1996;8:249–261.

42. Rauch SL, Jenike MA, Alpert NM, et al. Regional cerebral blood flow measured during symptom provocation in obsessive-compulsive disorder using oxygen 15-labeled carbon dioxide and positron emission tomography. *Arch Gen Psychiatry* 1994;51:62–70.

43. Abbruzzese M, Bellodi L, Ferri S, et al. Frontal lobe dysfunction in schizophrenia and obsessive-compulsive disorder. *Brain Cogn* 1995;27:202–212.

44. Rosenberg DR, Averbach DH, O'Hearn KM, et al. Oculomotor response inhibition abnormalities in pediatric obsessive-compulsive disorder. *Arch Gen Psychiatry* 1997;54:831–838.

45. Rosenberg DR, Dick EL, O'Hearn KM, et al. Response inhibition deficits in obsessive-compulsive disorder: an indicator of dysfunction in fronto-striatal circuits. *J Psychiatry Neurosci* 1997;22:29–38.

46. Schmidtke K, Schorb A, Winkelmann G, et al. Cognitive frontal lobe dysfunction in obsessive-compulsive disorder. *Biol Psychiatry* 1998;43:666–673.

47. Drewe EA. Go-no go learning after frontal lobe lesions in humans. *Cortex* 1975;11:8–16.

48. Eslinger PJ, Damasio AR. Severe disturbances of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology* 1985;35:1731–1741.

49. Chiocca EA, Martuza RL. Neurosurgical therapy of obsessive-compulsive disorder. In: Jenike MA, Baer L, Minichiello WE, eds. *Obsessive-conpulsive disorder: diagnosis and treatment*, 2nd ed. Chicago: Year Book, 1990.

50. Aggleton JP. *The amygdala: neurobiological aspects of emotion, memory and mental dysfunction*. New York: Wiley, 1992.

51. Bechara A, Tranel D, Damasio H, et al. Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* 1995;269:1115–1118.

52. Ketter TA, Andreason PJ, George MS, et al. Anterior paralimbic mediation of procaine-induced emotional and psychosensory experiences. *Arch Gen Psychiatry* 1996;53:59–69.

53. Irwin W, Davidson RJ, Lowe MJ, et al. Human amygdala activation detected with echo-planar functional magnetic resonance imaging. *Neuroreport* 1996;7:1765–1769.

54. Davis M. Neurobiology of fear responses: the role of the amygdala. *J Neuropsychiatry Clin Neurosci* 1997;9:382–402.

55. Davis M. The role of the amygdala in fear and anxiety. *Annu Rev Neurosci* 1992;15:353–375.

56. Nagy J, Zambo K, Decsi L. Anti-anxiety action of diazepam after intra-amygdaloid application in the rat. *Neuropharmacology* 1979;18:573–576.

57. Hodges H, Green S, Glenn B. Evidence that the amygdala is involved in benzodiazepine and serotonergic effects on punished responding but not on discrimination. *Psychopharmacology* 1987;92:491–504.

58. Costall B, Kelly ME, Naylor RJ, et al. Neuroanatomical sites of action of 5-HT-3 receptor agonist and antagonists for alteration of aversive behavior in the mouse. *Br J Pharmacol* 1989;96:325–332.

59. Gonzalez LE, Andrews N, File SE. 5-HT1A and benzodiazepine receptors in the basolateral amygdala modulate anxiety in the social interaction test, but not in the elevated plus-maze. *Brain Res* 1996;732:145–153.

60. Wise S, Rapoport JL. Obsessive-compulsive disorder: is it basal ganglia dysfunction? In: Rapoport JL, ed. *Obsessive-compulsive disorder in children and adolescents*. Washington, DC: American Psychiatric Press, 1989:327–347.

61. Jones EF. Cortical development and thalamic pathology. *Schizophr Bull* 1997;23:483–501.

62. Jones EG. *The thalamus*. New York: Plenum Press, 1985.

63. Bergmann F, Chaimovitz M, Pasternak V, et al. Compulsive gnawing in rats after implantation of drugs into the ventral thalamus. A contribution to the mechanism of morphine action. *Br J Pharmacol* 1974;51(2):197–205.

64. Sasaki K, Miyakawa M, Sudo T, et al. Evaluation of marble-burying behavior: induced alteration of monoamine metabolism in mouse brain. *Nippon Yakurigaku Zasshi* 1997;110(4):205–213.

65. Portenoy RK, Jarden JO, Sidtis JJ, et al. Compulsive thalamic self-stimulation: a case with metabolic, electrophysiologic and behavioral correlates. *Pain* 1986;27(3):277–290.

66. Saxena S, Brody AL, Schwartz JM, et al. Neuroimaging and frontal-subcortical circuitry in obsessive-compulsive disorder. *Br J Psychiatry* 1998;173(S35):26–38.

67. Rauch SL. Neuroimaging research and the neurobiology of obsessive-compulsive disorder: where do we go from here? *Biol Psychiatry* 2000;47(3):168–170.

68. Brody AL, Saxena S, Schwartz JM, et al. FDG-PET predictors of response to behavioral therapy versus pharmacotherapy in obsessive-compulsive disorder. *Psychiatry Res* 1998;84:1–6.

69. Andreason NC. *Brain imaging*. In Oldham JM, Riba MB, Tasman A, eds. Washington, DC: American Psychiatric Press, 1993:313–314.

70. Krishnan KRR, Doraiswamy PM. *Brain imaging in clinical psychiatry*. New York: Marcel Decker, 1997.

71. Keshavan MS, Krishnan RR. New frontiers in psychiatric neuroimaging. *Prog Neuropsychopharmacol Biol Psychiatry* 1997;21:1181–1183.

72. Johnstone EC, Crow TJ, Frith CD, et al. Cerebral ventricular and cognitive impairment in chronic schizophrenia. *Lancet* 1976;2(7992):924–926.

73. Shenton ME, Kikinis R, Jolesz FA, et al. Abnormalities of the left temporal lobe and thought disorder in schizophrenia. A quantitative magnetic resonance imaging study. *N Engl J Med* 1992;327:604–612.

74. Cendes F, Andermann F, Gloor P, et al. MRI volumetric measurement of amygdala and hippocampus in temporal lobe epilepsy. *Neurology* 1993;43:719.

75. Botteron KN, Figiel GS. The neuromorphometry of affective disorders. In: Krishnan KR, Doraiswamy PM, eds. *Brain imaging in clinical psychiatry*. New York: Marcel Decker, 1997:146–176.

76. Gilbert AR, Moore GJ, Keshavan MS, et al. Decrease in thalamic volumes of pediatric obsessive-compulsive disorder patients taking paroxetine. *Arch Gen Psychiatry* 2000;57(5):449–456.

77. Chakos MH, Lieberman JA, Bilder RM. Increase in caudate nuclei volumes of first-episode schizophrenic patients taking antipsychotic drugs. *Am J Psychiatry* 1994;151:1430–1436.

78. Keshavan MS, Bagwell WW, Haas GL, et al. Changes in caudate volume with neuroleptic treatment. *Lancet* 1994;344:1434.

79. Bartha R, Stein MB, Williamson PC, et al. A short echo 1H spectroscopy and volumetric MRI study of the corpus striatum in patients with obsessive-compulsive disorder and comparison subjects. *Am J Psychiatry* 1998;155(11):1584–1591.

80. Rosen BR, Buckner RL, Dale AM. Event-related functional MRI: past, present, and future. *Proc Natl Acad Sci USA* 1998;95:773–780.

81. Liu AK, Belliveau JW, Dale AM. Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. *Proc Natl Acad Sci USA* 1998;95:8945–8950.

82. Aylward EH, Schwartz J, Machlin S, et al. Bicaudate ratio as a measure of caudate volume on MR images. *Am J Neuroradiol* 1991;12:1217–1222.

83. Breiter HC, Filipek PA, Kennedy DN, et al. Retrocallosal white matter abnormalities in patients with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1994;51:663–664.

84. Jenike MA, Breiter HC, Baer L, et al. Cerebral structural abnormalities in obsessive-compulsive disorder: a quantitative morphometric magnetic resonance imaging study. *Arch Gen Psychiatry* 1996;53:625–632.

85. Robinson D, Wu H, Munne RA, et al. Reduced caudate nucleus volume in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1995;52:393–398.

86. Stein DJ, Coetzer R, Lee M. Magnetic resonance brain imaging in women with obsessive-compulsive disorder and trichotillomania. *Psychiatry Res* 1997;74:177.

87. Rosenberg DR, Keshavan MS, O'Hearn KM, et al. Fronto-striatal measurement of treatment-naive pediatric obsessive-compulsive disorder. *Arch Gen Psychiatry* 1997;54:824–830.

88. Rosenberg DR, Keshavan MS. Toward a neurodevelopmental model of obsessive-compulsive disorder. *Biol Psychiatry* 1998;43:623–640.

89. Behar D, Rapoport JL, Berg CJ, et al. Computerized tomography and neuropsychological test measures in adolescents with obsessive-compulsive disorder. *Am J Psychiatry* 1984;141:363–369.

90. Insel TR, Donnelly EF, Lalakea ML, et al. Neurological and neuropsychological studies of patients with obsessive-compulsive disorder. *Biol Psychiatry* 1983;18:741–751.

91. Rosenberg DR, Keshavan MS, Dick EL, et al. Corpus callosal morphology in treatment naive pediatric obsessive-compulsive disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 1997;21:1269–1283.

92. MacMaster FP, Dick EL, Keshavan MS, et al. Corpus callosal signal intensity in treatment naive pediatric obsessive-compulsive disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 1999;23(4):601–612.

93. Scarone S, Colombo C, Livian S, et al. Increased right caudate nucleus size in obsessive-compulsive disorder: detection with magnetic resonance imaging. *Psychiatry Res* 1992;45:115–121.

94. Aylward EH, Harris GJ, Hoehn-Saric R, et al. Normal caudate nucleus in obsessive-compulsive disorder assessed by quantitative neuroimaging. *Arch Gen Psychiatry* 1996;53:577–584.

95. Kellner CH, Jolley RR, Holgate RC, et al. Brain MRI in obsessive-compulsive disorder. *Psychiatry Res* 1991;36:45–49.

96. Luxenberg JS, Swedo SE, Flament MF, et al. Neuroanatomical abnormalities in obsessive-compulsive disorder determined with quantitative x-ray computed tomography. *Am J Psychiatry* 1988;145:1089–1093.

97. Stein DJ, Hollander E, Chan S, et al. Computed tomography and neurological soft signs in obsessive-compulsive disorder. *Psychiatry Res Neuroimag* 1993;50:143–150.

98. Rosenberg DR, MacMaster FP, Keshavan M, et al. Decrease in caudate glutamatergic concentrations in pediatric obsessive-compulsive disorder patients taking paroxetine. *J Am Acad Child Adolesc Psychiatry* 2000;39(9):1096–1103.

99. Peterson B, Riddle MA, Cohen DJ, et al. Reduced basal ganglia volumes in Tourette's syndrome using three-dimensional reconstruction techniques from magnetic resonance images. *Neurology* 1993;43:941–949.

100. Hebebrand J, Siemon P, Lutcke A, et al. A putaminal lesion in an adolescent with obsessive-compulsive disorder and atypical anorexia nervosa. *J Nerv Ment Dis* 1995;181:520–521.

101. Rothfield JM. Generalized dystonia and obsessive-compulsive disorder associated with bilateral circumscribed Magnetic Resonance signal changes in the putamen. *J Nerv Ment Dis* 1995;183:113–114.

102. Kiessling LS, Marcotte AC, Culpepper L. Antineuronal antibodies: tics and obsessive-compulsive symptoms. *J Dev Behav Pediatr* 1994;15:421–425.

103. Swedo SE, Leonard HL, Garvey M, et al. Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections: clinical description of the first 50 cases. *Am J Psychiatry* 1998;155(2):264–271.

104. Allen AJ, Leonard HL, Swedo SE. Case study: a new infection-triggered, autoimmune subtype of pediatric OCD and Tourette's syndrome. *J Am Acad Child Adolesc Psychiatry* 1995;34:307–311.

105. Grachev ID, Breiter HC, Rauch SL, et al. Structural abnormalities of frontal neocortex in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1998;55(2):181–182.

106. Caviness VS Jr, Kennedy DN, Richelme C, et al. The human brain age 7 to 11 years: a volumetric analysis based on magnetic resonance images. *Cereb Cortex* 1996;6(5):726–736.

107. Szeszko PR, Robinson D, Alvir JM, et al. Orbital frontal and amygdala volume reductions in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1999;56:913–919.

108. Rosvold HE, Mishkin M. Non-sensory effects of frontal lesions on discrimination learning and performance. In: Delafresnaye JF, ed. *Brain mechanisms and learning*. Oxford: Blackwell Scientific, 1961:555–576.

109. Passingham RE. Visual discrimination learning after selective prefrontal ablations in monkeys. *Neuropsychologia* 1972;10:27–39.

110. Diamond A. Developmental progression in human infants and infant monkeys, and the neural bases of, inhibitory control of reaching. In: Diamond A, ed. *The development and neural bases of higher cognitive functions*. New York: Academy of Science Press, 1990:267–317.

111. Luria AR. *Higher cortical function in man*, second ed. New York: Basic Books, 1996.

112. Stuss DT, Benson DF. Frontal lobe lesions and behavior. In: Kertesz A, ed. *Localization in neuropsychology*. New York: Academic Press, 1983:429–449.

113. Iversen SD, Mishkin M. Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp Brain Res* 1970;11:376–386.

114. Goodglass H, Kaplan E. *An assessment of aphasia and related disorders*. Philadelphia: Lea and Febiger, 1972.

115. Rosenkilde CE. Functional heterogeneity of the prefrontal cortex in the monkey: a review. *Behav Neural Biol* 1979;25:301–345.

116. Malloy P, Birhrle A, Duffy J, et al. The orbitomedial frontal syndrome. *Arch Clin Neuropsychol* 1993;8:185–201.

117. Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 1989;61:331–349.

118. Goldman PS, Rosvold HE. Localization of function within the dorsolateral prefrontal cortex of the rhesus monkey. *Exp Neurol* 1970;27:291–304.

119. deLacoste MC, Kirkpatrick JB, Ross ED. Topography of the human corpus callosum. *J Neuropathol Exp Neurol* 1985;44:578–591.

120. Seltzer B, Pandya DN. The topography of commissural fibers. In: Jasper HH, Lepore F, Ptito M, eds. *Two hemispheres, one brain. Functions of the corpus callosum*. New York: Liss, 1986:47–73.

121. Rakic P. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends Neurosci* 1995;18:383–388.

122. Yakovlev PI, Lecours AR. The myelogenetic cycles of regional maturation of the brain. In: Minkowski A, ed. *Regional development of the brain in early life*. Oxford: Blackwell Scientific Publications, 1967:3–70.

123. Papez JW. A proposed mechanism of emotion. *Arch Neurol Psychiatry* 1937;38:725–743.

124. Halgren E, Walter RD, Cherlow DG, et al. Mental phenomena evoked by electrical stimulation of the human hippocampal formation and amygdala. *Brain* 1978;10(1):83–117.

125. Lane RD, Reiman EM, Bradley MM, et al. Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia* 1997;35(11):1437–1444.

126. Jerslid AT. *The psychology of adolescence*. New York: MacMillan, 1963.

127. Diener E, Sandvik E, Larsen RF. Age and sex effects for affect intensity. *Dev Psychol* 1985;21:542–546.

128. Weinberger DR. A connectionist approach to the prefrontal cortex. *J Neuropsychiatry* 1993;5:241–253.

129. Bogerts B, Meertz E, Schonfeld-Bausch R. Basal ganglia and limbic system pathology in schizophrenia. *Arch Gen Psychiatry* 1985;42:784–791.

130. Ashtari M, Zito JL, Gold BI, et al. Computerized volume measurement of brain structure. *Invest Radiol* 1990;25:798–805.

131. Shedlack KJ, Lee E-K, Radtke RA, et al. Ipsilateral subcortical atrophy associated with temporal lobectomy. *Psychiatry Res* 1994;54:295–304.

132. Bergin PS, Raymond AA, Free SL, et al. Magnetic resonance volumetry. *Neurology* 1994;44(9):1770–1771.

133. Rosenberg DR, Benazon NR, Gilbert AR, et al. Thalamic volume in pediatric obsessive-compulsive disorder patients before and after cognitive behavioral therapy. *Biol Psychiatry* 2000;48:294–300.

134. Bennett-Clarke CA, Lane RD, Rhoades RW. Fenfluramine depletes serotonin from the developing cortex and alters thalamocortical organization. *Brain Res* 1995;702(1 and 2):255–260.

135. Rauch SL, Savage CR. Neuroimaging and neuropsychology of the striatum. Bridging basic science and clinical practice. *Psychiatr Clin North Am* 1997;20:741.

136. Baxter LR, Phelps ME, Mazziotta JC, et al. Local cerebral glucose metabolic rates in obsessive-compulsive disorder: a comparison with rates in unipolar depression and normal controls. *Arch Gen Psychiatry* 1987;44:211–218.

137. Baxter LR, Schwartz JM, Mazziotta JC. Cerebral glucose metabolic rates in non-depressed patients with obsessive-compulsive disorder. *Am J Psychiatry* 1988;145:1560–1563.

138. Nordahl TE, Benkelfat C, Semple WE, et al. Cerebral glucose metabolic rates in obsessive-compulsive disorder. *Neuropsychopharmacology* 1989;2:23–28.

139. Swerdlow NR. Serotonin, obsessive-compulsive disorder, and the basal ganglia. *Int Rev Psychiat* 1995;7:115–129.

140. McGuire PK, Bench CJ, Frith CD, et al. Functional anatomy of obsessive-compulsive phenomena. *Br J Psychiatry* 1994;164:459–468.

141. Breiter HC, Rauch SL, Kwong KK, et al. Functional magnetic resonance imaging of symptom provocation in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1996;53:595–606.

142. Mindus P, Ericson K, Greitz T, et al. Regional cerebral glucose metabolism in anxiety disorders studied with positron emission tomography before and after psychosurgical intervention. A preliminary report. *Acta Radiol Suppl (Stockh)* 1986;369:444.

143. Harris GJ, Hoehn-Saric R, Lewis RW, et al. Mapping of SPECT regional cerebral perfusion abnormalities in obsessive-compulsive disorder. *Hum Brain Map* 1994;1:237–248.

144. Machlin SR, Harris GJ, Pearlson GD, et al. Elevated medial-frontal cerebral blood flow in obsessive-compulsive patients: a SPECT study. *Am J Psychiatry* 1991;148:1240–1242.

145. Benkelfat C, Bradwejn J, Meyer E, et al. Functional neuroanatomy of CCK_4-induced anxiety in normal healthy volunteers. *Am J Psychiatry* 1995;152:1180–1184.

146. Rauch SL, Savage CR, Alpert NM, et al. The functional neuroanatomy of anxiety: a study of three disorders using positron emission tomography and symptom provocation. *Biol Psychiatry* 1997;42(6):446–452.

147. Saxena S, Brody AL, Maidment KM, et al. Localized orbitofrontal and subcortical metabolic changes and predictors of response to paroxetine treatment in obsessive-compulsive disorder. *Neuropsychopharmacology* 1999;21:683–693.

148. Schwartz JM, Stoessel PW, Baxter LR, et al. Systematic changes in cerebral glucose metabolic rate after successful behavior modification treatment of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1996;53:109–113.

149. Benkelfat C, Nordahl TE, Semple WE, et al. Local cerebral glucose metabolic activity in obsessive-compulsive disorder: patients treated with clomipramine. *Arch Gen Psychiatry* 1990;47:840–848.

150. Swedo SE, Pietrini P, Leonard HL. Cerebral glucose metabolism in childhood-onset obsessive-compulsive disorder: revisualization during pharmacotherapy. *Arch Gen Psychiatry* 1992;49:690–694.

151. Rubin RT, Ananth J, Villanueva-Meyer J, et al. Regional ^{133}Xenon cerebral blood flow and cerebral ^{99m}Tc-HMPAO uptake in patients with obsessive-compulsive disorder before and during treatment. *Biol Psychiatry* 1995;38:429–437.

152. Birken DL, Oldendorf WH. *N*-acetyl-L-aspartic acid: a literature review of a compound prominent in 1H-NMR spectroscopic studies of brain (review). *Neurosci Biobehavior Rev* 1989;13:23–31.

153. Ebisu T, Rooney WD, Graham SH, et al. N-acetylaspartate as an in vivo marker of neuronal viability in kainate-induced status epilepticus: 1H magnetic resonance spectroscopic imaging. *J Cereb Blood Flow Metab* 1994;14:373–382.

154. Ebert D, Speck O, Konig A, et al. 1-H magnetic resonance spectroscopy in obsessive-compulsive disorder: evidence for neuronal loss in the cingulate gyrus and the right striatum. *Psychiatry Res* 1997;74(3):173–176.

155. Fitzgerald KD, Moore GJ, Paulson LD, et al. Proton spectroscopic imaging of the thalamus in treatment-naive pediatric obsessive-compulsive disorder. *Biol Psychiatry* 2000;47:174–182.

156. Salt TE, Eaton SA. Functions of inotropic and metabotropic glutamate receptors in sensory transmission in the mammalian thalamus. *Prog Neurobiol* 1996;48(1):55–72.

157. Parent A, Cote PY, Lavoie B. Chemical anatomy of primate basal ganglia. *Prog Neurobiol* 1995;46:131–197.

158. Parent A, Hazrati LN. Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Res* 1995;20:91–127.

159. Kim JS, Hassler R, Haug P, et al. Effect of frontal cortex ablation on striatal glutamic acid level in rat. *Brain Res* 1977;132:370–374.

160. Becquet D, Faudon M, Hery F. In vivo evidence for an inhibitory glutamatergic control of serotonin release in the cat caudate nucleus: involvement of GABA neurons. *Brain Res* 1990;519:82–88.

161. Calabresi P, Pisani A, Mercuri NB, et al. The corticostriatal projection: from synaptic plasticity to dysfunction of the basal ganglia. *Trends Neurosci* 1996;19:279–280.

162. El mansari M, Bouchard C, Blier P. Alteration of serotonin release in the guinea pig orbito-frontal cortex by selective serotonin reuptake inhibitors: relevance to treatment of obsessive-compulsive disorder. *Neuropsychopharmacology* 1995;13:117–127.

163. Sibson NR, Dhankhar A, Mason GF, et al. In vivo 13C NMR measurements of cerebral glutamine synthesis as evidence for glutamate-glutamine cycling. *Proc Natl Acad Sci USA* 1997;94(6):2699–2704.

164. Grome JJ, Harper AM. Local cerebral glucose utilisation following indoleamine- and piperazine-containing 5-hydroxytryptamine agonists. *J Neurochem* 1986;46(1):117–124.

165. Grados M, Scahill L, Riddle MA. Pharmacotherapy in children and adolescents with obsessive-compulsive disorder. *Child Adolesc Psychiatr Clin N Am* 1999;8(3):617–634.

166. Marazziti D, Pfanner C, Palego L, et al. Changes in platelet markers of obsessive-compulsive patients during a double-blind trial of fluvoxamine versus clomipramine. *Pharmacopsychiatry* 1997;30:245.

167. Bastani B, Arora RC, Meltzer HY. Serotonin uptake and imipramine binding in the blood platelets of obsessive-compulsive disorder patients. *Biol Psychiatry* 1991;30:131.

168. Marazziti D, Hollander E, Lensi P, et al. Peripheral markers of serotonin and dopamine function in obsessive-compulsive disorder. *Psychiatry Res* 1992;42:41–51.

169. Weizman A, Carmi M, Hermesh H, et al. High-affinity imipramine binding and serotonin uptake in platelets of eight adolescent and ten adult obsessive-compulsive patients. *Am J Psychiatry* 1986;143:335.

170. Sallee FR, Richman H, Beach K, et al. Platelet serotonin transporter in children and adolescents with obsessive-compulsive disorder or Tourette's syndrome. *J Am Acad Child Adolesc Psychiatry* 1996;35:1647.

171. Black DW, Kelly M, Myers C, et al. Tritiated imipramine binding in obsessive-compulsive volunteers and psychiatrically normal controls. *Biol Psychiatry* 1990;2:319.

172. Insel TR, Mueller EA, Alterman I, et al. Obsessive-compulsive disorder and serotonin: is there a connection? *Biol Psychiatry* 1985;20:1174–1188.

173. Kim SW, Dysken MW, Pandey GN, et al. Platelet 3H-imipramine binding sites in obsessive-compulsive behavior. *Biol Psychiatry* 1991;30:467.

174. Vitiello B, Shimon H, Behar D, et al. Platelet imipramine binding and serotonin uptake in obsessive-compulsive patients. *Acta Psychiatr Scand* 1991;84:29.

175. Thoren P, Asberg M, Bertilsson L, et al. Clomipramine treatment of obsessive-compulsive disorder. II. Biochemical aspects. *Arch Gen Psychiatry* 1980;37:1289.

176. Leckman JF, Goodman WK, Anderson GM, et al. Cerebrospinal fluid biogenic amines in obsessive-compulsive disorder, Tourette's syndrome, and healthy controls. *Neuropsychopharmacology* 1995;12:73.

177. Swedo SE, Leonard HL, Kruesi MJP, et al. Cerebrospinal fluid neurochemistry in children and adolescents with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:29–36.

178. Asberg M, Thoren P, Bertilsson L. Clomipramine treatment of obsessive disorder: biochemical and clinical aspects. *Psychopharmacol Bull* 1982;18:13.

179. Insel TR, Winslow JT. Neurobiology of obsessive-compulsive disorder. *Psychiatr Clin North Am* 1992;15:813–823.

180. Graeff FG. Serotonergic systems. *Psychiatr Clin North Am* 1997;20:723.

181. Hollander E, Prohovnik I, Stein DJ. Increased cerebral blood flow during mCPP exacerbation of obsessive-compulsive disorder. *J Neuropsychiatry Clin Neurosci* 1995;7:485.

182. Pettibone DJ, Williams M. Serotonin-releasing effects of substituted piperazine in vitro. *Biochem Pharmacol* 1984;33:1531.

183. Curzon G, Kennett GA. m-CPP: a tool for studying behavioral responses associated with 5-HT1c receptors. *Trends Pharmacol Sci* 1990;11:181.

184. Seibyl JP, Krystal JH, Price LH, et al. Effects of ritanserin on the behavioral, neuroendocrine, and cardiovascular responses to meta-chlorophenylpiperazine in healthy human subjects. *Psychiatry Res* 1991;38:227.

185. Hollander E, DeCaria CM, Nitescu A, et al. Serotonergic function in obsessive-compulsive disorder. Behavioral and neuroendocrine responses to oral m-chlorophenylpiperazine and fenfluramine in patients and healthy volunteers. *Arch Gen Psychiatry* 1992;49:21–28.

186. Pigott TA, Hill JL, Grady TA, et al. A comparison of the behavioral effects of oral versus intravenous mCPP administration in OCD patients and the effect of metergoline prior to i.v. mCPP. *Biol Psychiatry* 1993;33:3.

187. Pigott TA, Zohar J, Hill JL, et al. Metergoline blocks the behavioral and neuroendocrine effects of orally administered m-chlorophenylpiperazine in patients with obsessive-compulsive disorder. *Biol Psychiatry* 1991;29:418.

188. Zohar J, Insel TR. Obsessive-compulsive disorder: psychobiological approaches to diagnosis, treatment, and pathophysiology. *Biol Psychiatry* 1987;22:667–687.

189. Goodman WK, McDougle CJ, Price LH, et al. m-Chlorophenylpiperazine in patients with obsessive-compulsive disorder: absence of symptom exacerbation. *Biol Psychiatry* 1995;38:138–149.

190. Ho Pian KL, Westenberg HGM, Den Boer JA, et al. Effects of meta-chlorophenylpiperazine on cerebral blood flow in obsessive-compulsive disorder and controls. *Biol Psychiatry* 1998;44:367.

191. Chugani DC, Muzik O, Chakraborty P, et al. Human brain serotonin synthesis capacity measured in vivo with alpha-(C-11) methyl-L-tryptophan. *Synapse* 1998;28:33–43.

192. Chugani DC, Muzik O, Rothermel R, et al. Altered serotonin synthesis in the dentatothalamocortical pathway in autistic boys. *Ann Neurol* 1997;42:666.

193. Rosenberg DR, Holttum J, Gershon S. *Textbook of pharmacotherapy for child and adolescent psychiatric disorder*. New York: Brunner/Mazel, 1994:104–150.

194. Rosenberg DR, Chugani DC, Muzik O, et al. Altered serotonin synthesis in fronto-striatal circuitry in pediatric obsessive-compulsive disorder. *Biol Psychiatry* 1998;S43:24.

195. Shoaf SE, Carson R, Hommer D, et al. Brain serotonin synthesis rates in rhesus monkeys determined by (11C) alpha-methyl-L-tryptophan and positron emission tomography compared to CSF 5-hydroxyindole-3-acetic acid concentrations. *Neuropsychopharmacology* 1998;19(5):345–353.

196. Goldman D, Enoch MA, Rotondo A, et al. The 5-HT2a-1438G: a polymorphism in anorexia nervosa and obsessive-compulsive disorder. *Biol Psychiatry* 1999;45:174S.

197. Kaye WH, Frank GK, Meltzer CC, et al. Anorexia and bulimia nervosa: altered serotonin modulation in the orbital frontal cortex? *Biol Psychiatry* 1999;45:173S.

198. Kaye WH, Weltzin TE, Hsu LK, et al. An open trial of fluoxetine in patients with anorexia nervosa. *J Clin Psychiatry* 1991;52:464–471.

199. Castellanos F, Lau E, Tayebi N, et al. Lack of an association between a dopamine-5 receptor polymorphism and attention-deficit/hyperactivity disorder: genetic and brain morphometric analyses. *Mol Psychiatry* 1998;3:431–434.

200. Baxter LR, Ackermann RF, Swerdlow NR, et al. Specific brain system mediation of obsessive-compulsive disorder responsive to either medication or behavior therapy. In: Goodman WK, Rudorfer MV, Maser JD, eds. *Obsessive-compulsive disorder: contemporary issues in treatment*. Mahwah, NJ: LEA, 2000:573–609.

114

Current and Experimental Therapeutics of Obsessive-Compulsive Disorder

Eric Hollander

Stefano Pallanti

Eric Hollander: Department of Psychiatry, Mount Sinai School of Medicine, New York, New York.

Stefano Pallanti: Institute for Neurosciences, Florence, Italy.

UNIQUE ASPECTS OF OCD PHARMACOTHERAPY

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

Obsessive-compulsive disorder (OCD) is a chronic disorder with substantial impact on quality of life (1) that affects 2% to 3% of the US population (2,3). Previously believed to be refractory to treatment, the symptoms of OCD are substantially reduced with the use of medications having potent effects on blocking the serotonin (5-HT) transporter. The response to serotonin reuptake inhibitors (SRIs) in OCD is somewhat unique compared to that in other mood and anxiety disorders, in that higher doses and a longer lag-time to therapeutic effect of SSRIs may be required, and a lack of response to other antidepressant/antianxiety agents in OCD is evident. Nevertheless, SRIs do not cure patients, and 30% to 50% remain treatment nonresponders; therefore, other pharmacologic approaches, augmentation strategies, and especially cognitive-behavioral treatments may become necessary. Recently, other more invasive treatment options have been studied in the refractory population as well. This chapter highlights the psychopharmacology of OCD, including the current state of the art and future directions. Cognitive behavioral treatments for OCD, which are also highly effective, are not extensively reviewed here.

DIFFERENTIAL DIAGNOSIS AND COMORBIDITY: IMPLICATIONS FOR THERAPEUTICS

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

Although the diagnosis of OCD is usually straightforward, presenting with classic obsessive and compulsive symptoms, sometimes OCD presents with atypical features. Conversely, other disorders may present with symptoms reminiscent of OCD. For example, the ruminations of depression, intrusive thoughts, or delusions of psychotic disorders, and stereotyped behaviors of developmental disorders, may mimic OCD. Thus, comprehensive clinical evaluation and careful differential diagnosis are essential.

Psychiatric comorbidity is the rule rather than the exception in OCD. Up to two-thirds of all patients with OCD have lifetime comorbidities (3 ,4). These comorbid conditions not only serve to cloud the diagnostic picture, but also can influence the selection of optimal treatments.

The prevalence of obsessive-compulsive symptoms in schizophrenic patients has been estimated to range from 7.8% to 46.6% (5 ,6 and 7). A recent paper reported that in the early phase of the disorder, 14% of schizophrenic patients fulfilled criteria for a diagnosis of OCD (8). Atypical neuroleptics have been associated with both new onset and exacerbation of obsessions and compulsions, with numerous reports for clozapine (9) and less for risperidone. Treatment of OCD symptoms in schizophrenic patients may take into account this possible effect of atypical neuroleptics, and dosage reduction and/or SRI augmentation may be recommended (10).

The prevalence of OCD in patients with bipolar disorder has been estimated to be around 30% (11 and 12), half of whom had one or two other associated anxiety disorders (13); therefore, previous manic or hypomanic episodes and subthreshold hyperthymic characteristics should be evaluated and treated accordingly (14). Because SRIs may sometimes precipitate hypomanic or manic episodes in adults (15 ,16 and 17) and adolescents (18) without previous manic episodes, low initial doses, gradual dose elevation, and addition of mood stabilizers may be required.

IMPACT OF COMORBIDITY/SUBTYPES, OUTCOME MEASURES, AND RESPONDER CRITERIA

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

In evaluating treatment response in OCD, the patient population under study and measurement of response can significantly

impact findings. For example, OCD patients who fail to respond to prior SSRI therapeutic trials may have only a 25% chance to respond to an additional SSRI trial (19). Likewise, patients with comorbid tics (20), delusional symptoms, or schizotypical personality disorder (21) may not respond to SSRIs, and may require neuroleptic augmentation strategies. OCD patients with neurological soft-signs may be poorly responsive to SSRIs (22). Also, specific OCD subtypes, such as hoarders, may be poorly responsive to SSRIs. Finally, related or OCD spectrum disorders share important features with OCD and may be comorbid with OCD, influencing treatment outcome (23).

TREATMENT RESPONSE

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

The degree of symptom resolution in response to treatment determines the need for dosage adjustment, augmentation, or switching to an alternative treatment. Treatment response may be assessed qualitatively via periodic clinical interviews or the regular use of validated scales such as the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) to quantify the ongoing severity of OCD symptoms. Approximately 60% of patients treated with SRIs experience at least a 25% to 35% decrease in symptoms in Y-BOCS (24), which is typically operationalized as a criterion for response.

Although the Y-BOCS score generally is an excellent gauge of symptomatic improvement, the overall change in quality of life also must be considered. The Y-BOCS does not reflect these quality of life issues, and may not be sensitive to subtle changes, such as going from 8 to 2 hours per day of rituals. The criteria set for response (i.e., 25% or 35% reduction in Y-BOCS, CGI improvement of 1 or 2, or a combination of the two) may markedly impact percentage of subjects who are considered responders in various trials, and thus studies that utilize different response criteria might yield very different response rates.

SEROTONIN AND DOPAMINE FUNCTION IN OCD

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

Evidence for serotonergic and dopaminergic function in OCD is discussed in detail elsewhere (see Chapter 112). Extensive research documenting the efficacy of SRIs has indicated that the antiobsessional properties of these drugs may be related to their 5-HT reuptake inhibition properties. Peripheral and central marker abnormalities have generally supported a proposed 5-HT hypothesis (25). Of note, in a subgroup of OCD patients, increased levels of CSF 5-hydroxyindoleacetic acid (5-HIAA) have been observed and a correlation has been found between response to treatment and a reduction of both 5-HIAA levels and platelet 5-HT concentration (26-27). The partial serotonin 5-HT2C agonist *m*-chlorophenylpiperazine (*m*-CPP) has been found to acutely exacerbate symptoms in a subgroup of OCD patients in some (28 ,29) but not all studies (30), and has generally demonstrated neuroendocrine blunting in these patients as compared to normal controls (29 ,31). Treatment with clomipramine or fluoxetine leads to cessation of this behavioral exacerbation and normalization of the neuroendocrine findings in response to repeat *m*-CPP challenge (32 ,33). There is some evidence for linkage disequilibrium of the 5-HT1DB receptor gene and OCD, with preferential transmission of the G allele to affected subjects (34). To date, a specific abnormality of the 5-HT system in OCD has not been identified and the strongest evidence in support of the serotonin hypothesis remains the preferential response to SRIs.

There is a debate regarding the nature of the SRI-induced changes to the 5-HT system. Administration of the SRIs results in an immediate inhibition of the 5-HT transporter, with the effect of increasing synaptic 5-HT; however, the full clinical response may not be seen for up to 8 to 12 weeks of SRI treatment. An understanding of the neuroadaptive changes that take place with treatment is helpful in clarifying the mechanism of action involved. It has been reported that desensitization of 5-HT-2 receptors is implicated in the antiobsessional effect of SRIs (35). Alteration of serotonin release in the orbito-frontal cortex has been found to occur only after 8 weeks of treatment (36). These adaptive changes seem to involve a reduction in the number of receptors and altered responsivity of second messengers (37). There are many subtypes of 5-HT receptors, each having a distinct pattern of brain localization, with those expressed in basal ganglia and orbitofrontal regions of particular interest in the etiology of OCD (38).

There are complex structural and functional interactions between dopamine (DA) and 5-HT in the brain. Evidence implicating DA in the neurobiology of OCD is derived from a number of areas. In animal models, amphetamines have been shown to induce stereotypies that are viewed as compulsive behaviors (39). An association of postencephalitic Parkinson syndrome with obsessive-compulsive symptoms has been found (40). The comorbidity of Tourette syndrome and OCD is well described (41), as well as the association of a variety of other basal ganglia disorders with OCD. There is also evidence of DRD2 and DRD3 receptor gene polymorphisms in OCD (42).

SEROTONIN REUPTAKE INHIBITORS: ACUTE TRIALS

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

SRIs include clomipramine (Anafranil), fluoxetine (Prozac), sertraline (Zoloft), paroxetine (Paxil), fluvoxamine (Luvox), and citalopram (Celexa). Of these, fluoxetine, sertraline, paroxetine, fluvoxamine, and citalopram are “selective serotonin-reuptake inhibitors” (SSRIs), characterized by minimal affinity or pharmacologic action at receptor sites other

than the serotonin transporter. Clomipramine, on the other hand, is a tricyclic antidepressant (TCA). Thus, although clomipramine's serotonin-reuptake-blocking properties are similar to those of the SSRIs, it also has pharmacologically significant affinity for cholinergic and adrenergic receptors, thereby influencing its side-effect profile.

Although all SRIs block reuptake of serotonin presynaptically, there are other pharmacologic characteristics that distinguish these agents from one another, including differences in metabolism, half-life, protein binding, and effects on other neurotransmitter systems (43 ,44 and 45). All SRIs undergo hepatic metabolism and renal excretion. SRIs may be divided into those with active and those with inactive metabolites. For clomipramine, fluoxetine, and sertraline, which have active metabolites, the half-life of the daughter compound must be considered when estimating duration of effect after dosing. In general, measuring serum levels of SRIs and/or their metabolites has not been useful in determining effective dosage or predicting clinical response. Protein binding may influence levels of available drug in at least two ways: (a) if the patient is on other highly protein-bound medicines, competition can produce elevated levels of the SRI and/or the other agent, and (b) hypoproteinemia, as seen in chronic medical illness, malnutrition, or advanced age, can lead to higher concentrations of unbound SRI (i.e., the active form). There are subtle differences in the relative affinities of the various SRIs for monoaminergic reuptake sites, and more prominent differences between clomipramine and the SSRIs with respect to anticholinergic and antiadrenergic postsynaptic effects. These pharmacologic distinctions do not appear to have major consequences in terms of differential efficacy, but they do influence side-effect profile.

Case reports first suggested that obsessions might be successfully treated with clomipramine 30 years ago (46). Since then, controlled trials demonstrated clomipramine's efficacy in OCD. In the past decade, the introduction of the SSRIs has greatly increased treatment options (Table 114.1) (47 ,48 ,49 ,50 ,51 ,52 ,53 ,54 ,55 ,56 ,57 ,58 ,59 ,60 ,61 ,62 ,63 ,64 and 65).

TABLE 114.1. CONTROLLED TRIALS OF SEROTONIN REUPTAKE INHIBITORS THERAPY FOR OBSESSIVE-COMPULSIVE DISORDER IN ADULT PATIENTS

Studies (Ref.)	Conditions	Results
Clomipramine (CMI) vs Placebo or Non-SRIs		
Karabanow, 1977 (47)	CMI vs placebo	CMI superior to placebo
Montgomery, 1980 (48)	CMI vs placebo crossover	CMI superior to placebo
Mavissakalin et al, 1985 (49)	CMI vs placebo	CMI superior to placebo
Jenike, 1989 (50)	CMI vs placebo	73% improved on CMI
Greist et al., 1990 (51)	CMI vs placebo	73% improved on CMI 6% improved on placebo
CMI collaborative group, 1991 (52)	CMI vs placebo	38%–44% decrease Sx with placebo 3%–5% decrease Sx with placebo
Thoren et al., 1980 (53)	CMI vs nort. vs placebo	CMI, but not nort., superior to placebo
Ananth et al., 1981 (54)	CMI vs amitriptyline	CMI superior to amitriptyline
Insel et al., 1983 (55)	CMI vs clorgyline	CMI effective, clorgyline not
Zahn et al., 1984 (56)	CMI vs clorgyline	CMI superior to clorgyline
Volavka et al., 1985 (57)	CMI vs imipramine	CMI superior to imipramine
Cui, 1986 (58)	CMI vs doxepin	78% improve on CMI 36% improve on doxepin
Lei, 1986 (59)	CMI vs imipramine crossover	CMI superior to imipramine
Zhao, 1991 (60)	CMI vs amitriptyline	95% improve on CMI 56% improve on amitriptyline
SSRIs vs Placebo or Non-SRIs		
Perse et al., 1987 (80)	Fluvoxamine vs placebo	Fluvoxamine superior to placebo
Goodman et al., 1989 (81)	Fluvoxamine vs placebo	Fluvoxamine superior to placebo
Goodman et al., 1990 (83)	Fluvoxamine vs desipramine	Fluvoxamine superior to desipramine
Chouinard et al., 1990 (85)	Setraline vs placebo	Sertraline superior to placebo
Jenike et al., 1990 (84)	Setraline vs placebo	Sertraline superior to placebo
Greist et al., 1992 (70)	Setraline vs placebo	Sertraline superior to placebo
Tollefson et al., 1994 (74)	Fluoxetine vs placebo	Fluoxetine superior to placebo
Montgomery et al., 1993 (78)	Fluoxetine vs placebo	Fluoxetine superior to placebo
CMI vs SSRIs		
Den Boer et al., 1987 (61)	CMI vs fluvoxamine	Comparable efficacy
Freeman et al., 1994 (103)	CMI vs fluvoxamine	Comparable efficacy
Pigott et al., 1990 (62)	CMI vs fluoxetine	Comparable efficacy
Koran et al., 1996 (63)	CMI vs fluvoxamine	Comparable efficacy
Lopez-Ibor et al., 1996 (102)	CMI vs fluoxetine	Comparable efficacy
Zohar et al., 1996 (64)	CMI vs paroxetine	Comparable efficacy

The tricyclic SRI clomipramine, and the four selective serotonin-reuptake inhibitors fluoxetine, fluvoxamine, sertraline, and paroxetine, are currently approved by the FDA for the treatment of OCD in adults; three, clomipramine, fluvoxamine, and sertraline, are approved for treatment of OCD in children and adolescents. SRIs have also been found to be effective in treating other obsessive-compulsive spectrum disorders (23 ,66 ,67).

Although 65% to 70% of OCD patients have a clinically meaningful response to their first SRI treatment (68), with sequential trials as many as 90% of OCD patients may ultimately respond (69). Nevertheless, most patients are left with notable residual symptoms, perhaps with a 25% to 60% improvement (69). Although this improvement is clinically significant, patients may remain significantly impaired. These five SRIs have been demonstrated to be effective and well tolerated in OCD in short-term large multicenter controlled trials (52 ,70 ,71 ,72 and 73). There is also evidence regarding long-term treatment response for some of these agents (71 ,74 ,75).

Clomipramine

Clomipramine is a tricyclic antidepressant that, in addition to being an SRI, is a potent norepinephrine (NE) inhibitor and has moderate dopamine (DA) receptor-blocking properties. It was the first medication found to be effective in the treatment of OCD and its efficacy has been firmly established over the past 30 years. The first two multicenter randomized controlled trials of clomipramine in the treatment of adult OCD were conducted in 1991 (52). CMI was administered in doses up to 300 mg per day for 10 weeks and resulted in reductions of 38% and 44% in OCD patients' total YBOCS scores compared to 3% and 5% in the placebo groups. These treatment effects were apparent at 6 weeks, and the YBOCS scores continued to improve over the course of the trial. Patients in the clomipramine versus placebo group were more likely to experience adverse events and discontinue treatment. Two notable adverse events on clomipramine were seizures (0.4%) and elevated aminotransferase levels (6.9%).

Fluoxetine

Fluoxetine, an SSRI, and norfluoxetine, its principal metabolite, have notably long half-lives. Early open trials suggested efficacy of fluoxetine in OCD (76 ,77). A large, double-blind placebo-controlled 8-week study of fluoxetine at three fixed doses (20 mg per day, 40 mg per day, and 60 mg per day) showed that the fluoxetine groups, combined, were superior to placebo in efficacy (72 ,78). Individually, the 40- and 60-mg dosage groups (but not the 20-mg group) were superior in efficacy to the placebo group. Fluoxetine was well tolerated and the dropout rate was low (<6%). In the multicenter study, which led to FDA approval, the same three doses of fluoxetine for 13 weeks were all significantly more effective than placebo in improving OCD (72). Defining response as a 35% improvement in total YBOCS score, response rates were 32.1%, 32.4%, and 35.1% for the 20-, 40-, and 60-mg groups, respectively, and 8.5% for placebo. Both obsessions and compulsions were found to respond to fluoxetine treatment independent of any antidepressant effect. Outcome on fluoxetine is unrelated to plasma levels of fluoxetine, norfluoxetine, or their sum (79).

Fluvoxamine

Several double-blind controlled studies have shown fluvoxamine to be effective in treatment of OCD (80 ,81 and 82). Perse and colleagues (80) reports on a multisite study that found fluvoxamine (100 to 300 mg per day) to be superior to placebo. In this study, 43% of patients receiving fluvoxamine

treatment responded after 6 weeks compared to 12% receiving placebo (defining response as much or very much improved on the improvement item of the Clinical Global Impression Scale). Of special interest, Goodman and associates (83) demonstrated the selective efficacy of SSRIs in OCD because fluvoxamine (up to 300 mg per day) was significantly more effective than the norepinephrine reuptake inhibitor desipramine in the reduction of obsessive-compulsive symptoms in an 8-week double-blind trial.

Sertraline

Although one early study did not find sertraline to be superior to placebo in the treatment of OCD (84), several subsequent studies have demonstrated its efficacy in OCD. Chouinard and associates (85) found sertraline (up to 200 mg per day) to be more effective than placebo on all outcome measures in an 8-week trial. A large 12-week, multicenter, placebo controlled, double-blind trial of sertraline in three fixed doses (50, 100, or 200 mg per day), found sertraline at 50 and 200 mg per day to be significantly more effective than placebo, but the 100-mg dose was not more effective than placebo. Clinical outcome was not correlated with sertraline plasma levels (86).

Paroxetine

A large multicenter placebo controlled study of paroxetine in three fixed doses (20, 40, and 60 mg per day) found that the 40- and 60-mg doses were significantly more effective than placebo, but the lower dose (20 mg) was not more effective than placebo (71). There was a suggestion of a dose–response relationship, and paroxetine was well tolerated in the acute dose study phase.

Citalopram

Citalopram, the most selective of the SSRIs, has not been granted FDA approval in OCD. In a 24-week open pilot study of 29 OCD patients treated with citalopram, 76% had reduction in Y-BOCS scores of more than 50% compared to baseline, with most doses between 40 and 60 mg per day (87). In a 10-week single-blind study of 30 patients with OCD who underwent randomized treatment with fluvoxamine, paroxetine, or citalopram, there were no significant differences found between the three treatments, although the study was underpowered to detect significant differences between active treatments (88). Recent controlled treatment trials suggest efficacy in OCD (89).

Zimeldine

Although early research with zimeldine in OCD was promising, it has been withdrawn from use owing to several reports of Guillain-Barre syndrome occurring during treatment.

Venlafaxine

Venlafaxine, a serotonin/norepinephrine reuptake inhibitor, has not been systematically studied in controlled trials in OCD, but open pilot data suggest potential efficacy and the need for controlled trials (90).

SEROTONIN REUPTAKE INHIBITORS: SIDE EFFECTS AND DRUG INTERACTIONS

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

Side Effects

Efficacy must be balanced against side effects in choosing treatment options, and side-effect profile is magnified in OCD because the treatment is likely to be chronic. All SRIs can cause side effects attributable to their serotonergic action. Clomipramine, a TCA, is most apt to cause anticholinergic and antiadrenergic side effects, whereas SSRIs tend to cause fewer side effects mediated via nonserotonergic receptor systems (43 ,44 and 45 ,91 ,92 ,93 ,94 and 95).

TCAs such as clomipramine have a quinidine-like antiarrhythmic effect that slows intracardiac conduction (95). Although generally an issue only in patients with known cardiac disease, occasional adverse effects may be seen in patients with no documented pre-existing condition. Furthermore, in overdose, the cardiac conduction effects of TCAs lead to much greater toxicity than the SSRIs. This is important, because there is an increased rate of suicide attempts in OCD, especially when associated with comorbid disorders. TCAs have anticholinergic effects that can cause tachycardia, blurred vision, constipation, urinary retention, and confusion (44). Orthostatic hypotension may occur as a result of alpha$_1$-adrenergic antagonism (44). Last, clomipramine lowers seizure threshold (94).

SSRIs are relatively safe compared to TCAs. Few, if any, deaths have been reported following overdose with SSRIs. Although side effects are generally less severe, one may see agitation, anxiety, nausea, headaches, weight gain (over time), and sexual dysfunction (43). Although any of these side effects can contribute to noncompliance, sexual dysfunction is seen in as many as one-third of patients (92) but may not be readily reported unless the clinician specifically inquires about it. There are little systematic data on the treatment of SSRI-induced sexual dysfunction, but case reports and clinical practice have shown that effective interventions may include lowering the dose of SSRI or adding yohimbine (an α_2-adrenergic antagonist), amantadine (a dopamine agonist), methylphenidate, cyproheptadine (an antihistaminic/antiserotonergic agent), buspirone, or sildenafil (69). There also exists a clinically significant discontinuation syndrome that occurs on abrupt discontinuation of an SSRI with short half-life (71).

Drug Interactions

Patients treated for OCD often take concurrent medications; therefore, potential drug interactions should be considered when selecting an antiobsessional agent. In addition to well-established drug interactions known to occur with clomipramine and other TCAs, individuals may also experience idiosyncratic reactions (91 ,93 ,95 ,96 ,97 ,98 and 99). Some medications interact with clomipramine by influencing its plasma concentration, whereas others potentiate clomipramine's side effects via synergy at relevant receptor sites. The hypotensive effects of clomipramine can be exacerbated by α-methyldopa, ß-adrenergic blockers, clonidine, diuretics, and low-potency antipsychotics. Quinidine and other class Ia antiarrhythmics as well as thioridazine, mesoridazine, and pimozide may add to cardiotoxic effects of TCAs. Common medications that have anticholinergic effects can synergize with TCAs to produce anticholinergic toxicity, including antihistamines, antiparkinsonians, low-potency antipsychotics, over-the-counter sleeping pills, and antispasmodics or antidiarrheals. Conversely, TCAs such as clomipramine can potentiate the effects of warfarin or block the effects of guanethidine.

SSRIs can participate in drug interactions as a consequence of effects on the hepatic cytochrome P-450 system (96 ,97 and 98). As each SSRI is metabolized by one or more isoenzymes of cytochrome P-450, they may either inhibit or induce the corresponding enzymatic activity, thereby affecting the metabolism of other drugs. Conversely, other medications can inhibit or induce the P-450 system, thereby modulating the metabolism of SRIs. There is tremendous individual variation in P-450 effects. In addition, because SSRIs are highly protein-bound, this can lead to drug interactions

that do not involve the cytochrome P-450 system per se. For example, SSRIs can compete with warfarin, carbamazepine, and valproate for protein-binding sites, leading to increased levels of these agents, with accompanying adverse effects. In general, these interactions do not represent absolute contraindications to coadministration, but may require necessary adjustments to the dose of SRIs or other medications.

COMPARATIVE STUDIES OF THE SRIs

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

Efficacy

Although metaanalyses have suggested that clomipramine is more effective than the SSRIs, head-to-head comparisons of clomipramine and SSRIs show no particular edge in efficacy to any of these medications. Three metaanalyses consisting of comparisons of previously published data sets found clomipramine superior to SSRIs in OCD treatment efficacy (24 ,100 ,101). In the Greist metaanalysis, the three SSRIs fluoxetine, fluvoxamine, and sertraline, were found to have similar efficacy, whereas clomipramine was found to be significantly more effective. It is important to realize that the studies on which these metaanalyses were based did not include head-to-head comparisons and patient samples for clomipramine may have differed from those for the SSRIs. The clomipramine studies, which were generally conducted in an earlier time period, included SRI naive patients, while the SSRI studies included patients who had previously failed to respond to clomipramine or fluoxetine trials; thus, the SRI trials were likely to include more treatment-refractory patients.

A few studies have evaluated the relative efficacy of the SSRIs in double-blind, head-to-head comparisons. No difference was found in the rate of response to fluoxetine (40 mg per day) compared to clomipramine (150 mg per day) when response was defined as a 35% reduction in total YBOCS scores (102) but fluoxetine was found to be better tolerated. Two comparisons of fluvoxamine and clomipramine have found then to be equally effective and, again, the SSRI was better tolerated (103 ,104). A comparison of paroxetine and clomipramine also found equal efficacy with superior tolerability for the SSRI, which also had a lower discontinuation rate (105). Additional comparative studies would be helpful because these had small sample sizes that were underpowered to show significant differences between two active medications, and did not include comparisons among the SSRIs.

Tolerability

In the studies noted in the preceding, the SSRIs were seen to have somewhat fewer side effects than clomipramine, but this did not necessarily impact compliance; however, in the CMI trials, there were no other alternatives available, which might have impacted discontinuation rates. Clomipramine has α-adrenergic, anticholinergic, and histaminergic effects and has quinidine-like cardiac properties that can increase the QTc interval. These effects can be a particular problem for OCD treatment, because it often requires doses higher than needed to treat depression. In a retrospective study comparing clomipramine and fluoxetine in the treatment of OCD, Jenike and colleagues (106) reported clomipramine treatment was associated with greater adverse events. For clomipramine, patients reported higher rates of sedation, dry mouth, nausea, dizziness, constipation, sweating, headache, and blurred vision. The adverse events reported for fluoxetine treatment were generally mild and transient. A metaanalysis of four large multicenter trials (52 ,70 ,72 ,86) found no difference in dropout rates owing to side effects: 8% for clomipramine, 10% for sertraline, 12% for fluoxetine, and 15% for fluvoxamine (24). Looking at the withdrawal rate for all causes, clomipramine had significantly fewer dropouts than the SSRIs despite having a greater rate of adverse events. However, as noted, this may be owing to the sample differences across studies discussed earlier, and lack of alternative treatment during the CMI trial. Typical tricyclic adverse effects were reported in the clomipramine collaborative study (52): dry mouth (80%), tremor (53%), dizziness (53%), sedation (49%), and male sexual dysfunction (41% of men). Fluoxetine was generally well tolerated in its multicenter study (72). The most commonly reported adverse effects were headache, nausea, insomnia, rhinitis, anorexia, dry mouth somnolence, anxiety, tremor, and diarrhea. There were significantly greater adverse effects with greater fluoxetine doses. In the sertraline study (24), the following adverse effects were reported significantly more frequently by the sertraline than the placebo group: diarrhea, insomnia, decreased libido, nausea, anorexia, ejaculation failure, tremor, increased sweating, and increased weight. Overall, 93% of the sertraline patients reported adverse effects with a correlation noted between higher dose and frequency of side effects; 77% of patients on placebo reported adverse effects. In the fluvoxamine study (70), the following adverse events were significantly more likely to be reported for fluvoxamine that placebo: insomnia, nausea, somnolence, asthenia, delayed ejaculation, nervousness, dry mouth, tremor, and anorexia. A recent review comparing controlled trials suggested that when compared with clomipramine, the SSRIs have equivalent efficacy and superior tolerability and lack anticholinergic side effects particularly (107).

Studies that compared clomipramine and an SSRI in double-blind, head-to-head comparisons, found the SSRIs to be better tolerated: Fluoxetine (40 mg per day) was better tolerated than clomipramine (150 mg per day) (102), as were fluvoxamine (103 ,104) and paroxetine, which also had a lower discontinuation rate (105). Efficacy and tolerability of clomipramine (100 to 250 mg per day) and fluvoxamine (100 to 300 mg per day) were compared in a double-blind, parallel group, randomized study (104). Both groups

showed steady improvement throughout the study; no differences were observed between the groups for any efficacy variable at any time and both clomipramine and fluvoxamine were equally effective in reducing OCD symptoms; they displayed differences in the profiles but not severity of the side effects.

LONG-TERM TREATMENT AND DISCONTINUATION

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

OCD is a chronic disorder; therefore, long-term treatment is often required. Although this research has methodologic limitations, data on long-term SSRI treatment (fluoxetine, sertraline, and paroxetine) suggests that efficacy is maintained and sometimes increases over time. Still more information is needed on the long-term efficacy and safety of SRIs in OCD treatment. SRIs have been proved effective in the acute treatment of OCD; double-blind substitution trials with clomipramine have shown that symptoms frequently recur after discontinuation of treatment (108 ,109).

A retrospective follow-up study of 85 patients with OCD reported that most of the patients treated with SSRIs for 1 to 3 years had maintained or increased symptom improvement (110). Two double-blind and/or placebo-controlled long-term SSRI continuation studies have reported continued efficacy and tolerability (74 ,75). Although valuable, these studies had no effective control group for the maintenance phase: Because only responders to an acute trial continued into long-term maintenance, there were only a handful of placebo patients in maintenance and there was no other control group. To date, no long-term, double-blind substitution trials have been published. An open-label discontinuation trial (111) followed 130 responders to 6 months of acute treatment with an SRI: clomipramine, fluoxetine, or fluvoxamine. This occurred for 2 years of treatment (or until they experienced a recurrence) with the same medication at the same dose, the same medication at half the dose, or no treatment. The study showed a superior therapeutic effect for both medication conditions compared to discontinuation of pharmacotherapy. One controlled, long term, double-blind substitution trial (71) has shown that paroxetine is superior to placebo in preventing OCD relapse during a 6-month discontinuation trial (which followed 12 weeks of acute treatment in a double-blind placebo-controlled paroxetine dose finding trial and 6 months of open-label paroxetine treatment). In addition, treatment effects appeared to be sustained and may have even increased over the time of the trial. In addition, following the double-blind discontinuation phase, subjects switched to placebo had a significantly greater rate of relapse, and time to relapse was significantly shorter on placebo versus paroxetine (71). One study suggested that long-term maintenance therapy might be provided with lower dosages of antiobsessional drugs, with a clear advantage for tolerability and compliance (112).

SRIs IN CHILDHOOD AND ADOLESCENT OCD

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

The clinical presentation of OCD in children and adolescents resembles that observed in adults (113). There has been increasing awareness of the frequency with which children and adolescents suffer from OCD (114). Clomipramine has been well studied in the treatment of OCD in this population. An 8-week multicenter double-blind study (115) found clomipramine to be effective and superior to placebo (mean Y-BOCS scores decreased by 37% versus 8%). These findings led to FDA approval of clomipramine for the treatment of OCD in children and adolescents. The side-effect profile in this sample included anticholinergic, antihistaminic, and α-adrenergic effects such as dry mouth, tremor, sedation, dizziness, sweating, and insomnia. No serious adverse events were reported; however, some patients were withdrawn owing to hepatic enzyme elevations and cardiac palpitations. Electrocardiographic monitoring should be undertaken because tricyclic antidepressants have the potential for being cardiotoxic (116).

Fluvoxamine is an FDA-approved SSRI for the treatment of pediatric OCD. In an 8-week open-label trial of fluvoxamine with adolescents (100 to 300 mg per day), the majority of patients showed significant improvement in their OCD symptomatology (117). The side-effect profile was similar to that reported in adults, including hyperactivity, anxiety, sedation, dizziness, headache, tremor, and nausea. A large (*N* = 120) double-blind placebo-controlled 10-week study of fluvoxamine (50 to 200 mg per day) with children and adolescents (ages 8 to 17) found the drug to be significantly superior to placebo (118). Side effects were described as mild; those that were more commonly reported in the fluvoxamine group included insomnia, agitation, hyperactivity, somnolence, and dyspepsia.

There have been four small prospective studies investigating the efficacy of fluoxetine in pediatric OCD. Three of these were open design (119 ,120 and 121). Riddle and co-workers (122) conducted a double-blind crossover trial of fluoxetine (*N* = 14; mean age 11.8; 20 mg per day, fixed daily dosing) and reported it to be superior to placebo as measured by CGI scores for global improvement. The most commonly reported adverse effects included insomnia, motor activation, fatigue, and nausea. In general, the side effects were reasonably well tolerated and did not result in withdrawal from the study, except for a 13-year-old boy who developed suicidal ideation in the third week of fluoxetine treatment. There were no significant changes in laboratory studies, EKG, weight, pulse, or blood pressure. Major limitations of the study were the small sample size and the fixed dosage.

In a retrospective chart review by Geller and colleagues (123), fluoxetine led to moderate to marked improvement in OCD symptoms in 74% of patients (*N* = 38); mean age 12.3). Relatively high doses of the drug were required, on average 50 mg per day (1.0 mg/kg per day). Long-term

efficacy was sustained over the follow-up period of, on average, 19 months. Fluoxetine was generally well tolerated and adverse effects were less marked than those associated with clomipramine. Despite this apparent tolerability of high doses, it has been recommended that, for children, an appropriate starting dose of fluoxetine is 10 mg per day or less (122).

The FDA has also approved sertraline for the treatment of child and adolescent OCD, which has been shown to be safe and effective in this population (124). In a randomized, double- blind, placebo-controlled trial, 107 children aged 6 to 12 years and 80 adolescents aged 13 to 17 years were randomized to receive either sertraline titrated to a maximum of 200 mg per day (53 children, 39 adolescent) or placebo (54 children, 41 adolescents). After 12 weeks, 42% of the patients receiving sertraline and 26% of those receiving placebo were very much or much improved on CGI ratings. The incidence of side effects was similar to that reported in adults; sertraline appears to be a safe and effective antiobsessional agent in children and adolescents.

In an open-label trial, the adverse effects and potential clinical efficacy of citalopram (10 to 40 mg) were examined in 23 children and adolescents (aged 9 to 18 years). After 10 weeks, over 75% of these youths showed a moderate or marked improvement and adverse effects appeared to be minor and transient, suggesting that citalopram might be well tolerated in children and adolescents (Table 114.2) (128).

Study (Ref.)	Conditions	Results
Flament et al., 1985 (125)	CMI vs placebo	CMI superior to placebo
Devaugh-Geiss et al., 1992 (115)	CMI vs placebo	37% decrease Sx on CMI 8% decrease Sx on placebo
Rapoport et al., 1980 (126)	CMI vs DMI vs placebo	No differences
Leonard et al., 1988 (127)	CMI vs DMI	CMI superior to DMI
Leonard et al., 1991 (109)	CMI substituted with DMI in 50%	18% on CMI relapsed
Riddle et al., 1992 (122)	Fluoxetine vs placebo	Fluoxetine superior to placebo
March et al., 1998 (124)	Sertraline vs placebo	Sertraline marked improved 42% vs 26% placebo
Riddle et al., 2001 (124a)		

TABLE 114.2. CONTROLLED TRIALS OF SEROTONIN REUPTAKE INHIBITOR THERAPY FOR OBSESSIVE-COMPULSIVE DISORDER IN CHILDREN AND ADOLESCENTS

Recent reviews on the treatment of children and adolescents with OCD suggest that adjunctive interventions including parent case management education and specific cognitive-behavioral should be considered in the majority of cases.

BEHAVIOR THERAPY AND PHARMACOTHERAPY

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

Prolonged exposure coupled with response prevention is a key element of behavior therapy for OCD. Exposure involves prolonged confrontation with stimuli that provoke obsessional thoughts or compulsive behaviors. Patients agree not to engage in their usual ritualistic behavior or cognitions and to remain in the situation until their discomfort wanes. Homework is generally prescribed outside the therapist contact. Success rates vary from 50% to 80% (129). Some studies suggest that up to 90% of patients have achieved clinically significant benefits from behavior therapy (130). Generally, 70% to 80% of these patients maintain their acute gains at 1-year follow-up (131).

Intensive behavior therapy over 3 weeks has produced promising results, even in a very symptomatic group. Foa and colleagues (132) reported that 51% of patients experience at least a 70% drop in symptoms, with 35% experiencing a 31% to 69% reduction. These gains are generally (75%) maintained at 12-month follow-up. It is often difficult to arrange such extensive exposure in an outpatient setting, however. Recent results in one trial suggest that providing this treatment in a group setting may be effective and cost effective (133). Relapse prevention programs have proved beneficial in maintaining the gains of acute treatment (134).

Comparisons

A metaanalysis comparing behavioral treatments and SSRIs in OCD involved 77 studies in 4,641 patients. Effect sizes for behavior therapy were not significantly different from those of the SSRIs. There were also no significant differences between the SSRIs, although CMI appeared to have some increased efficacy but not sufficient enough to warrant favoring it over other SSRIs given its side-effect spectrum and the dangers associated with overdose (135). The choice between these two modalities is often dominated by the relatively limited availability of behavioral therapists and by patient preference. Up to 25% of individuals refuse behavioral therapy, whereas others prefer a nonpharmacologic treatment. A computer program offering behavior therapy via 12 computer-controlled interactive phone calls was found effective in over two-thirds of the patients (136).

In clinical practice, a combination of pharmacotherapy and behavioral therapy is often employed, although available research is not yet clear on potential benefits of combined medication and behavior therapy, either on response or relapse prevention.

TREATMENT-RESISTANT OCD: CLINICAL AND BIOLOGICAL PREDICTORS OF NONRESPONSE

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

It has been recommended that a minimum of 12 weeks of SRI treatment is needed to evaluate treatment response. Long-term trials have also noted continued improvement beyond this period.

The relationship between plasma levels of the SSRIs and treatment outcome in OCD is unclear. Studies on fluoxetine and sertraline indicated that plasma levels correlated with dosage but there was no evidence of a concentration—response relationship (75 ,137). It was reported that brain levels of fluoxetine and fluvoxamine reach steady state after 6 months and 1 month, respectively, and correlate with plasma levels (138). Current research is investigating the connection between brain steady-state levels and treatment response.

The response of OCD to drug treatment is frequently partial and incomplete. Poorer outcome has been associated with the presence of compulsions, the chronicity of the illness, and a continuous as opposed to fluctuating course (139) as well as the coexistence of borderline, schizotypal, or avoidant personality disorders (140). Depression is the most common comorbid diagnosis with OCD. The presence of concurrent depression at the start of treatment was not found to predict whether SRIs were effective in reducing obsessive-compulsive symptoms (100). Hollander and associates (141) reported that select measures of serotonergic function may be useful in predicting response to SRI treatment in OCD. They found that nonresponders to SRIs were likely to experience worsening of OCD symptoms and to have a blunted prolactin response on being challenged with the partial 5-HT agonist *m*-chlorophenylpiperazine (*m*-CPP). Further work is needed in this area. Also, OCD patients with increased neurologic soft-signs, a measure of subtle neurologic dysfunction, had a worse response to SRIs (22).

AUGMENTATION STRATEGIES

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

Augmentation strategies play an important role in OCD pharmacotherapy for SRI partial and nonresponders, as 40% to 60% of patients with OCD will not respond to an adequate treatment trial of an SRI. Augmentation agents added to SRIs include both serotonin enhancers and agents involving other neurotransmitter systems, which may contribute to the pathophysiology of OCD.

Serotonergic Enhancers

Pharmacologic agents with serotonergic properties have shown varying utility in augmentation of SRIs.

Buspirone

Buspirone, a 5-HT1A agonist, has been reported effective in treating OC symptoms in augmentation to fluoxetine in some open-label studies (142 ,143) but not in a controlled study (144). A 10-week double-blind trial of buspirone addition to ongoing clomipramine treatment reported no significant further clinical improvement in OC symptoms (145). In a double-blind, placebo-controlled study adding buspirone to fluvoxamine-refractory OCD, buspirone was not found to be significantly better than placebo in reducing OCD symptoms (146).

Fenfluramine

D-L-Fenfluramine, an indirect 5-HT agonist, added open-label to ongoing SRI treatment led to improvement in OC symptoms in six of seven patients (147). D-Fenfluramine also has been cited as potentially effective in augmenting clomipramine in OCD (148); however, these agents have been removed from the market because of side-effect issues.

Tryptophan

Tryptophan, a 5-HT precursor, has shown varying degrees of effectiveness in case reports of SRI augmentation in OCD (68 ,149), but has been removed from the market because of side effects. 5-HTP has been reported helpful in anecdotal case reports.

Lithium

Lithium, which is thought to enhance presynaptic 5-HT release in the brain (150) and influence second messenger systems coupled to 5-HT receptors, was reported to improve OC symptoms in three out of four patients treated with open-label addition to ongoing fluoxetine treatment (151). However, in a 4-week double-blind study of lithium augmentation to 16 OCD partial responders on clomipramine, there was no further decrease in OC symptoms reported after lithium (152). In two double-blind, placebo-controlled trials of lithium, addition to ongoing fluvoxamine treatment in OCD nonresponders (2-week study [20 points] and 4-week study [10 points]), only a small statistically significant reduction in OCD symptoms was reported from the 2-week trial, but not from the 4-week trial (153).

Clonazepam

Clonazepam, a benzodiazepine with unique serotonin properties, has been efficacious in augmentation to SRIs in the patients with OCD in case series (154) and in a double-blind, controlled augmentation trial with fluoxetine or clomipramine (155).

Trazodone

Trazodone, a 5-HT2 and α-adrenergic blocker with weak 5-HT reuptake properties, which has *m*-CPP as a minor metabolite, was recently reported effective in augmentation to various SRIs in five cases of refractory OCD (156).

Pindolol

Pindolol is a B-adrenergic blocker with 5-HT1B and 5-HT1A receptor antagonist activity. Recent reports indicate that adjunctive pindolol may shorten the latency to antidepressant response to SRIs (157); however, data in OCD patients are mixed. One study of pindolol augmentation to paroxetine in resistant OCD resulted in mild improvement (158), whereas another report found that pindolol did not shorten the latency of fluvoxamine antiobsessional response (159). Thus, the utility of pindolol in non-depressed OCD patients is in doubt.

Dopaminergic Agents

The dopamine system has also been implicated in the pathophysiology of OCD; therefore, it is potentially useful in augmentation studies. Dopamine antagonist/SRI combinations have been reported to be effective in OCD.

Haloperidol

In a double-blind, placebo-controlled study of OCD patients on fluvoxamine monotherapy, haloperidol augmentation was found to be significantly more effective than that of placebo, and 100% of patients with a concurrent chronic tic disorder (8/8) responded to ongoing fluvoxamine-haloperidol treatment (160).

Pimozide

Open case series have shown the effectiveness of pimozide/SRI combinations in OCD patients with and without comorbid tic-related disorders (161).

Risperidone

In an open trial of risperidone/SRI combination treatment, 87% (14/16) of patients with refractory OCD had substantial reduction in OC symptoms (162). In another open trial, after the addition of risperidone, seven out of 14 patients had clinical improvement after failing to respond to SSRI treatment alone (163). Other cases of risperidone augmentation reported effectiveness in reducing OC symptoms in patients who had failed SRI trials (164 ,165). A chart review, including eight OCD patients treated with the combination of an SRI and risperidone, reported three (37.5%) patients much or very much improved in OCD symptoms (166).

Olanzapine

Case series (167 ,168) and open-label trials (169) have reported benefits of adding titrated doses of olanzapine (10 or 15 mg per day) to an SSRI in refractory OCD patients. Thus, olanzapine might also be considered in the augmentation of severe forms of OCD, although careful monitoring of the potential interaction with SSRIs is suggested because this combination may produce idiosyncratic effect on plasma levels.

GABA/Second Messenger Systems

Gabapentin, a γ-aminobutyric acid (GABA) analogue, was reported to improve OC symptoms in 5/5 partial responders in a 6-week pilot study of fluoxetine augmentation within 2 weeks of treatment (170). In an open-label augmentation trial of SRIs with inositol, a precursor in the phosphatidyl-inositol cycle, 3/10 refractory OCD patients reported clinically significant responses on the CGI improvement scale (171); therefore, agents affecting other neurotransmitter or second messenger systems may play a role in augmentation strategies.

NOVEL PHARMACOTHERAPIES

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

Some individuals with OCD remain refractory even to augmentation strategies. For them, alternate pharmacotherapy may provide relief.

Intravenous Clomipramine

Intravenous clomipramine was first reported successful in treating obsessive symptoms in 1967 (46). A review of the literature found over 100 cases of successful treatment with intravenous clomipramine (172). In a double-blind, placebo-controlled trial of intravenous versus oral pulse loading of clomipramine in 15 OCD patients, six of seven patients treated with intravenous clomipramine were responders compared to one of eight patients with oral clomipramine (4.5 days after pulse loading), suggesting greater immediate improvement with intravenous pulses (173). This may be owing to avoidance of the first-pass liver metabolism of clomipramine, resulting in a greater clomipramine/desmethyl clomipramine ratio, more potent central 5-HT effects,

and fewer side effects. The superiority of the intravenous route compared to the oral one has been reported in a placebo-controlled study (174). After 1 month of intravenous clomipramine (doses up to 250 mg per day) 58% of 21 patient nonresponders to oral administration randomized to receive intravenous clomipramine showed a marked clinical improvement rated by both CGI and Y-BOCS, without serious adverse events.

Monoamine Oxidase Inhibitors

Monoamine oxidase inhibitors (MAOIs), which block the catabolism of serotonin as well as norepinephrine and dopamine, have been reported successful for refractory OCD in some but not all studies. Case reports of successful treatment are available for iproniazid (175) and most substantially for phenelzine, which was one of the first pharmacologic agents tried for OCD. These reports primarily combined phenelzine with other medications (176 ,177 and 178). However, in a 10-week placebo-controlled comparison trial of fluoxetine versus phenelzine in 54 patients with OCD, patients treated with fluoxetine significantly improved more than those in the phenelzine or placebo groups, except for a subgroup of patients with symmetry obsessions who responded to phenelzine (179).

Serotonergic Agents

Buspirone monotherapy was reported to be ineffective in an 8-week open trial of 14 OCD patients (180) and Hewlett's survey found seven reported successes and 21 reported failures (172). Trazodone has been found successful in both cases and open trials of comorbid depression with OCD (181 ,182 and 183); however, it was found ineffective in reducing OC symptoms in a double-blind, placebo-controlled trial (184). The first treatment response to tryptophan was described in seven patients with OCD (185), with no further controlled studies. The role of other serotonergic agents in treating OCD including 5-HT3 agonists (ondansetron) and 5-HT2 agonists requires further research.

Noradrenergic Agents

The noradrenergic system has also been explored as a novel individual pharmacotherapy in OCD. A case report of clonidine, an α2-agonist, documented improvement in OCD (186). Intravenous clonidine was reported to markedly reduce obsessions in six OCD patients (187). However, in a double-blind controlled crossover trial of clomipramine, clonazepam, and clonidine in 28 OCD patients, clonidine was found ineffective in reducing OCD symptoms (188).

Stimulants and Dopamine Releasers

Although most attention has focused on the blockade of dopaminergic receptors in OCD, there are also reports that agents that release dopamine and dopamine receptor agonists may also have efficacy in OCD. Insel and associates (189) reported that two patients treated with amphetamines (10 to 20 mg) achieved a "persistent benefit" for a period of several weeks. He reported an additional two patients treated with "low-dose" amphetamines for several months who reported a decrease in obsessional symptoms. Ceccherini-Nelli and Guazzelli (190) described three cases of OCD with concurrent depression responding to the dopaminergic agonist bromocriptine (15 to 30 mg). These reports must be reconciled with reports that chronic administration of methylphenidate and amphetamine may induce ritualized behaviors and other OCD-like symptoms (191 ,192).

There are seven cases of reported improvement and one reported failure on chronic dopaminergic agonists. In a survey by Hewlett (172), five of 28 subjects (19%) achieved a good response with chronic amphetamine. Neither of these two patients treated with bromocriptine had any improvement. Bupropion, which inhibits dopaminergic reuptake, was associated with 15 failed treatments.

The OCD-like behaviors induced by chronic amphetamine may represent stereotypies, or complex tics, rather than compulsions that are performed to reduce anxiety. Clomipramine is a potent inhibitor of dopamine uptake and is antiobsessional. Conceivably, the response to low-dose neuroleptics in OCD may stem from their effects on blocking presynaptic dopamine D2 receptors, increasing dopamine release (190).

Benzodiazepines

Case reports of clonazepam have also shown its efficacy (172 ,193). In a double-blind crossover trial, 40% of the patients who had failed clomipramine treatment had clinically significant responses to clonazepam treatment, and clonazepam was significantly more effective than clomipramine during the first 3 weeks of treatment (188). However, a double-blind multicenter placebo-controlled trial of clonazepam demonstrated no efficacy (194). Other benzodiazepines, such as alprazolam, have also not shown efficacy in treating OCD (195).

Anticonvulsants

There are three case reports of significant response to nonbenzodiazepine-related anticonvulsants (two of whom had clinical epilepsy) all with carbamazepine, and 23 treatment failures (172). Two patients with OCD and clinical epilepsy responded to clonazepam treatment (196); however, clinical experience with anticonvulsants may be more positive. Two of 12 patients (17%) at two sites had successful trials of carbamazepine, and six of 26 patients (23%) at three sites had positive outcomes with sodium valproate (172).

Gonadal Steroids

There have been six reported cases of improvement associated with antiandrogen treatment, four of which were menstrually related, and one menstrually related improvement with an anovulatory medication (172). Only one antiandrogen failure has been reported, and one failure of estrogen treatment alone. An open trial with flutamide, an androgen receptor antagonist, in eight OCD patients, demonstrated a lack of response (197). None of the OCD centers has reported using this modality in treating OCD. As such, this treatment has not been well studied. In practice, the feminizing effects of these treatments in males limit their use in this population. In females, it is unclear whether treatment efficacy, if present at all, is limited to specific phases of the menstrual cycle. This treatment is unlikely to become a mainstream therapy for OCD.

Second Messenger System Agents

Agents that affect second messenger systems may also be effective in OCD. In a 6-week double-blind controlled crossover trial of inositol versus placebo in 13 OCD patients, YBOCS scores with inositol treatment were significantly lower than scores when on placebo (198).

Peptides

The role of peptide hormones has been studied in OCD. In one study, elevated CSF levels of oxytocin were found in a subtype of OCD patients compared to Tourette syndrome patients and controls, and levels were correlated with OCD severity, suggesting a possible role for oxytocin in the neurobiology of OCD (199). Oxytocin is a hormone released by the posterior pituitary that regulates uterine and lactiferous duct contraction. It has been implicated in certain ritualized behaviors and the extinction of active avoidance behavior in animals (200). In all, there have been three reported cases of improvement with chronic intranasal oxytocin and 12 failures
(172 ,173 ,174 ,175 ,176 ,177 ,178 ,179 ,180 ,181 ,182 ,183 ,184 ,185 ,186 ,187 ,188 ,189 ,190 ,191 ,192 ,193 ,194 ,195 ,196 ,197 ,198 ,199 ,200 and 201). There have been no reports of improvement with vasopressin; however, this medication has not been administered chronically. None of eight patients at two OCD sites improved with oxytocin treatment. Although experience is limited with this modality, it does not appear to be a promising treatment.

Opiates

Warneke (202) reported that oral morphine in doses of 20 to 40 mg every 5 to 8 days produced marked benefit in five very severe cases of OCD. In contrast, naloxone has proved unsatisfactory and controversial (203 ,204). Tramadol, an analgesic that binds to opioid receptors and inhibits the reuptake of norepinephrine and serotonin, has been reported to produce a significant decrease in Y-BOCS scores in a 6-week open-label study in seven treatment refractory OCD patients (205). Controlled trials are required, although it has low abuse potential, low physical dependency, and mild tolerance.

AUTOIMMUNE TREATMENTS

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

Autoimmune mechanisms may also play a role in at least a subtype of patients with OCD, particularly those who manifest a sudden onset of OCD symptoms following infection by group A B-hemolytic streptococci. Studies with oral penicillin were not found to be effective in such patients (206). Plasmapheresis and intravenous immunoglobulin (IVIG) have been reported effective in case studies (207). Recently, both IVIG and plasma-exchange groups were associated with significant improvement at 1 month in 30 children with infection-triggered exacerbations of OCD or tic disorders who received IVIG, plasma exchange, or placebo (208).

INVASIVE PROCEDURES

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

Finally, novel nonpharmacologic treatments may also play a role in treating some OCD patients. Right prefrontal repetitive transcranial magnetic stimulation has been shown to significantly decrease compulsive urges, compared to nonsignificant increases in urges with stimulation to midoccipital (control) areas and a nonsignificant reduction of urges in left lateral prefrontal areas (209). Further studies are warranted. Recent work with vagal nerve stimulation, deep brain stimulation, and various forms of neurosurgery suggest promise, but require controlled trials.

Patients who have failed all pharmacologic and behavioral treatments (and probably ECT) may be candidates for neurosurgical treatments of OCD (210). Various procedures in intractable cases have been successful, with 25% to 30% of patients experiencing significant benefit without undue side effects. The most common current neurosurgical procedure is cingulotomy, either performed via craniotomy or with the γ-knife procedure. In a United States trial, 18 patients underwent cingulotomy. At follow-up 2 years later, 28% met conservative criteria as responders, with 17% more meeting criteria as partial responders (210). Neurosurgery should be considered for a small percentage of truly refractory patients.

OCD SPECTRUM DISORDERS

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

An OCD spectrum has been proposed, consisting of various disorders that overlap with OCD in several features, including clinical symptoms (repetitive thoughts and behaviors),

course of illness, comorbidity, family history, neurobiology, and treatment response (selective efficacy of SSRIs) (25).

Three key clusters of disorders have been identified: (a) disorders with preoccupation with body image, body weight, or body sensations; (b) impulsive disorders in which repetitive behaviors are driven by pleasure; and (c) neurologically based disorders with repetitive behaviors. Clusters 2 and 3 are dealt with elsewhere in the volume. Here we focus on one disorder from the first cluster where there exist double-blind controlled pharmacologic data.

Body dysmorphic disorder (BDD), the distress of imagined ugliness, is a somatoform disorder in which patients are obsessed with an imagined imperfection or deformity in their appearance, and repeatedly check their appearance in mirrors or engage in cosmetic surgery to change their appearance. There is often poor insight or delusional conviction, and secondary depression and social phobia.

A recent double-blind crossover trial compared the SRI clomipramine to the noradrenergic reuptake inhibitor desipramine (DMI) (8 weeks of each phase) in 40 BDD patients (211). Desipramine, the active control, was chosen to control for nonspecific antidepressant and antidepressants, and because it has a similar side-effect profile to CMI, enhancing the blind. The SRI CMI resulted in significantly greater improvement in all primary outcome measures of BDD severity than did DMI, and also improved measures of functional impairment to a significantly greater extent than did DMI. Subjects with delusional conviction regarding body defect improved on CMI but not DMI, and severity of delusional conviction improved with CMI. Thus, BDD, like OCD, but in contrast to other mood or anxiety disorders, demonstrates a selective efficacy to SRIs, but not to NRI treatment. The delusional conviction in BDD appears secondary to obsessive preoccupation, and also responds to SRI treatment.

CONCLUSION

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

In summary, SSRIs and the tricyclic antidepressant clomipramine are currently the first-line treatment for OCD, with the SSRIs' side-effect profile being more favorable than that of clomipramine. However, 40% to 60% of patients with OCD may not respond to an adequate treatment trial of an SRI. Furthermore, not all patients tolerate SSRIs, and there is often a time delay in seeing a full therapeutic response. Thus, other pharmacologic approaches to treating OCD have been investigated, and certainly combinations of pharmacotherapy and cognitive behavioral therapy are considered the treatment of choice. Augmentation and novel monotherapy strategies have been explored in refractory patients, with serotonergic enhancers, dopamine/serotonin antagonists, enhancers of second messenger systems, and GABAergic agents, with varying efficacy. Recently, immunomodulatory and invasive procedures have been explored as well, but require further study.

ACKNOWLEDGMENTS

Part of "114 - Current and Experimental Therapeutics of Obsessive-Compulsive Disorder "

The authors acknowledge the support of the PBO Foundation. Dr. Hollander has received research support and/or served as a consultant or on a speaker's bureau for the following companies: Solvay, Abbott, SmithKline Beecham, Lilly, Wyeth-Ayerst, and Bristol Myers Squibb.

REFERENCES

1. Hollander E, Stein DJ, Kwon JH, et al. Psychosocial function and economic costs of obsessive-compulsive disorder. CNS spectrums. *Int J Neuropsychiatric Med* 1997;2(10):16–25.

2. *Diagnostic and Statistical Manual of Mental Disorders*, fourth ed. Washington, DC: American Psychiatric Association, 1994.

3. Karno M, Golding JM, Sorenson SB, et al. The epidemiology of obsessive-compulsive disorder in five US communities. *Arch Gen Psychiatry* 1988;45:1094–1099.

4. Rasmussen SA, Eisen JL. The epidemiology and differential diagnosis of obsessive-compulsive disorder. *J Clin Psychiatry* 1994;55(Suppl):5–14.

5. Fenton WS, McGlashan TH. The prognostic significance of obsessive-compulsive symptoms in schizophrenia. *Am J Psychiatry* 1986;143:437–441.

6. Eisen JL, Beer DA, Pato MT, et al. Obsessive-compulsive disorder in patients with schizophrenia or schizoaffective disorder. *Am J Psychiatry* 1997;154:271–273.

7. Berman I, Kalinoswski A, Berman SM, et al. Obsessive and compulsive symptoms in chronic schizophrenia. *Comprehen Psychiatry* 1995;36:6–10.

8. Poyurovsky M, Fuchs C, Weizman A. Obsessive-compulsive disorder in patients with first-episode schizophrenia. *Am J Psychiatry* 1999;156:1998–2000.

9. de Haan L, Linszen DH, Gorsira R. Clozapine and obsession in patients with recent-onset schizophrenia and other psychotic disorders. *J Clin Psychiatry* 1999;60:364–365.

10. Sasson Y, Bermanzhon PC, Zohar J. Treatment of obsessive-compulsive syndromes in schizophrenia. *CNS Spectrum* 1997;2:34–36, 45.

11. Kruger S, Cooke RG, Hasey GM, et al. Comorbidity of obsessive-compulsive disorder in bipolar disorder. *J Affect Disord* 1995;34:117–120.

12. Chen YW, Dilsaver SC. Comorbidity for obsessive-compulsive disorder in bipolar and unipolar disorders. *Psychiatry Res* 1995;29(1 and 2):57–64.

13. Cassano GB, Pini S, Saettoni M, et al. Multiple anxiety disorder comorbidity in patients with mood spectrum disorder with psychotic features. *Am J Psychiatry* 1999;156:474–476.

14. Perugi G, Akiskal HS, Pfanner C, et al. The clinical impact of bipolar and unipolar affective comorbidity in obsessive-compulsive disorder. *J Affect Disord* 1997;46(1):15–23.

15. Steiner W. Fluoxetine-induce mania in a patient with obsessive-compulsive disorder. *Am J Psychiatry* 1992;148(10):1403–1404.

16. Vieta E, Bernardo M. Antidepressant-induced mania in obsessive-compulsive disorder. *Am J Psychiatry* 1992;149(9):1282–1283.

17. Rihmer Z, Barsi J, Belso N, et al. Antidepressant induced hypomania in obsessive-compulsive disorder. *Int Clin Psychopharmacol* 1996;11(3):203–205.

18. Go FS, Malley EE, Birmaher B, et al. Manic behaviors associated with fluoxetine in three 12- to 18-year-olds with obsessive-compulsive disorder. *J Child Adoles Psychopharmacol* 1998;8(1):73–80.

19. Goodman WK, McDougle CJ, Price LH. Pharmacotherapy of obsessive-compulsive disorder. *J Clin Psychiat* 1992;53(Suppl 4):29–37.

20. McDougle CJ, Goodman WK, Leckman JF, et al. Haloperidol addition in fluvoxamine-refractory obsessive-compulsive disorder. *Arch Gen Psychiatry* 1994;51:302–308.

21. Jenike MA, Rauch SL. Managing the patient with treatment-resistant obsessive-compulsive disorder. *J Clin Psychiat* 1994;55(Suppl 3):11–17.

22. Hollander E, DeCaria C, Saoud J, et al. Neurological soft-signs in obsessive-compulsive disorder. (in reply). *Arch Gen Psychiatry* 1991;48:278–279.

23. Hollander E (ed). *Obsessive-compulsive related disorders*. Washington, DC: American Psychiatric Press, 1993.

24. Greist JH, Jefferson JW, Kobak KA, et al. Efficacy and tolerability of serotonin transport inhibitors in obsessive-compulsive disorder: a meta-analysis. *Arch Gen Psychiatry* 1995;52:53–60.

25. Marazziti D, Hollander E, Lensi P, et al. Alteration of serotonin and dopamine function in obsessive-compulsive disorder. *Psychiatry Res* 1992;42(1):41–51.

26. Thoren P, Asberg M, Bertilsson L, et al. Clomipramine treatment of obsessive-compulsive disorder. II. Biochemical aspects. *Arch Gen Psychiatry* 1980;37:1289–1294.

27. Flament MF, Rapoport JL, Murphy DL, et al. Biochemical changes during clomipramine treatment of childhood obsessive-compulsive disorder. *Arch Gen Psychiatry* 1987;44:219–225.

28. Zohar J, Insel T. Obsessive-compulsive disorder: psychobiological approaches to diagnosis, treatment, and pathophysiology. *Biol Psychiatry* 1987;22:667–687.

29. Hollander E, DeCaria CM, Nitescu A, et al. Serotonergic function in obsessive-compulsive disorder: behavioral and neuroendocrine responses to oral *m*-CPP and fenfluramine in patients and healthy volunteers. *Arch Gen Psychiatry* 1992;49:21–28.

30. Goodman W, McDougle CJ, Price LH, et al. m-Chlorophenylpiperazine in patients with obsessive-compulsive disorder: absence of symptom exacerbation. *Biol Psychiatry* 1995;38:138–149.

31. Charney DS, Goodman WK, Price LH, et al. Serotonin function in obsessive-compulsive disorder. A comparison of the effects of tryptophan and m-chlorophenylpiperazine in patients and healthy subjects. *Arch Gen Psychiatry* 1988;45(2):177–185.

32. Zohar J, Insel TR, Zohar-Kadouch RC. Serotonergic responsivity in obsessive-compulsive disorder: effects of chronic clomipramine treatment. *Arch Gen Psychiatry* 1988;45:167–172.

33. Hollander E, Decaria C, Gully R, et al. Effects of chronic fluoxetine treatment on behavioral and neuroendocrine responses to meta-chlorophenylpiperazine in obsessive-compulsive disorder. *Psychiatry Res* 1991;36(1):1–17.

34. Mundo E, Richter MA, Sam F, et al. Is the 5-$HT_{1D\beta}$ receptor gene implicated in the pathogenesis of obsessive-compulsive disorder? *Am J Psychiatry* 2000;157:1160–1161.

35. Kennett GA, Lightowler S, De Biasi V, et al. Effect of chronic administration of selective 5-hydroxytryptamine and noradrenaline uptake inhibitors on a putative index of 5-HT2C/2B receptor function. *Neuropharmacology* 1994;33:1581–1588.

36. Mansari ME, Bouchard C, Blier P. Alteration of serotonin release in the guinea pig orbito-frontal cortex by selective serotonin reuptake inhibitors: relevance to treatment of obsessive-compulsive disorder. *Neuropsychopharmacology* 1995;13:117–127.

37. Erzegovesi S, Ronchi P, Smeraldi E. 5-HT-2 receptor and fluvoxamine effect in obsessive-compulsive disorder. *Hum Psychopharmacol* 1992;7:287–289.

38. Blier P, de Montigny C. Serotonin and drug-induced therapeutic responses in major depression, obsessive-compulsive and panic disorder. *Neuropsychopharmacology* 1999;21(2 Suppl):91S–98S.

39. Pitman RK. Animal models of compulsive behavior. *Biol Psychiatry* 1989;26:189–198.

40. Jenike MA. Theories of etiology. In: Jenike MA, Baer L, Minichiello WE, eds. *Obsessive-compulsive disorders: theory and management*. Littleton, MA: John Wright-PSG, 1986:11–21.

41. Robertson MM. The Gilles de la Tourette syndrome: the current status. *Br J Psychiatry* 1989;154:147–169.

42. Nicolini H, Cruz C, Camarena B, et al. DRD2, DRD3 and 5-HT2A receptor genes polymorphisms in obsessive-compulsive disorder. *Mol Psychiatry* 1996;1:461–465.

43. Grimsley SR, Jann MW. Paroxetine, sertraline, and fluvoxamine: new selective serotonin reuptake inhibitors. *Clin Pharmacol* 1992;11:930–957.

44. Richelson E. The pharmacology of antidepressants at the synapse: focus on newer compounds. *J Clin Psychiatry* 1994;55(Suppl):34–39.

45. Tollefson GD. Selective serotonin reuptake inhibitors. In: Schatzberg AF, Nemeroff CB, eds. *Textbook of psychopharmacology*. Washington, DC: American Psychiatric Press, 1995:161–182.

46. Fernandez-Cordoba E, Lopez-Ibor J. Use of monochlorimipramine in psychiatric patients who are resistant to other therapy. *Actas Luso Esp Neurol Psiquiatr* 1967;26:119–147.

47. Karabanov O. Double-blind controlled study in phobias and obsessions. *J Int Med Res* 1977;5(Suppl 5):42–48.

48. Montgomery SA. Clomipramine in obsessional neurosis: a placebo controlled trial. *Phamaceut Med* 1980;1:189–197.

49. Mavissakalian MR, Jones B, Olson S, et al. Tricyclic antidepressants in obsessive-compulsive disorder: antiobsessional or antidepressant agents? II. *Am J Psychiatry* 1985;142(5):572–576.

50. Jenike Ma, Baer L, Summergard P, et al. Obsessive-compulsive disorder: a double-blind placebo-controlled trial of clomipramine in 27 patients. *Am J Psychiatry* 1989;146(10):1328–1330.

51. Greist JH, Jefferson JW, Rosenfeld R, et al. Clomipramine and obsessive-compulsive disorder: a placebo-controlled double-blind study of 32 patients. *J Clin Psychiatry* 1990;51(7):292–297.

52. The clomipramine collaborative study group. Clomipramine in the treatment of patients with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1991;48(8):730–738.

53. Thoren P, Asberg M, Cronholm B, et al. Clomipramine treatment of obsessive-compulsive disorder. I. A controlled clinical trial. *Arch Gen Psychiatry* 1980;37(11):1281–1285.

54. Ananth J, Pecknold JC, van den Steen N, et al. Double blind comparative study of clomipramine and amitriptyline in obsessive neurosis. *Prog Neuropharmacol* 1981;5(3):257–262.

55. Insel TR, Murphy DI, Cohen RM, et al. Obsessive-compulsive disorder. A double- blind trial of clomipramine and clorgyline. *Arch Gen Psychiatry* 1983;40(6):605–612.

56. Zahn TP, Insel Tr, Murphy DL. Psychophysiological changes during pharmacologic treatment of patients with obsessive-compulsive disorder. *Br J Psychiatry* 1984;145:39–44.

57. Volavka J, Neziroglu F, Yaryura-Tobias JA. Clomipramine and imipramine in obsessive-compulsive disorder. *Psychiatry Res* 1985;14(1):85–93.

58. Cui YH. A double-blind trial of chlorimipramine and doxepin in obsessive-compulsive neurosis. *Chung Hua Shen Ching Ching Shen Ko Tsa Chih (in Chinese)* 1986;19(5):279–281.

59. Lei Bs. A cross-over treatment of obsessive-compulsive neurosis with imipramine and cholimipramine. *Chung Hua Shen Ching Ching Shen Ko Tsa Chih (in Chinese)* 1986;19(2):275–278.

60. Zhao JP. A control study of clomipramine and amitriptyline for treating obsessive-compulsive study *Chung Hua Shen Ching Ching Shen Ko Tsa Chih (in Chinese)* 1991;24(2):67–70, 123.

61. den Boer JA, Westenberg HG, et al. Effect of serotonin uptake inhibitors in anxiety disorder: a double blind comparison of clomipramine and fluvoxamine. *Int Clin Psychopharmacol* 1987;2(1):21–32.

62. Pigott Ta, Pato MT, Bernstein SE, et al. Controlled comparison of clomipramine and fluoxetine in the treatment of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1990;47(19):926–932.

63. Koran LM, McElroy, Davidson JR, et al. Fluvoxamine versus clomipramine for obsessive-compulsive disorder: a double-blind comparison. *J Clin Psychopharmacol* 1996;16(2):121–129.

64. Zohar J, Judge R, the OCD Paroxetine Study Investigators. Paroxetine versus clomipramine in the treatment of obsessive-compulsive disorder. *Br J Psychiatry* 1996;169:468–474.

65. Mundo E, Bianchi L, Bellodi L. Efficacy of fluvoxamine, paroxetine, and citalopram in the treatment of obsessive-compulsive disorder: a single blind study. *J Clin Psychopharmacol* 1997;17:267–271.

66. Hollander E, Allen A, Kwon J, et al. Clomipramine vs. desipramine crossover trial in body dysmorphic disorder: selective efficacy of a serotonin reuptake inhibitor in imagined ugliness. *Arch Gen Psychiatry* 1999;56:1033–1039.

67. Hollander E. *Obsessive-compulsive related disorders*. Washington, DC: American Psychological Association Press, 1993.

68. Rasmussen SA, Eisen JL, Pato MT. Current issues in the pharmacologic management of obsessive-compulsive disorder. *J Clin Psychiatry* 1993;54:4–9.

69. Jenike MA. Drug treatment of obsessive-compulsive disorders. In: Jenike MA, Baer L, Minichiello, eds. Obsessive-compulsive disorders: practical management. New York: Mosby, 1998.

70. Greist JH. Fluvoxamine in OCD: a multicenter parallel design double-blind placebo-controlled trial. Presented at the 18th Collegium Internationale Neuro-Psychopharmacologicum Congress, Nice, France, June, 1992.

71. Hollander E, Allen A, Steiner M, et al. Acute and long-term treatment and prevention of relapse of obsessive-compulsive disorder with paroxetine. *Am J Psychiatry* in press.

72. Tollefson GD, Rampey AH, Potvin JH, et al. A multicenter investigation of fixed-dose fluoxetine in the treatment of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1994;51:559–567.

73. Greist J, Chouinard G, Duboff E, et al. Double-blind parallel comparison of three dosages of sertraline and placebo in outpatients with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1995;52:289–295.

74. Tollefson GD, Birkett M, Koran L, et al. Continuation treatment of OCD: double-blind and open label experience with fluoxetine. *J Clin Psychiatry* 1994;55:69–76.

75. Greist JH, Jefferson JW, Kobak KA, et al. A 1 year double-blind placebo-controlled fixed dose study of sertraline in the treatment of obsessive-compulsive disorder. *Int Clin Psychopharmacol* 1995;10:57–65.

76. Jenike MA, Buttolph L, Baer L, et al. Open trial of fluoxetine in obsessive-compulsive disorder. *Am J Psychiatry* 1989;147:923–928.

77. Liebowitz MR, Hollander E, Schneier F, et al. Fluoxetine treatment of obsessive-compulsive disorder: an open clinical trial. *J Clin Psychopharmacol* 1989;9:423–427.

78. Montgomery SA, McIntyre A, Osterheider M, et al. A double-blind, placebo-controlled study of fluoxetine inpatients with DSM-III-R obsessive-compulsive disorder. *Eur Neuropsychophamacol* 1993;3:143–152.

79. Kornan LM, Cain JW, Dominguez RA, et al. Are fluoxetine plasma levels related to outcome in obsessive-compulsive disorder? *Am J Psychiatry* 1996;153:1450–1454.

80. Perse TL, Greist JH, Jefferson JW, et al. Fluvoxamine treatment of obsessive-compulsive disorder. *Am J Psychiatry* 1987;144:1543–1548.

81. Goodman WK, Price LH, Rasmussen, et al. Efficacy of fluvoxamine in obsessive-compulsive disorder. *Arch Gen Psychiatry* 1989;46:36–44.

82. Jenike MA, Hymna S, Baer L, et al. A controlled trial of fluvoxamine in obsessive-compulsive disorder: implications for a serotonergic theory. *Am J Psychiatry* 1990;147:1209–1215.

83. Goodman WK, Price LH, Delgado PL, et al. Specificity of serotonin reuptake inhibitors in the treatment of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1990;47:577–585.

84. Jenike MA, Baer L, Summergrad P, et al. Sertraline in obsessive-compulsive disorder: a double-blind comparison with placebo. *Am J Psychiatry* 1990;147:923–928.

85. Chouinard G, Goodman W, Greist JH, et al. Results of a double-blind placebo controlled trial of a new serotonin uptake inhibitor, sertraline, in the treatment of obsessive-compulsive disorder. *Psychopharmacol Bull* 1990;26:279–284.

86. Greist J, Chouinard G, Duboff E, et al. Double-blind parallel comparison of three dosages of sertraline and placebo in outpatients with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1995;52:289–295.

87. Koponen H, Lepola U, Leinonen E, et al. Citalopram in the treatment of obsessive-compulsive disorder: an open pilot study. *Acta Psychiatr Scand* 1997;96:343–346.

88. Mundo E, Bianchi L, Bellodi L. Efficacy of fluvoxamine, paroxetine, and citalopram in the treatment of obsessive-compulsive disorder. A single blind study. *J Clin Psychopharmacology* 1997;17:267–271.

89. Stein DJ, Maud CM, Bouwer C. Use of the serotonin selective reuptake inhibitor citalopram in obsessive-compulsive disorder. *J Serotonin Res* 1996;1:29–33.

90. Grossman R, Hollander E. Treatment of obsessive-compulsive disorder with venlafaxine. *Am J Psychiatry* 1996;153:576–577.

91. Glassman AH, Preudhomme XA. Review of the cardiovascular effects of heterocyclic antidepressants. *J Clin Psychiatry* 1993;54(Suppl):16–22.

92. Herman JB, Brotman AW, Pollack MH, et al. Treatment emergent sexual dysfunction with fluoxetine. *J Clin Psychiatry* 1990;51:25–27.

93. Pollack MH, Smoller JW. Management of antidepressantinduced side effects. In: Pollack MH, Otto MW, Rosenbaum JF, eds. *Challenges in clinical practice*. New York: Guilford Press, 1996:451–480.

94. Rosenstein DL, Nelson C, Jacobs SC. Seizures associated with antidepressants: a review. *J Clin Psychiatry* 1993;54:289.

95. Baldessarini RJ. Drugs and the treatment of psychiatric disorders: depression and mania. In: Hardman JG, Limbird LE, Molinoff PB, eds. *Goodman and Gilman's the pharmacologic basis of therapeutics*, ninth ed. New York: McGraw-Hill, 1996:431–459.

96. Shen WW, Lin KM. Cytochrome P-450 monooxygenases and interactions of psychotropic drugs. *Int J Psychol Med* 1991;21:47–56.

97. von Moltke LL, Greenblatt DJ, Harmatz JS, et al. Cytochromes in psychopharmacology. *J Int Clin Psychopharmacol* 1994;14:1–4.

98. Nemeroff CB, DeVane L, Pollack BG. Newer antidepressants and the cytochrome P-450 system. *Am J Psychiatry* 1996;153:311–320.

99. Stoudamire A, Fogel BS. Psychopharmacology in the medically ill. In: Stoudamire A, Fogel BS, eds. *Principles of medical psychiatry*. Orlando, FL: Grune & Stratton, 1987:79–112.

100. Piccinelli M, Pini S, et al. Efficacy of drug treatment in obsessive-compulsive disorder: a meta-analytic review. *Br J Psychiatry* 1995;166:424–443.

101. Stein DJ, Spadaccini E, Hollander E. Metaanalysis of pharmacotherapy trials for obsessive-compulsive disorder. *Int Clin Psychopharmacol* 1995;10:11–18.

102. Lopez-Ibor JJ, Saiz J, Cottraux J, et al. Double-blind comparison of fluoxetine versus clomipramine in the treatment of obsessive-compulsive disorder. *Eur Neuropsychopharmacol* 1996;6:111–118.

103. Freeman CPL, Trimble MR, Deakin JFW, et al. Fluvoxamine versus clomipramine in the treatment of obsessive-compulsive disorder: a multi-center, randomized, double-blind, parallel group comparison. *J Clin Psychiatry* 1994;55:301–305.

104. Koran LM, Cain JW, Dominguex RA, et al. Fluvoxamine versus clomipramine for obsessive-compulsive disorder: a double-blind comparison. *J Clin Psychopharmacol* 1996;16:121–129.

105. Zohar J, Judge R. Paroxetine versus clomipramine in the treatment of obsessive-compulsive disorder. *Br J Psychiatry* 1996;169:468–474.

106. Jenike MA, Baer L, Greist JH. Clomipramine versus fluoxetine in obsessive-compulsive disorder: a retrospective comparison of side effects and efficacy. *J Clin Psychopharmacol* 1990;10:122–124.

107. Pigott TA, Seay SM. A review of the efficacy of selective serotonin reuptake inhibitors in obsessive-compulsive disorder. *J Clin Psychiatry* 1999;60(2):101–106.

108. Pato MT, Zohar-Kadouch R, Zohar J, et al. Return of symptoms after discontinuation of clomipramine in patients with obsessive-compulsive disorder. *Am J Psychiatry* 1988;145:1521–1525.

109. Leonard HL, Swedo SE, Lenane MC, et al. A double-blind desipramine substitution during long-term clomipramine treatment in children and adolescents with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1991;48:922–927.

110. Orloff LM, Battle MA, Baer L, et al. Long-term follow-up of 85 patients with obsessive-compulsive disorder. *Am J Psychiatry* 1994;151:441–442.

111. Ravizza L, Barzega G, Bellino S, et al. Drug treatment of obsessive-compulsive disorder (OCD): long-term trial with clomipramine and selective serotonin reuptake inhibitors (SSRIs). *Psychopharmacol Bull* 1996;32:167–173.

112. Mundo E, Bareggi SR, Pirola R, et al. Long-term pharmacotherapy of obsessive-compulsive disorder: a double blind controlled study. *J Clin Psychopharmacol* 1997;17:4–10.

113. Rapoport JL, Swedo SE, Leonard HL. Childhood obsessive-compulsive disorder. *J Clin Psychiatry* 1992;53:11–16.

114. Piacentini J, Jaffer M, Gitow A, et al. Psychopharmacologic treatment of child and adolescent obsessive-compulsive disorder. *Psychiatr Clin North Am* 1992;15:87–107.

115. DeVeaugh-Geiss J, Moroz G, Biederman J, et al. Clomipramine hydro-chloride in childhood and adolescent obsessive-compulsive disorder—a multicenter trial. *J Am Acad Child Adolesc Psychiatry* 1992;31:45–49.

116. March JS, Leonard HL, Swedo SE. Pharmacotherapy of obsessive-compulsive disorder. *Child Adolesc Psychiatr Clin North Am* 1995;4:217–236.

117. Apter A, Ratzoni G, King RA, et al. Fluvoxamine open-label treatment of adolescent inpatients with obsessive-compulsive disorder or depression. *J Am Acad Child Adolesc Psychiatry* 1994;33:342–348.

118. Riddle MA, Landbloom R, Yaryura-Tobias J, et al. Fluvoxamine for OCD in children and adolescents: a controlled trial. Presented at the Annual Meeting of the American Psychiatric Association, New York, May 1996.

119. Liebowitz MR, Hollander E, Fairbanks J, et al. Fluoxetine for adolescents with obsessive-compulsive disorder (Letter). *Am J Psychiatry* 1990;147:370–371.

120. Riddle MA, Hardin MT, King R, et al. Fluoxetine treatment of children and adolescents with Tourette's and obsessive-compulsive disorders: preliminary clinical experience. *J Am Acad Child Adolesc Psychiatry* 1990;29:45–48.

121. Como PG, Kurlan R. An open-label trial of fluoxetine for obsessive-compulsive disorders in Gilles de la Tourette's syndrome. *Neurology* 1991;41:872–874.

122. Riddle MA, Scahill L, King RA, et al. Double-blind, crossover trial of fluoxetine and placebo in children and adolescents with obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry* 1992;31:1062–1069.

123. Geller DA, Biederman J, Reed ED, et al. Similarities in response to fluoxetine in the treatment of children and adolescents with obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry* 1995;34:36–44.

124. March JS, Biderman J, Wolkow R et al. Sertraline in children and adolescents with obsessive-compulsive disorder: a multicenter randomized controlled trial. *JAMA* 1998 25;280(20):1752–1756 (published erratum appears in *JAMA* 2000;8:283(10):1293). 124a. Riddle MA, Reeve EA, Yaryura-Tobias JA, et al. Fluvoxamine for children and adolescents with obsessive-compulsive disorder: a randomized, controlled, multicenter trial. *J Am Acad Child Adolesc Psychiatry* 2001;40(2):222–229.

125. Flament MF, Rapoport JL, Berg CJ, et al. Clomipramine treatment of childhood obsessive-compulsive disorder. A double-blind controlled study. *Arch Gen Psychiatry* 1985;42(10):977–983.

126. Rapoport J, Elkins R, Mikkelsen E. Clinical controlled trial of chlorimipramine in adolescent with obsessive-compulsive disorder *Psychopharmacol Bull* 1980;16(3):61–63.

127. Leonard H, Swedo S, Rapoport JL, et al. Treatment of childhood obsessive-compulsive disorder with clomipramine and desmethylimipramine: a double-blind cross-over comparison. *Psychopharmacol Bull* 1988;24(1):93–95.

128. Thomsen PH. Child and adolescent obsessive-compulsive disorder treated with citalopram: finding from an open trial of 23 cases. *J Child Adolesc Psychopharmacol* 1997;7:157–166.

129. Hand I. Ambulante Verhaltenstherapie bei Zwangsstorungen. *Fortschr Neurol Psychia* 1995;63(1):12–18.

130. Greist JH. Behavior therapy for obsessive-compulsive disorder. Review. *J Clin Psychiatry* 1994;55(Suppl):60–68.

131. Piacentini J. Cognitive behavioral therapy of childhood OCD. *Child Adolesc Psychiatr Clin North Am* 1999;8(3):599–616.

132. Foa EB, Steketee GS, Ozarow BJ. Behavior therapy with obsessive-compulsives: from theory to treatment. In: Mavissakalian M, Turner SM, Michelson L, ed. *Obsessive-compulsive disorder: psychological and pharmacologic treatment*. New York: Plenum Press, 1985:49–129.

133. Fals-Stewart W, Marks AP, Schafer J. A comparison of behavioral group therapy and individual behavior therapy in treating obsessive-compulsive disorder. *J Nerv Ment Dis* 1993;181:189–193.

134. Hiss H, Foa EB, Kozak MJ. Relapse prevention program for treatment of obsessive-compulsive disorder. *J Consult Clin Psychol* 1994;62:801–808.

135. Kobak KA, Greist JH, Jefferson JW, et al. Behavioral versus pharmacologic treatments of obsessive-compulsive disorder: a meta-analysis. *Psychopharmacology* 1997;136:205–216.

136. Marks IM, Baer L, Greist JH, et al. Home self-assessment of obsessive-compulsive disorder: use of a manual and a computer-conducted telephone interview: two UK-US studies. *Br J Psychiatry* 1998;172:406–412.

137. Koran LM, Cain JW, Dominguez RA, et al. Are fluoxetine plasma levels related to outcome in obsessive-compulsive disorder? *Am J Psychiatry* 1996;153:1450–1454.

138. Strauss WL, Layton ME, Hayes CE, et al. 19F magnetic resonance spectroscopy investigation in vivo of acute and steady-state brain fluvoxamine levels in obsessive-compulsive disorder. *Am J Psychiatry* 1997;154:516–522.

139. Ravizza L, Barzega G, Bellino S, et al. Predictors of drug treatment response in obsessive-compulsive disorder. *J Clin Psychiatry* 1993;56:368–373.

140. Baer L, Jenike MA, Black DW, et al. Effect of Axis II diagnoses on treatment outcome with clomipramine in 55 patients with obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:862–866.

141. Hollander E, Stein DJ, DeCaria M, et al. A pilot study of biological predictors of treatment outcome in obsessive-compulsive disorder. *Biol Psychiatry* 1993;33:747–749.

142. Markovitz PJ, Stagno SJ, Calabrese JR, et al. Buspirone augmentation of fluoxetine in obsessive-compulsive disorder. *Am J Psychiatry* 1990;147:798–800.

143. Jenike MA, Baer L, Buttolph L, et al. Buspirone augmentation of fluoxetine in patients with obsessive-compulsive disorder. *J Clin Psychiatry* 1991;52:13–14.

144. Grady TA, Pigott TA, L'Heureux F, et al. Double-blind study of adjuvant buspirone for fluoxetine-treated patients with obsessive-compulsive disorder. *Am J Psychiatry* 1993;150:819–921.

145. Piggott TA, L'Heureux F, Hill JL, et al. A double-blind study of adjuvant buspirone hydrochloride in clomipramine-treated patients with obsessive-compulsive disorder. *J Clin Psychopharmacol* 1992;12(1):11–18.

146. McDougle CJ, Goodman WK, Leckman JF, et al. Limited therapeutic effect of addition of buspirone in fluvoxamine-refractory obsessive-compulsive disorder. *Am J Psychiatry* 1993;150:647–673.

147. Hollander E, DeCaria CM, Schneier FR, et al. Fenfluramine augmentation of serotonin reuptake blockade antiobsessional treatment. *J Clin Psychiatry* 1990;51:119–123.

148. Judd FK, Chua P, Lynch C, et al. Fenfluramine augmentation of clomipramine treatment of obsessive-compulsive disorder. *Aust NZ J Psychiatry* 1991;25:412–414.

149. Mattes JA. A pilot study of combined trazodone and tryptophan in obsessive-compulsive disorder. *Int Clin Psychopharmacol* 1986;1:170–173.

150. deMontigny C. Enhancement of the 5-HT neurotransmission by antidepressant treatments. *J Physiol* 1981;77:455–461.

151. Ruegg RG, et al. Lithium plus fluoxetine treatment of obsessive-compulsive disorder. New Research Abstr 92, 143rd Annual Meeting of the American Psychiatric Association, New York, 1990.

152. Pigott TA, Pato M, L'Heureux F, et al. A controlled comparison of adjuvant lithium carbonate or thyroid hormone in clomipramine-treated patients with obsessive-compulsive disorder. *J Clin Psychopharmacol* 1991;11(4):242–248.

153. McDougle CJ, Price LH, Goodman WE, et al. A controlled trial of lithium augmentation in fluvoxamine refractory obsessive-compulsive disorder: lack of efficacy. *J Clin Psychiatry* 1991;11:175–184.

154. Leonard HL, Topol D, Bukstein O, et al. Clonazepam as an augmenting agent in the treatment of childhood-onset obsessive-compulsive disorder. *J Am Acad Child Adolesc Psychiatry* 1994;33(6):792–794.

155. Pigott TA, et al. A controlled trial of clonazepam augmentation in OCD patients treated with clomipramine or fluoxetine. New Research Abstr 144. 145th Annual Meeting of the American Psychiatric Association, Washington, DC, May 4, 1992.

156. Marazziti D, Gemignani A, Dell'Osso L. Trazodone augmentation in OCD: a case series report. *CNS Spectrum* 1999;4(12):48–49.

157. Zanardi R, Franchini L, Gasperini M, et al. Faster onset of action of fluvoxamine in combination with pindolol in the treatment of delusional depression: a controlled study. *J Clin Psychopharamcol* 1998;18(6):441–446.

158. Dannon PN, Sasson Y, Hirschmann S, et al. Pindolol augmentation in treatment-resistant obsessive-compulsive disorder: a placebo controlled trial. *Eur Neurol Psychopharmacol* 2000;10(3):165–169.

159. Mundo E, Guglielmo E, Bellodi L. Effect of adjuvant pindolol on the antiobsessional response to fluvoxamine: a double-blind, placebo-controlled study. *Int Clin Psychopharmacol* 1998;13(5):219–224.

160. McDougle CJ, Goodman WK, Leckman JF, et al. Haloperidol addition in fluvoxamine-refractory obsessive-compulsive disorder: a double-blind, placebo-controlled study in patients with an without tics. *Arch Gen Psychiatry* 1994;51:302–308.

161. McDougle CJ, Goodman WK, Price LH, et al. Neuroleptic addition in fluvoxamine-refractory obsessive-compulsive disorder. *Am J Psychiatry* 1990;147:652–654.

162. Saxena S, Wandg D, Bystritsky A, et al. Risperidone augmentation of SSRI treatment for refractory obsessive-compulsive disorder. *J Clin Psychiatry* 1996;57(7):303–306.

163. Ravizza L, Barzega G, Bellino S, et al. Therapeutic effect and safety of adjunctive risperidone in refractory obsessive-compulsive disorder (OCD). *Psychopharmacol Bull* 1996;32(4):677–682.

164. McDougle CJ, Fleischmann RL, Epperson CN, et al. Risperidone addition in fluvoxamine-refractory obsessive-compulsive disorder: three cases. *J Clin Psychiatry* 1995;56(11):526–528.

165. Jacobsen FM. Risperidone in the treatment of affective illness and obsessive-compulsive disorder. *J Clin Psychiatry* 1995;56(9):423–429.

166. Stein DJ, Bouwer C, Hawkridge S, et al. Risperidone augmentation of serotonin reuptake inhibitors in obsessive-compulsive and related disorders. *J Clin Psychiatry* 1997;58:119–122.

167. Weiss EL, Potenza M, et al. Olanzapine addiction in obsessive-compulsive disorder refractory to selective serotonin reuptake inhibitors: an open label case series. *J Clin Psychiatry* 1999;60:524–527.

168. Marazziti D, Pallanti S. Effectiveness of olanzapine treatment for severe obsessive-compulsive disorder. *Am J Psychiatry* 1999;156:1834–1835.

169. Koran LM, Ringold AL, Elliot MA. Olanzapine augmentation for treatment-resistant obsessive-compulsive disorder. *J Clin Psychiatry* 2000;61(7):514–517.

170. Cora-Locatelli G, Greenberg BD, Martin J, et al. Gabapentin augmentation for fluoxetine-treated patients with obsessive-compulsive disorder. *J Clin Psychiatry* 1998;59:480–481.

171. Seedat S, Stein DJ. Inositol augmentation of serotonin reuptake inhibitors in treatment-refractory obsessive-compulsive disorder: an open trial. *Int Clin Psychopharmacol* 1999;14(6):353–356.

172. Hewlett WA. Novel pharmacologic treatments of OCD. In: Hollander E, Stein DJ, eds. *Obsessive-compulsive disorders*. New York: Marcel Dekker, 1997:162–201.

173. Koran LM, Sallee FR, Pallanti S. Rapid benefit of intravenous pulse loading of clomipramine in obsessive-compulsive disorder. *Am J Psychiatry* 1997;154(3):396–401.

174. Fallon BA, Liebowitz MR, Campeas R, et al. Intravenous clomipramine for obsessive-compulsive disorder refractory to oral clomipramine: a placebo-controlled study. *Arch Gen Psychiatry* 1998;55(10):918–924.

175. Joel S. Twenty month study of iproniazid therapy. *Dis Nerv Sys* 1959;20:521–524.

176. Annesley PT. Nardil response in a chronic obsessive compulsive. *Br J Psychiatry* 1969;115:748.

177. Annath J. Treatment of obsessive-compulsive neurosis: pharmacologic approach. *Psychosomatics* 1976;17:180–184.

178. Jain V, Swinson RP, Thomas JC, et al. Phenelzine in obsessional neurosis. *Br J Psychiatry* 1970;117:237–238.

179. Jenike MA, Baer L, Minichiello WE, et al. Placebo-controlled trial of fluoxetine and phenelzine for obsessive-compulsive disorder. *Am J Psychiatry* 1997;144(9):1261–1264.

180. Jenike MA, Baer L. An open trial of buspirone in obsessive-compulsive disorder. *Am J Psychiatry* 1988;145:1285–1286.

181. Prasad AJ. Obsessive-compulsive disorder with depression responsive to trazodone (letter). *Am J Psychiatry* 1984;141:612–613.

182. Baxter L. Two cases of obsessive-compulsive disorder with depression responsive to trazodone. *J Nerv Ment Dis* 1985;173:432–433.

183. Baxter LRJ, et al. Trazodone treatment response in obsessive-compulsive disorder-correlated with shifts in glucose metabolism in the caudate nuclei. *Psychopathology* 1987;1:114–122.

184. Piggott TA, et al. A double-blind, placebo controlled study of trazodone in patients with obsessive-compulsive disorder. *J Clin Psychopharmacol* 1992;12:156–162.

185. Yaryura-Tobias JA, Bhagavan TN. L-tryptophan in obsessive-compulsive disorder. *Am J Psychiatry* 1977;134:1298–1299.

186. Knesevich JW. Successful treatment of obsessive-compulsive disorder with clonidine hydrochloride. *Am J Psychiatry* 1982;139:364–365.

187. Hollander E, Fay M, Cohen B, et al. Serotonergic and noradrenergic sensitivity in obsessive-compulsive disorder: behavioral finding. *Am J Psychiatry* 1988;145(8):1015–1017.

188. Hewlett WA, Vinogradov S, Agras WS, et al. Clomipramine, clonazepam, and clonidine treatment of obsessive-compulsive disorder. *J Clin Psychopharmacol* 1992;12:420–430.

189. Insel TR, et al. D-amphetamine in obsessive-compulsive disorder. *Psychopharmacology (Berlin)* 1983;80:231–235.

190. Ceccherini-Nelli A, Guazzelli M. Treatment of refractory OCD with the dopamine agonist bromocriptine [letter]. *J Clin Psychiatry* 1994;55:415–416.

191. Koizumi HM. Obsessive-compulsive symptoms following stimulants [letter]. *Biol Psychiatry* 1985;20:1332–1333.

192. Frye PE, Arnold LE. Persistent amphetamine-induced compulsive rituals: response to pyridoxine (B6). *Biol Psychiatry* 1981;16:583–587.

193. Bodkin JA, White K. Clonazepam in the treatment of obsessive-compulsive disorder associated with panic disorder in one patient. *J Clin Psychiatry* 1989;40:265–266.

194. Kaplan A, Hollander E, Stahl SM, et al. Clonazepam in obsessive-compulsive disorder: a double-blind, placebo controlled trial. *J Clin Psychiatry* submitted for publication.

195. Stein DJ, Hollander E, Mullen LS, et al. Comparison of clomipramine, alprazolam and placebo in the treatment of obsessive-compulsive disorder. *Hum Psychopharmacol* 1992;7:389–395.

196. Hewlett WA, et al. Clonazepam treatment of obsessions and compulsions. *J Clin Psychiatry* 1990:51:158–161.

197. Altemus M, et al. Open trial of flutamide for treatment of obsessive-compulsive disorder. *J Clin Psychiatry* 1999;60(7):442–445.

198. Fux M, Levine J, Aviv A, et al. Inositol treatment of obsessive-compulsive disorder. *Am J Psychiatry* 1996;153:1219–1221.

199. Leckman JF, Goodman WK, North WG, et al. Elevated cerebrospinal fluid levels of oxytocin in obsessive-compulsive disorder: comparison with Tourette's syndrome and healthy controls. *Arch Gen Psychiatry* 1994;51(10):782–792.

200. Kovacs G, Telegdy G. Role of oxytocin in memory and amnesia. *Pharmacol Ther* 1987;18:375–395.

201. Ansseau M, Legros JJ, Mormont C, et al. Intranasal oxytocin in obsessive-compulsive disorder. *Psychoneuroendocrinology* 1987;12(3):231–236.

202. Warneke L. A possible new treatment approach to obsessive-compulsive disorder. *Can J Psychiatry* 1997;42:667–668.

203. Insel Tr, Pickar D. Naloxone administration in obsessive-compulsive disorder: report of two cases. *Am J Psychiatry* 1983;140:1219–1220.

204. Keuler DJ, Altemus M, Greenberg B, et al. Behavioral effects of naloxone infusion in obsessive-compulsive disorder. *Biol Psychiatry* 1996;40(2):154–156.

205. Goldsmith TB, Shapira Na, Keck PE Jr. Rapid remission of OCD with tramadol hydrochloride. *Am J Psychiatry* 1999;156(4):660–661.

206. Garvey MA, Perlmutter SJ, Allen AJ, et al. A pilot study of penicillin prophylaxis for neuropsychiatric exacerbations triggered by streptococcal infections. *Biol Psychiatry* 1999;45(12):1564–1571.

207. Allen AJ, et al. Case study: a new infection-triggered, autoimmune subtype of pediatric OCD and Tourette's syndrome. *J Am Acad Child Adoles Psychiatry* 1995;34:307–311.

115

The Behavioral Neuroscience of Eating

Gerard P. Smith

Nori Geary

Gerard P. Smith and Nori Geary: Department of Psychiatry, Weill Medical College of Cornell University, New York-Presbyterian Hospital, Westchester Division, White Plains, New York.

Clinical syndromes stimulate basic science by providing unexpected combinations or dissociations of phenomena that basic science did not predict or cannot explain. That clinical eating disorders in which abnormally large meals can occur in patients with low, normal, or high body weight contradicts the assumption that the only function of eating is to provide energy intake for nutritional homeostasis. In the past decade the basic science of eating has responded to this problem in such a fundamental way that it has undergone a paradigm shift. Instead of seeking the neurobiological mechanisms of eating solely in the molecular transformations of energy homeostasis, eating is now seen as a problem in behavioral neuroscience. This shift promises for the first time an adequate basic science of eating because the new view includes genetic and sexual vulnerabilities, learning, and a coherent neural system composed of peripheral feedbacks and central integrations that use amines, peptides, and steroids. The shift has been driven by the recent progress in a "top-down" analysis of meals, the functional unit of eating (1). This analysis has included the application of molecular genetics that revealed a central cascade of neuropeptides, the recognition that the neurology of eating was a system that integrated peripheral feedback and central information to turn a central pattern generator for oromotor movements on and off, the realization that learned controls of eating developed rapidly and acted frequently, and a renewed attack on the mechanisms by which estrogen inhibits eating. We review these areas in this chapter.

MOLECULAR GENETICS AND CENTRAL NEUROPEPTIDE CASCADE

Part of "115 - The Behavioral Neuroscience of Eating "

Although peptides have been implicated in the control of eating since 1957 (2) and more than 20 peptides had been shown to have effects before 1990, the use of molecular genetic techniques to discover agouti protein in 1993 and leptin in 1994 galvanized an intensive search for new brain peptides relevant to the control of food intake and metabolism. The search succeeded (Table 115.1).

TABLE 115.1. CHRONOLOGY OF PEPTIDE EFFECTS ON FOOD INTAKE AFTER CENTRAL OR PERIPHERAL ADMINSTRATION

Peripheral	Central
1957 Glucagon	1974 Opioids
1973 Cholecystokinin	1977 Thyrotropin-releasing hormone
1979 Bombesin	1979 Cholecystokinin, insulin
1980 Insulin	1981 Bombesin
1981 Somatostatin	1982 Neurotensin
1983 Neurotensin	1983 Corticotropin-releasing factor
1984 Calcitonin gene-related peptide	1984 Calcitonin gene-related peptide, somatostatin
1987 Tumor necrosis factor	1985 Neuropeptide Y
1988 Enterostatin	1986 Galanin, alpha melanocyte-stimulating hormone
1989 Interleukin 1-beta	1988 Interleukin-1 beta, tumor necrosis factor alpha
1991 Amylin	1991 Amylin, enterostatin
1992 Apolipoprotein AIV	1992 Tumor necrosis factor beta
1998 Glucagon-like peptide -1	1993 Apolipoprotein AIV
	1994 Agouti protein
	1995 Leptin
	1996 Urocortin, Glucagon-like peptide 1, melanin concentrating hormone
	1997 Agouti related protein
	1998 Orexins, cocaine and amphetamine regulated transcript

The year listed is the first report of the effect according to a search of the literature using Pub Med in 1999 and references in the reviews that accompanied this article. Smith GP. Introduction to the reviews on peptides and the control of food intake and body weight. *Neuropeptides* 1999;33:323–328.
Reproduced from Smith GP. Introduction to the reviews on peptides and the control of food intake and body weight. *Neuropeptides* 1999;33:323–328, with permission of the publisher.

Of the new peptides, leptin was the most intensively investigated because it was hoped that it was the long sought negative-feedback signal synthesized from and released by adipose tissue that was hypothesized by Kennedy in 1953 to be the crucial link between food intake and energy storage (3). The hyperphagia and obesity that occurred in mice that had a genetic deficit in leptin production (*ob/ob*) or in leptin receptors (*db/db*) apparently substantiated the importance of leptin as a negative-feedback signal. However, the euphoria evaporated when it was discovered that circulating leptin was abnormally high, rather than low, in almost all obese humans as well as in mice and rats that became obese on a high-fat diet.

The nature of this leptin "resistance" is under intensive investigation. It appears to involve decreased transport into the brain and decreased intracellular signal transduction after leptin binds to its receptor (4). Its common occurrence shows that the negative-feedback effect of leptin is easily overcome by diets that increase eating. This is a compelling demonstration of a central theme in behavioral neuroscience: Reinforcement frequently overcomes regulation.

Although leptin was not a "magic bullet" for the treatment of obesity, a crude idea driven by commercial hopes rather than scientific knowledge, the analysis of its inhibition of food intake and increased metabolism has stimulated an enormous amount of work that can be summarized briefly. (See refs. 4 and 5 for extensive reviews.) The arcuate nucleus in the ventromedial hypothalamus is a nodal point for leptin's action. Leptin stimulates a lateral population of proopiomelanocortin (POMC) neurons and inhibits a medial population of neurons that express neuropeptide Y (NPY), a potent stimulant of eating, and agouti-related peptide (AGRP), an antagonist of melacortin (MC) receptors, especially MC_4, that also stimulates eating. The MC_4 receptor

is necessary for leptin's inhibition of food intake because the effect of leptin is abolished when an antagonist blocks MC_4. Furthermore, MC_4 knockout mice are hyperphagic and obese despite high circulating leptin. The endogenous agonist for the MC_4 receptor is α-melanocyte-stimulating hormone (MSH), a translational product of the POMC neurons. The roles of other peptides, such as orexin A, cocaine and amphetamine-regulated transcript (CART), glucagon-like peptide-1 (GLP-1), apolipoprotein-IV, melanin concentrating hormone (MCH), and mahogany protein in the physiologic control of eating, and their relationship (if any) to leptin's action remain to be determined.

Identification of these peptides sparked anatomic investigations that used immunocytochemical techniques to trace connections from NPY-AGRP neurons to orexin neurons in the lateral hypothalamus, and projections of NYP-AGRP neurons and POMC neurons to the paraventricular nucleus in the anteromedial hypothalamus and to the hindbrain, especially to the region of the nucleus tractus solitarius (NTS). Discussion of these results has attempted to extract their meaning from the medial and lateral hypothalamic syndromes. This is not illuminating because those syndromes never clarified the normal control of eating and the extent of the anatomic damage was not determined; they were problems, not explanations.

The current status of this work on central neuropeptides can be described as "a few small islands of scientific understanding surrounded by a vast area of uncertain phenomena" (2). The progress represented by the work with these peptides has been real and has been trumpeted loudly in the scientific literature and lay press. Its limitations have been less emphasized. They include the facts that leptin resistance is a significant problem in dietary-induced obesity and that high-fat diets decrease the potency of a variety of peptides. Furthermore, the central neuropeptide cascade defined by leptin action has been described in the unusual situation of 24 or 48 hours of food deprivation. This makes the relevance of this cascade to the controls of food intake and body weight under more normal conditions problematic. It is particularly important to understand that most of the progress in the field has been horizontal, that is, it has added new peptides to the list that affect energy intake, storage, and expenditure. Relatively little progress has been made in the vertical problems of physiologic function, interactions, and generalizations. Despite their difficulty, the vertical problems must be pursued in order to decipher the meaning of these molecules. To say that a peptide increases or decreases food intake is to pose a problem for scientific investigation rather than to state a conclusion about physiologic function. The physiologic function of a peptide is learned only when we can specify the function of that peptide in the central neural networks that control food intake and body weight.

The example of cholecystokinin makes this point. Cholecystokinin (CCK) released from the small intestine during a meal provides a negative-feedback signal for the control of the size of that meal in animals (6) and humans (7). An important part of the evidence for this was that administration of a specific antagonist of CCK_A receptors produced a significant increase in meal size in rodents, pigs, monkeys, and humans under a variety of conditions. This satiating action of CCK is mediated by CCK_A receptors on vagal afferent fibers that project to the medial and caudal NTS. The disconnected caudal brainstem of the chronic decerebrate rat has sufficient neural complexity to process this visceral information into a stop signal to the central pattern generator (cpg) that controls oromotor movements (8), but in the intact brain, forebrain structures, such as the paraventricular nucleus, are also involved (6). Thus in this case, we know the biological meaning of CCK in the control of food intake because we can define it as one of the stimuli of the peripheral negative feedbacks that control meal size.

The experimental history of the extensive and convergent results required to prove that the satiating effect of CCK was a physiologic function of the peptide is a case study for

those pursuing the meaning of other peptides (9). Behavioral specificity, receptor mechanism, predictable results with reversible antagonists, afferent neural mediation, and the effect of experimental context (genetic, dietary, metabolic, and prior experience) had to be assessed and interpreted. The experience teaches that physiologic meaning is not read out directly from molecular structure or from an increase or decrease of food intake.

NEURAL CONTROL OF EATING

Part of "115 - The Behavioral Neuroscience of Eating "

Because the biological meaning of a peptide for the control of eating is defined by its role in the neural network that integrates peripheral and central stimuli into oromotor output, we now review the important progress that has been made in that area in the past decade.

Although it is common to refer to this area of research as the Neural Control of Food Intake, this is imprecise and misleading because food intake denotes a measurement made by investigators, not a movement made by animals or people. The somatic nervous system controls the oromotor movements of eating; the autonomic nervous system controls movements of the digestive tract through its effects on the enteric nervous system, intraluminal digestion, neuroendocrine release, and metabolic transformations. The sensory stimuli from these efferent effects are integrated in the central nervous system to affect somatic and visceral efferent output. The major advance in the understanding of this vast, complicated neural system has been in the somatic nervous system's control of eating.

Eating consists of rhythmic oromotor movements, such as licking, lapping, and mastication. These movements have a relatively fixed rate. For example, rats make five to eight licks per second (the range in individual rats is less). This is the motor signature of a group of neurons acting as a cpg. The cpg for licking in the rat is in the medial, intermediate, and lateral reticular formation of the medulla (10). A network of premotor neurons extends forward from the caudal brainstem to the region of the substantia nigra (11). Thus, the neural control of eating can be reduced to what turns the cpg on and off (12).

Eating can be initiated by a variety of external stimuli, such as visual, social, olfactory, and auditory. Internal stimuli, such as a slight decrease in plasma glucose (13), a rise in liver temperature (14), and a decrease in basal metabolism (15) are also effective. The efficacy of most, if not all, of these stimuli can be modified by experience. It is important to note that the adequate stimuli for the initiation of eating do not determine the duration or size of the subsequent meal. These aspects of a meal are determined by the mechanisms that maintain eating.

The fact that the initiation of eating does not determine how long eating will continue or how much will be ingested means that eating a meal is not produced by a ballistic control system. Eating, once initiated, is under feedback control. Positive feedback is stimulated by orosensory stimuli and negative feedback is stimulated by gastric and small intestinal stimuli. The positive feedback turns the cpg on and the negative feedback turns it off.

The brain processes these feedbacks within the neural context of other stimuli that are relevant to the control of eating, but are not produced by ingested food stimuli acting on the mucosa of the gastrointestinal tract. A network that compares the relative potency of positive and negative feedbacks analyzes the result of this distributed processing of the peripheral feedback information. Eating is maintained as long as positive feedback exceeds negative feedback; eating stops and the meal ends when negative feedback exceeds positive feedback for a considerable time (Fig. 115.1).

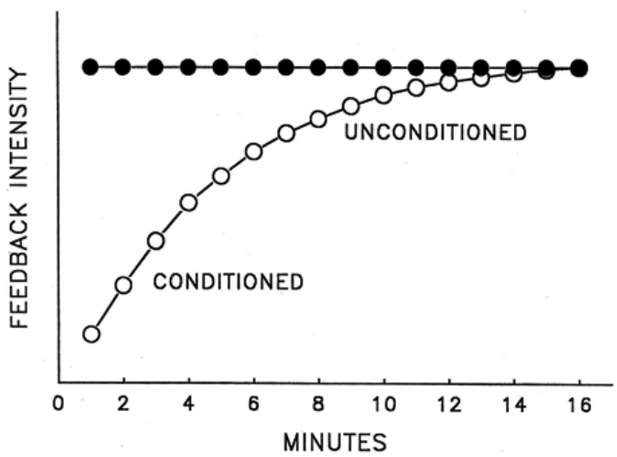


FIGURE 115.1. This depicts the temporal interaction of positive and negative afferent feedbacks produced by ingested carbohydrate solutions during a representative meal. Note that the meal ends when the potencies of the positive and negative feedbacks are judged to be equal by a comparator function(s) of the central networks for the control of ingestion. •, positive feedback; ô, negative feedback. Reproduced from Smith GP. Feeding: control of eating. In: Adelman G, Smith BH, eds. *Elsevier's encyclopedia of neuroscience*. New York: Elsevier, 1999:711–714, with permission of the publisher.

The oromotor output of this continuous integration of positive and negative feedbacks is a sequence of clusters of licks separated by short intervals of nonlicking (16). The number of licks in a cluster is a measure of orosensory positive feedback (palatability). The number of clusters is a measure of the relative potency of the positive and negative feedbacks (1 ,17). The meal ends when the animal no longer reinitiates licking for a relatively long time (15 to 120 minutes in the rat).

There are two important points about these feedback effects: First, the site of action of the adequate stimuli is preabsorptive. In addition to its classic motor and secretory functions, the gastrointestinal tract is a sensory sheet from the tip of the tongue to the end of small intestine. It is peppered with mechanical and chemical receptors; their dispersion

over large areas provides for the effect of stimulus load (i.e., concentration and volume of stimuli).

The second point is that all of the afferent fibers from the mouth, stomach, and small intestine project to the caudal brainstem.

The direct preabsorptive stimulation by the stimuli of ingested food and its digestive products that provide feedback control during a meal is a criterion for distinguishing these feedback controls from all other controls. These feedback controls are direct controls of meal size (Fig. 115.2) and all other controls are indirect controls (Table 115.2).

Categories	Examples
Rhythmic	Diurnal, estrogen
Metabolic	Changes in leptin, insulin, and fatty acids
Thermal	Environmental and fever
Conditioned	Preferences, aversions, and satiations
Cognitive	Social and, in humans, cultural and esthetic
Ecological	Relative densities of predators and foods

[a]The list of categories is neither mutually exclusive nor exhaustive; this is particularly true for conditioned, cognitive, and ecological.

TABLE 115.2. INDIRECT CONTROLS OF MEAL SIZE[a]

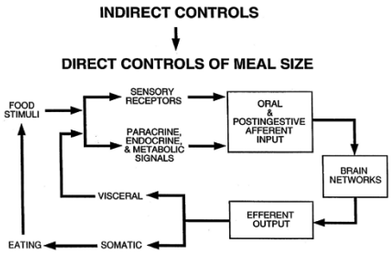


FIGURE 115.2. Flow diagram of the direct controls of meal size stimulated by ingested food acting on preabsorptive receptors of the gastrointestinal tract. Note that food stimuli activate afferent neurons providing positive and negative feedbacks directly and indirectly through effects on paracrine, endocrine, and metabolic signals. The efferent output of the central networks for the control of eating is carried over somatic efferent fibers. Because some of the direct controls are stimulated by ingested food in every meal, indirect controls of meal size exert their effects by modulating direct controls. (See the unidirectional arrow between indirect and direct controls.) Reproduced from Smith GP. Feeding: control of eating. In: Adelman G, Smith BH, eds. *Elsevier's encyclopedia of neuroscience*. New York: Elsevier, 1999:711–714, with permission of the publisher.

This is not an arbitrary classification because it is based on a biological criterion of site of action. The classification also carries neurologic meaning. That meaning comes from experiments in the chronic decerebrate rat. Because the caudal brainstem contains the cpg and all of the projections of the afferent nerves mediating peripheral feedback effects, the decerebrate rat responds to direct controls (18 ,19). In contrast, none of the indirect controls that have been tested affect eating in the chronic decerebrate rat. Because indirect controls require the forebrain to be connected to the caudal brainstem in order to control eating, the reciprocal connections between forebrain and hindbrain are necessary for the modulation of the direct controls by the indirect controls. This theory asserts that indirect controls have no direct action on the cpg during a meal in the absence of direct controls activated by ingested food. Specifying the peptidergic and aminergic connections that mediate an indirect control's effect on the direct controls is the next step and it is the place where the recent advances in central peptides and the neural control of eating converge.

The identification of the importance of the positive and negative feedbacks from the periphery in the direct controls of eating that are modulated by the indirect controls facilitates the investigation of human eating disorders in three ways.

1. The peripheral, preabsorptive sites of action are accessible to controlled stimulation in the conscious human before, during, and after test meals.
2. An increase or a decrease in meal size can be explained by changes in feedback potency (Table 115.3).

Change of Meal Size	Afferent Feedback	
	Positive	Negative
Increase	Increase	Decrease
Increase	Increase	No change
Increase	Increase	Smaller increase
Increase	No change	Decrease
Decrease	Decrease	Increase
Decrease	Decrease	No change
Decrease	Decrease	Smaller decrease
Decrease	No change	Increase

[a]Some changes of afferent feedbacks responsible for increased or decreased meal size. Identification of the mechanisms of a specific change(s) in potency of feedbacks is an experimental problem. (See Table 115.4 for candidates.)
Reproduced from Smith GP. The controls of eating: a shift from nutritional homeostasis to behavioral neuroscience. *Nutrition* 2000;17:10–20, with permission of the publisher.

TABLE 115.3. CHANGES IN POTENCY OF AFFERENT FEEDBACKS THAT DETERMINE CHANGES IN MEAL SIZE[a]

3. Identifying which combination of changes in feedback underlies the change in meal size focuses the search for neurobiological mechanism because the feedbacks have

different mechanisms in the rat and are likely to be similarly differentiated in the human (Table 115.4).

Direct controls	Peripheral	Central
Orosensory	Gustatory and olfactory transducers	Dopamine[a] Opioids[a]
Gastric	CCK[a] at CCK_A vagal mechanoreceptors, other mechanoreceptors, and bombesin-like peptides	Amino acids from gastric vagal afferent terminals in NTS Serotonin[a]
Small intestinal	CCK[a] at CCK_A receptors on vagal mechanoreceptors and chemoreceptors; glucagon,[a] amylin, enterostatin, apolipoprotein IV, and insulin released by nutrient or digestive stimuli through contact with mucosal receptors or by the release of incretins	Amino acids from duodenal and hepatic vagal afferent terminals in NTS Serotonin[a]

NTS, nucleus tractus solitarius.
[a]Indicates that a molecule has been demonstrated to be a physiologic mechanism. The physiologic status of the other molecules is uncertain.
Reproduced from Smith GP. The controls of eating: brain meanings of food stimuli. In: Mayer EA, Saper CB, eds. *The biological basis for mind body interactions.* New York: Elsevier, 2000:173–186, with permission of the publisher.

TABLE 115.4. MOLECULAR MECHANISMS OF DIRECT CONTROLS OF MEAL SIZE

An example of the use of this theory of the control of meal size is the recent work concerning the pathophysiology of the abnormally large meals that characterize patients with bulimia nervosa. Since the 1970s there has been evidence that these patients do not feel as full as normal after the same size meal. This has been confirmed more precisely in recent work that showed that bulimics require more food to report equal fullness (20). This suggests a defect in the satiating process (21), specifically a defect in the potency of negative feedback. This hypothesis was strengthened when orosensory positive feedback measured psychophysically did not reveal large increased responses to carbohydrate or fat stimuli (22). Thus, using Table 115.3 , the hypothesized combination was decreased negative feedback and no change in positive feedback. The decreased negative feedback could involve peripheral mechanisms or central modulation. Two peripheral abnormalities have been found: an enlarged stomach capacity (23) and a decreased release of CCK (24). The decreased release of CCK was ameliorated when binge eating stopped in one experiment (25), but further experiments are required to evaluate this phenomenon.

There may also be a defect in the central processing of the decreased peripheral negative feedback information owing to abnormal function of the central serotonin system. If central serotonin function is decreased in bulimia patients, they should be more vulnerable than controls to a further decrease in serotonin function produced by serotonin depletion. This prediction has been confirmed: Acute tryptophan depletion that probably decreased central serotonin activity increased meal size in patients with bulimia (26).

The combination of decreased central serotonergic processing with decreased peripheral negative feedback could be particularly disruptive of satiation because the satiating potency of CCK in rats is synergistic with gastric distension and is reduced by decreased central serotonergic function, particularly at $5\text{-}HT_{2C}$ receptors (6).

In addition to decreased negative feedback, bulimia patients also have an abnormal cognitive indirect control. They eat much larger meals when they are instructed to binge compared to when they are instructed not to binge (27).

LEARNING AND EATING

Part of "115 - The Behavioral Neuroscience of Eating "

Numerous regions of the brain can be implicated in eating by a variety of techniques in animals and humans. This reflects the fundamental biological importance of eating to individual life and reproduction, and the functional requirements of a foraging omnivore. From this viewpoint, it is not surprising that learning and memory are important processes in the control of eating. Three important types of learning have been identified using Pavlovian procedures and theory: conditioned preference, conditioned aversion and avoidance, and conditioned satiation.

Conditioned preferences are formed by flavor–flavor associations or flavor–postingestive associations (28). Once formed, the preferences increase the size of meals. When the postingestive unconditioned stimulus is omitted, the conditioned preference persists for months, but its effect on

intake extinguishes rapidly. The acquisition of a conditioned flavor–postingestive preference requires dopamine acting at D_1 receptors, perhaps in the nucleus accumbens (29). Opioid mechanisms are apparently not necessary.

Conditioned aversions and avoidance are formed by associations between orosensory stimuli (especially gustatory) and aversive postingestive stimuli. Nausea is commonly reported in humans who have conditioned aversions.

The anorexia that accompanies amino acid imbalance may be an example of a conditioned avoidance. It involves a serotonergic mechanism because a 5-HT_3 antagonist abolishes it (30). The same 5-HT_3 mechanism is observed in the conditioned aversions and avoidances observed in cancer patients undergoing chemotherapy (31). The recent report of successful treatment of binge eating with a 5-HT_3 antagonist suggests that conditioned avoidance may also be involved in that eating disorder (32).

Certainly, conditioned avoidance of food characterizes patients with anorexia nervosa. This aversive stimulus appears to be cognitive and part of the morbid fear of fatness. The strong potency of this psychopathological inhibitory control of eating can be appreciated when it is remembered that the low circulating leptin of the emaciated patient (4) disinhibits the central cascade of peptides so that the neurological drive to eat is intense (see section on molecular genetics and central neuropeptide cascade).

The *c-fos* technique has been used to detect changes in the neural network that underlies the acquisition and expression of a conditioned taste aversion (CTA) (33 ,34). The most significant changes occur in the NTS in the hindbrain and central nucleus of the amygdala in the forebrain. The increased C-Fos in the NTS correlates with the acquisition, extinction, and forgetting of the CTA (35). The meaning of this correlation is under active investigation.

The changes in the NTS depend on connections with the forebrain because they are abolished ipsilateral to surgical hemidecerebration at the level of the superior colliculus (36). This is a nice example of how an indirect control, learning, requires connections between the forebrain and hindbrain in order to affect eating. (See ref. 37 for other examples.)

There is evidence that D_1 receptor mechanisms are necessary for the acquisition of a CTA as well as a conditioned preference. Injection of a D_1 antagonist into the lateral hypothalamus blocked the acquisition of a CTA (38).

The third type of learning is conditioned satiation. Like conditioned preference of the flavor–postingestive type, it depends on the association between orosensory stimuli and a postingestive stimulus. Unlike conditioned preferences and aversions, conditioned satiation is hedonically neutral. Its function is to decrease the rate of eating concentrated liquids during the early part of a meal (39). It can be acquired or extinguished within one or two meals. The postingestive stimulus acts in the stomach and beyond the pylorus (40). Increases of plasma glucose are not a sufficient UCS to form a conditioned satiation (41). Nothing is known about the mechanisms that mediate this type of learning.

It is interesting that the way to extinguish conditioned satiation in the rat is to prevent the accumulation of ingested food in the stomach and small intestine by draining the gastric contents out through a chronic gastric fistula. This form of sham feeding leads to a significant increase in meal size owing to the removal of unconditioned negative feedback from the stomach and small intestine. After three to five consecutive sham-meals, conditioned satiation is extinguished and meal size is maximal. If real-feeding meals are given between sham-feeding meals, however, the size of a sham-fed meal is larger than normal, but not maximal, because some conditioned satiation is present (42). These phenomena in the sham-feeding rat (Fig. 115.3) may be relevant

to the abnormally large meals that occur with repetitive bingeing followed by vomiting or purging.

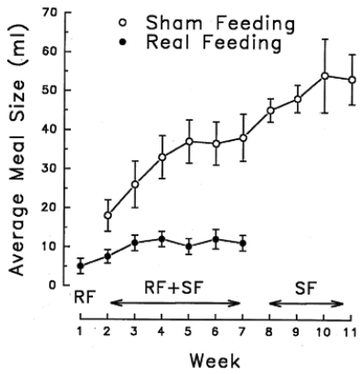


FIGURE 115.3. The potency of learned controls of meal size based on postingestive food stimuli is revealed by the progressive increase in test meal size during repeated sham-feeding trials. During sham feeding, liquid food drains from open gastric cannulas without significant accumulation in the stomach or small intestine, so that learned controls based on gastrointestinal food stimuli extinguish. In this experiment 13 rats were offered a sweet liquid diet once daily, after 3 hours of deprivation of their maintenance diet. During week 1, rats fed normally (real feeding, RF); during weeks 2 to 7, sham-feeding (SF) tests alternated with real-feeding tests, and during weeks 8 to 11, rats were only sham fed. The figure shows the average real and sham meal sizes in each week. During the first sham-feeding test, rats still ate well-defined meals terminated by behavioral signs of normal satiety, indicating that after this short period of food deprivation, pregastric food stimuli can elicit satiety. However, meal size nearly doubled during this test, because of the absence of direct, unconditioned gastric and postgastric controls of eating. Sham meal size doubled during weeks 3 to 7, when sham- and real-feeding tests were alternated, whereas real meal size increased only a small amount, and sham meal size almost doubled a third time during weeks 8 to 11, when there were no real feeding tests. These further increases during the last 4 weeks reflect the extinction of conditioned satiety. Reproduced from Geary N, Smith GP. Appetite. In: Sadock BJ, Sadock VA, eds. *Kaplan & Sadock's comprehensive textbook of psychiatry*. Philadelphia: Lippincott Williams & Wilkins, 2000:209–218, with permission of the publisher.

ESTROGEN AND EATING

Part of "115 - The Behavioral Neuroscience of Eating "

Given the high incidence of eating disorders in women and the frequent onset of them when the ovarian rhythm begins, the recent renewed interest in the control of eating by estrogen in rats is most welcome.

Estrogen has two inhibitory effects on eating. The first occurs during the periovulatory phase of the estrus cycle in rats. The decrease in food intake is owing to a decrease in meal size (43). The combination of feedback potencies (Table 115.3) that accounts for the decrease in meal size is no change in positive feedback and increased negative feedback (44). The increased negative feedback is the result of estrogen increasing the potency of endogenous CCK released from the small intestine (Fig. 115.4) (45 ,46). Presumably this synergism is a central action of estrogen changing the processing of the vagal afferent stimulation of the NTS in response to CCK acting on CCK_A receptors of vagal afferent terminals in the upper small intestine, but it is not known where this synergism occurs or where the receptors are that mediate it (47 ,48 and 49).

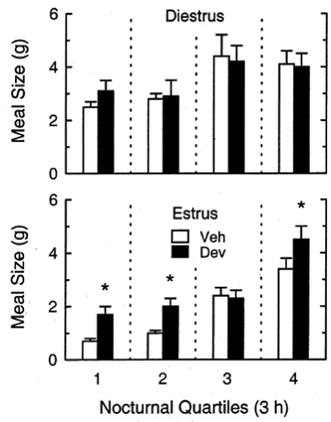


FIGURE 115.4. Antagonism of CCK_A receptors with devazepide (Dev) increased rats' spontaneous meal size throughout the nocturnal period of estrus, but not during diestrus. Thus, an increase of endogenous CCK's satiating potency contributes to the cyclic decrease in meal size during estrus that is mediated by estradiol. Furthermore, endogenous CCK does not appear to contribute to the tonic inhibitory effect of estradiol on meal size because devazepide did not increase meal size during diestrus. Data are mean meal sizes per 3-hour quartile of the nocturnal phase. Note that meal size increased across the dark phase in both estrus and diestrus and that meal size after vehicle (Veh) treatment during the fourth quartile of estrus, when devazepide increased meal size, was as large as control meal size during the first two quartiles of diestrus, so that devazepide's selective effect during estrus was not an artifact of the cyclic differences in meal size. *Nocturnal meal size after devazepide larger than after vehicle, $P < .05$. Reproduced from Eckel LA, Geary N. Endogenous cholecystokinin's satiating action increases during estrus in female Long-Evans rats. *Peptides* 1999;20(4):451–456, with permission of the publisher.

The inhibitory effect of estrogen on food intake during the periovulatory phase has been reported in women (50 ,51). The role of CCK in this effect has not been investigated.

The facts that the ovarian rhythm is disrupted in anorexia nervosa and circulating estrogen is low adds a further disinhibition to the central network that controls eating in these patients.

The second inhibitory effect of estradiol on eating is a tonic inhibition of meal size that acts throughout the ovarian cycle in rats. Release from this inhibition by ovariectomy causes a sustained increase in meal size and obesity. This effect of estrogen, however, does not appear to be mediated by a change in the satiating potency of CCK.

Both effects of estrogen depend on binding to the estrogen receptor α because mice with this receptor knocked out do not show either effect (52).

There are sex differences in the incidence or clinical course of many diseases associated with anorexia as well as in the anorectic response to many immune-system mediators, such as IL-1 and α-TNF. Some of these sex differences appear to be related to estrogenic function. Crohn disease, an inflammatory bowel disease in which anorexia is an early sign (53), is one such. The incidence of Crohn disease is higher in women than men (54) and use of estrogen-containing contraceptives increases women's risk further (55).

Anorexia caused by Gram-negative bacterial infection is also estradiol-sensitive. The effect of estradiol to increase the anorexia produced in rats by intraperitoneal administration of bacterial lipopolysaccharide is expressed by a decrease in meal frequency without a change in meal size (56), indicating that this effect of estrogen is separate from the effects on meal size.

CONCLUSION

Our understanding of the controls of eating in rodents has been transformed in the past 5 years. Although the relationship of eating to nutritional and energetic homeostasis continues to be investigated, particularly in relationship to the new peptides that have been discovered with molecular genetic techniques, the investigation of eating has been broadened in several ways. More attention to behavioral analysis has paid off, especially the microstructure of eating. It has revealed the operation of a cpg as the final common path for the neural control of oromotor movements and provided a continuous measure of the integrated output of the central neural network controlling the cpg during a meal, the functional unit of eating behavior.

The recognition that the size of a meal is under positive and negative feedback controls has been exploited. Specific aminergic and peptidergic mechanisms have been demonstrated to be involved in these feedbacks. The afferent nerves that carry the peripheral information generated along the preabsorptive surface of the gut from the tip of the tongue to the small intestine have been identified. Because some of these peripheral mechanisms are activated in every meal, all controls of eating not related to the food being ingested during a meal act on eating by modulating the central processing of the peripheral feedback stimuli. This has led to a new theory of the controls of eating that is more biological, comprehensive, quantitative, and testable than previous ones (1).

The widely distributed processing of information relevant to the control of eating in the brain reflects the importance and complexity of eating in omnivores such as rodents and humans.

What to eat? Where? When? With whom? These are pressing questions for rodents as well as humans. The ability to answer them with apparent ease requires learning and memory. Recognition of this fact increasingly affects research on eating.

The paradigm shift that the study of eating has undergone, that is, from viewing eating as serving only nutrient and energetic homeostasis to a recognition that the search for the controls of eating is a fundamental problem in behavioral neuroscience (1), makes the basic science more useful for and more relevant to the investigation of clinical eating disorders. Using the similarity in eating behavior, gastrointestinal function, and peripheral visceral afferent neurons as a bridge, the transfer of new information from the laboratory to the clinic should accelerate markedly in the next 5 years.

ACKNOWLEDGMENTS

Part of "115 - The Behavioral Neuroscience of Eating "

We thank Laurel Torres for assistance with processing this manuscript. We are supported by NIH grants MH40010 (GPS), and MH51135 and DK54523 (NG).

REFERENCES

1. Smith GP. The controls of eating: a shift from nutritional homeostasis to behavioral neuroscience. *Nutrition* 2000;17:10–20.

2. Smith GP. Introduction to the reviews on peptides and the control of food intake and body weight. *Neuropeptides* 1999;33(5):323–328.

3. Kennedy GC. The role of depot fat in the hypothalamic control of food intake in the rat. *Proc R Soc Lond, Ser B* 1953;140:578–592.

4. Schwartz MW, Woods SC, Porte D Jr, et al. Central nervous system control of food intake. *Nature* 2000;404:661–671.

5. Woods SC, Schwartz MW, Baskin DG, et al. Food intake and the regulation of body weight. *Annu Rev Psychol* 2000;51:255–277.

6. Smith GP, Gibbs J. The satiating effect of cholecystokinin and bombesin-like peptides. In: Smith GP, ed. *Satiation: from gut to brain*. New York: Oxford University Press, 1998:97–125.

7. Matzinger D, Gutzwiller JP, Drewe J, et al. Inhibition of food intake in response to intestinal lipid is mediated by cholecystokinin in humans. *Am J Physiol* 1999;277(6 Pt 2):R1718–R1724.

8. Grill HJ, Smith GP. Cholecystokinin decreases sucrose intake in chronic decerebrate rats. *Am J Physiol* 1988;254:R853–R856.

9. Smith GP. Cholecystokinin—the first twenty-five years. In: Bray GA, Ryan DH, eds. *Nutrition, genetics, and obesity*. Baton Rouge, LA: Louisiana State University Press, 1999:227–245.

10. Travers JB, Dinardo LA, Kariminamazi H. Motor and premotor mechanisms of licking. *Neurosci Biobehav Rev* 1997;21:631–647.

11. Fay RA, Norgren R. Identification of rat brainstem multisynaptic connections to the oral motor nuclei using pseudorabies virus. III. Lingual muscle motor systems. *Brain Res Rev* 1997;25:291–311.

12. Smith GP. Feeding: control of eating. In: Adelman G, Smith BH, eds. *Elsevier's encyclopedia of neuroscience*. New York: Elsevier, 1999:711–714.

13. Campfield LA, Smith FJ. Systemic factors in the control of food intake. In: Stricker EM, ed. *Handbook of behavioral neurobiology*. New York: Plenum, 1990:183–206.

14. De Vries J, Strubbe JH, Wildering WC, et al. Patterns of body temperature during feeding in rats under varying ambient temperatures. *Physiol Behav* 1993;53(2):229–235.

15. Even P, Coulaud H, Nicolaidis S. Integrated metabolic control of food intake after 2-deoxy-D-glucose and nicotinic acid injection. *Am J Physiol* 1988;255(1 Pt 2):R82–R89.

16. Davis JD, Smith GP. Analysis of the microstructure of the rhythmic tongue movements of rats ingesting maltose and sucrose solutions. *Behav Neurosci* 1992;106:217–228.

17. Davis JD. A model for the control of ingestion, 20 years later. In: Fluharty SJ, Morrison AR, eds. *Progress in psychobiology and physiologic psychology*. New York: Academic Press, 1998:127–173.

18. Grill HJ, Kaplan JM. Caudal brainstem participates in the distributed neural control of feeding. In: Stricker EM, ed. *Handbook of behavioral neurobiology*. New York: Plenum, 1990:125–149.

19. Smith GP. Control of food intake. In: Shils ME, Olson JA, Shike M, et al, eds. *Modern nutrition in health and disease*. Baltimore: Williams & Wilkins, 1999:631–644.

20. Kissileff HR, Wentzlaff TH, Guss JL, et al. A direct measure of satiety disturbance in patients with bulimia nervosa. *Physiol Behav* 1996;60:1077–1085.

21. Smith GP. *Satiation: from gut to brain*. New York: Oxford Press, 1998.

22. Sunday SR, Einhorn A, Halmi KA. Relationship of perceived macronutrient and caloric content to affective cognitions about food in eating-disordered, restrained, and unrestrained subjects. *Am J Clin Nutr* 1992;55(2):362–371.

23. Geliebter A, Melton PM, McCray RS, et al. Gastric capacity, gastric emptying, and test-meal intake in normal and bulimic women. *Am J Clin Nutr* 1992;56:656–661.

24. Devlin MJ, Walsh BJ, Guss JL, et al. Postprandial cholecystokinin release and gastric emptying in patients with bulimia nervosa. *Am J Clin Nutr* 1997;65:114–120.

25. Geracioti TD Jr, Liddle RA. Impaired cholecystokinin secretion in bulimia nervosa. *N Engl J Med* 1988;319:683–688.

26. Weltzin TE, Fernstrom MH, Fernstrom JD, et al. Acute tryptophan depletion and increased food intake and irritability in bulimia nervosa. *Am J Psychiatry* 1995;152:1668–1671.

27. Kissileff HR, Walsh BT, Kral JG, et al. Laboratory studies of eating behavior in women with bulimia. *Physiol Behav* 1986;38(4):563–570.

28. Sclafani A. Macronutrient-conditioned flavor preferences. In: Berthoud H-R, Seeley RJ, eds. *Neural control of macronutrient selection*. Boca Raton, FL: CRC Press, 1999:93–106.

29. Mark GP, Smith SE, Rada PV, et al. An appetitively conditioned taste elicits a preferential increase in mesolimbic dopamine release. *Pharmacol Biochem Behav* 1994;48(3):651–660.

30. Aja SM, Barrett JA, Gietzen DW. CCK(A) and 5-HT3 receptors interact in anorectic responses to amino acid deficiency. *Pharmacol Biochem Behav* 1999;62(3):487–491.

31. Edelman MJ, Gandara DR, Meyers FJ, et al. Serotonergic blockade in the treatment of the cancer anorexia-cachexia syndrome. *Cancer* 1999;86(4):684–688.

32. Faris PL, Kim SW, Meller WH, et al. Effect of decreasing afferent vagal activity with ondansetron on symptoms of bulimia nervosa: a randomised, double-blind trial. *Lancet* 2000;355(9206):792–797.

33. Swank MW, Bernstein IL. c-Fos induction in response to a conditioned stimulus after single trial taste aversion learning. *Brain Res* 1994;636:202–208.

34. Houpt TA, Philopena JM, Wessel TC, et al. Increased c-fos expression in nucleus of the solitary tract correlated with conditioned taste aversion to sucrose in rats. *Neurosci Lett* 1994;172:1–5.

35. Houpt TA, Philopena JM, Joh TH, et al. c-Fos induction in the rat nucleus of the solitary tract correlates with the retention and forgetting of a conditioned taste aversion. *Learn Mem* 1996;3:225–230.

36. Schafe GE, Seeley RJ, Bernstein IL. Forebrain contributions to the induction of a cellular correlate of conditioned taste aversion in the nucleus of the solitary tract. *J Neurosci* 1995;15:6789–6796.

37. Smith GP. The controls of eating: brain meanings of food stimuli. In: Mayer EA, Saper CB, eds. *The biological basis for mind body interactions*. New York: Elsevier, 2000:173–186.

38. Caulliez R, Meile MJ, Nicolaidis S. A lateral hypothalamic D1 dopaminergic mechanism in conditioned taste aversion. *Brain Res* 1996;729(2):234–245.

39. Davis JD, Smith GP, Singh B, et al. Increase in intake with sham feeding experience is concentration dependent. *Am J Physiol* 1999;277:R565–R571.

40. Davis JD, Smith GP, Miesner J. Postpyloric stimuli are necessary for the normal control of meal size in real feeding and sham feeding rats. *Am J Physiol* 1993;265:R888–R895.

41. Gowans SE, Weingarten HP. Elevations of plasma glucose do not support taste-to-postingestive consequence learning. *Am J Physiol* 1991;261(6 Pt 2):R1409–R1417.

42. Davis JD, Smith GP. Learning to sham feed: behavioral adjustments to loss of physiologic postingestional stimuli. *Am J Physiol* 1990;259:R1228–R1235.

43. Drewett RF. The meal patterns of the oestrous cycle and their motivational significance. *Quart J Exp Psychol* 1974;26:489–494.

44. Hrupka BJ, Smith GP, Geary N. Ovariectomy and estradiol affect postingestive controls of sucrose licking. *Physiol Behav* 1997;61:243–247.

45. Eckel LA, Geary N. Endogenous cholecystokinin's satiating action increases during estrus in female Long-Evans rats. *Peptides* 1999;20(4):451–456.

46. Asarian L, Geary N. Cyclic estradiol treatment phasically potentiates endogenous cholecystokinin's satiating action in ovariectomized rats. *Peptides* 1999;20(4):445–450.

47. Butera PC, Xiong M, Davis RJ, et al. Central implants of dilute estradiol enhance the satiety effect of CCK-8. *Behav Neurosci* 1996;110:823–830.

48. Dagnault A, Richard D. Involvement of the medial preoptic area in the anorectic action of estrogens. *Am J Physiol* 1997;272:R311–R317.

49. Hrupka B, Smith GP, Geary N. Central implants of dilute estradiol fail to reduce feeding in ovariectomized rats. *Physiol Behav*, submitted.

50. Lyons PM, Truswell AS, Mira M, et al. Reduction of food intake in the ovulatory phase of the menstrual cycle. *Am J Clin Nutr* 1989;49:1164–1168.

51. Gong EJ, Garrel D, Calloway DH. Menstrual cycle and voluntary food intake. *Am J Clin Nutr* 1989;49:252–258.

52. Geary N, Asarian L, Korach KS, et al. Deficits in estradiol-dependent control of feeding, weight gain, and CCK-satiation in estradiol receptor-α null mice. *Endocrinology* 2001;142: in press.

53. Rigaud D, Angel LA, Cerf M, et al. Mechanisms of decreased food intake during weight loss in adult Crohn's disease patients without obvious malabsorption. *Am J Clin Nutr* 1994;60(5):775–781.

54. Fonager K, Sorensen HT, Olsen J. Change in incidence of Crohn's disease and ulcerative colitis in Denmark. A study based on the National Registry of Patients, 1981–1992. *Int J Epidemiol* 1997;26(5):1003–1008.

55. Logan RF, Kay CR. Oral contraception, smoking and inflammatory bowel disease—findings in the Royal College of General Practitioners Oral Contraception Study. *Int J Epidemiol* 1989;18(1):105–107.

56. Sheahan J, Asarian L, Langhans W, et al. Bacterial lipopolysaccharide anorexia in female rats reveal a novel control of feeding by estradiol. *Soc Neurosci Abstr* 2000;26:2261.

57. Geary N, Smith GP. Appetite. In: Sadock BJ, Sadock VA, eds. *Kaplan & Sadock's comprehensive textbook of psychiatry*. Philadelphia: Lippincott Williams & Wilkins, 2000:209–218.

116

Psychopharmacology of Eating Disorders

Walter H. Kaye

B. Timothy Walsh

Walter H. Kaye: Western Psychiatric Institute & Clinics, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania.

B. Timothy Walsh: Columbia University College of Physicians & Surgeons, New York, New York.

Anorexia nervosa (AN) and bulimia nervosa (BN) are disorders characterized by aberrant patterns of feeding behavior and weight regulation, and disturbances in attitudes toward weight and shape and the perception of body shape. In AN, there is an inexplicable fear of weight gain and unrelenting obsession with fatness even in the face of increasing cachexia. BN usually emerges after a period of dieting, which may or may not have been associated with weight loss. Either self-induced vomiting, or some other means of inappropriate compensation for the excess of food ingested follows binge eating. The majority of people with BN have irregular feeding patterns and satiety may be impaired. Although current AN is an exclusion for the diagnosis of BN, some 25% to 30% of individuals with BN presenting to treatment centers have a prior history of AN; however, all BN subjects have pathologic concern with weight and shape. Common to individuals with AN or BN are low self-esteem, depression, and anxiety.

In certain respects, both diagnostic labels are misleading. Individuals affected with AN rarely have complete suppression of appetite, but rather exhibit a volitional and often an ego syntonic resistance to feeding drives, eventually becoming preoccupied with food and eating rituals to the point of obsession. Similarly, BN may not be associated with a primary pathologic drive to overeat; rather, individuals with BN, like those with AN, have a seemingly relentless drive to restrain their food intake, an extreme fear of weight gain, and often a distorted view of their actual body shape. Loss of control with overeating usually occurs intermittently and typically only some time after the onset of dieting behavior. Episodes of binge eating ultimately develop in a significant proportion of people with AN (1), whereas 5% of those with BN eventually develop AN (2). Considering that restrained eating behavior and dysfunctional cognitions relating weight and shape to self-concept are shared by patients with both of these syndromes, and that transitions between these syndromes occur in many, it has been argued (3) that AN and BN share at least some risk and liability factors in common.

The etiology of AN and BN is presumed to be complex and multiply influenced by developmental, social, and biological processes (4 ,5); however, the exact nature of these interactive processes remains incompletely understood. Certainly, cultural attitudes toward standards of physical attractiveness have relevance to the psychopathology of eating disorders (EDs), but it is unlikely that cultural influences in pathogenesis are prominent. Dieting behavior and the drive toward thinness are quite commonplace in industrialized countries throughout the world, yet AN and BN affect only an estimated .3% to .7%, and 1.7% to 2.5%, respectively, of women in the general population. Moreover, the fact that numerous clear descriptions of AN date from the middle of the nineteenth century suggests that factors other than our current culture play an etiologic role. Also, both syndromes (particularly AN) have a relatively stereotypic clinical presentation, sex distribution, and age of onset. This supports the possibility that there are significant biologic vulnerabilities to developing an ED.

PHENOMENOLOGY

Part of "116 - Psychopharmacology of Eating Disorders "

Variations in feeding behavior have been used to subdivide individuals with AN into two meaningful diagnostic subgroups that differ in other psychopathologic characteristics (6 ,7). In the restricting subtype of AN, subnormal body weight and an ongoing malnourished state are maintained by unremitting food avoidance; in the binge eating/purging subtype of AN, there is comparable weight loss and malnutrition, yet the course of illness is marked by episodes of binge eating, and/or by some type of compensatory action such as self-induced vomiting or laxative abuse. Individuals with the binge eating/purging subtype of AN are also more

likely to exhibit histories of behavioral dyscontrol, substance abuse, and overt family conflict in comparison to those with the restricting subtype of AN. Personality traits of marked perfectionism, conformity, obsessionality, constriction of affect and emotional expressiveness, and reduced social spontaneity are particularly common in individuals with AN. These traits typically appear in advance of the onset of illness and persist even after long-term weight recovery, indicating they are not simply epiphenomena of acute malnutrition and disordered eating behavior (8 ,9 ,10 and 11).

Individuals with BN remain at normal body weight, although many aspire to ideal weights far below the range of normalcy for their age and height. The core features of BN include repeated episodes of binge eating followed by compensatory self-induced vomiting, laxative abuse, or pathologically extreme exercise, as well as abnormal concern with weight and shape. The DSM-IV has specified a distinction within this group between those individuals with BN who engage in self-induced vomiting or abuse of laxatives, diuretics, or enemas (purging type), and those who exhibit other forms of compensatory action such as fasting or exercise (nonpurging type). Beyond these differences, it has been speculated (12) that there are two clinically divergent subgroups of individuals with BN differing significantly in psychopathologic characteristics: a so-called multi-impulsive type, in whom BN occurs in conjunction with more pervasive difficulties in behavioral self-regulation and affective instability; and a second type whose distinguishing features include self-effacing behaviors, dependence on external rewards, and extreme compliance. Individuals with BN of the multi-impulsive type are far more likely to have histories of substance abuse and display other impulse control problems such as shoplifting and self-injurious behaviors.

Most cases of AN emerge during the period of adolescence, although the condition can be observed in children. Whether or not prepubertal onset of the illness confers a more or less ominous prognosis is not known. Recovery from the illness tends to be protracted, but studies of long-term outcome reveal the illness course to be highly variable: Roughly 50% of individuals eventually have reasonably complete resolution of the illness, whereas 30% have lingering residual features that wax and wane in severity long into adulthood. Ten percent of people with AN pursue a chronic, unremitting course; the remaining 10% of those affected eventually die from the disease.

BN is usually precipitated by dieting and weight loss, yet it can occur in the absence of apparent dietary restraint. The frequency of binge episodes, their duration, and the amount of food consumed during any one episode all vary considerably among patients. Age of onset is somewhat more variable in BN than AN, with most cases developing during the period from mid- to late adolescence through the mid-twenties. Follow-up studies of clinical samples 5 to 10 years after presentation show 50% of patients recovered, whereas nearly 20% to 30% continued to meet full criteria for BN (13). Following onset disturbed eating behavior waxes and wanes over the course of several years in a high percentage of clinic cases. Approximately 30% of remitted women relapse into BN symptoms.

PERSISTENT PSYCHOLOGICAL DISTURBANCES AFTER RECOVERY

Part of "116 - Psychopharmacology of Eating Disorders "

People who have an ED often have a variety of symptoms aside from pathologic eating behaviors. Physiologic symptoms include an abundance of neuroendocrine, autonomic, and metabolic disturbances (see the following). Psychological symptoms include depression, anxiety, substance abuse, and personality disorders. Determining whether such symptoms are a consequence or a potential cause of pathologic feeding behavior or malnutrition is a major methodologic issue in the study of EDs. It is impractical to study EDs prospectively because of their low incidence, early age of onset, and difficulty of premorbidly identifying those who will develop an ED. However, subjects can be studied after long-term recovery from an ED. The assumed absence of confounding nutritional influences in recovered ED women raises a possibility that persistent psychobiological abnormalities might be trait-related and potentially contribute to the pathogenesis of this disorder. A limited number of studies have investigated people who have recovered from AN and BN. Although the definition of recovery from an ED has not been formalized, investigators tend to include people formerly ill with AN after they are at a stable and healthy body weight for months or years and have not been malnourished or engaged in pathologic eating behavior during that period of recovery. For BN, investigators tend to include subjects who have been abstinent from binge eating and purging for months or years. Some investigators include criteria of normal menstrual cycles and a minimal duration of recovery, such as 1 year of time.

Investigators (8 ,9 ,10 and 11) have found that women who were long-term recovered from AN had a persistence of obsessional behaviors as well as inflexible thinking, restraint in emotional expression, and a high degree of self- and impulse-control. In addition, they have social introversion, overly compliant behavior, and limited social spontaneity as well as greater risk avoidance and harm avoidance. Moreover, individuals who are long-term recovered from AN had continued core ED symptoms, such as a drive for thinness, and significant psychopathology related to eating habits. Similarly, people who have recovered from BN continue to be overly concerned with body shape and weight, display abnormal eating behaviors, and report dysphoric mood (14 ,15 ,16 and 17). Recovered AN and BN women have increased perfectionism; their most common obsessional target symptoms are the need for symmetry and ordering/arranging. Considered together, these residual behaviors can be characterized as over-concerns with body image and thinness, obsessionality

with symmetry, exactness, and perfectionism, and dysphoric/negative affect. In general, pathologic eating behavior and malnutrition appears to exaggerate the magnitude of these concerns. Thus, the intensity of these symptoms is less after recovery but the content of these concerns remains unchanged. The persistence of these symptoms after recovery raise the possibility that the disturbances are premorbid traits that contribute to the pathogenesis of AN and BN.

PHARMACOLOGIC TREATMENT OF ANOREXIA NERVOSA

Part of "116 - Psychopharmacology of Eating Disorders "

Most medication trials in AN have been conducted with inpatients in an attempt to accelerate restoration of weight. Some studies also examined the impact of medication on mood or anorectic attitudes. A wide variety of psychoactive medications, such as L-dopa (18), phenoxybenzamine (19), diphenylhydantoin (20 ,21), stimulants (22), and naloxone (23), have been administered to people with anorexia nervosa in open, uncontrolled trials. In many of these trials, medications have been claimed to be beneficial, but none of these observations has been confirmed under double-blind, controlled conditions.

Few studies of medication using rigorous double-blind placebo-controlled trials have been reported in patients with AN. In contrast to the positive claims from open trials, results from double-blind trials have been limited, for the most part. Double-blind studies, at most, report marginal success in treatment of specific problems such as improving the rate of weight gain during refeeding, and disturbed attitudes toward food and body image, depression, or gastrointestinal discomfort.

One problem with determining the efficacy of pharmacotherapy in AN is that often medications have been given in association with other therapies. Thus, it may be unclear whether it was the medication or therapy that resulted in improvement. Furthermore, the primary criterion for improvement has often been weight gain, not a normalization of thinking and reduction in fears of being fat. It is important to emphasize that treatment in structured settings, such as inpatient units, even without medication, succeeds in restoring the weight of over 85% of underweight patients (24). Thus, it may be difficult to prove that an active medication is effective in such a setting. However, relapse within 1 year *after* successful inpatient weight restoration is very common (25). For example, the Maudsley study (26) reported that only 23% of the patients had a good outcome at 1 year after discharge despite intensive outpatient individual or family therapy.

Controlled trials of the neuroleptics pimozide (27) and sulpiride (28) have suggested limited effects in accelerating weight gain or altering anorectic attitudes for some patients for part of the study, but overall drug effect was marginal. Recently, there has been clinical interest in “atypical” neuroleptics for AN because of their notoriety for causing weight gain in other patient populations (29). A recent case report suggested that olanzapine administration was associated with weight gain and maintenance as well as reduced agitation and resistance to treatment in 2 women with AN (30). Several drugs have been tested because of anecdotal reports of their effects on stimulating appetite. Tetrahydrocannabinol (THC) was not useful and, in fact, may have been detrimental because it increased dysphoria in some patients (31). Clonidine was also found to have no therapeutic effect on increasing weight restoration as compared to placebo (32), even with doses that affected hemodynamic parameters.

When underweight, patients with anorexia nervosa have delayed gastric emptying (33), which improves with refeeding. Still, delayed gastric emptying could perpetuate the disorder in some patients by limiting the quantity of food that may be comfortably eaten. Most studies of prokinetic drugs in AN have been limited to parenteral preparations or experiments with small uncontrolled groups of patients (34 ,35 and 36). In a controlled trial, cisapride (37) was no better than placebo in improving gastric emptying, although some subjective measures of distress during meals and measures of hunger improved more in the group on cisapride.

In summary, these medication trials have been of short duration and focused on whether medication produces additive benefit to an established treatment program. Few follow-up studies have examined whether medication treatment produces lasting benefit. A new generation of studies has begun to focus on whether medication can prevent relapse after patients leave to a structured treatment setting.

Use of Antidepressants in AN

There has been controversy as to whether AN and major depressive disorders share a common diathesis; however, critical examination of clinical phenomenology, family history, antidepressant response, biological correlates, course and outcome, and epidemiology yield limited support for this hypothesis (38 ,39 and 40). Still, the high frequency of mood disturbances associated with this disorder resulted in trials of drugs such as amitriptyline (41 ,42 and 43), and lithium (44). Neither medication appears to significantly improve mood compared with the effects of placebo.

For more that 50 years (45), investigators have suggested that AN shares similarities with obsessive-compulsive disorder (OCD). In fact, patients with AN have a high prevalence of obsessive-compulsive symptoms or disorders (46 ,47 and 48), as well other anxiety disorders (49). More over, adult women with OCD have an increased incidence of prior AN (50). Individuals with a past history of AN display evidence of increased serotonin (51) activity that persists after long-term weight recovery. In addition, women who recover from AN continue to have modest, but significant, increases in negative

mood, obsessionality, perfectionism, and core eating disorder symptoms. Similarly (10), personality characteristics associated with AN, such as introversion, self-denial, limited spontaneity, and a stereotyped thinking style, may also persist after weight recovery. Studies in humans and animals suggest that serotonin activity is related to behavioral inhibition. Together, these data raise the possibility that increased CSF 5-HIAA could be associated with inhibition and an obsessive need with exactness and perfectionism. A disturbance of this neurotransmitter system has been implicated in OCD (52) and only serotonin-specific medication has been found to be useful in treating OCD.

There are suggestions that medications that affect the serotonin system may impact the clinical characteristics of patients with AN. Initial reports on cyproheptadine, a drug that is thought to act on the serotonergic and histaminergic system (53), indicated that it might have beneficial effects on weight gain, mood, and attitude in some patients (54 ,55). Cyproheptadine data from comparison trials with amitriptyline and placebo found cyproheptadine to significantly improve weight gain in the restricting subtype of AN, whereas amitriptyline was more effective in those patients with bulimic behavior (56).

Several groups (57 ,58) reported that an open trial of fluoxetine, a highly specific serotonin reuptake inhibitor might help patients with AN gain and/or maintain a healthy body weight. Recently, the Pittsburgh group reported a double-blind placebo-controlled trial of fluoxetine in 35 patients with restrictor-type AN (59). Subjects were started on fluoxetine after they achieved weight restoration (approximately 90% of ideal body weight) during a hospitalization. Patients were randomly assigned to fluoxetine (*N* = 16) or placebo (*N* = 19) after inpatient weight-restoration and then were followed as outpatients for 1 year. After 1 year of outpatient follow-up, 10 of 16 (63%) subjects had a good outcome on fluoxetine, whereas only three of 19 (16%) had a good outcome on placebo (*P* = .006) (Fig. 116.1). Aside from improved outcome, fluoxetine administration was associated with a significant reduction in obsessions and compulsions and a trend toward a reduction in depression. These data suggest that fluoxetine may help prevent relapse in some patients with AN.

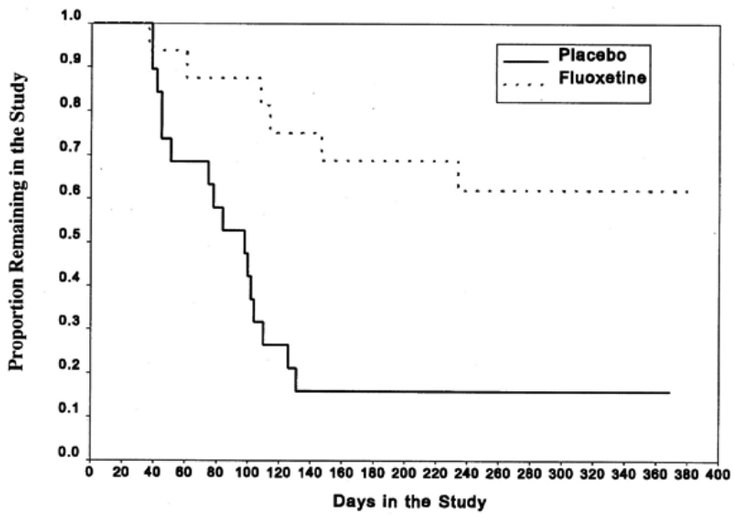


FIGURE 116.1. Survival of subjects with anorexia nervosa treated with fluoxetine or placebo.

It is important to note that SSRIs appear to have little effect on reducing symptoms and preventing hospitalization in malnourished, underweight anorexics (60 ,61). Women with AN, when malnourished and underweight, have reduced plasma tryptophan availability (62) and reduced CSF 5-HIAA (63), the major metabolite of serotonin in the brain. In addition, low estrogen values during the malnourished state may reduce serotonin activity by effects on gene expression for serotonin receptors (64) or the serotonin transporter (65). SSRIs are dependent on neuronal release of serotonin for their action. If malnourished patients have compromised release of serotonin from presynaptic neuronal storage sites and reduced synaptic serotonin concentrations, then a clinically meaningful response to an SSRI might not occur (66). The possibility that fluoxetine is only effective for patients after weight restoration is supported by the fact that a change of serotonin activity is associated with weight gain. For example, CSF 5-HIAA levels are low in underweight anorexics, normal in short-term weight-restored anorexics, and elevated in long-term weight-restored anorexics (67). If CSF 5-HIAA levels accurately reflect CNS serotonin activity, then these data imply that a substantial increase in serotonin activity occurs after weight gain.

The use of serotonin-specific medications in the treatment of AN is promising but many questions remain. First, only one double-blind placebo-controlled study has been completed in a relatively small number of restrictor-type patients. Thus, it will be important to replicate this work

in a larger group of patients. Second, more data are needed to determine if there are differential effects in the restricting of binge eating/purging subtypes of AN. Third, it needs to be determined whether certain features are especially responsive to serotonin-specific medications: core anorexic symptoms, depression, anxiety, obsessionality, or eating behavior.

Guidelines for Clinical Treatment

The first line of treatment for underweight patients with AN should be refeeding and weight restoration. As noted, although difficult, most patients will gain weight in a structured eating disorders treatment program without the use of medication. Weight gain alone tends to reduce exaggerated obsessionality and dysphoric mood in many patients (68). There is limited evidence that fluoxetine and possibly other serotonergic medications help prevent relapse after weight restoration. It is important to emphasize that some physiologic and cognitive alterations persist for months after achieving goal weight, including increased energy needs, menstrual disturbances, several neurotransmitter disturbances, urges to engage in disordered eating patterns, and body image distortions. Thus, treatment should continue for at least 3 to 6 months after achieving goal weight, preferably until there is resumption of menstrual periods, normalization of caloric needs, remediation of any physical complications, and sufficient remission of pathologic eating and body-image distortions so that daily activities are not disturbed. We strongly support use of the recent American Psychiatric Association (APA) guidelines for eating disorders (69), which describe comprehensive treatment of AN.

PHARMACOLOGIC TREATMENT OF BN

Part of "116 - Psychopharmacology of Eating Disorders "

As summarized in *Neuropsychopharmacology, the Fourth Generation of Progress,* a substantial body of work was published during the 1980s and early 1990s demonstrating that antidepressants are more effective than placebo in the treatment of BN (70). In 1996, the FDA approved the use of fluoxetine (71 ,72) for this disorder, the only medication to receive such an official indication to date. Although the notion of using antidepressants for BN emerged because of the high frequency of symptoms of depression and anxiety, the utility of antidepressants does not appear confined to patients with concurrent major depression, suggesting that these medications may exert their effects, at least in part, via alterations in the neural systems underlying the control of appetite. The notion that antidepressants may be useful in BN via mechanisms other than those that are responsible for their antidepressant activity is also suggested by the observations that a higher daily dose of fluoxetine (60 mg per day) appears superior to the standard antidepressant dose (20 mg per day) in the treatment of BN and that the onset of benefit from antidepressant treatment typically is quite rapid (Fig. 116.2).

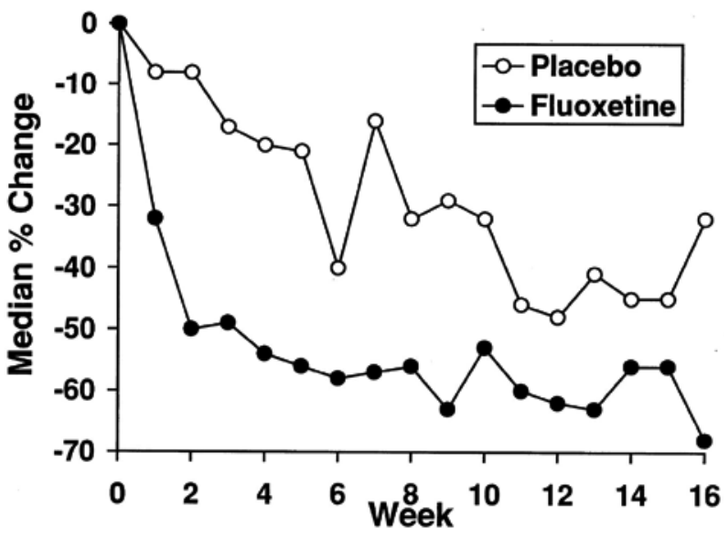


FIGURE 116.2. Median percentage change in the number of binge-eating episodes among patients with bulimia nervosa receiving fluoxetine or placebo.

No trials have been published in which the efficacy of one antidepressant is compared directly to that of another. In the absence of such data, although virtually all classes of antidepressants appear superior to placebo in reducing binge frequency, SSRIs are generally preferred because of their relatively benign side-effect profile; however, aside from fluoxetine, only fluvoxamine has been formally examined in BN. Fichter and colleagues reported a study of novel design in which patients were randomly assigned to fluvoxamine or placebo following successful completion of inpatient treatment (73). Although fluvoxamine was associated with a dropout rate of 38% over 19 weeks compared to 14% on placebo, the active drug was superior to placebo in reducing the re-emergence of bulimic behaviors and attitudes. In light of these results, it is surprising that a large European trial has been reported to find no difference between the response to fluvoxamine and placebo in the initial treatment of outpatients with bulimia (Freeman, personal communication, 1999). Thus, although most clinicians expect sertraline, paroxetine, citalopram, and venlafaxine to be useful, the efficacy and ideal dose of SSRIs other than fluoxetine for the treatment of BN have not been established.

Since the clear recognition of bulimia as a syndrome in 1979, effective psychotherapeutic approaches, have also been developed, most of which utilize cognitive behavioral therapy (CBT). CBT is generally believed to be more effective than a single course of an antidepressant medication (69). This fact, coupled with reasonable evidence of sustained benefit following CBT and the reluctance of many patients to take psychotropic medications, has led to CBTs being generally considered the treatment of first choice for BN. Several studies have examined whether it is beneficial

to combine psychological treatment with antidepressant medication.

The earliest studies of the combination of medication and psychotherapy utilized tricyclic antidepressants. Mitchell and associates (74) found that imipramine was associated with a greater reduction in measures of anxiety and depression than was placebo when combined with an intensive group psychotherapy program; however, imipramine did not augment the impact of the psychological treatment on the salient behavioral symptom, binge eating. Agras and colleagues (75) compared five treatments for BN: individual CBT alone, desipramine alone for 16 or 24 weeks, and CBT plus desipramine for either 16 or 24 weeks. As was also true of the study of Mitchell and colleagues, Agras and co-workers reported that the outcome of psychological treatment alone was clearly superior to that of a course of tricyclic antidepressant. There were a few hints of a small advantage for the combination of medication and CBT, but these were not impressive. Leitenberg and co-workers (76) attempted to compare CBT to a course of desipramine and to a combination, but terminated the study prematurely because of a high dropout rate, primarily caused by medication side effects.

More recent studies have examined the potential advantages of combining an SSRI (fluoxetine) and psychological treatment. The Columbia group has reported the results of a study that compared two forms of individual psychological treatment (CBT and supportive psychotherapy) combined either with placebo or a two-stage medication intervention (77). Patients assigned to receive active medication received desipramine; if desipramine was either ineffective or intolerable, the medication was changed to fluoxetine under double-blind conditions. In this study, CBT was clearly superior to supportive psychotherapy in reducing the key behavioral symptoms of BN. In addition, compared to placebo, active medication added modestly but significantly to improvement in binge eating and depression.

Goldbloom and colleagues (78) compared individual CBT to a course of fluoxetine and the combination. Unfortunately, interpretation of the results is limited by a high dropout rate, which resulted in few significant differences among the three treatments. Beumont and colleagues (79) reported a comparison of fluoxetine versus placebo when combined with nutritional counseling, which presumably included many elements of CBT. The nutritional counseling program was impressively effective, and at the conclusion of treatment, there were no significant differences between the fluoxetine and placebo groups in binge frequency; however, the fluoxetine-treated group exhibited less dietary restraint and fewer concerns about body shape and weight.

Combined, these data suggest that adding medication to structured psychological treatment for BN does provide added benefit, but of small magnitude. Clinically, guidelines to identify patients who are particularly likely to benefit from one treatment approach or another would be extremely useful. Unfortunately, attempts to identify such predictors of treatment response have been impressively unsuccessful. Because those patients who derive the greatest benefit from treatment typically exhibit an early response (80), it may useful to initiate treatment with CBT, for example, and to add another intervention such as medication if the initial response is not satisfactory. Recent data demonstrate that medication can be useful for patients who do not respond adequately to psychological treatment or who relapse following the end of treatment (18).

Despite the progress in developing treatment approaches for bulimia in the last 20 years, a major current problem is the absence of treatments of established efficacy other than CBT and antidepressant medication. Even in the best hands, only about 50% of patients achieve remission with these treatments, and a significant number relapse following the conclusion of the initial intervention. Clinicians and investigators have considered the use of other psychotropic medications that are believed to reduce appetite, such as topiramate, but no controlled data are available to date about its utility in BN. Recently, Faris and associates (82) have reported that the antiemetic medication ondansetron, a 5-HT3 antagonist, was more effective than placebo during a 4-week trial in reducing binge eating and vomiting in a group of chronically and severely ill bulimic patients. More data regarding the side effects of ondansetron and its impact on psychological features of the disorder are required to assess the clinical utility of this agent, but the exploration of novel medication interventions for BN is overdue.

PHARMACOLOGIC TREATMENT OF BINGE EATING DISORDER

Part of "116 - Psychopharmacology of Eating Disorders "

During the development of DSM-IV, interest grew in defining another eating disorder characterized by frequent binge eating but without the recurrent use of inappropriate compensatory behavior required for the diagnosis of BN. Out of these discussions, criteria for binge eating disorder (BED) evolved, and were included in an appendix of DSM-IV as a criteria set for further study. Significant interest in the characteristics and treatment of BED has since developed, and the results of several psychopharmacologic interventions have been published. Although obesity is not required by the criteria for BED suggested in DSM-IV, the studies to date have generally focused on overweight or obese individuals.

Even before the delineation of BED, McCann and Agras (83) reported that desipramine was superior to placebo in reducing binge frequency among a group of “nonpurging bulimics.” Most of the participants were overweight, but neither desipramine nor placebo was associated with significant weight loss in this short-term study. In contrast, Marcus

and associates (84) found that when combined with a behavioral weight loss program, fluoxetine was associated with greater weight loss than was placebo during a 1-year study of obese binge eaters; unfortunately, information on binge frequency was not obtained. In contrast, Alger and colleagues (85) reported that neither imipramine nor naltrexone was more effective than placebo in reducing binge frequency or weight in obese binge eaters.

More recently, studies have focused on patients meeting the criteria of DSM-IV for BED and have examined serotonergic agents. Stunkard and co-workers (86) found that the appetite suppressant *d*-fenfluramine, which has since been withdrawn from the market, was more effective than placebo among 28 obese women with BED in reducing binge frequency; however, surprisingly, not in promoting weight loss. Nevertheless, fluvoxamine compared with placebo was associated with significant reductions in both binge frequency and body mass index among 85 patients with BED (87).

Obesity is a major and growing public health problem in the industrialized world in general and the United States in particular. Approximately one-third of obese individuals presenting to weight loss clinics meet diagnostic criteria for BED; therefore, effective treatments for this disorder may be of widespread clinical utility. Research conducted to date suggests that pharmacotherapy may play a role, but a number of important issues are unresolved. Individuals with BED have disturbances in eating behavior by definition, and are typically overweight and exhibit symptoms of anxiety and depression in clinical samples. Effective interventions should lead to improvements in all three spheres. Yet, it is surprising that the response of these presumably related symptoms to medication is at least somewhat inconsistent, so that patients may report a decrease in binge frequency but no change in weight. A major problem in the development of effective treatment strategies is an impressively high response of binge eating to nonspecific interventions, including placebo. In part for this reason, the effects of medication treatment have been modest in size. Furthermore, the impact of medication in BED appears to fade rapidly once medication has been discontinued. These issues leave the role of pharmacotherapy for BED currently unresolved, and underline the need for additional research, including studies to examine the potential benefits of combining medication with psychological treatment, especially CBT.

REFERENCES

1. Halmi KA, Eckert E, Marchi P. Comorbidity of psychiatric diagnoses in anorexia nervosa. *Arch Gen Psychiatry* 1991;48(8):712–718.

2. Hsu LKG, Sobkiewicz TA. Bulimia nervosa: a four-to six-year followup-study. *Psychol Med* 1989;19:1035–1038.

3. Schweiger U, Fitcher M. Eating disorders: clinical presentation, classification and aetiological models. In: Jimerson D, Kaye WH, eds. *Bailliere's clinical psychiatry international practice and research, eating disorders*. London: Balliere's Tindall, 1997:199–216.

4. Garner DM. Pathogenesis of anorexia nervosa. *Lancet* 1993;341:1631–1635.

5. Treasure J, Campbell I. The case for biology in aetiology of anorexia nervosa. *Psychol Med* 1994;24(1):3–8.

6. Garner DM, Garfinkel PE, O'Shaughnessy M. The validity of the distinction between bulimia with and without anorexia nervosa. *Am J Psychiatry* 1985;142(5):581–587.

7. Strober M, Salkin B, Burroughs J, et al. Validity of the bulimia-restricter distinction in anorexia nervosa: parental personality characteristics and family psychiatric morbidity. *J Nerve Ment Dis* 1982;170(6):345–351.

8. Casper RC. Personality features of women with good outcome from restricting anorexia nervosa. *Psychosom Med* 1990;52(2):156–170.

9. O'Dwyer AM, Lucey JV, Russell GF. Serotonin activity in anorexia nervosa after long-term weight restoration: response to *d*-fenfluramine challenge. *Psychol Med* 1996;26(2):353–359.

10. Srinivasagam NM, Kaye WH, Plotnicov KH, et al. Persistent perfectionism, symmetry, and exactness after long-term recovery from anorexia nervosa. *Am J Psychiatry* 1995;152(11):1630–1634.

11. Strober M. Personality and symptomatological features in young, nonchronic anorexia nervosa patients. *J Psychosom Res* 1980;24(6):353–359.

12. Vitousek K, Manke F. Personality variables and disorders in anorexia nervosa and bulimia nervosa. *J Abnorm Psychol* 1994;103(1):137–147.

13. Keel PK, Mitchell JE. Outcome in bulimia nervosa. *Am J Psychiatry* 1997;154(3):313–321.

14. Collings S, King M. Ten-year follow-up of 50 patients with bulimia nervosa. *Br J Psychiatry* 1994;164:80–87.

15. Johnson-Sabine E, Reiss D, Dayson D. Bulimia nervosa: a 5-year follow-up study. *Psychol Med* 1992;22:951–959.

16. Kaye WH, Greeno CG, Moss H, et al. Alterations in serotonin activity and psychiatric symptomatology after recovery from bulimia nervosa. *Arch Gen Psychiatry* 1998;55:927–935.

17. Norring CE, Sohlberg SS. Outcome, recovery, relapse and mortality across six years in patients with clinical eating disorders. *Psychiatr Scand* 1993;87(6):437–444.

18. Johanson AJ, Knorr N J. L-Dopa as treatment for anorexia nervosa. In: Vigersky RA, ed. *Anorexia nervosa*. New York: Raven Press, 1977:363–372.

19. Redmond DE, Swann A, Heninger GR. Phenoxybenzamine in anorexia nervosa. *Lancet* 1976;2:397.

20. Green RS, Rau JH. Treatment of compulsive eating disturbances with anticonvulsant medication. *Am J Psychol* 1974;131:428–432.

21. Green RS, Rau JH. The use of diphenylhydantoin in compulsive eating disorders: further studies in anorexia nervosa. In: Vigersky RA, ed. *Anorexia nervosa*. New York: Raven Press, 1977:377–385.

22. Wulliemier JF, Rossel F, Sinclair K. La therapie comportementale de l'anorexia. *J Psychsom Res* 1975;19:267–272.

23. Moore R, Mills IH, Forster A. Naloxone in the treatment of anorexia nervosa: effect on weight gain and lipolysis. *J Roy Soc Med* 1981;74:129–131.

24. Halmi KA. Anorexia nervosa. In: Kaplan HI, Freedman AM, Sadock RJ, eds. *Comprehensive textbook of psychiatry*, vol 2, 3rd ed. Baltimore: Williams & Wilkins, 1980.

25. Hsu LKG. The treatment of anorexia nervosa. *Am J Psychiatry* 1986;143:573–581.

26. Russell GFM, Szmukler GI, Dare C, et al. An evaluation of family therapy in anorexia nervosa and bulimia nervosa. *Arch Gen Psychiatry* 1987;44:1047–1056.

27. Vandereychken W, Pierlott R. Pimozide combined with behavior therapy in the short-term treatment of anorexia nervosa. *Acta Psych Scand* 1982;66:445–450.

28. Vandereychken W. Neuroleptics in the short-term treatment of anorexia nervosa: a double-blind, placebo controlled study with sulpiride. *Br J Psychol* 1984;144:288–292.

29. Casey DE. Side effect profiles of new antipsychotics. *J Clin Psychol* 1996;57(S11):38–45.

30. La Via MC, Gray N, Kaye WH. Case reports of olanzapine treatment of anorexia nervosa. *Int J Eating Dis* 2000;27(3):363–366.

31. Gross H, Ebert MH, Faden VB. A double-blind trial of delta-9-tetrahydrocannabinol in primary anorexia nervosa. *J Clin Psychopharmacol* 1983;3:165–171.

32. Casper RC, Schlemmer RF, Javaid JI. A placebo-controlled, crossover study of oral clonidine in acute anorexia nervosa. *Psychol Res* 1987;20:249–260.

33. Szmukler GI, Young GP, Lichenstein M, et al. A serial study of gastric emptying in anorexia and bulimia nervosa. *Aust NZ J Med* 1990;20:220–225.

34. Domstad PA, Shih WJ, Humphries L, et al. Radionuclide gastric emptying studies in patients with anorexia nervosa. *J Nucl Med* 1987;28:816–819.

35. Russell DM, Freedman ML, Feiglin DH, et al. Delayed gastric emptying and improvement with domperidone in a patient with anorexia nervosa. *Am J Psychiatry* 1983;140:1235–1236.

36. Saleh JW, Lebwohl P. Metoclopramide-induced gastric emptying in patients with anorexia nervosa. *Am J Gastroenterol* 1980;74:127–132.

37. Stacher G, Bergmann H, Wiesnagrotzki S, et al. Intravenous cisapride accelerates delayed gastric emptying and increases antral contraction amplitude in patients with primary anorexia nervosa. *Gastroenterology* 1987;92:1000–1006.

38. Strober M, Katz JL. Depression in the eating disorders: a review and analysis of descriptive, family, and biological findings. In: Garner DM, Garfinkel PE, eds. *Diagnostic issues in anorexia nervosa and bulimia nervosa*. New York: Brunner/Mazel, 1988.

39. Rothenberg A. Differential diagnosis of anorexia nervosa and depressive illness: a review of 11 studies. *Comp Psychiatry* 1988;29:427–432.

40. Swift WJ, Andrews D, Barklage NE. The relationship between affective disorder and eating disorders: a review of the literature. *Am J Psychol* 1986;143:290–299.

41. Moore R. Amitriptyline therapy in anorexia nervosa. *Am J Psychol* 1977;13:1303–1304.

42. Needlemen HL, Waber D. The use of amitriptyline in anorexia nervosa. In: Vigersky RA, ed. *Anorexia nervosa*. New York: Raven Press, 1977.

43. Beiberman J, Herzog DB, Rivinus TM, et al. Amitriptyline in the treatment of anorexia nervosa: a double-blind placebo-controlled study. *J Clin Psychopharmacol* 1985;5:10–16.

44. Gross HA, Ebert MH, Faden VB, et al. A double-blind controlled trial of lithium carbonate in primary anorexia nervosa. *J Clin Psychopharmacol* 1981;1:376–381.

45. Kaye WH, Weltzin T, Hsu LKG. Anorexia nervosa. In: Hollander E, ed. *Obssessive-compulsive related disorders*. Washington, DC: American Psychiatric Press, 1993:49–70.

46. Strober M. Personality and symptomatological features in young, nonchronic anorexia nervosa patients. *J Psychosom Res* 1980;24:353–359.

47. Hudson JI, Pope HG, Jonas JM, et al. Phenomenologic relationship of eating disorders to major affective disorder. *Psychiatry Res* 1983;9:345–354.

48. Solyom LS, Freeman RJ, Miles JE. A comparative psychometric study of anorexia nervosa and obsessive neurosis. *Can J Psychiatry* 1982;27:282–286.

49. Toner BB, Garfinkel PE, Garner DM. Long term follow-up of anorexia nervosa. *Psychosom Med* 1986;48:520–529.

50. Kasvikis YG, Tsakiris F, Marks IM. Past history of anorexia nervosa in women with obsessive compulsive disorder. *Int J Eating Dis* 1986;5:1069–1075.

51. Kaye WH, Gwirtsman HE, George DT, et al. Altered serotonin activity in anorexia nervosa after long-term weight restoration. *Arch Gen Psychol* 1991;48:556–562.

52. Zohar J, Insel TR. Obsessive-compulsive disorder: psychobiological approaches to diagnosis, treatment, and pathophysiology. *Biol Psychiatry* 1987;22:667–687.

53. Stone CA, Wenger HC, Ludden CT. Antiserotonin-antihistamine properties of cyproheptadine. *J Pharmacol Exp Ther* 1961;131:73–81.

54. Goldberg SC, Halmi KA, Eckert ED. Cyproheptatadine in anorexia nervosa. *Br J Psychiatry* 1979;134:67–70.

55. Halmi KA, Eckert E, Falk JR. Cyproheptadine for anorexia nervosa. *Lancet* 1982;1:1357–1358.

56. Halmi KA, Eckert ED, LaDu, TJ, et al. Anorexia nervosa: treatment efficacy of cyproheptadine and amitriptyline. *Arch Gen Psychiatry* 1986;43:177–181.

57. Gwirtsman HE, Guze BH, Yager J, et al. Fluoxetine treatment of anorexia nervosa: an open clinical trial. *J Clin Psychiatry* 1990;51:378–382.

58. Kaye WH, Weltzin TE, Hsu LKG, et al. An open trial of fluoxetine in patients with anorexia nervosa. *J Clin Psychol* 1991;52:464–471.

59. Kaye WH, Hsu LKG, Weltzin TE, et al. Double-blind placebo-controlled administration of fluoxetine in restricting- and restricting-purging-type anorexia nervosa. *Biol Psychiatry* 2001;49(7):644–652.

60. Ferguson CP, La Via MC, Crossan PJ, et al. Are SSRIs effective in underweight anorexia nervosa? *Int J Eating Dis* 1999;25(1):11–17.

61. Attia E, Haiman C, Walsh BT, et al. Does fluoxetine augment the inpatient treatment of anorexia nervosa? *Am J Psychiatry* 1998;155(4):548–551.

62. Schweiger V, Warnoff M, Pahl J, et al. Effects of carbohydrate and protein meals on plasma large neutral amino acids, glucose and insulin plasma levels of anorectic patients. *Metabolism* 1986;35(2):938–943.

63. Kaye WH, Gwirtsman HE, George DT, et al. CSF 5-HIAA concentrations in anorexia nervosa: reduced values in underweight subjects normalize after weight gain. *Biol Psychiatry* 1988;23(1):102–105.

64. Fink G, Summner BEH. Oestrogen and mental state. *Nature* 1996;383:306.

65. McQueen JK, Wilson H, Dow RC, et al. Oestradiol-17 Beta increases serotonin transporter (SERT) binding sites and SERT mRNA expression in discrete regions of female rat brain. *J Physiol (Lond)* 1996;495:114.

66. Tollefson GD. Selective serotonin reuptake inhibitors. In: Schatzberg AF, Memeroff CB, eds. *Textbook of psychopharmacology*. Washington, DC: American Psychiatric Press, 1995.

67. Kaye WH, Ebert MH, Raleigh M, et al. Abnormalities in CNS monoamine metabolism in anorexia nervosa. *Arch Gen Psychiatry* 1984;41(4):350–355.

68. Pollice C, Kaye WH, Greeno CG, et al. Relationship of depression, anxiety, and obsessionality to state of illness in anorexia nervosa. *Int J Eating Dis* 1997;21:367–376.

69. American Psychiatric Association. Practice guideline for the treatment of patients with eating disorders (revision). *Am J Psychiatry* 2000;157:139.

70. Walsh BT, Devlin M. Psychopharmacology of anorexia nervosa, bulimia nervosa, and binge eating. In: Bloom FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1993.

71. Flouxetine Bulimia Nervosa Collaborative Study Group. Fluoxetine in the treatment of bulimia nervosa: a multicenter, placebo-controlled, double-blind trial. *Arch Gen Psychiatry* 1992;49:139–147.

72. Goldstein DJ, Wilson MC, Thompson VL, et al. Long-term fluoxetine treatment of bulimia nervosa. *Br J Psychiatry* 1995;166:660–666.

73. Fichter MM, Kruger R, Rief W, et al. Fluvoxamine in prevention of relapse in bulimia nervosa: effects on eating—specific psychopathology. *J Clin Psychopharmacol* 1996;16:9–18.

74. Mitchell JE, Pyle RL, Eckert ED, et al. A comparison study of antidepressants and structured group psychotherapy in the treatment of bulimia nervosa. *Arch Gen Psych* 1990;47:149–157.

75. Agras WS, Arnow B, Schneider JA, et al. Pharmacologic and cognitive-behavioral treatment for bulimia nervosa: a controlled comparison. *Am J Psychiatry* 1992;49:82–87.

76. Leitenberg H, Rosen JC, Wolf J, et al. Comparison of cognitive-behavior therapy and desipramine in the treatment of bulimia nervosa. *Behav Res Ther* 1994;32:37–45.

77. Walsh BT, Wilson GT, Loeb KL, et al. Medication and psychotherapy in the treatment of bulimia nervosa. *Am J Psychiatry* 1997;154:523–531.

78. Goldbloom DS, Olmsted MP, Davis R, et al. A randomized controlled trial of fluoxetine and cognitive behavioral therapy for bulimia nervosa: short-term outcome. *Behav Res Ther* 1997;9:803–811.

79. Beumont PJV, Russel JD, Touyz SW, et al. Intensive nutritional counseling in bulimia nervosa: a role for supplementation with fluoxetine? *Aust NZ J Psychiatry* 1997;31:514–524.

80. Wilson GT, Loeb KL, Walsh BT, et al. Psychological versus pharmacological treatments of bulimia nervosa: predictors and processes of change. *J Consult Clin Psychol* 1999;67:451–459.

81. Walsh BT, Agras WS, Devlin M, et al. Fluoxetine for bulimia nervosa following poor response to psychotherapy. *Am J Psychiatry* 2000;157(8):1332–1334.

82. Faris PL, Kim SW, Meller WJ, et al. Effect of decreasing afferent vagal activity with ondansetron on symptoms of bulimia. *Lancet* 2000;355(9206)792–797.

83. McCann UD, Agras WS. Successful treatment of nonpurging bulimia nervosa with desipramine: a double-blind, placebo-controlled study. *Am J Psychiatry* 1990;147:1509–1513.

84. Marcus MD, Wing RR, Ewing L, et al. A double-blind, placebo-controlled trial of fluoxetine plus behavior modification in the treatment of obese binge eaters and non-binge eaters. *Am J Psychiatry* 1990;147:876–881.

85. Alger SA, Schwalberg MD, Bigaouette JM, et al. Effect of a tricyclic antidepressant and opiate antagonist on binge-eating in normoweight bulimic and obese, binge-eating subjects. *Am J Clin Nutr* 1991;53:865–871.

86. Stunkard A, Berkowitz R, Tanrikut C, et al. *d*-Fenfluramine treatment of binge eating disorder. *Am J Psychiatry* 1996;153:1455–1459.

87. Hudson JI, McElroy SL, Raymond NC, et al. Fluvoxamine in the treatment of binge-eating disorder: a multicenter placebo-controlled, double-blind trial. *Am J Psychiatry* 1998;155:1756–1762.

117

Tourette Syndrome and Related Tic Disorders

Neal R. Swerdlow

James F. Leckman

Neal R. Swerdlow: Department of Psychiatry, University of California, San Diego School of Medicine, La Jolla, California.

James F. Leckman: Child Study Center, Yale University School of Medicine, New Haven, Connecticut.

> Each movement is preceded by certain preliminary sensory signals and is in turn followed by sensory impressions at the end of the action. Each movement is the result of a voluntary capitulation to a demanding and relentless urge accompanied by an extraordinarily subtle sensation that provokes and fuels the urge. Successively sharper movements build up to a climax—a climax that never comes (1).

A MODEL NEUROPSYCHIATRIC DISORDER

Part of "117 - Tourette Syndrome and Related Tic Disorders "

Biological models allow investigators to extrapolate from simple to complex systems, to generate and test hypotheses, and to grasp schema that are within range of our intellect, as we reach to conceptualize things beyond this range. Tourette syndrome (TS) is a “model neuropsychiatric disorder” (2 ,3) that seems tantalizing in its simplicity. The genetic basis is stronger than any common neuropsychiatric disorder other than Huntington disease. The age of onset and the sex distribution of TS are strong clues that neurodevelopmental and hormonal processes are causative in TS. Emerging evidence suggests a role of epigenetic (e.g., hypoxia) or “environmental” (e.g., group A streptococcal infections) factors in the origin of at least some cases of TS, again making TS an ideal model for “nature-nurture” interactions in the pathogenesis of brain disorders. The clinical presentation and responsivity to dopamine antagonists provide strong clues that the critical substrates of TS fall within the basal ganglia, a system implicated in an increasing number of neuropsychiatric disorders and perhaps the one most studied and best understood in the neuropsychiatric literature. The familial and phenomenologic links to obsessive-compulsive disorder (OCD) have led many investigators to conceptualize tics in TS as “movement-equivalents” of obsessions and compulsions, and the apparent connections with OCD and attention-deficit/hyperactivity disorder (ADHD) raise hope that by solving the TS “model,” we will understand a family of disorders that collectively affects close to 10% of the population. By all accounts, the TS model should be readily solvable, like a practice question before the really tough questions on an examination. Although we still lack clear answers for many of the complex questions raised by this syndrome, this chapter reviews the current state of progress in understanding the clinical features and neurobiology of TS and related tic disorders.

TICS: MOTOR, PHONIC, AND BEYOND

Part of "117 - Tourette Syndrome and Related Tic Disorders "

The DSM-IV describes tics as “sudden, rapid, recurrent, nonrhythmic, stereotyped movements or vocalizations,” but the self-assessments by Dr. Joseph Bliss (quoted earlier) and others (4) make it clear that tics in TS have a depth and dimension far beyond their motor or vocal components. Tics can be characterized by their anatomic location, frequency, intensity, and “complexity.” Facial and upper torso muscle groups are commonly active in motor tics, but virtually any motor group can be involved, including diaphragmatic muscles, in which rapid contractions can generate sounds by the expulsion of air through the upper airways. *Simple tics* are brief, circumscribed movements or sounds that are fragments or “chunks” of behavior or speech, rather than “self-contained,” meaningful motor sequences or utterances. These may include blinking, facial grimacing, mouth movements, head jerks, shoulder shrugs, and arm and leg jerks. *Complex tics* are more elaborate, sustained actions or linguistically meaningful sounds that often give the appearance of an intentional, “willful” event. Examples include facial gestures and movements such as brushing hair back, possibly in combination with head jerk, and body shrugs. Although most tics can be distinguished from chorea and dystonic movements, this differential diagnosis is not

always easily made, and some patients manifest symptoms across the entire range of involuntary movements. Audible tics may or may not involve the vocal cords, but true vocal tics can range from grunts and barking sounds to complete, grammatically correct phrases. Only about 10% of patients with TS express vocal tics with obscene content, termed *coprolalia* (5).

Tics can often be willfully suppressed for brief periods. Unfortunately, voluntary tic suppression can be associated with a buildup of inner tension, so when the tics are expressed, they are more forceful than they would otherwise be. Tics are also diminished during periods of goal-directed behavior that requires focused attention. Jim Eisenreich, who was a lifetime .300 hitter despite having significant motor tics for his entire professional baseball career, explained, “you only have to concentrate for 5 seconds on each pitch” (6). Tics can also be “suggestible,” activated by a verbal suggestion, or they can mimic or “echo” behavior or sounds from other people or the surrounding environment, analogous to stimulus-dependent behaviors in some posttraumatic or vascular orbitofrontal syndromes.

The degree of impairment associated with particular tics is partly dependent on their frequency, intensity, complexity, and duration. For example, a very frequent simple motor wrist tic may be less impairing than an infrequently occurring, forceful obscene (copropraxic) gesture. Very commonly, functional impairment in TS is strongly related to the severity of associated symptoms, including obsessions, compulsions, and attention deficits, as discussed later.

Many patients with tic disorders report a variety of sensory and mental states associated with their tics. Simple sensory tics, like simple motor or phonic tics, are rapid, recurrent, and stereotyped, and they are experienced as a sensation at or near the skin. The sensations are typically bothersome or uncomfortable, like an “itch” or a “crawling” feeling. Patients may be unusually aware, distracted, and distressed by particular sensory stimuli that most persons would not notice. One patient explained, “you know the scratchy feeling of a tag on your neck when you put on a new shirt? I have tags on every part of every shirt, all the time.” This site sensitization to certain forms of sensory information is a relatively understudied phenomenon that may provide important clues to the neurobiology of this disorder (7). *Premonitory urges* are more complex phenomena, which often include both sensory and psychic discomfort that may be momentarily relieved by a tic (7 ,8 and 9). An extension of the sensory-psychic dimension of tics may include a sense of discomfort or distress if sensory information (typically visual, but also tactile) is not experienced as “just right”; the assessment that something is “just right” can reflect complex stimulus properties, including balance and symmetry, texture, or context. The full elaboration of tics therefore can include sequential experience: (a) a sensory event or premonitory urge, (b) a complex state of inner conflict over whether and when to yield to the urge, (c) the motor or phonic production, and (d) a transient sensation of relief.

Further underscoring the importance of the sensory dimensions of TS, many patients with tic disorders are remarkably sensitive to perceptions arising both from within themselves (of somatic origin) and from the external world. Patients may unconsciously mirror the behavior and speech of others as well as of themselves. A related phenomenon is *triggering perceptions* in which some patients report urges to perform dangerous, forbidden, or simply senseless acts, such as to touch a hot iron, to jump from heights, to put the car in reverse gear while driving down a highway, or to shout in a quiet church service (10).

Diagnosis

The diagnosis of TS is based exclusively on the history obtained from the patient, parents, or other family members and on direct examination. Diagnostic criteria for “Tourette’s disorder” (DSM-IV) (11) and “definite Tourette syndrome” (TS Classification Study Group) (12) differ only slightly: both diagnoses require the frequent occurrence of multiple motor tics and one or more vocal tics, over a continuous interval that involves most of a full year, with the onset of symptoms early in life (before age 18 to 21 years). The controversial requirement that tics cause “marked distress or significant impairment in social, occupational, or other areas of functioning” has been challenged; as currently written in the DSM-IV, this criterion excludes persons who have adjusted well to the presence of tics, because these persons are not considered to have Tourette disorder if the syndrome is not a major source of distress.

The DSM-IV lists two specific tic disorders other than Tourette disorder. The diagnosis of *chronic motor or vocal tic disorder* is made when tics are limited to one or the other domain, but the patient otherwise meets criteria for Tourette disorder. Chronic motor tic disorder is the more common of these two conditions, and both are often viewed as part of the “broader phenotype” of TS. One clue that TS and related tic disorders reflect some aberration in a normal developmental process is that various tics are exhibited at some point in early development by most children. To bridge this gradient of “normal” versus “abnormal” tic behaviors in childhood, and to span the temporal gap between symptom onset and the 1-year “duration” requirement for the diagnosis of Tourette disorder, a diagnosis of *transient tic disorder* can be made if childhood tics, either motor or vocal, are frequent and cause distress, and they last between 1 and 12 months. As many as one in ten children may meet criteria for this diagnosis (13), and thus by extrapolation, in at most 10% of these children will symptoms continue beyond a year, thereby meeting criteria for one of the chronic tic disorders.

Comorbid Conditions

Although some persons experience “pure” tic disorders, functional impairment is often more directly related to the partial or full manifestation of comorbid conditions such as OCD and ADHD. Clinical and epidemiologic studies indicate that more than 40% of patients with TS experience recurrent obsessive-compulsive symptoms (8 ,14). Compared with “pure” TS, patients with comorbid TS and OCD experience more lifetime functional impairment, as rated by standardized scales, and specifically in areas such as employment and social relations. Overall, the level of lifetime impairment in these persons correlates significantly with the severity of OCD symptoms (15). Specific obsessive-compulsive symptoms associated with most impairment in patients with TS include aggressive and sexual obsessions and repeating and counting compulsions. Aggressive and sexual obsessions are also associated with more severe motor and phonic tics, even in patients with TS who do not meet full diagnostic criteria for OCD (15). These relationships underscore the importance of early recognition and treatment of comorbid obsessive-compulsive symptoms in TS.

Studies sharply differ on the rates of ADHD seen among patients with TS. Clinical studies vary according to setting and established referral patterns, but it is not uncommon to see reports of 50% or more of referred children with TS diagnosed with comorbid ADHD (16). In contrast, epidemiologic studies typically indicate a much lower incidence of comorbidity (14 ,16). Similar to the functional impact of comorbid OCD, patients with comorbid ADHD and TS experience significantly greater impairment than patients with “pure” TS. Longitudinal studies confirm that children with comorbid TS and ADHD are at high risk of anxiety and mood disorders, oppositional defiant disorder, and conduct disorder (17 ,18), whereas children who only have TS tend to fare better (17 ,19 ,20). Levels of tic severity are less predictive of peer acceptance than is the presence of ADHD (19 ,20), and rates of subsequent psychiatric morbidity in comorbid TS and ADHD are nearly identical to those seen in prior cross-sectional and longitudinal studies of “pure” ADHD (21 ,22).

Natural History and Epidemiology

Tics typically begin between 3 and 8 years of age. For persons who go on to develop TS, the tics typically follow a waxing and waning course, usually with a progressive pattern of tic worsening. On average, the period of greatest tic severity occurs between 8 and 12 years of age. The onset of puberty is not associated with either the timing or the severity of tics. The early teens are generally followed by a steady decline in tic severity, and by 18 years of age, perhaps as many as 50% of patients with TS are nearly tic free (23). Symptoms in adulthood may typically settle into a more predictable, yet idiosyncratic repertoire, with increases in tic frequency and forcefulness during periods of stress or emotional excitement.

Estimates of the prevalence of TS vary considerably across studies. In studies relying on identified treated cases, prevalence estimates were 0.046 (Minnesota) (24) to 5.2 in 10, 000 (North Dakota) (25). School-based surveys yielded much higher estimates, up to 23.4 in 10,000 (26). Community surveys yielded prevalence rates that range from 2.9 in 10,000 in Monroe County, NY (27) to 299 in 10,000 in the United Kingdom (16 ,28). Thus, although a population prevalence of 5 in 10,000 is commonly cited, considerable evidence suggests that the true prevalence of TS may be considerably greater than this.

NEUROBIOLOGY OF TOURETTE SYNDROME

Part of "117 - Tourette Syndrome and Related Tic Disorders "

Interconnected cortico-striato-pallido-thalamic (CSPT) circuitry has long been implicated in the regulation of movement, thought, and affect; abnormalities within this circuitry have also been proposed to contribute to the pathophysiology of many neuropsychiatric disorders (29 ,30). The “motor” loops of CSPT circuitry are known to be the locus of disease in primary movement disorders such as Huntington disease (31), Parkinson disease (32) and hemiballism (33), and the “limbic” CSPT loops have been proposed as the source of disease in schizophrenia (29), depression (29), OCD (34 ,35), ADHD (36), substance abuse disorders (37), temporal lobe epilepsy (38), and many other seemingly disparate forms of psychopathology (39). Important models have been proposed to account for a wide range of CSPT disorders, based on variations in responses to “generalized” or “epigenetic” early developmental insults, from neonatal hypoxia to bacterial infections. Despite this departure from the notion of discrete, specific “lesions” and circumscribed clinical presentations, the prevailing model in the search for the pathophysiology of TS is perhaps closest to that previously applied to Huntington disease, in which a unique mutation in a single gene causes—by as yet unknown cellular mechanisms—a characteristic disease that leads to a fairly predictable clinical presentation and course. Yet the actual pathophysiology of TS remains quite elusive. The scant tangible evidence of the pathophysiology of TS comes primarily from studies in neuropathology and neuroimaging; supportive evidence comes from other fields of investigation, including neuropsychology and psychophysiology, and from ties between TS and other disorders, such as OCD and ADHD, providing indirect evidence based on what is known about the pathophysiology of these other disorders.

The published literature of TS neuropathology studies now includes seven presumed TS cases; of these, informative clinical and histologic data are available from five cases. Interpretation of the findings from even these five cases is

clouded by issues of diagnostic uncertainty, comorbidity, and potentially confounding neurologic insults. Preliminary findings have identified four different locations of potential pathology within CSPT circuitry: (a) intrinsic striatal neuron abnormalities, including increased packing density of neurons in the striatum (n = 1) (40); (b) a diminished striatopallidal "direct" output pathway, with reduced dynorphin-like immunoreactivity in the lenticular nuclei (n = 5) (41 ,42); (c) increased dopaminergic innervation of the striatum, with increased density of dopamine transporter sites (n = 3);(43); and (d) reduced glutamatergic output from the subthalamic nucleus, based on reduced lenticular glutamate content (n = 4) (44). Thus, in a manner more reminiscent of neuropathologic findings in schizophrenia than, for example, Huntington disease, these preliminary neuropathologic findings in TS do not converge to identify a specific, circumscribed "hole" in CSPT connections, but instead they suggest a range of disturbances or imbalances that affect the "whole" circuit (45). Many other measures of CSPT biology in TS, including amine levels and receptors, have been reported to be normal, also in these preliminary, small studies. Clearly, postmortem studies are hampered by limitations in the nature and number of the brains that have been studied (46). Efforts by the Tourette Syndrome Association (TSA) to secure adequate material for neuropathologic studies are currently under way and should allow a new generation of tissue-based research during the next decade.

Neuroimaging findings may ultimately provide information critically important to our understanding of the pathophysiology of TS. Volumetric imaging studies demonstrate minimal, if any consistent, abnormalities in persons with TS. Among reports of enlarged corpus callosum volume (47), reduced caudate volume (48), or diminished right-to-left asymmetry for the caudate nucleus (49) and left-to-right asymmetry for putamen and lenticular nucleus (48), the magnitude of such changes is small, typically on the order of 5%; furthermore, some reports fail to support even these small abnormalities. The specific cellular or structural processes that may be responsible for these anatomic abnormalities are unknown. Concerns regarding sample heterogeneity, comorbidity, and effects of chronic medication exposure, described earlier in relation to neuropathologic studies, are equally applicable to neuroimaging studies in TS.

Neuroimaging studies of regional perfusion or glucose uptake presumably measure features indicative of neuronal metabolism. In general, these studies in TS report reduced glucose uptake in orbitofrontal cortex, caudate, parahippocampus, and midbrain regions (52), as well as reduced blood flow in the caudate nucleus, anterior cingulate cortex, and temporal lobes (53 ,54 ,55 and 56). Regional glucose uptake patterns may reflect distributed CSPT dysfunction, as suggested by the observed covariate relationships between reduced glucose uptake in striatal, pallidal, thalamic, and hippocampal regions (56). The single greatest consistency across metabolic imaging studies in TS—that of distributed hypometabolism—contrasts sharply with the observed corticostriatal hypermetabolism reported by many groups in patients with OCD (33 ,34). The only suggestion of regional activation in TS comes during active tic suppression, which is associated with increased right caudate neuronal activity, as measured by functional magnetic resonance imaging (fMRI) (57); however, tic suppression is also accompanied by bilaterally diminished neuronal activity on fMRI measures, in the putamen, globus pallidus, and thalamus. The most analogous paradigm in OCD—obsession provocation—is associated with increased metabolic activity at every level of CSPT circuitry (35), in sharp contrast to the pattern observed in TS.

Neurochemical imaging studies have reported relatively subtle abnormalities in levels of dopamine receptors (58), dopamine release (59), and dopamine transporter (60) in the striatum of some patients with TS. Some of these findings have not been replicated (61 ,62), others await replication, even internally (60), and others are evident only in a small subgroup of TS, such as four of 20 patients (58), issues raising concern about their generalizability to the pathophysiology of TS. A potentially important report of 17% greater caudate D2 receptor binding among more symptomatic TS identical twins (63) was based on five twin pairs and reached statistical significance at the $p < .04$ level, by nonparametric analyses. Significant correlations between symptom severity and D2 binding were obtained using aggregate symptom scores from three clinical measures. These latter findings do not directly address the brain mechanisms that distinguish persons with TS from those without TS, but rather, point to the need to understand specific factors that contribute to the heterogeneity of the TS phenotype among affected persons.

Various techniques have been used to demonstrate a wide array of abnormalities in the levels of many major neurotransmitters, precursors, metabolites, biogenic amines, and hormones in blood, cerebrospinal fluid and urine of patients with TS, compared with controls (64 ,65 and 66). Attempts to understand the relation of these abnormalities to the pathophysiology of TS have ranged from a proposed causal role ascribed to a single metabolic abnormality, such as the reported cerebrospinal fluid elevation of the potential excitotoxin kynurenine (67), to models for imbalances in norepinephrine, dopamine, and serotonin (5-HT) systems similar to "imbalance" models proposed for other complex forms of psychopathology. One qualitatively different finding, reported by Singer et al. (68), is that of approximately 40% elevations of serum antiputamen antibodies in children with TS. This finding may have particular importance, based not only on its magnitude and specific linkage to basal ganglia circuitry, but also on converging evidence of autoimmune contributions to at least some forms of TS (see later).

Additional evidence of disorders in CSPT circuitry in

TS comes from neuropsychological and psychophysiologic studies (69 ,70 ,71 ,72 ,73 ,74 ,75 ,76 and 77). Even these findings generally suggest mild deficits, at most: response distributions overlap greatly among patients with TS and control subjects, with most patients with TS performing within the normal range. The most consistently observed deficits occur on tasks requiring the accurate copy of geometric designs, that is, “visual-motor integration” or “visual-graphic” ability (69 ,77); somewhat similar deficits are reported in patients with OCD (78). No compelling evidence links these deficits in TS and OCD with a specific frontal or frontal corticostriatal territory, although visuospatial functions have generally been conceptualized to be regulated by dorsolateral prefrontal cortex and descending cerebrospinal fluid inputs to the head of the caudate nucleus (79). Neurophysiologic studies have documented a reduced cortical silent period after repeated *transcranial magnetic stimulation* (rTMS) in TS (80). This increased cortical excitability could result from impaired inhibition through disinhibited thalamocortical inputs or through abnormalities intrinsic to cortex, or both. Further support for the role of basal ganglia circuitry in TS comes from anecdotal reports of symptom exacerbation or reduction in patients with tumors within, or transections of, CSPT elements, respectively (81 ,82).

GENETICS OF TOURETTE SYNDROME

Part of "117 - Tourette Syndrome and Related Tic Disorders "

TS may be the most clearly inheritable common neuropsychiatric disorder. First-degree relatives of TS probands appear to be 20 to 150 times more likely to develop TS, compared with unrelated persons (83). Concordance rates for TS among monozygotic twins approach 90%, if the phenotypic boundaries include chronic motor or vocal tics, versus 10% to 25% concordance for dizygotic twins, across the same boundaries (84). The mode of inheritance remains elusive, even after more than 15 years of studies. Some segregation analyses have supported transmission through an incompletely penetrant autosomal dominant major locus (85 ,86), but in other studies, more complex models could not be excluded (87). The impact of assortive mating on inheritance may be particularly strong in TS, based on higher-than-predicted rates of bilineal transmission (88 ,89 and 90). Other findings suggest different inheritance patterns for maternal versus paternal transmission (91). Approaches using candidate genes or chromosomal translocations have offered results that were exciting but thus far not generally informative (92 ,93 and 94). Perhaps the most conservative assessment is that susceptibility to TS may be determined by a major gene in some families and by multiple genes of small relative effect in others, with a “dose-effect” of greater susceptibility for individuals homozygous versus heterozygous for these genes.

The TSA International Genetics Consortium completed the first genome-wide scan in an affected sibling-pairs (sib-pair) study. The *sib-pair design* relies on the comparison of the number of alleles at a given locus that are shared by two affected siblings, across all families in the sample. If the number of affected siblings sharing an allele or alleles is significantly higher than that expected by chance, it suggests a gene or genes of etiologic importance for TS. Using 76 affected sib-pair families with a total of 110 sib-pairs, the multipoint maximum-likelihood scores (MLS) for two regions (4q and 8p) were suggestive of high sharing (MLS greater than 2.0). Four additional regions also gave multipoint MLS scores between 1.0 and 2.0 (95). Collection of a second, replication set of approximately 100 sib-pairs is nearly complete, and it will be used to map these broad chromosomal regions more narrowly.

ENVIRONMENTAL FACTORS

Part of "117 - Tourette Syndrome and Related Tic Disorders "

Evidence of nongenetic environmental factors in the genesis of TS supports an interactive role for at least three sets of environmental factors: adverse prenatal and perinatal events, acute and chronic psychosocial stressors, and postinfectious autoimmune mechanisms.

Retrospective studies have identified an association between adverse events during the prenatal and perinatal period and an increased risk for the development of TS. Although the strongest evidence points to chronic mechanisms that influence the supply of nutrients by the placenta (96 ,97), other risk factors have been proposed including severe nausea and vomiting during the first trimester, severe maternal stress during pregnancy, exposure to high levels of androgenic steroids, and chronic or acute hypoxic and ischemic injury (98 ,99 ,100 and 101). Although no specific mechanism is known to connect these early life events and the development of TS, preclinical studies have shown that various neural insults during the prenatal and perinatal period result in the delayed emergence of pathology within interconnected CSPT circuitry and in specific behavioral abnormalities that are also manifested by individuals with TS, such as reductions in sensorimotor gating of the startle reflex (102 ,103 ,104 ,105 and 106). These early insults may also set the stage for a heightened stress response in adulthood and altered immune function (107 ,108 and 109).

As with several other disorders associated with CSPT dysfunction, including OCD, schizophrenia, and affective disorders, increased life stressors are associated with symptom exacerbation in TS. More than 98% of patients with TS report worsening of tic symptoms during periods of stress and anxiety (110). A direct assessment of the relationship between stress and tics revealed that anticipation of a stressful medical procedure, a lumbar puncture, has been shown to produce greater elevations in plasma adrenocorticotropic hormone in patients with TS than in control subjects and an elevated urinary excretion of catecholamines that correlated with symptom severity (66). Patients with

TS also have been reported to have elevated levels of cerebrospinal fluid norepinephrine and corticotropin-releasing factor (66 ,111). Importantly, although stress clearly alters CSPT dynamics and increases symptoms of numerous different neuropsychiatric disorders, no existing data implicate a specific etiologic relationship between stress and TS.

It is well known that *group A ß-hemolytic streptococcal (GABHS) infections* can trigger immune-mediated disease in genetically predisposed persons (112). Acute rheumatic fever (RF) can occur approximately 3 weeks after an inadequately treated GABHS infection. In addition to inflammatory lesions involving the heart (rheumatic carditis) and joints (polymigratory arthritis), rheumatic fever can be accompanied by CSPT disease responsible for the manifestation of Sydenham chorea. Patients with Sydenham chorea frequently display motor and vocal tics, obsessive-compulsive, and ADHD symptoms suggesting the possibility that at least in some instances these disorders share a common origin (113 ,114). As in Sydenham chorea, antineuronal antibodies have been reported to be elevated in the sera of patients with TS (68).

It has been proposed that *pediatric autoimmune neuropsychiatric disorder associated with streptococcal infection* (PANDAS) represents a distinct clinical entity and includes Sydenham chorea and some cases of TS and OCD (115). The most compelling evidence that acute exacerbations of TS and OCD can be triggered by GABHS comes from two independent reports that most patients with childhood-onset TS or OCD have elevated expression of a stable B-cell marker (116 ,117). The D8/17 marker identifies close to 100% of patients with rheumatic fever (with or without Sydenham chorea), but it is present at low levels of expression in healthy control populations. Dr. Susan Swedo and her colleagues reported that in children who met PANDAS criteria, GABHS infection was likely to have preceded neuropsychiatric symptom onset for 44% of the children, whereas pharyngitis (no culture obtained) preceded onset for another 28% of the children. In a minority of cases (31%), neuropsychiatric symptom exacerbations were associated with documented GABHS infection, and in another 42%, they were associated with symptoms of pharyngitis or upper respiratory infection (no throat culture obtained) (115). Although these results are intriguing, they are not compelling with regard to specific immunologic mechanisms linking PANDAS and TS. Clearly, independent replication and systematic study of this intriguing phenomenon may provide a basis for the rational design of therapeutic and preventative interventions. The need for accurate and complete information in this area is underscored by the finding that, based solely on the existing minimal data, parents are actively seeking for their children invasive treatments such as plasmapheresis and intravenous immunoglobulins. With emerging preliminary findings suggesting possible links between streptococcal infections and some aspects of OCD and ADHD, this area of investigation has become a major public health issue.

TREATMENT

Part of "117 - Tourette Syndrome and Related Tic Disorders "

Therapeutic models of TS emphasize the importance of flexible, integrated biopsychosocial strategies. Flexibility is important because the nature of the disorder and its impact on patients and families change dramatically across its course. It is often the case that, when families first present for assessment of TS, confusion, fear, anger, and embarrassment fill an “information void” and are exacerbated by the very public outward manifestations of tics, their deceptive “volitional” appearance, and their sometimes socially unacceptable content. Although parents may worry about a child’s counting rituals or may be exasperated by a child’s continued disruption of a school class, the child’s tics often evoke in them a more intense, visceral sense of desperation. Such reactions reverberate throughout the family and affect the child. The familial nature of the illness means that, almost invariably, when a child first manifests symptoms, close relatives (often parents or siblings) are, or were once, also affected; this sets the stage for a range of “generational” psychological consequences for parents, as painful memories are rekindled.

Education

Initially, much of the distress associated with TS can result from a lack of understanding of the illness. Education about the natural history of TS, emphasizing the involuntary, “no-fault” nature of certain brain-behavior relationships, is an essential part of the early treatment of this disorder. This process can begin in the diagnostic assessment: faced with a set of simple, matter-of-fact questions about tic symptoms, many of which are found in printed, standardized scales such as the Yale Global Tic Severity Scale (118), parents and patients recognize that other people must have had experiences much like their own. As information displaces the “mystery” of a child’s tic behaviors, the urgency to make immediate somatic interventions can diminish.

Physicians, parents, and patients should be aware of the waxing and waning nature of the illness. An initial clinical visit may be precipitated by a recent exacerbation of previously subclinical or tolerable symptoms. Given the cyclic pattern of TS, such periods are often followed by a gradual diminution of symptoms, even in the absence of a specific biological intervention. One danger of rapidly initiating medication treatment in TS is that a “false-positive” response, based on the normal cyclic fluctuation of symptoms, will convince patient, family, or physician that a particular medication is “effective.” This false-positive response, and the false hope that it creates, can result in unnecessary medication exposure and side effects and, ultimately, in heightened

frustration when the illness follows its natural course toward the next phase of exacerbation. Observing the cyclic fluctuations, while the patient is free of medications, can provide a more clear impression of a patients' natural illness course and can thereby be a useful basis for interpreting future medication effects.

The "bigger picture" of the natural history of TS is also critically important: persons with TS can and should be expected to live full, productive lives; as many as half of these persons will be largely symptom free by the time they enter their twenties; and every persons has strengths that must be nurtured and developed and that will ultimately be more significant determinants of life quality and character than are tics or other TS symptoms. Within this broader context, parents and patients should understand that, at present, the benefits of medication treatments of TS are relatively modest, and the potential social, psychological, and biological side effects are not trivial.

Pharmacotherapy

Medications can play an important role in the treatment of TS. Because functional impairment in this disorder is most closely linked to comorbid conditions such as OCD and ADHD, symptoms of these disorders (and even their "subclinical" manifestations) are often the first targets of pharmacotherapy in TS. Effective treatment of these comorbid conditions can often markedly diminish tic severity. The basis for this therapeutic interaction is not well understood, but it may include some or all of the following: (a) simple interactions at a neural level, such as a direct medication effect within multiple or interacting CSPT "open loops," including, for example, those that mediate both obsessions and tics; (b) interactions related to the stressful nature of comorbid conditions; for example, tics are reduced by a diminution of stress that follows successful treatment of OCD or ADHD symptoms; (c) interactions at a level of cognition resources, such as improved volitional "suppressibility" of tics because of improved attentional allocation; and (d) interactions at a symptom level, such as diminished need for repetition or complex rituals that could otherwise make tics more elaborate. Regardless of the mechanism, treatment of comorbid OCD or ADHD (as well as mood disorders, discussed later) should be a high priority because these conditions are responsible for significant functional impairment, they are generally responsive to appropriate pharmacotherapy, and their treatments are relatively free of significant side effects. Even the prolonged use of stimulants in comorbid TS and ADHD, once avoided because of fears of stimulant-potentiation of tics, was shown to be safe and effective in a large TSA-funded study with a 2-year longitudinal design (119). Clearly, close clinical monitoring is important in all pharmacotherapy, particularly in children.

Dopamine Antagonists

When tic-suppressing agents are necessary, the cost-to-benefit ratio differs among medications and across clinical conditions. *Dopamine antagonists*, particularly high potency, D2-preferential blockers such as haloperidol, fluphenazine, and pimozide, are the most potent and rapid-acting tic-suppressing agents that have been studied in controlled trials. These medications may be most useful in patients with severe, intractable tics, but they also have undesired side effects, causing blunting of cognitive skills, mood, and motivation (120 and 121); when discontinued, these high potency D2 blockers can precipitate withdrawal dyskinesia and significant worsening of tics (120). In adults, these drugs are clearly linked to an increased risk of tardive dyskinesia, although in children, this relationship has not been as clearly defined. One newer, "atypical" antipsychotic, risperidone, is a mixed dopamine/5-HT receptor blocker that is proving to be a useful anti-tic medication, with a side effect profile somewhat preferable to that of haloperidol or pimozide (122). Controlled studies of risperidone efficacy in TS are in progress; however, experience suggests that significant, undesired weight gain with this drug is not infrequent. Dopamine antagonists are also useful in conjunction with a primary antiobsessional agent (e.g., selective serotonin reuptake inhibitors) in treatment refractory OCD with comorbid tics; risperidone is frequently used in this capacity. A double-blind trial of ziprasidone in TS yielded encouraging short-term results without the weight gain associated with risperidone (123). The utility of other "atypical" antipsychotics such as olanzapine or quetiapine in the treatment of TS or tic-related OCD is not known.

α_2-Agonists

The α_2-adrenergic agonists, clonidine and guanfacine, are often used as first-line anti-tic agents, both because of their relatively favorable side effect profile and because of some evidence linking these drugs to improved attentional abilities in children with ADHD (124 ,125). These drugs have relatively weaker anti-tic abilities, compared with dopamine antagonists, and their benefit generally evolves more gradually than with dopamine antagonists. Still, there are advantages to prescribing these medications in lieu of antipsychotics; aside from the somatic side effects of neuroleptic agents, it is often not inconsequential for a child when his or her teacher researches these drugs in the *Physician's Desk Reference* and reads "for severe psychotic states only."

Newer Pharmacologic Approaches

Several new therapeutic avenues for TS are being explored in controlled studies. Preliminary studies suggest that one severely impairing feature of some forms of TS—rage attacks—may be sensitive to treatment with selective serotonin

reuptake inhibitors (126). One double-blind trial suggested some anti-tic benefit from low doses of dopamine agonists such as pergolide (127), an effect attributed to “autoreceptor” actions that suppress activity in midbrain dopamine nuclei; other dopamine agonists are also being explored in this capacity. Nicotinic manipulations, ranging from nicotine patches to the nicotinic antagonist mecamylamine, may offer significant anti-tic benefit, if controlled studies replicate the impressive series of case reports with these agents (128 ,129 and 130). Tic reduction after nicotine patches has been reported to be sustained for several weeks after exposure to the patches for only a few hours. The seemingly paradoxical effectiveness of both nicotine and mecamylamine may suggest that nicotine’s beneficial effects in TS reflect a rapid and sustained desensitization of nicotine receptors. Studies in adults suggest that sustained use of the nicotine patch does not increase the liability for nicotine abuse or dependence (131), but relevant studies in children, with intermittent brief patch exposure, have not been attempted. Δ^9-tetrahydrocannabinol was also reported to diminish tic severity in TS, in a large case series (132); controlled studies with this agent are in progress. Obviously, such putative therapeutic effects may provide clues relevant to the pathophysiology of TS. Various other agents, ranging from opiate antagonists to androgen antagonists, have been reported to have beneficial effects on tic symptoms; generally, these effects have been modest (133) or accompanied by a worrisome liability for side effects (134).

Certain intractable and localized tics have also been treated successfully with repeated injections of botulinum toxin. Generally, tic location and type shift across the course of an illness, so such a “peripheral” approach may be effective only for short periods of an illness. Still, substantial case series evidence suggests the utility of this approach for some patients, and even vocal tics have been treated effectively with botulinum toxin (135). At another extreme, *habit reversal therapy* involves the application of cognitive and behavioral therapy principles to TS (136), analogous to the successful use of these therapies in the treatment of OCD. Although studies with habit reversal therapy are only now being completed, initial results appear promising, offering the possibility of a true, “nonpharmacologic” approach to this disorder. Conceptually, once it is effectively learned, habit reversal therapy may be easy to apply over a period of years, to a variety of different types of tics.

One novel approach being studied in TS, as well as OCD and depression, involves the therapeutic use of rTMS. Although clearly at an investigational phase, technical advances in rTMS may ultimately allow localized, stereotaxic-guided activation of frontal and striatal CSPT elements implicated in the pathophysiology of TS. Proponents of rTMS emphasize that the procedure can be administered on an outpatient basis, at relatively low cost, without apparent significant side effects. More heroic efforts for intractable TS and OCD have included psychosurgical interventions, particularly anterior cingulotomy or capsulotomy (80 ,81 and 82). Even this more extreme approach is now being accomplished on an outpatient basis, through the use of a gamma knife. However, the most promising psychosurgical approach may involve high-frequency electrical stimulation after the placement of deep brain electrodes (137).

Alternative Therapies

The void created by a lack of fully effective medication treatments for TS has been filled by *alternative therapies*, as parents and patients use self-experimentation to identify treatment strategies that work for themselves and then share this information with others in the TS community. The dissemination of this information has been greatly facilitated by internet communications. Proponents of alternative approaches, which generally include nutritional and vitamin supplementation, believe that these interventions are natural, safe, and effective and may be targeting a metabolic or biochemical defect that underlies TS, but that has thus far eluded detection by scientific investigations. Just as for conventional therapies, symptom reduction with alternative treatments must be interpreted in the context of the natural waxing and waning course of TS. Perhaps even more so than for conventional therapies, these alternative approaches carry a potential for improvement based on expectation or “placebo” effects. For an individual patient, a placebo effect is not necessarily problematic, but tremendous cost and hardship are created if such treatments are applied on a larger clinical scale and ultimately fail to deliver the promised therapeutic benefit. Nutritional or related interventions may be “natural,” but they ultimately cause biochemical changes in the body and can impart significant side effects. Other “side effects” occur when patients substitute these alternative approaches for treatments that are known to offer some efficacy and that keep them in contact with clinicians who can monitor their overall medical condition. Ultimately, it is critically important for these alternative treatments to be tested in controlled trials, in which their tolerability, safety, and efficacy can be established in an objective manner. Unfortunately, the cost of such studies is often prohibitive, and they are generally not a high priority for funding from the pharmaceutical industry. Clinicians should ask patients and parents whether alternative therapies are being used, particularly if medications are being added or changed, to avoid potentially harmful drug–drug interactions.

Treating Tourette Syndrome and Depression

As is true with most neuropsychiatric conditions, comorbidity of TS with affective disorders deserves special attention, owing to the insidious and often profound morbidity imparted by depression. Comorbid affective illness accompanying TS is generally sensitive to standard pharmacotherapies

for these disorders. These can often be used in combination with anti-tic regimens, but the possibility of iatrogenic depression from dopamine antagonists should always be considered, because this may dictate a reduction in neuroleptic dose rather than the addition of an antidepressant. As with comorbid OCD or ADHD, it appears that *comorbid depression* can often cause a worsening of tics in TS, and there are reports of severe, refractory, mood-dependent tics in comorbid TS and depression that show dramatic sensitivity to electroconvulsive therapy (138 ,139). Whereas more representative epidemiologic data are necessary, the lifetime prevalence of comorbid mood disorders in TS patients seen in specialty clinics may be as high as 70%, comparable to that reported in patients with OCD (140).

FUTURE DIRECTIONS

Part of "117 - Tourette Syndrome and Related Tic Disorders "

Several lines of inquiry are positioned to make major advances in our understanding of the etiology and treatment of TS, based on the tremendous progress that has already been made in each of these areas:

The *clinical "phenotype"* of TS has been particularly well characterized. However, compared with motor and phonic tics, sensory phenomena in TS are relatively less well understood and are more difficult to study. A full understanding of the TS phenotype will clearly enhance research efforts, by permitting stratification in measures from all levels of analysis, from genetics to neuropsychology. Detailed clinical characterization, and the experimental analysis that it facilitates, will also be important in clarifying the potentially critical distinction between the form of TS that undergoes substantial remission by the early twenties and that which is lifelong and unremitting. This issue has tremendous importance, because these two "types" of TS are not generally (and often cannot be) segregated in genetic, neuroimaging, or other biological measures in this disorder. For example, it may not be appropriate to generalize to all forms of TS the biology that is described by neurochemical brain imaging studies, which, because of ethical concerns, exclusively involve adults with TS. At another level, current TS studies that include children will likely involve both "types" of TS, without any clear way to stratify this heterogeneous sample. A full understanding of the TS phenotype may also be of value in developing symptom "clusters" or "factors," like those that provide meaningful segregating strategies in OCD.

Future efforts in *TS neuropathology* will result in the collection of an adequate number of well-phenotyped, optimally prepared brains, to permit meaningful studies. Through the efforts of the TSA Neuropathology Program and the Harvard Brain Tissue Resource Center, approximately three new, optimal TS brains are collected each year, and in 1998, eight TS brains were cut for neurochemical and neuroanatomic studies. A common "library" of matched control tissue is also being collected, to diminish variability across different studies. These systematic efforts, combined with the judicious but timely distribution of brain tissue for studies, should help to overcome many of the limitations of past TS neuropathologic studies. Our ability to utilize this material effectively and to interpret the information that it provides depends entirely on our progress in understanding the complex interconnections within CSPT circuitry. Although we have useful maps of the major thoroughfares within CSPT circuitry, we will need to be equally knowledgeable regarding the detailed input-output relationships of functionally and neurochemically distinct striatal subterritories. Important pathologic findings can be lost if they are unknowingly averaged across heterogeneous anatomic domains.

Neuroimaging efforts in TS will focus on two strategies that have been so informative in studies of OCD: pretreatment versus posttreatment repeated measures and on-line behavioral or psychophysiologic probes in conjunction with functional imaging. Nonpharmacologic treatment effects on regional brain metabolism or brain activation may be studied before and after habit reversal therapy, similar to such studies using cognitive and behavior therapy in OCD (141). Appropriate on-line probes for fMRI studies must be carefully developed. Optimally, these probes should either (a) detect deficits in patients with TS, or (b) in healthy persons, selectively activate relevant brain substrates, including the basal ganglia or frontal circuitries. Event-related fMRI techniques appear especially promising as investigators seek to identify the sequence of neural events that precede and follow tics.

Replication of the initial *genome scan* from the TSA International Sib-Pair Study is in progress, and fine mapping of the most promising regions is planned. Verification and extension of this work within several extended TS families are also anticipated. Parallel efforts will continue in targets of opportunity, including informative chromosomal translocations. Ultimately, genes conferring a risk of TS will be identified and cloned. Experience from similar efforts that are already completed with the Huntington disease gene suggest that identification of the TS genes will be followed by a substantial amount of work designed to understand their normal function and the pathologic impact of their abnormal products.

Even with the most detailed description of the "clinical phenotype" of TS, it is possible that a full understanding of the functional relevance of the "TS genes" will require the *use of endophenotypes*—physiologic or neuropsychological markers of gene function that can be identified in patients with TS and that could also be abnormal in unaffected, first-degree relatives. These markers could greatly enhance the power of linkage analyses, by allowing a "physiologic" parsing of the phenotype of affected and at-risk individuals. A variety of such markers has been explored in TS, including antisaccade measures (74), eyeblink measures of prepulse

inhibition of startle (73) or condition-test paired pulse paradigms (76), and measures of cortical silent periods after rTMS (80). The neural bases of these various measures are consistent with our present conceptualization of CSPT substrates of TS, but the effect sizes of existing measures are small or moderate and thus are suboptimal for use in genetic analyses. This area of work in TS, however, has been relatively understudied.

Until recently, TS research suffered from the relative lack of informative *animal models*. Mutant rodent models may prove useful (142), particularly in understanding the behavioral and physiologic consequences of the selective loss of a specific neural element, such as the dopamine transporter (143). Essentially, such models provide a nonpharmacologic, nonsurgical means of studying the effects of a chronic perturbation in basal ganglia circuitry. Animal models can also be used, however, in guiding the development of candidate endophenotypes; particularly valuable would be measures that can be assessed across species and that can be shown in preclinical studies to have predictive or construct validity for TS.

One animal model that promises to be particularly informative regarding the neurobiology of TS involves the manipulation of the relative activity of medium spiny projection neurons within the striosomal and matrix compartments of the striatum (144). These two compartments differ with respect to their cortical inputs: striosomal neurons receive limbic and prelimbic inputs, and neurons in the matrix receive inputs from ipsilateral primary motor and sensory motor cortices. Dopamine projection neurons from the pars compacta of the substantia nigra serve to “tune” this system as it determines responsivity to certain cues (interoceptive or exteroceptive). As noted by Leckman and Riddle (145), this model may provide a meaningful integration of knowledge drawn from different perspectives that may be directly relevant to certain important clinical issues, including (a) the stress responsiveness of tics, (b) the presence of premonitory sensory urges and “just-right” perceptions, (c) the need to “even-up” sensory and motor stimuli in a bilaterally symmetrical fashion, (d) the reduction of tics when a person is engaged in acts that require selective attention and guided motor action, and (e) the timing of tics and the course of tic disorders. Such a model may also guide hypothesis-driven studies in other areas of TS research. For example, based on the distinct input-output characteristics of neurons in the striatal matrix versus striosomes, it may be useful to determine whether there is a differential involvement of these neurons in Sydenham chorea and postinfectious forms of TS. Even more subtle clinical distinctions within the “TS spectrum” may be predicted to reflect a differential involvement of these striatal components, with more affectively laden symptoms common to “tic-related OCD” reflecting the involvement of striosomal elements and more affectively “neutral” motor and sensory symptoms of “pure” TS reflecting the involvement of the striatal matrix (15).

Although several new therapeutic approaches to TS are being developed, as discussed earlier, there is clearly a significant need to *understand the proposed role of streptococcal infections in the pathogenesis of TS* and to assess the potential role of antistreptococcal or immunosuppressive therapies. Penicillin prophylaxis is gaining increasing use in patients with childhood tic disorders and OCD, without any controlled studies demonstrating efficacy for this approach; an analogous patterns has been seen in the recent rush for unproven, experimental treatments for autism (146). Certainly, repeated injections or even oral treatment with antibiotics can have detrimental consequences in children. Even more concerning is the increasing use of plasmapheresis or intravenous immunoglobulin therapies in affected children, with only preliminary data from controlled studies demonstrating efficacy for these costly and invasive interventions (147). Given the expanding clinical boundaries of streptococcal-associated conditions, controlled, large sample studies of penicillin prophylaxis in TS, and perhaps of other related therapies as well, should be a major priority from a public health standpoint.

SCIENTIFIC HURDLES

Part of "117 - Tourette Syndrome and Related Tic Disorders "

Major obstacles impede the search for the pathophysiology of TS. Serious ethical issues complicate invasive studies in children and affect TS research tools ranging from the use of radioisotopes in neuroimaging studies to the acquisition of meaningful age-matched control samples. The longevity of patients with TS is certainly a blessing, but it also precludes the availability of neuropathologic material for systematic studies. The lifelong interval between diagnosis and study increases the likelihood that the “lesion” will melt into the compensatory milieu of nervous tissue or will be camouflaged by the many other insults that befall an aging brain. The medical care of patients with TS is rarely focused on “end-of- the-life-spectrum” issues, and thus pediatricians and families rarely consider procedures for brain donation. Indeed, the optimal therapeutic approach to TS is often to downplay the notion of TS as a “disorder,” to not “pathologize,” but rather to emphasize the ways to survive and thrive with TS. Neuropathologic material is thus most readily available from persons whose TS has remained severe throughout adulthood and who are identified by family and physicians as being particularly impaired by this disorder. As noted earlier, refractory TS may reflect pathologic features quite distinct from the more common forms of this illness, and thus neuropathologic studies with tissue acquired from elderly patients with TS may not be generally informative about TS and potentially could be scientifically misleading.

Despite these hurdles, with centralized research coordination by the TSA, and increased attention to critical issues of clinical heterogeneity and comorbidity, future studies will

be much better equipped to overcome the obstacles that have left us, to date, with wide gaps in our understanding of the pathophysiology of TS. The formidable hurdles facing TS research are not unique to this disorder, but are equally relevant to conditions that span the range from primary movement disorders (148) to disorders with primary psychiatric manifestations (149). In this manner, the quest to understand the neurobiology of TS does serve as a practice question for more complex disorders, and the strategies developed to overcome challenging scientific hurdles will be invaluable lessons, with applications across a wide range of neuropsychiatric disorders.

ACKNOWLEDGMENTS

Part of "117 - Tourette Syndrome and Related Tic Disorders "

Topics covered in this chapter have been reviewed in previous reports by the authors. NRS was supported by MH01436. Portions of the research described in this review were supported by grants to JFL from the National Institutes of Health: MH18268, MH49351, HD03008, NS16648, and RR00125, as well as by the concerted and visionary leadership of the Tourette Syndrome Association.

REFERENCES

1. Bliss J. Sensory experiences of Gilles de la Tourette syndrome. *Arch Gen Psychiatry* 1980;37:1343–1347.

2. Cohen DJ, Leckman JF, Pauls D. Neuropsychiatric disorders of childhood: Tourette's syndrome as a model. *Acta Paediatr* 1997;422:106–111.

3. Hyde TM, Weinberger DR. Tourette's syndrome: a model neuropsychiatric disorder. *JAMA* 1995;273:498–501.

4. Hollenbeck P. How life imitates Tourette syndrome: reflections of an afflicted neuroscientist. *CNS Spectrums* 1998;4:22–23.

5. Erenberg G, Cruse RP, Rothner AD. Tourette syndrome: an analysis of 200 pediatric and adolescent cases. *Cleve Clin Q* 1986;53:127–131.

6. Fuerst ML. An interview with Jim Eisenreich. *CNS Spectrums* 1999;4:18.

7. Cohen AJ, Leckman JF. Sensory phenomena associated with Gilles de la Tourette's syndrome. *J Clin Psychiatry* 1992;5:319–323.

8. Leckman JF, Walker DE, Cohen DJ. Premonitory urges in Tourette's syndrome. *Am J Psychiatry* 1993;150:98–102.

9. Leckman JF, Walker DE, Goodman WK, et al. "Just right" perceptions associated with compulsive behavior in Tourette's syndrome. *Am J Psychiatry* 1994;151:675–680.

10. Leckman JF, Cohen DJ, Goetz CG, et al. Tourette syndrome: pieces of the puzzle. In: Cohen DJ, Goetz C, Jankovic J, eds. *Tourette syndrome and associated disorders*. New York: Lippincott Williams & Wilkins, 2000.

11. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington, DC: American Psychiatric Association, 1994.

12. Tourette Syndrome Classification Study Group. Definitions and classification of tic disorders. *Arch Neurol* 1993;50:1013–1016.

13. Shapiro AK, Shapiro ES, Braun RD, et al. *Gilles de la Tourette syndrome*. New York: Raven, 1978.

14. Apter A, Pauls DL, Bleich A, et al. An epidemiological study of Gilles de la Tourette's syndrome in Israel. *Arch Gen Psychiatry* 1993;50:734–738.

15. Swerdlow NR, Zinner S, Farber RH, et al. Symptoms in obsessive compulsive disorder and Tourette syndrome: a spectrum? *CNS Spectrums* 1999;4:21–33.

16. Kadesjo B, Gillberg C. Tourette's disorder: epidemiology and comorbidity in primary school children. *J Am Acad Child Adolesc Psychiatry* 2000;39:548–555.

17. Carter AS, O'Donnell DA, Schultz RT, et al. Social and emotional adjustment in children affected with Gilles de la Tourette's syndrome: associations with ADHD and family functioning. attention deficit hyperactivity disorder. *J Child Psychol Psychiatry* 2000;41:215–223.

18. Peterson BS, Pine DS, Cohen P, et al. A prospective longitudinal study of tic, obsessive-compulsive and attention deficit-hyperactivity disorders in an epidemiological sample. *J Am Acad Child Adolesc Psychiatry* 2001;40:685–695.

19. Bawden HN, Stokes A, Camfield CS, et al. Peer relationship problems in children with Tourette's disorder or diabetes mellitus. *J Child Psychol Psychiatry* 1998;39:663–668.

20. Carter AS, Pauls DL, Leckman JF, et al. A prospective longitudinal study of Gilles de la Tourette's syndrome. *J Am Acad Child Adolesc Psychiatry* 1994;33:377–385.

21. Greene RW, Biederman J, Faraone SV, et al. Adolescent outcome of boys with attention-deficit/hyperactivity disorder and social disability: results from a 4-year longitudinal follow-up study. *J Consult Clin Psychol* 1997;65:758–767.

22. Mannuzza S, Klein RG, Bessler A, et al. Adult psychiatric status of hyperactive boys grown up. *Am J Psychiatry* 1998;155:493–498.

23. Leckman JF, Zhang H, Vitale A, et al. Course of tic severity in Tourette syndrome: the first two decades. *Pediatrics* 1998;102:14–19.

24. Lucas AR, Beard CM, Rajput AH, et al. Tourette syndrome in Rochester, Minnesota, 1968–1979. *Adv Neurol* 1982;35:267–269.

25. Burd L, Kerbeshian L, Wikenheiser M, et al. A prevalence study of Gilles de la Tourette's syndrome in North Dakota school-aged children. *J Am Acad Child Adolesc Psychiatry* 1986;25:552–553.

26. Debray-Ritzen P, Dubois H. Simple tic disease in children: a report on 93 cases. *Rev Neurol (Paris)* 1980;136:15–18.

27. Caine ED, McBride MC, Chiverton P, et al. Tourette's syndrome in Monroe County school children. *Neurology* 1988;38:472–475.

28. Mason A, Banerjee S, Eapen V, et al. The prevalence of Tourette syndrome in a mainstream school population. *Dev Med Child Neurol* 1998;40:292–296.

29. Swerdlow NR, Koob GF. Dopamine, schizophrenia, mania and depression: toward a unified hypothesis of cortico-striato-pallido-thalamic function. *Behav Brain Sci* 1987;10:197–245.

30. Penny JB, Young AB. Speculations on the functional anatomy of basal ganglia disorders. *Annu Rev Neurosci* 1983;6:73–94.

31. Reiner A, Albin RL, Anderson KD, et al. Differential loss of striatal projection neurons in Huntington disease. *Proc Natl Acad Sci USA* 1988;85:5733–5737.

32. Javoy-Agid F, Ruberg M, Taquet H, et al. Biochemical neuropathology of Parkinson's disease. *Adv Neurol* 1984;40:189–198.

33. Marsden CD. Motor disorders in basal ganglia disease. *Hum Neurobiol* 1984;2:245–255.

34. Baxter LR, Schwartz JM, Bergman, et al. Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:681–689.

35. Rauch SL, Jenike MA, Alpert NM, et al. Regional cerebral blood flow measured during symptom provocation in obsessive-compulsive disorder using oxygen 15–labeled carbon dioxide and positron emission tomography. *Arch Gen Psychiatry* 1994;51:62–70.

36. Castellanos FX. Toward a pathophysiology of attention-deficit/hyperactivity disorder. *Clin Pediatr* 1997;36:381–393.

37. Modell JG, Mountz JM, Beresford TP. Basal ganglia/limbic striatal and thalamocortical involvement in craving and loss of control in alcoholism. *J Neuropsychiatry Clin Neurosci* 1990;2:123–144.

38. Van Paesschen W, Revesv T, Duncan JS, et al. Quantitative neuropathology and quantitative magnetic resonance imaging of the hippocampus in temporal lobe epilepsy. *Ann Neurol* 1997;42:756–766.

39. Swerdlow NR. Cortico-striatal substrates of cognitive, motor and sensory gating: Speculations and implications for psychological function and dysfunction. In: Panksepp J, ed. *Advances in biological psychiatry*, vol 2. Greenwich, CT: JAI, 1996;179–208.

40. Balthasar K. Uber das anatomische Substrat der geralisierten Tic-Krankheit (maladie des tics, Gilles de la Tourette): Entwicklungshemmung des Corpus striatum. *Archiv Psychiatr Nervenkr (Berl)* 1957;195:531–549.

41. Haber SN, Kowall NW, Vonsattel JP, et al. Gilles de la Tourette's syndrome: a postmortem neuropathological and immunohistochemical study. *J Neurol Sci* 1986;75:225–241.

42. Haber SN, Wolfer D. Basal ganglia peptidergic staining in Tourette syndrome: a follow-up study. *Adv Neurol* 1992;48:145–150.

43. Singer HS, Hahn IH, Moran TH. Abnormal dopamine uptake sites in postmortem striatum from patients with Tourette's syndrome. *Ann Neurol* 1991;30:558–562.

44. Anderson GM, Pollak ES, Chatterjee D, et al. Postmortem analysis of subcortical monoamines and amino acids in Tourette syndrome. *Adv Neurol* 1992;58:123–133.

45. Swerdlow NR. Neuropsychology of schizophrenia: the "hole" thing is wrong. *Behav Brain Sci* 1991;14:51–53.

46. Swerdlow NR, Young AB. Neuropathology in Tourette syndrome. *CNS Spectrums* 1999;4:65–74.

47. Baumgardner TL, Singer HS, Denckla MG, et al. Corpus callosum morphology in children with Tourette syndrome and attention deficit hyperactivity disorder. *Neurology* 1996;47:477–482.

48. Peterson B, Riddle MA, Cohen DJ, et al. Reduced basal ganglia volumes in Tourette's syndrome using three-dimensional reconstruction techniques from magnetic resonance images. *Neurology* 1993;43:941–949.

49. Singer HS, Reiss AL, Brown JE, et al. Volumetric MRI changes in basal ganglia of children with Tourette's syndrome. *Neurology* 1993;43:950–956.

50. Moriarty J, Varma AR, Stevens J, et al. A volumetric MRI study of Gilles de la Tourette's syndrome. *Neurology* 1997;49:410–415.

51. Castellanos FX, Giedd JN, Hamburger SD, et al. Brain morphometry in Tourette's syndrome: the influence of comorbid attention-deficit/hyperactivity disorder. *Neurology* 1996;47:1581–1583.

52. Braun AR, Stoetter B, Randolph C, et al. The functional neuroanatomy of Tourette's syndrome: an FDG-PET study. I. Regional changes in cerebral glucose metabolism differentiating patients and controls. *Neuropsychopharmacology* 1993;9:277–291.

53. Moriarty J, Costa DC, Schmitz B, et al. Brain perfusion abnormalities in Gilles de la Tourette's syndrome. *Br J Psychiatry* 1995;167:249–254.

54. Sieg KG, Buckingham D, Gaffney GR, et al. Tc-99m HMPAO SPECT brain imaging of Gilles de la Tourette's syndrome. *Clin Nucl Med* 1993:18:255.

55. Riddle MA, Rasmusson AM, Woods SW, et al. SPECT imaging of cerebral bloodflow in Tourette syndrome. *Adv Neurol* 1992;58:207–211.

56. Eidelberg D, Moeller JR, Antonini A, et al. The metabolic anatomy of Tourette's syndrome. *Neurology* 1997;48:927–934.

57. Peterson BS, Skudlarski P, Anderson AW, et al. A functional magnetic resonance imaging study of tic suppression in Tourette syndrome. *Arch Gen Psychiatry* 1998;55:326–333.

58. Wong DF, Singer HS, Brandt J, et al. D2-like dopamine receptor density in Tourette syndrome measured by PET. *J Nucl Med* 1997;38:1243–1247.

59. Syzmanski S, Galiano J, Yokoi F, et al. Basal ganglia dopamine release in Tourette syndrome. *Proc Soc Nucl Med* 1997;37:12P.

60. Malison RT, McDougle CJ, van Dyck CH, et al. [^{123}I]beta-CIT SPECT imaging of striatal dopamine transporter binding in Tourette's disorder. *Am J Psychiatry* 1995;152:1359–1361.

61. George MS, Robertson MM, Costa DC, et al. Dopamine receptor availability in Tourette's syndrome. *Psychiatry Res* 1994;55:193–203.

62. Turjanski N, Sawle GV, Playford ED, et al. PET studies of the presynaptic and postsynaptic dopaminergic systems in Tourette's syndrome. *J Neurol Neurosurg Psychiatry* 1994;57:688–692.

63. Wolf SS, Jones DW, Knable MB, et al. Tourette syndrome: prediction of phenotypic variation in monozygotic twins by caudate nucleus D2 receptor binding. *Science* 1996;273:1225–1227.

64. Leckman JF, Goodman WK, Anderson GM, et al. Cerebrospinal fluid biogenic amines in obsessive compulsive disorder, Tourette's syndrome, and healthy controls. *Neuropsychopharmacology* 1995;12:73–86.

65. De Groot CM, Bornstein RA, Baker GB. Obsessive-compulsive symptom clusters and urinary amine correlates in Tourette syndrome. *J Nerv Ment Dis* 1995;183:224–230.

66. Chappell P, Riddle M, Anderson G, et al. Enhanced stress responsivity of Tourette syndrome patients undergoing lumbar puncture. *Biol Psychiatry* 1994;36:35–43.

67. Dursun SM, Farrar G, Handley SL, et al. Elevated plasma kynurenine in Tourette syndrome. *Mol Chem Neuropathol* 1994;21:55–60.

68. Singer HS, Giulano JD, Hansen BH, et al. Antibodies against human putamen in children with Tourette syndrome. *Neurology* 1998;50:1618–1624.

69. Schultz RT, Carter AS, Gladstone M, et al. Visual-motor integration functioning in children with Tourette syndrome. *Neuropsychology* 1998;12:134–145.

70. Georgiou N, Bradshaw JL, Phillips JG, et al. The Simon effect and attention deficits in Gilles de la Tourette's syndrome and Huntington's disease. *Brain* 1995;118:1305–1318.

71. Georgiou N, Bradshaw JL, Phillips JG, et al. Advance information and movement sequencing in Gilles de la Tourette's syndrome. *J Neurol Neurosurg Psychiatry* 1995;58:184–191.

72. Georgiou N, Bradshaw JL, Phillips JG, et al. The effect of Huntington's disease and Gilles de la Tourette's syndrome on the ability to hold and shift attention. *Neuropsychologia* 1996;34:843–851.

73. Castellanos FX, Fine EJ, Kaysen D, et al. Sensorimotor gating in boys with Tourette's syndrome and ADHD: preliminary results. *Biol Psychiatry* 1996;39:33–41.

74. Farber RH, Swerdlow NR, Clementz BA. Saccadic performance characteristics and the behavioral neurology of Tourette syndrome. *J Neurol Neurosurg Psychiatry* 1999;66:305–312.

75. Swerdlow NR, Magulac M, Filion D, et al. Visuospatial priming and latent inhibition in children and adults with Tourette's disorder. *Neuropsychology* 1996;10:485–494.

76. Smith SJ, Lees AJ. Abnormalities of the blink reflex in Gilles de la Tourette syndrome. *J Neurol Neurosurg Psychiatry* 1989;52:895–898.

77. Sutherland RJ, Kolb B, Schoel WM, et al. Neuropsychological assessment of children and adults with Tourette syndrome: a comparison with learning disabilities and schizophrenia. *Adv Neurol* 1982;35:311–322.

78. Savage CR, Baer L, Keuthen NJ, et al. Organizational strategies mediate nonverbal memory impairment in obsessive-compulsive disorder. *Biol Psychiatry* 1999;45:905–916.

79. Cummings JL. Frontal-subcortical circuits and human behavior. *Arch Neurol* 1993;50:873–880.

80. Ziemann U, Paulus W, Rothenberger A. Decreased motor inhibition in Tourette's disorder: evidence from transcranial magnetic stimulation. *Am J Psychiatry* 1997;154:1277–1284.

81. de Divitiis E, D'Errico A, Cerillo A. Stereotactic surgery in Gilles de la Tourette syndrome. *Acta Neurochir* 1977;24:73.

82. Korzen AV, Pushkov VV, Kharitonov RA, et al. Stereotaxic thalamotomy in the combined treatment of Gilles de la Tourette disease. *Zh Nevropatol Psikhiatr Im S S Korsakova* 1991;91:100–101.

83. Pauls DL, Raymond CL, Stevenson JM, et al. A family study of Gilles de la Tourette syndrome. *Am J Hum Genet* 1991;48:154–163.

84. Price RA, Kidd KK, Cohen DJ, et al. A twin study of Tourette syndrome. *Arch Gen Psychiatry* 1985;42:815–820.

85. Pauls DL, Leckman JF. The inheritance of Gilles de la Tourette's syndrome and associated behaviors: evidence for autosomal dominant transmission. *N Engl J Med* 1986;315:993–997.

86. Pauls DL, Pakstis AJ, Kurlan R, et al. Segregation and linkage analyses of Tourette's syndrome and related disorders. *J Am Acad Child Adolesc Psychiatry* 1990;29:195–203.

87. Seuchter SA; Hebebrand J; Klug B; et al. Complex segregation analysis of families ascertained through Gilles de la Tourette syndrome. *Genet Epidemiol* 2000;18:33–47.

88. Kurlan R, Eapen V, Stern J, et al. Bilineal transmission in Tourette's syndrome families. *Neurology* 1994;44:2336–2342.

89. Hanna PA, Janjua FN, Contant CF, et al. Bilineal transmission in Tourette syndrome. *Neurology* 1999;53:813–818.

90. McMahon WM, van de Wetering BJ, Filloux F, et al. Bilineal transmission and phenotypic variation of Tourette's disorder in a large pedigree. *J Am Acad Child Adolesc Psychiatry* 1996;35:672–680.

91. Eapen V, O'Neill J, Gurling HM, et al. Sex of parent transmission effect in Tourette's syndrome: evidence for earlier age at onset in maternally transmitted cases suggests a genomic imprinting effect. *Neurology* 1997;48:934–937.

92. Comings DE, Gonzalez N, Wu S, et al. Studies of the 48 bp repeat polymorphism of the DRD4 gene in impulsive, compulsive, addictive behaviors: Tourette syndrome, ADHD, pathological gambling, and substance abuse. *Am J Med Genet* 1999;88:358–368.

93. Gelernter J, Vandenbergh D, Kruger SD, et al. The dopamine transporter protein gene (SLC6A3): primary linkage mapping and linkage studies in Tourette syndrome. *Genomics* 1995;30:459–463.

94. Hebebrand J, Nothen MM, Ziegler A, et al. Nonreplication of linkage disequilibrium between the dopamine D4 receptor locus and Tourette syndrome. *Am J Hum Genet* 1997;61:238–239.

95. The Tourette Syndrome Association International Consortium for Genetics. A complete genome screen in sib pairs affected by Gilles de la Tourette syndrome. *Am J Hum Genet* 1999;65:1428–1436.

96. Leckman JF, Price RA, Walkup JT, et al. Nongenetic factors in Gilles de la Tourette's syndrome. *Arch Gen Psychiatry* 1987;44:100.

97. Hyde T, Aaronson B, Randolph C, et al. Relationship of birth weight to the phenotypic expression of Gilles de la Tourette's syndrome in monozygotic twins. *Neurology* 1992;42:652–658.

98. Leckman JF, Hardin MT, Dolansky ES, et al. Perinatal factors in the expression of Tourette's syndrome. *J Am Acad Child Adolesc Psychiatry* 1990;29:220–226.

99. Peterson BS, Leckman JF, Scahill L, et al. Hypothesis: steroid hormones and sexual dimorphisms modulate symptom expression in Tourette's syndrome. *Psychoneuroendocrinology* 1992;17:553–563.

100. Santangelo SL, Pauls DL, Goldstein J, et al. Tourette's syndrome: what are the influences of gender and comorbid obsessive-compulsive disorder? *J Am Acad Child Adolesc Psychiatry* 1994;33:795–804.

101. Whitaker AH, Van Rossem R, Feldman JF, et al. Psychiatric outcomes in low birth-weight children at age 6 years: relation to neonatal cranial ultrasound abnormalities. *Arch Gen Psychiatry* 1997;54:847–856.

102. Lipska BK, Swerdlow NR, Geyer MA, et al. Neonatal excitotoxic hippocampal damage in rats causes post-pubertal changes in prepulse inhibition of startle and its disruption by apomorphine. *Psychopharmacology (Berl)* 1995;122:35–43.

103. Engel JA, Zhang J, Bergstrom T, et al. Neonatal herpes simplex virus type 1 brain infection affects the development of sensorimotor gating in rats. *Brain Res* 2000;863:233–240.

104. Ellenbroek BA, van den Kroonenberg PT, Cools AR. The effects of an early stressful life event on sensorimotor gating in adult rats. *Schizophr Res* 1998;30:251–260.

105. Rothschild DM, O'Grady M, Wecker L. Neonatal cytomegalovirus exposure decreases prepulse inhibition in adult rats: implications for schizophrenia. *J Neurosci Res* 1999;57:429–434.

106. Schwarzkopf SB, Mitra T, Bruno JP. Sensory gating in rats depleted of dopamine as neonates: potential relevance to findings in schizophrenic patients. *Biol Psychiatry* 1992;31:759–773.

107. Weinstock M. Does prenatal stress impair coping and regulation of hypothalamic-pituitary-adrenal axis? *Neurosci Biobehav Rev* 1997;21:1–10.

108. Kay G, Tarcic N, Poltyrev T, et al. Prenatal stress depresses immune function in rats. *Physiol Behav* 1998;63:397–402.

109. Vizek M, Dostal M, Soukupova D. Perinatal hypoxia suppresses immune response of adult rats. *Physiol Res* 1993;42:201–204.

110. Bornstein RA, Stefl ME, Hammond L. A survey of Tourette syndrome patients and their families: the 1987 Ohio Tourette Survey. *J Neuropsychiatry Clin Neurosci* 1990;2:275–281.

111. Chappell PB, Leckman JF, Goodman WK, et al. Elevated cerebrospinal fluid corticotropin releasing factor in Tourette's syndrome. *Biol Psychiatry* 1996;39:776–783.

112. Bisno AL. Group A streptococcal infections and rheumatic fever. *N Engl J Med* 1991;325:783–793.

113. Swedo S, Rapoport J, Cheslow D, et al. High prevalence of obsessive-compulsive symptoms in patients with Sydenham's chorea. *Am J Psychiatry* 1989;146:246–249.

114. Mercadante MT, Campos MCR, Marques-Dias MJ, et al. Vocal tics in Sydenham s chorea. *J Am Acad Child Adolesc Psychiatry* 1997;36:305–306.

115. Swedo SE, et al. Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections: clinical description of the first 50 cases. *Am J Psychiatry* 1998;155: 264–271.

116. Swedo SE, Leonard HL, Mittleman BB, et al. Identification of children with pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections by a marker associated with rheumatic fever. *Am J Psychiatry* 1997;154:110–112.

117. Murphy TK, Goodman WK, Fudge MW, et al. B lymphocyte antigen D8/17: a peripheral marker for childhood-onset obsessive-compulsive disorder and Tourette's syndrome. *Am J Psychiatry* 1997;154:402–407.

118. Leckman JF, Riddle MA, Hardin MT, et al. The Yale Global Tic Severity Scale: initial testing of a clinician rated scale of tic severity. *J Am Acad Child Adolesc Psychiatry* 1989;28:566–573.

119. Gadow KD, Sverd J, Sprafkin J, et al. Long-term methylphenidate therapy in children with comorbid attention-deficit hyperactivity disorder and chronic multiple tic disorder. *Arch Gen Psychiatry* 1999;56:330–336.

120. Kurlan R. Treatment of tics. *Neurol Clin North Am* 1997;15:403–409.

121. Erenberg G, Cruse RP, Rothner AD. The natural history of Tourette syndrome: a follow-up study. *Ann Neurol* 1987;22:383–385.

122. Bruun RD, Budman CL. Risperidone as a treatment for Tourette's syndrome. *J Clin Psychiatry* 1996;57:29–31.

123. Sallee FR, Kurlan R, Goetz CG, et al. Ziprasidone treatment of children and adolescents with Tourette's syndrome: a pilot study. *J Am Acad Child Adolesc Psychiatry* 2000;39:292–299.

124. Leckman JF, Hardin MT, Riddle MA, et al. Clonidine treatment of Gilles de la Tourette's syndrome. *Arch Gen Psychiatry* 1991;48:324–328.

125. Chappell PB, Riddle MA, Scahill L, et al. Guanfacine treatment of comorbid attention-deficit hyperactivity disorder and Tourette's syndrome: preliminary clinical experience. *J Am Acad Child Adolesc Psychiatry* 1995;34:1140–1146.

126. Bruun RD, Budman CL. Paroxetine treatment of episodic rages associated with Tourette's disorder. *J Clin Psychiatry* 1998;59:581–584.

127. Gilbert DL, Sethuraman G, Sine L, et al. Tourette's syndrome improvement with pergolide in a randomized, double-blind, crossover trial. *Neurology* 2000;54:1310–1315.

128. Silver AA, Shytle RD, Philipp MK, et al. Case study: long-term potentiation of neuroleptics with transdermal nicotine in Tourette's syndrome. *J Am Acad Child Adolesc Psychiatry* 1996;35:1631–1636.

129. Dursun SM, Reveley MA, Bird R, et al. Longlasting improvement of Tourette's syndrome with transdermal nicotine. *Lancet* 1994;344:1577.

130. Dursun SM, Reveley MA. Differential effects of transdermal nicotine on microstructured analyses of tics in Tourette's syndrome: an open study. *Psychol Med* 1997;27:483–487.

131. West R, Hajek P, Foulds J, et al. A comparison of the abuse liability and dependence potential of nicotine patch, gum, spray and inhaler. *Psychopharmacology (Berl)* 2000;149:198–202.

132. Muller-Vahl KR, Kolbe H, Schneider U, et al. Cannabis in movement disorders. *Forschr Komplementarmed* 1999;6[Suppl 3]:23–27.

133. Kurlan R, Majumdar L, Deeley C, et al. A controlled trial of propoxyphene and naltrexone in patients with Tourette's syndrome. *Ann Neurol* 1991;30:19–23.

134. Peterson BS, Zhang H, Anderson GM, et al. A double-blind, placebo-controlled, crossover trial of an antiandrogen in the treatment of Tourette's syndrome. *J Clin Psychopharmacol* 1998;18:324–331.

135. Trimble MR, Whurr R, Brookes G, et al. Vocal tics in Gilles de la Tourette syndrome treated with botulinum toxin injections. *Mov Disord* 1998;13:617–619.

136. Woods DW, Miltenberger RG, Lumley VA. Sequential application of major habit-reversal components to treat motor tics in children. *J App Behav Anal* 1996;29:483–493.

137. Vandewalle V, van der Linden C, Groenewegen HJ, et al. Stereotactic treatment of Gilles de la Tourette syndrome by high frequency stimulation of thalamus. *Lancet* 1999;353:724.

138. Swerdlow NR, Berkowitz A, Giertz M, et al. Treatment of a patient with a tic and a depression. *J Clin Psychiatry* 1990;51:34–35.

139. Rapoport M, Feder V, Sandor P. Response of major depression and Tourette's syndrome to ECT: a case report. *Psychosom Med* 1998;60:528–529.

140. Coffey BJ, Park KS. Behavioral and emotional aspects of Tourette syndrome. *Neurol Clin North Am* 1997;15:277–289.

141. Schwartz JM, Stoessel PW, Baxter LRJ, et al. Systematic changes in cerebral glucose metabolic rate after successful behavior modification treatment of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1996;53:109–113.

142. Campbell KM, de Lecea L, Severynse DM, et al. OCD-like behaviors caused by a neuropotentiating transgene targeted to cortical and limbic D1+ neurons. *J Neurosci* 1999;19:5044–5053.

143. Dumartin B, Jaber M, Gonon F, et al. Dopamine tone regulates D1 receptor trafficking and delivery in striatal neurons in dopamine transporter-deficient mice. *Proc Natl Acad Sci USA* 2000;15;97:1879–1884.

144. Canales JJ, Graybiel AM. A measure of striatal function predicts motor stereotypy. *Nat Neurosci* 2000;3:377–383.

145. Leckman JF, Riddle MA. Tourette's syndrome: when habit forming systems form habits of their own? *Neuron* 2000;28:349–354.

146. Sandler AD, Sutton KA, DeWeese J, et al. Lack of benefit of a single dose of synthetic human secretin in the treatment of autism and pervasive developmental disorder. *N Engl J Med* 1999;341:1801–1806.

147. Perlmutter SJ, Leitman SF, Garvey MA, et al. Therapeutic plasma exchange and intravenous immunoglobulin for obsessive-compulsive disorder and tic disorders in childhood. *Lancet* 1999;354:1153–115.

148. Gasser T, Fahn S, Breakefield XO. The autosomal dominant dystonias. *Brain Pathol* 1992;2:297–230.

149. Insel TR. Toward a neuroanatomy of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;49:739–744.

118
Animal Models of Aggression

Berend Olivier

Larry J. Young

Berend Olivier: Department of Psychopharmacology, Faculty of Pharmacy, Utrecht University, Utrecht, Netherlands; Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut.

Larry J. Young: Departments of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia.

Over the past several decades, various animal models have been used extensively to characterize the activity of various drugs and drug classes, and from these results, to anticipate their activity in humans. However, the value of animal models that purport to predict the potential therapeutic value of new drugs is often accepted with considerable reservation and, when the therapeutic objective involves psychiatric disease, is perhaps viewed with outright suspicion.

In general, the animal models most readily accepted as a basis for predicting responses in humans are those that are homologous, that is, those in which both the condition being observed and its origin are similar to those in humans. Examples may include suppression of bacterial infections by antibiotics or hypertension in monkeys. Few, if any, models of psychiatric dysfunction, however, can be considered homologous, if only because the origin of the psychiatric condition is unknown.

In the absence of homologous models, isomorphic models (in which the observed condition is apparently similar even if the cause is not) may be fairly rapidly accepted. An example may be amphetamine-induced psychosis as a putative model for schizophrenia. Finally, there are many models in which neither the condition nor the origin can be clearly linked with the disease being modeled, but in which there is empiric evidence of some predictive validity either for the disease or some aspect of its therapy. In psychopharmacology, the evidence is usually the discovery that agents with known therapeutic activity in humans consistently correlate with some response in an animal model. It can be argued that before animal models are developed for any disorder, the essential features of the disorder should be known. Here researchers run into trouble because the essential features of many disorders in humans are unclear. Pathologic aggression is not a DSM-IV disorder for which criteria are set for determining what is normal or abnormal (1). Moreover, there is little understanding of the biological factors underlying pathologic aggression in humans, so it is difficult to formulate a rational research program. A discussion of the characteristics of pathologic aggression is needed for the development of animal models of this disorder.

Despite the drawbacks adherent to aggression research, there is an increasing knowledge of the effects of psychoactive drugs on aggressive behavior, both in animals and in patients. Thus, two roads are emerging: one studying the fundamental causes of aggression and dysfunctions, the other studying the modification of behavior by pharmacologic interventions.

CLASSIFICATION OF AGGRESSION

Part of "118 - Animal Models of Aggression "

Despite many attempts, a generally acceptable definition of *aggression*, particularly as it applies to individual human behavior, has not yet emerged. This failure arises in part from the following: (a) the varying theoretic or philosophic persuasions of those offering definitions; (b) the inherent difficulty in capturing the essence of a multifaceted behavior; and (c) the attempt to include within the definitions elements of motivation that cannot readily be observed or elicited. However, a generally acceptable working definition from the perspective of animal research is somewhat easier to obtain and could read “any overt behavior that produces aversive or noxious stimuli or harm to another organism.” In this definition, the “motivation” of the behavior is not an essential element, but it may be deduced directly from the stimuli that elicit the behavior and from the overt behavior itself. Different types of animal aggression can be distinguished based on the environmental situations eliciting those behaviors. Moyer was the first to describe such a classification (2), later followed by alternative classifications (3).

All these classifications have their own inherent problems, such as to link the aggression to connotations of offensive and defensive aggression, using the ethology-derived term *agonistic behavior* as an important distinction to classify aggression models.

The distinction of aggression into offensive and defensive models is universal (4), is functional, and seems to be paralleled in brain mechanisms involved in the behavior. *Offensive behavior* is characterized by the initiative of the aggressor and intended damage to the opponent (5 ,6). In contrast, *defensive behavior* lacks active approach (initiative), and the defensive animal (6) inflicts no intentional damage. *Predatory aggression* represents separate classification of aggression that seems to be primarily driven by appetite mechanisms and apparently has a distinct brain system involved.

Neither in humans nor in animals is agonistic behavior pathologic. In the framework of evolutionary theory, these behaviors are understood to encourage survival of the fittest, to disperse populations, to aid adaptation to threatening environments, and generally to improve the probability of individual and species survival. In humans, agonistic behavior is considered acceptable or not based on certain predetermined rules. Although “aggressive” behavior is associated with certain somatic and psychiatric disease states, and such behavior may be considered in establishing a diagnosis, there is no diagnostic category of “aggressive disease” or “offensive syndrome” *per se*. In the clinical literature, such behaviors may be referred to as “violent,” “hostile,” “agitated,” “impulsive,” or “pathologic aggressive.” Although animal models of aggression try to simulate the human conditions as much as possible, this is difficult because we know so little about the underlying mechanisms of aggression in disease states. By studying several paradigms in animals with the expectation that they have at least some predictive validity for human disorders with pathologic aggression, we hope (a) to develop new drugs for treatment of patients and (b) to gain insights into the underlying mechanisms resulting in the disorder. With the emergence of molecular genetic technologies, we increasingly understand the roles of certain genes in aggression, which may ultimately lead to development of novel treatment strategies for pathologic aggression.

Instead of being exhaustive, the present chapter focuses on some selected animal models of aggression with some bearing for human pathologic conditions. Specific examples of drug effects and underlying mechanisms are also discussed.

MODELS OF OFFENSIVE BEHAVIORS

Part of "118 - Animal Models of Aggression "

Several paradigms are used to study offensive aggression, such as isolation-induced offensive behavior (mouse), resident-intruder offensive behavior (rat/mouse/hamster), offensive behavior after electrical stimulation of the brain (rat), maternal offensive behavior (mouse/rat), offensive play-fighting among juvenile rats, and offensive behavior among piglets.

Some of these models are described, and some relevant pharmacology is outlined (benzodiazepines, neuroleptics, psychostimulants, antidepressants, serenics). The putative face, construct, and predictive validity are discussed. In addition, several models providing insights into novel neural mechanisms of aggression as well as interactions between genes and early environment are presented. Moreover, new data are introduced regarding mutant mice showing phenotypic changes in aggression, such as the serotonin ($5\text{-}HT_{1B}$) and nitric oxide synthase (NOS) knockouts. The information coming from these new genetic models can be of help in understanding possible causes of human pathologic aggression.

Isolation-Induced Offensive Behavior

A manipulation often used to induce aggression is isolation of male animals, typically mice, for several weeks. Many such isolated animals, on encountering another male, will reliably exhibit attack behavior (7). The effect of isolation on aggressive behavior is strain dependent (8). This isolation-induced aggression paradigm in mice is one of the most frequently used aggression models in behavioral pharmacology (9), and it has engendered an extensive pharmacology mainly described in median effective dose (ED_{50}) values. Because there are many ways to affect aggression in a nonspecific way (sedation, motor disturbances, psychostimulation) an ED_{50} value is not at all helpful in delineating how and why drugs reduce aggression.

Because isolated male mice show a full repertoire of agonistic behaviors (10), ethologic techniques have been used to detect very specific drug effects (11). Although most tests are performed in the home cage of the isolated male mouse, performing the test in a neutral arena is attractive because the situation delivers a mix of offensive-defensive and flight behaviors, which are not seen or are infrequently seen in the home cage confrontation (11).

The latter model is very interesting because it shows properties of drugs that are revealed only partially, or not at all, by common pharmacologic test models (10 ,12 ,13). An extensive literature exists about the effects of γ-aminobutyric acid ($GABA_A$)-benzodiazepine agonists in this paradigm. Interestingly, classic non–subunit-selective benzodiazepine-receptor agonists show inverse U-shaped dose–response curves in this model. At lower doses, increases in aggression are seen, whereas at higher doses, no effects or decreases are observed (14), probably because of nonspecific effects such as muscle relaxation or sedation (11 ,15). A similar pattern of activity can be found after alcohol administration or consumption. Miczek et al. extensively investigated the effects of various doses of alcohol on aggressive behavior of male mice, rats, and monkeys and consistently found that individual animals respond differentially to alcohol (14 ,16 ,17 and 18). Approximately 25% of mice show heightened aggression after receiving low doses of alcohol (AHA mice), whereas the remainder show no increase in aggression (ANA mice). Interestingly, this alcohol-heightened aggression is attenuated by pretreatment of $5\text{-}HT_{1A}$

and 5-HT_{1B} receptor agonists (19 ,20), a finding indicating an essential role the serotonergic system in the modulation of offensive aggressive behavior. Enhanced aggression after benzodiazepines and alcohol treatment in some, but not all, animals is highly similar to the pattern found in humans (21 ,22 and 23) and supports the predictive validity of this kind of animal models for some types of human aggression.

Psychostimulants also disturb the normal agonistic behavior, although the resulting behavior is clearly differentially affected (10 ,11 ,19). D-Amphetamine–treated animals show aberrant "stimulated" behavior, which severely interferes with normal agonistic behavior (11 ,15 ,24). Neuroleptics (chlorpromazine, haloperidol) exert antiaggressive effects, but nonspecific effects, such as motor disturbances (catalepsy) (11 ,15), cause this. Serenics (serotonergic 5-$HT_{1B/1A}$-receptor agonists) have a highly selective antiaggressive profile in this test, reducing aggression specifically without dramatically affecting other behaviors dramatically and certainly not causing any unwanted side effects (25 ,26).

In the 1990s, molecular biological techniques provided us with potentially very exciting ways of studying aggression. Several gene knockout mice were generated, and some have been tested on their aggressive behavior. Several knockouts seem to be more aggressive than their wild types, including, among others, the 5-HT_{1B} receptor (27), the neural form of nitric oxide synthase (nNOS) (28), monoamine oxidase A (MAO A) (29) and calcium-calmodulin kinase II (CAMKII) (30). Interestingly, mice with a deletion in the endothelial form of the nitric oxide synthase (eNOS) (31) exhibit a virtual elimination of aggressive behavior. One has to be careful to consider the hyperaggression obtained after the mutation directly caused by the absence of the gene. Genetic background effects may cloud a clear interpretation, whereas adaptational processes over time may also influence the outcome. Most studies reported do not use extensive description of the behavioral phenotype, and conclusions whether the observed "aggressive" phenotype of the mutant is directly caused by the absence of the gene or results from maladaptation of the mutant to external stimuli have to be investigated before a mutant can be considered as a putative model for a certain kind of aggression. Nonetheless, studies screening knockout mice for alterations in aggressive behavior should prove useful in identifying novel mechanisms involved in aggression and provide useful models for development of novel drug intervention targets.

The 5-HT_{1B}-receptor knockout mouse (27) has been evaluated most extensively on different aspects of its hyperaggressiveness and has been proposed as an animal model of impulsivity (32 ,33). The latter study investigated territorial aggression in 5-HT_{1B} knockout males and corresponding wild types (in a 129SV background) while equipped with telemetric senders to record heart rate and body temperature during the experiment. Ethologic analysis of the behavior showed that the knockouts where more aggressive than the wild types. The significant findings from these studies were that, although these animals displayed higher levels of offensive aggression with a faster onset, other behaviors, including social investigation, defense, and exploration were completely normal (Table 118.1).

	Frequency		Duration	
Behavior	**Wild-Type**	**Knockout**	**Wild-Type**	**Knockout**
Nonsocial activity				
Attention	28 (19–33)	22.5 (17–26)	74.2 (53–137)	61.1 (50–90)
Rear	0.5 (0–2)	4.5 (0–9)	0.5 (0–4)	7.1 (0–28)
Sniff	22 (22–26)	17 (14–21)	60.3 (55–83)	42.4* (33–49)
Walk	23.5 (17–25)	22 (14–27)	39.0 (36–51)	28.7 (21–47)
Body care	1 (1–2)	4* (3–6)	2.7 (0.8–8.5)	18.0* (14–24)
In nest	9 (8–12)	0.5* (0–2)	86.3 (78–235)	0.5* (0–6.9)
Social activity				
Approach	8.5 (3–12)	18* (17–24)	6.6 (2.3–9.9)	19.0* (13–24)
Follow	1.5 (1–3)	1.0 (0–3)	2.0 (0.6–6.9)	1.1 (0–2.9)
Walk away	1.5 (0–4)	10* (8–13)	1.2 (0–3.4)	15.6* (11–17)
Social sniff	22.5 (16–30)	42.5* (35–48)	86.7 (55–114)	162* (145–186)
Genital sniff	10 (5–12)	10 (7–14)	29.8 (19–69)	35.7 (31–68)
Mount	0 (0–2)	1 (0–4)	0.0 (0–5.2)	2.8 (0–10)
Aggression				
Tail rattle	0 (0–2)	2* (1–7)	0.0 (0–1.5)	2.8* (0.9–11)
Lateral threat	0 (0–1)	2* (0–8)	0.0 (0–0.9)	2.3 (0–12)
Bite	2 (1–8)	9 (4–14)	1.9 (0.9–7.0)	9.6 (3.2–14)
Clinch/fight	0.5 (0–4)	7* (1–10)	0.4 (0–25)	22.1 (3.3–37)
In tube				
In tube	14.5 (13–17)	16.5 (8–25)	64.8 (47–108)	82.9 (46–108)

[a]Male wild-type and 5-HT_{1B} receptor knockout mice were singly housed (residents) for several months. These mice (n = 12 per genotype) were equipped with telemetric devices to record heart rate and body temperature. Mice were subjected to a 10-min encounter with a group-housed male intruder, and the behavior of the resident was scored, using an ethologic method. Data are given as median (using twenty-fifth to seventy-fifth percentiles) frequencies and duration (seconds) per 10 min. Mann-Whitney U test: the asterisk means significantly different from wild-type mice (p .< 05).

TABLE 118.1. ANIMAL BEHAVIOR ANALYSIS

Telemetric data on heart rate and body temperature showed no obvious abnormalities during the fight, although 5-HT_{1B} knockouts responded faster to all types of sensory stimuli such as opening the cage, handling, and injection (33). This pattern of reactivity was in line with the presumed impulsivity of this mutant and lends support to the use of this animal model in research into the mechanisms underlying impulse and aggression disorders, and it will be of help in screening new potential antiaggressive or antiimpulsive drugs. Much more fundamental work, including pharmacology, has to be done on this mutant, but the appearance of new animal models of human diseases seems a realistic option (34 ,35).

Isolation-induced aggression in mice is an animal model of offensive aggression with excellent predictive validity toward human aggression. Although some face validity is clearly present (offensive impulsive aggression in human aggression), the construct validity is as yet largely unknown.

Resident-Intruder Offensive Behavior

This model, very frequently used in psychopharmacology, uses the resident animal's response to a conspecific intruder (24 ,36 ,37). In the *resident-intruder paradigm*, a male rat is housed with a female, a situation resembling the natural situation in which animals establish and defend territories (38). When resident or territorial males meet an unfamiliar male intruder in their territory, heavy fighting may ensue, considered natural fighting (39 ,40). The attacking male performs a complete agonistic repertoire including both appetitive and consummatory behaviors Aggressive behavior in this situation may consist of searching (patrolling), approach, investigation, threats fighting, chasing, and dominant posturing The nature of such interactions between an attacking resident and an opponent varies with the quality of the intruder, especially age and hormonal status, and the resident's experience. The types of behaviors displayed by the resident toward the intruder are not random but follow certain rules (15), a strong indicator of the neural substrates involved (4).

The resident-intruder model differs both from isolation-induced aggression in mice and intermale aggression in rats, because there is no isolation, which may lead to behavioral abnormalities (8). Moreover, resident-intruder paradigms have a very wide species generality (41), including humans (42). Isolation-induced aggression, in contrast, is far more restricted to certain species (10). This model discriminates effectively the quality and behavioral mechanisms of action

of several drugs with proaggressive and antiaggressive actions (13 ,15 ,43).

Benzodiazepines at low doses enhance aggression (14 ,44), whereas at higher doses they clearly cause ataxia, which interferes with the behavioral performance. Alcohol, as in mice, enhances aggression in some rats, but not in others (14 ,18). Interestingly, this increased aggression in a subpopulation of the resident males was observed both after experimenter-administered ethanol and after self-administered ethanol (18). Understanding the underlying neurochemical mechanisms responsible for the individual differences in behavioral response to ethanol in these two subpopulations of rats or mice should prove valuable for understanding the factors resulting in pathologic aggression in humans.

Neuroleptics, like psychostimulants, alter the display of agonistic behavior (15), although in different ways (25). Serenics display a highly specific antiaggressive profile. This effect is caused by the activation of postsynaptic 5-HT_{1B} receptors because ligands affecting other 5-HT receptors have quite different antiaggressive profiles.

The resident-intruder paradigm has also been used in hamsters to elucidate a novel neurochemical pathway involved in aggression. Hamsters, which are territorial, quickly attack intruders. The neuropeptide vasopressin has been shown to act in the preoptic area and anterior hypothalamus to stimulate both displays of dominance and aggression (45). Vasopressin receptor antagonists injected into this area are potent inhibitors of aggression. Fluoxetine, a selective serotonin uptake inhibitor, decreases offensive aggression in male hamsters and prevents vasopressin-induced aggression. This finding has led to the hypothesis that there is an interaction between vasopressinergic and serotonergic systems in the regulation of offensive aggression. Hamsters subjected to social subjugation as juveniles displayed elevated levels of aggression toward smaller hamsters as adults (46). As adults, these subjugated hamsters had altered levels of both vasopressin and serotonin in the anterior hypothalamus, a finding providing a potential mechanism by which environmental influences may permanently alter the neural circuits regulating aggression. Interestingly, elevated levels of vasopressin in the cerebrospinal fluid has been correlated with indices of aggression in personality-disordered patients (47). This model provides an example in which discovering the neural mechanisms underlying aggression could potentially lead to new targets of intervention for therapy in pathologic aggression.

The resident-intruder paradigm has a very good predictive validity toward human aggression. Both proaggressive (alcohol and benzodiazepines) and antiaggressive effects of

psychoactive drugs are highly similar in rodents and humans. Because of the species generality of this type of aggression, the model also has considerable face validity. Construct validity is as yet less clear, but the brain mechanisms involved, the hormonal sequelae, and the behavior-evoking stimuli support reasonable construct validity. As such, this paradigm seems an excellent choice in screening for potential antiaggressive compounds (serenics), but it also indicates other drug effects such as sedation and sensory and motor impairment (15).

Offensive Behavior after Electrical Brain Stimulation

Behavior largely similar to that of offensive territorial males can be elicited by electrical stimulation in the medial-lateral hypothalamus of male and female rats (48 ,49 and 50). Hypothalamic aggression in male rats is sensitive to manipulations of androgen levels (51), and it can be induced in an area (52) roughly coinciding with the areas where levels of circulating sex hormones are regulated. Moreover, stimulation of this area is accompanied by elevated levels of stress hormones (adrenocorticotropic hormone, corticosterone, and prolactin) resulting from activation of the area itself and not caused by the stress of fighting (53). In female rats, aggression can be elicited in this same area (54 ,55). This behavior is readily reproduced under controlled circumstances, thereby meeting an important requirement for a model to study aggression. The aggressive behavior induced by the stimulation can be explosive. Depending on the stimulus intensity, extreme forms of offensive attack and severe damage to the opponent can be observed (50). The attack behavior is not purely driven by internal stimulation of the hypothalamic substrate. The animal's response is still dependent on external cues such as the age and sex of the opponent. In addition to aggressive behavior, stimulation in this area of the hypothalamus also stimulates other behaviors, including locomotion and teeth chattering (54 ,56), thereby allowing for the determination of the specificity of the drug. In this paradigm, the effects of drugs are measured by the changes in the current thresholds required to evoke the respective behavior (56). Increases in the current thresholds for aggression indicate antiaggressive effects, considered specific if simultaneously the drug does not affect thresholds for locomotion. Several drugs have been analyzed in this model, including benzodiazepines, neuroleptics, psychostimulants, alcohol, 5-HT_{1A}-receptor agonists, serenics (5-$HT_{1A/1B}$-receptor agonists) and selective serotonin reuptake inhibitors (56 ,57 ,58 and 59).

Chlordiazepoxide, in contrast to its effects in the isolation-induced and resident-intruder paradigms, had no effect on aggression and teeth-chattering thresholds at lower doses and enhanced the thresholds for both aggression and locomotion only at high doses, presumably reflecting the muscle relaxant properties at these doses. Alcohol, up to a dose of 2 g/kg, had no effects on any parameter, a finding suggesting that the proaggressive actions of alcohol and also the benzodiazepines seen in the territorial and isolated male paradigms are probably related to variables (anxiety?) other than aggression *per se*.

Haloperidol enhanced aggression thresholds simultaneously with locomotion, again indicative of nonspecific effects on aggression. Because thresholds for teeth chatter, which accompanies normal aggression, were not affected, it was concluded that aggression was not at all influenced by haloperidol, in line with earlier findings that antiaggressive actions of neuroleptics result from their side effects (catalepsy). D-Amphetamine had no effect on aggression and teeth chattering, but it decreased the locomotor threshold, a finding illustrating its stimulatory action without having specific effects on aggression. Scopolamine, a (muscarinic) anticholinergic drug, had effects similar to those of D-amphetamine, again illustrating that activation of substrates for locomotor activity is independent from activation or inhibition of aggression substrates in the brain. Naloxone, an opiate antagonist, did not influence any aspect of the brain stimulation–induced behaviors, in line with its absence on spontaneous aggression (11 ,15). Manipulation of various serotonergic mechanisms showed that activation of the 5-HT_{1B} receptor, by eltoprazine, fluprazine, meta-chlorophenylpiperazine, DL-propranolol and other phenylpiperazines (25 ,26), induces a highly specific effect on aggression. Aggression and teeth-chattering thresholds were enhanced, although aggression still could be evoked, but locomotor activity was not affected or was even somewhat decreased. The profiles of drugs that modulate other serotonergic receptors, including 5-HT_{1A}, 5-HT_3, and the serotonin transporter, demonstrate the specificity of the 5-HT_{1B} receptor in aggression, a finding suggesting that the nonspecific effects of serotonergic drugs are mediated through other 5-HT-receptor mechanisms.

This hypothalamic-induced aggression model is highly relevant for modeling certain kinds of human aggression. By directly stimulating neural substrates in the brain involved in offensive aggression, this model has great potential to predict violent, pathologic aggression in humans. In contrast to the more natural models (isolation-induced, resident-intruder, maternal aggression), this model is not sensitive to certain intervening variables present in the other paradigms (anxiety, fear, sedation, and motor and sensory disturbances) and directly reflects antiaggressive properties of drugs. In addition, this model is not completely artificial or pathologic in the sense that attacking animals do not respond to nonsalient stimuli in preparation of or during the attack. For example, such animals do not attack rats that previously have defeated them or females in estrus. The predictive validity of this model seems to be somewhat less than the other models; nonetheless, the model is useful in determining how drugs bring about the antiaggressive effect.

Maternal Aggression

Although aggression is often considered a male-related phenomenon, females can be quite aggressive under certain conditions, such as in hypothalamically induced aggression in rats (54 ,55), aggression in nonestrus hamsters (58), and maternal aggression in several rodent species (59 ,60). The use of a female aggression paradigm to model human (female?) aggression has been quite uncommon, particularly for psychopharmacologic purposes. The maternal aggression model seems to constitute such a model because it shows very wide species generality, including humans (61), and it has clear neural and hormonal determinants. Maternal aggression is highly purposeful, providing protection to the offspring. The maternal aggression paradigm is based on the finding that a lactating female rat or mouse with pups will exhibit offensive behaviors toward a wide variety of intruders. This behavior is most pronounced during the first part of the lactating period (62 ,63). Because the critical stimulus is clearly the proximity of some threatening object to the female's young, some authors (2) consider this behavior mainly defensive. However, the behavior of the lactating female toward an intruder is clearly self-initiated, proactive, and not necessarily reactive to any threat initiated by the intruder. Although the paradigm is labor intensive and needs extensive planning, several psychoactive drugs have been tested in it and have led to a model with a comparable predictive validity as the male offensive paradigms in rats (resident-intruder) and mice (isolation-induced). The psychopharmacology of maternal aggression has been mainly studied in rats, although some work in mice has been performed (64).

Although the topography of the aggressive behavior of the attacking female is clearly different from male aggression (63), the effects of several classes of drugs were remarkably similar to those found in the male paradigms. Benzodiazepines and alcohol showed, at low doses, proaggressive effects, which faded at higher doses because of nonspecific effects (sedation, ataxia, and muscle relaxation). D-Amphetamine reduced aggression at higher doses, but this was clearly attributable to interfering effects of motor stimulation, whereas haloperidol inhibited aggression by its highly nonspecific side effects. Serenics (eltoprazine and related phenylpiperazines) have a highly selective effect in this paradigm, reducing aggression without affecting other behaviors, including pup care. The critical role of the $5\text{-}HT_{1B}$ receptor in this effect was again demonstrated by additional pharmacology showing that modulating other serotonergic receptors did not have such effects.

MODELS OF DEFENSIVE BEHAVIORS

Part of "118 - Animal Models of Aggression "

Those forms of agonistic behavior in which elements of initiative and approach prevail belong to the offensive repertoire, characterized by initiative, attack, and similar proactive behavior. This sharply contrasts with the defensive repertoire, which is characterized by submission, flight, and similar reactive behaviors. Fighting, when it occurs in a defensive animal, is merely a reaction to attack. Other defensive behaviors, such as flight or submission, are apparently intended to escape from or prevent further agonistic interactions (65 when it occurs in a defensive animal, is merely a reaction to attack. Other defensive behaviors, such as flight or submission, are apparently intended to escape from or prevent further agonistic interactions (65when it occurs in a defensive animal, is merely a reaction to attack. Other defensive behaviors, such as flight or submission, are apparently intended to escape from or prevent further agonistic interactions (65). Some of the drugs known to suppress offensive behaviors have highly undesirable effects on defensive behavior; for example, neuroleptics inhibit all activities including defensive and flight reactions.

Pain- or Shock-Induced Defensive Behavior

Delivering electric shock to the hind paws of a pair of rats or mice evokes so-called *foot shock*– or pain-induced aggression (66). Similar behavioral responses can be found when certain drugs (apomorphine, mescaline) are given to pairs of animals. Whereas either or both animals may attack, the behavior is conceived as defensive (5), in part because the animals mutually exhibit typical upright defensive postures and squealing and the behavior is clearly reactive; without switching on the current, no agonistic interactions will occur. Although this paradigm was extensively used in the past to assess antiaggressive activity of drugs, a confounding factor in the model is that the behavior-releasing factor (pain) can be masked by analgesic properties of drugs. This fact, together with the limited behavioral repertoire exhibited in this paradigm, limits its utility considerably.

A useful application of foot shock–induced defense behavior is to determine the ED_{50} to lower the amount of fighting episodes by 50% and concurrently to determine an ED_{50} for paralysis, the ability of mice, hanging by their forelimbs from a thin bar, to bring their hind limbs on to the bar within a certain time. Specificity of the antidefense effects is calculated as the ratio between these two ED_{50} values. A high ratio indicates good antidefense specificity; low values suggest strong interfering effects.

Neuroleptics show very low ratios (less than 1), tricyclic antidepressants have ratios of approximately 1, whereas benzodiazepines also have ratios less than 1 because of their muscle-relaxing effects (26 ,67). Serenics (eltoprazine, fluprazine, and others) appear highly selective; ratios of more than 20 (fluprazine) or, in the case of eltoprazine, not determinable have been found. Eltoprazine did, up to a very high dose, not inhibit this foot shock–induced behavior, thereby illustrating its highly selective antioffense character.

Defensive Behavior in Rats (Intruder Model)

A more natural model of defensive behavior is the behavior shown by an intruder in the resident-intruder or maternal aggression situation. Defending animals in these paradigms

use special tactics to protect the more vulnerable parts of their bodies. In unconstrained conditions, animals on the defense usually flee from the territory of the residential male or lactating female, but when this is impossible, as in laboratory conditions, they defend themselves by flight, crouching, upright defensive postures, emission of ultrasounds, and submissive postures. Generally, these behaviors aim at protecting the back, the area where most wounds are inflicted by attacking rats (68).

Although this model involves at least two animals, the offender and the defender, it provides an opportunity for various drug manipulations and to study direct and indirect drug effects (69). Drugging the offender and changing its offensive behavior have clear effects on the behavior of the intruder (26 ,70). The quality of the intruder (i.e., age and size) also determines the behavioral outcome, and interactions between drug effects and the intruder quality have been observed for D-amphetamine and chlordiazepoxide (44 ,70). This finding illustrates the construct and face validity of this model because it has high resemblances to the human (psychiatric) situation.

This model has been of limited use in psychopharmacology mainly because of the complexity in interpreting the interaction between drug effect and intruder quality. However, using standardized circumstances in which the intruder is basically not threatening the role of the resident, that is, by using young or inexperienced inexperienced intruders, the direct effects of drugs on the behavior of the intruder can be studied. It appears that influencing the sensory or motor capabilities of the intruder (by neuroleptics, alcohol, benzodiazepines, or psychostimulants) leads to changes in the defense or flight responses of the animals indicating enhanced flight (D-amphetamine) or impaired defense or flight (haloperidol). This may lead to enhanced attacks on the intruder in the case of D-amphetamine or diminished interest in the intruder by the resident in the case of neuroleptics (26). Serenics do not affect the defense or flight capabilities of the intruders (26), an effect in line with their specific antioffense qualities.

The resident-intruder model is a unique animal model for different aspects of social interactions and provides an opportunity to determine not only what drugs are doing directly to an organism, but also the indirect effects on the partner. Human pathologic aggression is often associated with complex interpersonal interactions, and the resident-intruder interaction model may be particularly relevant to predict what drugs may do in humans in particular circumstances.

MISCELLANEOUS MODELS

Part of "118 - Animal Models of Aggression "

Predatory aggression, such as mouse killing (muricide) in rats or locust killing (insecticide) in mice, occurs spontaneously in a proportion of individuals, depending on the strain used (71). There has been a great deal of dispute about the nature of muricide in rats, resulting in various descriptions of the behavior including, interspecies aggression (72), predatory aggression (73), or simply predatory behavior (74). This model is clearly different from those described earlier, and its human equivalent is questionable, although predatory aggression has been described in relation to pathologic aggression in humans. Predatory attack clearly differs from intraspecies attack with regard to neuroanatomic, physiologic, and hormonal mechanisms (75), but the pharmacology is less developed. The model is rarely used anymore because of several ethical constraints, and therefore most data are from before the 1990s (26). Benzodiazepines and alcohol, unless given at extremely high doses, do not inhibit muricide. Neuroleptics and psychostimulants do inhibit muricide but clearly not in a very specific way (cataleptic, motor stimulation). Antidepressants (tricyclics and selective serotonin reuptake inhibitors) inhibit muricide in a quite specific way, and in the 1950s up until the 1970s, the muricide test was in use in the pharmaceutical drug discovery process as an antidepressant screen. 5-HT_{1A}-receptor agonists do not inhibit muricide, whereas 5-HT_{1B}-receptor agonists (serenics) inhibit it, but only at much higher doses than the antioffense effects. Therefore, this model is not believed to have potential qualities to predict certain human pathologic aggression situations.

DISCUSSION

Part of "118 - Animal Models of Aggression "

The present contribution has suggested a limited number of animal models for different forms of aggression, namely, offensive aggression, defensive aggression, and predatory aggression. This is absolutely not an exhaustive coverage of the field and shows a logical way to frame the existing animal tests and paradigms into meaningful categories especially for predicting the effects of psychoactive drugs for human pathologic conditions. This approach led to the development in the 1970s and 1980s of a group of serotonergic agonists, serenics, as potential antiaggressive agents to treat certain types of human pathologic aggression (26 ,76).

Moreover, the animal models outlined appeared to have predictable validity and enable us to predict putative outcomes when applied in humans. Good examples are the benzodiazepines, with which, at low doses, aggression-enhancing effects were often found that corresponded to the so-called “paradoxic” aggression seen after human use (22 ,23). The antiaggressive effects of neuroleptics, often clinically used as first treatment in emergencies, are not specific at all but result from interference with vital functions (cognition, motor, sensory). Antiaggressive effects can be the result of many (side) effects of drugs, and the proposed animal models are capable of detecting and describing these effects. Consequently, they have a high predictable validity

toward their effects in humans, although it is very difficult to predict for which human disorder or symptom (77).

One of the biggest obstacles in the study of the psychopharmacology of aggression and the predictability for the human situation is the lack of consensus on definitions. The only primary aggression disorder in DSM-IV is intermittent explosive disorder, but all other aggression and impulsivity occurring in various disorders are considered as symptoms of different underlying disorders; this situation makes it extremely difficult to compare human with animal aggression directly. From the animal research, the evidence is very strong that there are specific neural substrates in the brain subserving these different functions in agonistic behavior, and it is more than likely that similar mechanisms are available in the human brain (78 ,79). The fundamental research in animals suggests that serotonin, actually only a subset of the system, such as the postsynaptic 5-HT_{1B} receptor (80), is an important neurotransmitter in at least part of this brain circuitry. Genetic modification of this system (5-HT_{1B}-receptor knockout mouse) has added considerable evidence for the importance of this system, although it is clear that the latter is only a small part of a much bigger and very complex circuitry in the brain involved in agonistic behavior.

Preclinical aggression research is under considerable pressure because of ethical and societal constraints on doing “biologically” oriented research in understanding the neurobiology of aggression and possible disturbances of the systems involved in the case of pathologic aggression (81). However, further research, using animal models of aggression, is needed to discover new treatments for pathologic aggression and violence. Analysis of the behavioral profiles of genetically altered mice with targeted gene deletions holds great promise over the next decade for discovering novel neurochemical pathways in the brain involved in the control of aggression. New developments in the molecular biology area, generating inducible, and brain region–specific mutants will engender exciting tools to study the role of genes, environment, and their interaction in the causation of aggression, and important new clues for the study and treatment of pathologic aggression in humans will emerge.

REFERENCES

1. Eichelman B, Hartwig A. The clinical psychopharmacology of violence: towards a nosology of human aggressive behavior. *Psychopharmacol Bull* 1993;29:57–63.

2. Moyer KE. Kinds of aggression and their physiological basis. *Commun Behav Biol* 1968;2:65–87.

3. Huntingford F, Turner A. *Animal conflict*. London: Chapman & Hall, 1987.

4. Adams DB. Brain mechanisms for offense, defense and submission. *Behav Brain Sci* 1979;2:201–241.

5. Blanchard RJ, Blanchard DC, Takahashi T, et al. Attack and defensive behavior in the albino rat. *Anim Behav* 1977;25:622–634.

6. Olivier B, Van Dalen D, Hartog J. A new class of psychoactive drugs: serenics. *Drugs Future* 1990;11:473–499.

7. Valzelli L. Aggressive behaviour induced by isolation. In: Garattini S, Siggs SB, eds. *Aggressive behaviour*. Amsterdam: Excerpta Medica, 1969:70–76.

8. Valzelli L. The “isolation syndrome” in mice. *Psychopharmacologia* 1973;31:305–320.

9. Malick JB. The pharmacology of isolation-induced aggressive behaviour in mice. *Curr Dev Psychopharmacol* 1979;5:1–27.

10. Miczek KA, Krsiak M. Drug effects on agonistic behavior. In: Thompson T, Dews PB, eds. *Advances in behavioral pharmacology*, vol. 2. New York: Academic Press 1979:87–162.

11. Olivier B, Van Dalen D. Social behaviour in rats and mice: an ethologically based model for differentiating psychoactive drugs. *Aggress Behav* 1982;8:163–168.

12. Miczek KA, Barry HIII. Pharmacology of sex and aggression. In: Glick SD, Goldfarb J, eds. *Behavioral pharmacology*. St. Louis: CV Mosby, 1976:176–257.

13. Olivier B. Selective anti-aggresive properties of DU27725: ethological analyses of intermale and territorial aggression in the male rat. *Pharmacol Biochem Behav* 1981;14[Suppl 1]:61–77.

14. Miczek KA, DeBold JF, Van Erp AM, et al. Alcohol, $GABA_A$-benzodiazepine receptor complex, and aggression. *Recent Dev Alcohol* 1997;13:139–171.

15. Olivier B, Van Aken H, Jaarsma I, et al. Behavioural effects of psychoactive drugs on agonistic behaviour of male territorial rats (resident-intruder paradigm). In: Miczek KA, Kruk MR, Olivier B, eds. *Ethopharmacological aggression research*. New York: Alan R. Liss, 1984:137–156.

16. Miczek KA, Barros HM, Sakoda L, et al. Alcohol and heightened aggression in individual mice. *Alcohol Clin Exp Res* 1998;22:698–705.

17. Weerts EM, Tornatzky W, Miczek KA. “Anxiolytic” and “anxiogenic” benzodiazepines and beta-carbolines: effects on aggressive and social behavior in rats and squirrel monkeys. *Psychopharmacology (Berl)* 1993;110:451–459.

18. Van Erp AM, Miczek KA. Increased aggression after ethanol self-administration in male resident rats. *Psychopharmacology (Berl)* 1997;131:287–295.

19. Miczek KA, Hussain S, Faccidomo S. Alcohol-heightened aggresion in mice: attenuation by 5-HT_{1A}receptor agonists. *Psychopharmacology (Berl)* 1998;139:160–168.

20. Fish EW, Faccidomo S, Miczek KA. Aggression heightened by alcohol or social instigation in mice: reduction by the 5-HT(1B) receptor agonist CP-94,253. *Psychopharmacology (Berl)* 1999;146:391–399.

21. Ratey J, Gordon A. The psychopharmacology of aggression: toward a new day. *Psychopharmacol Bull* 1993;29:165–173.

22. Bond A, Lader M. Differential effects of oxazepam and lorazepam on aggressive responding. *Psychopharmacology (Berl)* 1988;95:369–373.

23. Weisman AM, Berman ME, Taylor SP. Effects of clorazepate, diazepam, and oxazepam on a laboratory measurement of aggression in men. *Int Clin Psychopharmacol* 1998;13:183–188.

24. Miczek KA. Intraspecies aggression in rats: effects of D-amphetamine and chlordiazepoxide. *Psychopharmacology (Berl)* 1974;39:275–301.

25. Olivier B, Mos J, Hartog J, et al. Serenics: a new class of drugs for putative treatnment selective treatment of pathological destructive behavior. *Drugs News Perspect* 1990;3:261–271.

26. Olivier B, Mos J, Raghoebar M. et al. Serenics. *Prog Drug Res* 1994;42:167–308.

27. Saudou F, Amara DA, Dierich A, et al. Enhanced aggressive behavior in mice lacking 5-HT_{1B} receptor. *Science* 1994;265:1875–1878.

28. Nelson RJ, Demas GE, Huang PL, et al. Behavioral abnormalities in male mice lacking neuronal nitric oxide synthase. *Nature* 1995;378:383–386.

29. Cases O, Seif I, Grimsby J, et al. Aggressive behavior and altered amounts of brain serotonin and norepinephrine in mice lacking MAOA. *Science* 1995;268:1763–1766.

30. Chen C, Rainnie DG, Greene RW, et al. Abnormal fear response and aggressive behavior in mutant mice deficient for calcium-calmodulin kinase II. *Science* 1994;266:291–294.

31. Demas GE, Kriegsfield LJ, Blackshaw S, et al. Elimination of aggressive behavior in male mice lacking endothelial nitric oxide synthase. *J Neurosci* 1999;19:19–30.

32. Brunner D, Hen R. Insights into the neurobiology of impulsive behavior from serotonin receptor knockout mice. *Ann NY Acad Sci* 1997;836:81–105.

33. Bouwknecht JA, Hijzen TH, Van der Gugten J. et al. Absence of 5-HT_{1B} receptors is associated with impaired motor impulse control in male 5-HT_{1B} knockout mice. *Biol Psychiatry* 2001;49:557–568.

34. Murphy DL, Wichems C, Li Q, et al. Molecular manipulations as tools for enhancing our understanding of 5-HT neurotransmission. *Trends Pharmacol Sci* 1999;20:246–252.

35. Rudolph U, Möhler H. Genetically modified animals in pharmacological research: future trends. *Eur J Pharmacol* 1999;375:327–337.

36. Adams DB. The relation of scent-marking, olfactory investigation, and specific postures in the isolation-induced fighting of rats. *Behaviour* 1976;56:286–297.

37. Olivier B. The ventromedial hypothalamus and aggressive behaviour in rats. *Aggress Behav* 1977;3:47–56.

38. Barnett SA. *The rat: a study in behaviour*. Chicago: University of Chicago Press, 1975.

39. Blanchard RJ, Blanchard DC. Alcohol and aggression in animal models. In: Olivier B, Mos J, Brain PF, eds. *Ethopharmacology of agonistic behaviour in animals and man*. Dordtrecht: Martinus Nijhoff, 1977:145–161.

40. Miczek KA. A new test for aggression in rats without aversive stimulation: differential effects of D-amphetamine and cocaine. *Psychopharmacology (Berl)* 1979;60:253–259.

41. Wilson EA. *Sociobiology*. Cambridge: Belknap Press of Harvard University Press, 1975.

42. Malmberg T. *Human territoriality: survey of behavioural territories in man with preliminary analysis of meaning*. The Hague, Mouton, 1980.

43. Mos J, Olivier B, Tulp MThM. Ethopharmacological studies differentiate the effects of various serotonergic compounds on aggression in rats. *Drug Dev Res* 1992;26:343–360.

44. Mos J, Olivier B, Van der Poel AM. Modulatory actions of benzodiazepine receptor ligands on agonistic behaviour. *Physiol Behav* 1987;41:265–278.

45. Ferris CF, Melloni RH, Koppel G, et al. Vasopressin/serotonin interactions in the anterior hypothalamus control aggressive behavior in golden hamsters. *J Neurosci* 1997;17:4331–4340.

46. Delville Y, Melloni RH, Ferris CF. Behavioral and neurobiological consequences of social subjugation during puberty in golden hamsters. *J Neurosci* 1998;18:2667–2872.

47. Coccaro EF, Kavoussi RJ, Hauger RI, et al. Cerebrospinal fluid vasopressin levels: correlates with aggression and serotonin function in personality-disordered subjects. *Arch Gen Psychiatry* 1998;55:708–714.

48. Koolhaas JM. Hypothalamically induced intraspecific aggressive behaviour in the rat. *Exp Brain Res* 1978;32:365–375.

49. Kruk MR, Van der Poel AM. Is there evidence for a neural correlate of an aggressive behavioural system in the hypothalamus of the rat? *Prog Brain Res* 1980;53:385–390.

50. Kruk MR, Van der Poel AM, De Vos-Frerichs TP. The induction of aggressive behaviour by electrical stimulation in the hypothalamus of male rats. *Behaviour* 1979;70:292–322.

51. Bermond B, Mos J, Meelis W, et al. Aggression induced by stimulation of the hypothalamus: effects of androgens. *Pharmacol Biochem Behav* 1982:16:41–45.

52. Kruk MR, Van der Poel AM, Melis W, et al. Discriminant analysis of the localization of aggression inducing electrode placements in the hypothalamus of male rats. *Brain Res* 1983;260:61–79.

53. Kruk MR, Westphal KGC, Van Erp AMM, et al. The hypothalamus: cross-roads of endocrine and behavioural regulation in grooming and aggression. *Neurosci Biobehav Rev* 1998;23:163–177.

54. Kruk MR, Van der Laan CE, Mos J, et al. Comparison of aggressive behaviour induced by electrical stimulation in the hypothalamus of male and female rats. *Prog Brain Res* 1984;61:303–314.

55. Mos J, Olivier B, Lammers JHMC, et al. Pregnancy and lactation do not interact with current thresholds for brain stimulation induced aggression in female rats. *Brain Res* 1987;404:263–266.

56. Van der Poel AM, Olivier B, Mos J, et al. Anti-aggressive effect of a new phenylpiperazine compound (DU27716) on hypothalamically induced behavioural activities. *Pharmacol Biochem Behav* 1982;17:147–153.

57. Olivier B, Mos J, Rasmussen D. Behavioural pharmacology of the serenic eltoprazine. *Drug Metab Drug Interact* 1990;8:31–83.

58. Floody OR, Pfaff DW. Aggressive behavior among female hamsters: the hormonal basis for fluctuations in female aggressiveness correlated with estrus state. *J Comp Physiol Psychol* 1977;91:443–446.

59. Siegel A, Roeling TAP, Gregg TR, et al. Neuropharmacology of brain-stimulation-evoked aggression. *Neurosci Biobehav Rev* 1999;23:359–389.

60. Floody OR. Hormones and aggression in female animals. In: Svare BB, ed. *Hormones and aggressive behavior*. New York: Plenum, 1983:39–89.

61. Svare B, Mann M. Hormonal influence on maternal aggression. In: Svare B, ed. *Hormones and aggressive behavior*. New York: Plenum, 1983:91–104.

62. Archer J. *The behavioural biology of aggression*. Cambridge: Cambridge University Press, 1988.

63. Erskine MS, Barfield RJ, Goldman BS. Intraspecific fighting during late pregnancy and lactation in rats and effects of litter removal. *Behav Biol* 1978;23:206–208.

64. Olivier B, Mos J. A female aggression paradigm for use in psychopharmacology: maternal agonistic behaviour in rats. In: Brain PF, Ramirez MJ, eds. *Cross-disciplinary studies on aggression*. Seville: University of Seville Press 1986:73–111.

65. Yoshimura H. Ethopharmacology of agonistic behaviour in male and female mice. In: Olivier B, Mos J, Brain PF, eds. *Ethopharmacology of agonistic behaviour in animals and man*. Dordtrecht: Martinus Nijhoff, 1987:94–109.

66. Dixon AK, Kaesermann HP. Ethopharmacology of flight behaviour. In: Olivier B, Mos J, Brain PF, eds. *Ethopharmacology of agonistic behaviour in animals and man*. Dordrecht: Martinus Nijhoff, 1987:46–79.

67. Ulrich RE, Azrin NH. Reflexive fighting in response to aversive stimulation. *J Exp Anal Behav* 1962;5:511–520.

68. Tedeschi RE, Tedeschi DH, Mucha A, et al. Effect of various centrally acting drugs on fighting behavior of mice. *J Pharmacol Exp Ther* 1959;125:28–34.

69. Mos J, Olivier B, Van Oorschot R, et al. Different test situations for measuring offensive aggression in male rats do not result in the same wound pattern. *Physiol Behav* 1984;32:453–456.

70. Blanchard DC, Griebel G, Rodgers RJ et al. Benzodiazepine and serotonergic modulation of antipredator and conspecific defense. *Neurosci Biobehav Rev* 1998;22:597–612.

71. Mos J, Olivier B, Van Oorschot R. Maternal aggression towards different sized male opponents: effect of chlordiazepoxide treatment of the mothers and D-amphetamine treatment of the intruders. *Pharmacol Biochem Behav* 1987;26:577–584.

72. Karli P. The Norway's rat killing response to the white mouse: an experimental analysis. *Behaviour* 1956;10:81–103.

73. Karli P, Vergnes M, Didiergeorges F. Rat-mouse interspecific aggressive behaviour and its manipulation by brain ablation and by brain stimulation. In: Garattini S, Sigg EB, eds. *Aggressive behaviour*. Amsterdam: Excerpta Medica, 1969:47–55.

74. Potegal M. The reinforcing value of several types of aggressive behavior: a review. *Aggress Behav* 1979;5:353–373.

75. Adamec RE, Himes M. The interaction of hunger, feeding, and experience in alteration of topography of the rat's predatory response to mice. *Behav Biol* 1978;22:230–243.

76. Huntingford FA, The relationship between intra- and inter-specific aggression. *Anim Behav* 1976;24:485–497.

77. Mak M, De Koning P, Mos J, et al. Preclinical and clinical studies on the role of 5-HT_1 receptors in aggression. In: Hollander E, Stein DJ, eds. *Impulsivity and aggression*. Chichester, UK: John Wiley, 1995:289–311.

78. Kavoussi R, Coccaro EF. Psychopharmacological treatment of impulsive aggression. In: Maes M, Coccaro EF, eds. *Neurobiology and clinical views on aggression and impulsivity*. Chichester, UK: John Wiley, 1998:197–211.

79. New AS, Novotny SL, Buchsbaum MS, et al. Neuroimaging in impulsive-aggressive personality disorder patients. In: Maes M, Coccaro EF, eds. *Neurobiology and clinical views on aggression and impulsivity*. Chichester, UK: John Wiley, 1998:81–93.

80. Stein DJ, Van Heerden B, Hugo F. Neuropsychiatry of aggression and impulsivity. In: Maes M, Coccaro EF, eds. *Neurobiology and clinical views on aggression and impulsivity*. Chichester, UK: John Wiley, 1998:95–107.

81. Olivier B, Mos J, Van Oorscht R, et al. Serotonin receptors and animal models of aggressive behaviour. *Pharmacopsychiatry* 1995;28[Suppl]:80–90.

82. Enserink M. Searching for the mark of Cain. *Science* 2000;289:575–579.

119

Pathophysiology and Treatment of Aggression

Emil F. Coccaro

Larry J. Siever

Emil F. Coccaro: Department of Psychiatry, University of Chicago, Pritzker School of Medicine, Chicago, Illinois.

Larry J. Siever: Psychiatry Service, Bronx Veterans Affairs Medical Center, Bronx, New York.

Human aggression constitutes a multidetermined act that often results in physical (or verbal) injury to others or self (or objects). It appears in several forms and may be defensive, premeditated (e.g., “predatory”), or impulsive (e.g., “nonpremeditated”). *Defensive aggression* is generally seen within the normal range of human behavior. However, *premeditated and impulsive aggressive behaviors* are commonly viewed as pathologic.

Aggression may be measured as both a dimensional and a categoric variable. However, whereas aggressive behavior (or the tendency to behave aggressively) is truly dimensional, it is difficult to estimate the societal relevance of aggressive behavior using dimensional assessments. As a categoric variable, aggressive subjects (“cases”) may be counted in populations of interest. For example, the current age-adjusted rate for homicide in the United States is 0.01% (1). With respect to physical assault, approximately one-fourth of all men and approximately half as many women report a history of physical fighting after 18 years of age (2). From these figures, it may be estimated that approximately 35 million adults in the United States have engaged in at least one *serious* act of aggression. The rate of individuals demonstrating “recurrent, problematic, aggression” is lower, of course. Using DSM-IV intermittent explosive disorder (IED) as a proxy for “recurrent, problematic, aggression,” Zimmerman et al. reported that nearly 7% of psychiatric outpatients meet criteria for IED at any point in their lives (3). Including borderline and antisocial personality-disordered patients, who are also quite aggressive, increases this number to about 13% of psychiatric outpatients. Regardless, these data suggest that the lifetime prevalence of recurrent, problematic, aggressive behavior may be 1% or higher (at least 2.5 million) in the general community (4).

Research into the etiologic determinants of both premeditated and impulsive human aggression has focused on various genetic, biological, and psychosocial factors and reveals a rich and complex picture of human aggression involving both constitutional and nonconstitutional elements. Twin, adoption, and family studies all suggest a genetic influence underlying aggression (5), with heritability estimates for dimensional measures of aggression ranging from 44% to 72% in adults. A metaanalysis of more than 20 twin studies confirmed a substantial role for a genetic influence underlying aggression (6). Although behavioral genetic studies to date have not attempted to distinguish among aggression subtypes, impulsive aggression appears to be quite distinct from premeditated aggression. Overall, the recurring theme emerging from more than 20 years of empiric research is that “impulsive aggression” demonstrates the most consistent and noteworthy findings with respect to both biological correlates (7 ,8) and psychopharmacologic treatment (9 ,10). Biological factors include a variety of neurotransmitter and neuromodulator systems. Most data involve the central serotonin (5-hydroxytryptamine or 5-HT) system, although limited data is now emerging for a role for other central systems involving catecholamines, steroids, neuropeptides, and cholesterol and fatty acids. This chapter reviews the neuropsychopharmacologic data relevant to these systems and concludes with a discussion of the psychopharmacology of aggression.

NEUROPSYCHOPHARMACOLOGY OF AGGRESSION

Part of "119 - Pathophysiology and Treatment of Aggression "

Serotonin

The hypothesis that 5-HT modulates impulsive aggressive behavior in humans was first advanced in the same year by both Sheard et al. (9) and Asberg et al. (11). The former group demonstrated that treatment with lithium carbonate (an agent with putative 5-HT–enhancing properties), but not placebo, reduced impulsive aggressive behavior in prison inmates; the latter group demonstrated that violent (and

eventually lethal) suicidal behavior was a characteristic of patients with low lumbar cerebrospinal fluid (CSF) concentrations of 5-hydroxyindolacetic acid (5-HIAA). This work led to a large number of subsequent studies, using a variety of 5-HT measures, designed to test the 5-HT hypothesis of aggression further.

5-HT and Aggression: CSF 5-HIAA Studies

Brown et al. first reported a strong inverse correlation between CSF 5-HIAA and a life history of aggression (r = -.78) in male personality-disordered adults (12). This finding was replicated in a follow-up study using another variable reflective of aggression (i.e., "psychopathic deviance," defined as defiance of authority and impulsivity) (13). In both samples, a trivariate relationship between reduced CSF 5-HIAA and aggression and suicidal behavior was demonstrated. Later reports demonstrated that reduced CSF 5-HIAA was specific to impulsive rather than premeditated (i.e., nonimpulsive) aggression (7 ,14). Linnoila et al. first demonstrated that CSF 5-HIAA among violent offenders whose index crime was classified as "impulsive" (i.e., no apparent plan) were significantly lower than those among violent offenders whose index crime was classified as "nonimpulsive" (i.e., knew the victim and planned the crime) (7). The hypothesis that "impulsiveness" is the key behavioral correlate of reduced 5-HT activity was later advanced in a series of studies from Virkkunen et al. (8 ,15 ,16). In these studies, CSF 5-HIAA concentrations of "impulsive" arsonists were reduced to the same degree as in "impulsive violent offenders" and were significantly lower than those observed in healthy volunteers. Because both impulsive arsonists and impulsive violent offenders had (theoretically) "impulsivity" in common, these investigators proposed that the key correlate to reduced CSF 5-HIAA was impulsivity as opposed to aggression. Closer inspection of the data reveals, however, that impulsive "arsonists" and impulsive "violent offenders" also shared a general history of impulsive aggression (e.g., similar rates of IED and suicide attempts), and, accordingly, it may be premature to conclude that the key behavioral correlate of reduced central 5-HT function is impulsivity rather than a combined construct of impulsive aggression. An inverse relationship between CSF 5-HIAA and aggression or impulsivity has also been reported in male patients with alcoholism (17), in behaviorally disruptive male children and adolescents (18), and in rhesus (19) and pigtailed macaques (20).

Despite these data, some studies have not replicated the finding of an inverse relationship between CSF 5-HIAA and aggression (21 ,22 ,23 ,24 and 25). The reason probably is the presence of subjects who are less severe in their aggressive behavior. Because the relationship between CSF 5-HIAA and aggression appears to be direct (rather than inverse) in several of these studies, it is possible that the direction, as well as the magnitude (17), of this relationship is a function of the severity of aggression. A direct relationship between lumbar CSF 5-HIAA concentration and aggression in these subjects suggests that aggressiveness may be associated with increased (rather than decreased) intrasynaptic concentrations of 5-HT. These data could be consistent with a deficiency hypothesis of 5-HT for aggression if postsynaptic 5-HT function is reduced by compensation. If postsynaptic 5-HT function is unchanged (or increased), however, these data would suggest increased "net" 5-HT function in these subjects. Evidence examining these hypotheses is discussed in the following sections.

5-HT and Aggression: Pharmacochallenge Studies

Coccaro et al. first reported a blunted prolactin D,L-fenfluramine (PRL[D,L-FEN]) response in drug-free mood-disordered and personality-disordered patients compared with healthy volunteers (24 ,26 ,27). In addition, patient subjects with a history of a suicide attempt displayed blunted PRL[D,L-FEN] responses compared with those without this history. Personality-disordered, but not mood-disordered, patients also displayed an inverse relationship between various measures of impulsive aggression (but not depression) and PRL[D,L-FEN] responses. Because experimental reduction in norepinephrine (NE) activity has been shown to eliminate the expected aggressive behavior of animals with reduced 5-HT (28), a reduction in NE system function in the mood-disordered (29), but not the personality-disordered (30), subjects could have mitigated the influence reduced 5-HT function could be expected to have on the expression of aggressive behavior. For the depressed patient, reduced NE system function may be associated with a reduction in efficiency to attend to novel (e.g., aversive) stimuli. If so, only the most potent stimuli (e.g., suicidal ideation) would be likely to trigger a behavioral action that could be poorly inhibited by a dysfunctional central 5-HT system. Further study noted that PRL[D,L-FEN] responses were inversely related to CSF 5-HIAA and were directly related to PRL[meta-chlorophenylpiperazine or m-CPP] responses, an index of postsynaptic receptor activation (25). When these 5-HT measures were examined in the same personality-disordered subjects, a relationship between 5-HT and aggression was noted for both PRL[D,L-FEN] and PRL[m-CPP] responses (which were also directly correlated) but not for CSF 5-HIAA, a finding suggesting that the 5-HT–aggression relationship, as detected by PRL[D, L-FEN] response, may be caused by a reduction in the sensitivity of the postsynaptic 5-HT receptor. Available data suggest that PRL responses to FEN reflect the activation of central 5-HT_{2C}, but not 5-HT_{1A} or 5-HT_{3}, receptors (27 ,31 ,32); other 5-HT receptors have not been studied in this regard. 5-HT_{1A} receptors may still play a role in human aggression as evidenced by inverse relationships noted between

aggression and physiologic responses to buspirone (33 ,34) and ipsapirone (35 ,36) challenge.

Support for these pharmacochallenge findings has been reported in patients with personality disorders (37), patients with alcoholism (38 ,39), suicidal patients (40), violent offenders (41), healthy volunteers from the community (42), and macaques (43). Nonreplication studies involve subjects with history of primarily nonalcoholic substance abuse (44 ,45) and children with disruptive behavior disorders (46 ,47), in whom positive correlations between D,L-FEN challenge and aggression variables have been reported in several, although not all, studies (48 ,49 and 50). In substance abusers, it is possible that nonalcoholic drugs of abuse modify the neurobiological substrate of subjects so correlations between 5-HT and measures of impulsive aggression are direct rather than inverse, as are seen in patients with alcoholism (38 ,39). In children, two studies reported a positive correlation between aggression and PRL[D,L-FEN] response (46 ,47), and one reported a negative correlation between aggression and thermal [D,L-FEN] responses (50). In adolescents or older children, two studies reported no correlation between PRL[D,L-FEN] and aggression (48 ,49). It is possible that changes in the 5-HT system occurring over development affect the nature of the 5-HT–aggression relationship in that this relationship is positive in some 5-HT–mediated pathways, such as the PRL[D,L-FEN] response, in prepubertal children, is absent in postpubertal children, and is inverse in adults. The neurobiological mechanisms underlying this hypothesis are unknown, although, in animal models, the overexpression of 5-HT early in development has been shown to lead to the down-regulation of the 5-HT system later in life (52).

5-HT and Aggression: Peripheral Marker Studies

Studies examining 5-HT–aggression relationships using peripheral indices of 5-HT are relatively limited. The platelet 5-HT transporter (or 5-HT uptake activity) has been assessed in regard to aggression in children, adolescents, and adults and is the one peripheral 5-HT measure to demonstrate some consistency across studies. Positive findings in children and adolescents include the studies of Stoff et al. (53) and Birmaher et al. (54), in which platelet ^{3}H-imipramine (B_{max}) binding was lower in aggressive subjects. In adult subjects, three of four studies reported inverse relationships between platelet 5-HT transporter binding and aggression or impulsivity in personality-disordered subjects (55 ,56) and aggressive institutionalized adults (57); the fourth study reported an increase in platelet 5-HT transporter binding in criminal offenders when compared with normal control subjects (58). Two studies examined the function of the platelet 5-HT transporter, with one demonstrating a reduction of platelet 5-HT uptake in aggressive adult subjects and an inverse relationship with impulsivity (59), and a second study demonstrating no differences in platelet 5-HT uptake in boys with and without disruptive behavior disorders in a study in which no difference was found in the number of platelet ^{3}H-imipramine binding sites between these two groups of subjects (60). Studies examining the platelet 5-HT_{2A} receptor in aggression are few, and results are mixed, with one study reporting no relationship (61), another reporting a negative relationship (62), and yet another reporting a positive relationship (63) between this 5-HT receptor and aggression.

Published studies of blood and platelet 5-HT and plasma tryptophan in humans are relatively few and inconsistent in their results. Whole-blood 5-HT concentrations have been reported as elevated in juvenile offenders compared with normal control subjects (64) and as a function of age of onset (65). A positive correlation between platelet 5-HT concentration and measures of aggression in adult depressed patients (66) has also been reported. Negative studies, however, include those performed in mentally retarded adults (67) and in children with attention-deficit/hyperactivity disorder (68). The ratio of plasma tryptophan to other competing neutral amino acids was lowest among patients with alcoholism with a history of depression or aggression and was lowest among those patients with alcoholism with a history of both depression and aggression in two studies (69 ,70). Other studies reported elevated levels of plasma tryptophan (or the tryptophan ratio to neutral amino acids) in violent offenders (71 ,72) or positive correlations with aggression in healthy volunteers (73).

5-HT and Aggression: Behavioral Studies

Behavioral measures of aggression have been available for many years and include paradigms in which subjects are instructed to deliver a noxious stimulus (e.g., electric shock or loud noise; Taylor Aggression Paradigm) (74) to, or to take monetary points away from (Point Subtraction Aggression Paradigm or PSAP) (75), “confederate” subjects under specific social conditions. In these paradigms, the amount of the noxious stimuli delivered to (or points subtracted from) the confederate represents the subject’s tendency to behave aggressively. In a study of 14 male personality-disordered subjects (56), both the PSAP and a life history measure of aggression correlated inversely with PRL[$_{d}$-FEN] responses. Another study reported a similar finding with respect to 5-HT_{1A} receptor function (36). In this study, high PSAP (“high aggressive”) responders had blunted thermal responses to ipsapirone challenge compared with low PSAP (“low aggressive”) responders.

Experimental studies in which 5-HT activity is manipulated and aggressive responding is monitored have been conducted in research volunteers without documented psychopathology. Four studies in which brain 5-HT was putatively manipulated by tryptophan depletion, supplementation, or both (76 ,77 ,78 and 79) reported data consistent with an inverse relationship between 5-HT activity and aggressive responding

in the laboratory, although one suggested that this effect is restricted to a subgroup of aggressive subjects (79). The one negative study in this area did not use a laboratory paradigm in which provoked aggression could be assessed (80). Studies in which 5-HT activity was acutely increased by using either single doses of D,L-FEN or of the 5-$HT_{1A/1B}$ agonist eltoprazine (34 ,81) reported a reduction in aggressive responding on behavioral paradigms. For D,L-FEN, but not eltoprazine, this result was specific to aggressive responding (as opposed to both aggressive and nonaggressive responding in the case of eltoprazine). These data are consistent with clinical trial data using 5-HT enhancing agents, discussed later (82).

Catecholamines

Based on animal studies, increased noradrenergic (NE) and dopaminergic (DA) activity could be hypothesized to facilitate aggressive responding in humans (83 ,84). Brown et al. reported a positive correlation between CSF methoxyhydroxyphenylglycol (MHPG), but not CSF homovanillic acid (HVA), concentrations and life history of aggression (12). Further analysis revealed, however, that CSF 5-HIAA accounted for 80% of the variance in aggression scores. Plasma NE was modestly, but positively, correlated with self-reported impulsivity in male personality-disordered subjects in another study (85). However, significant reductions in CSF MHPG in impulsive violent offenders was reported in one study (15), although not in a later study by the same investigators with a much larger sample (8). NE pharmacologic challenge studies in this area have been limited and include a positive correlation between the growth hormone response to the α_2-NE agonist clonidine and self-reported “irritability” (a correlate of aggression) in a small sample of male personality-disordered and healthy volunteer subjects (30). A role of α_2-NE receptors in aggression has been suggested by animal data in which the intrahypothalamic injection of α_2-NE agents enhances aggressive responding in the cat (86). The authors suggested that one putative mechanism underlying this finding could involve stimulation of α_2-heteroceptors on presynaptic 5-HT neurons thereby inhibiting 5-HT outflow.

Support for a DA hypothesis of human aggression is also limited. Although some studies reported no relationship between CSF HVA and aggression (12 ,15), other studies suggested the presence of an inverse relationship between these variables. A reduction in CSF HVA in antisocial, though not “explosive,” impulsive violent offenders was reported in one study (7). A reduction in CSF HVA has also been reported among recidivist violent offenders in comparison with nonrecidivist violent offender controls (16), a finding suggesting that reduced dopaminergic function plays a role in predicting future aggressive behavior. These findings must be taken with caution, however, because an inverse relationship between CSF 5-HIAA and aggression was also present in each of these studies. Because CSF 5-HIAA may “drive” CSF HVA (87), it is possible that these findings are related to a more primary relationship between CSF 5-HIAA and aggression. Conversely, an imaging study of striatal dopamine transporters in human subjects reported greater heterogeneity in these receptors in impulsive violent offenders compared with control subjects (88), a finding suggesting that a reduction in CSF HVA may not be secondary to alterations in 5-HT function.

Neurosteriods

Testosterone

Testosterone and related androgens generally play a facilitative role in aggressive behaviors (see refs. 89 and 90 for review). Positive correlations between plasma testosterone concentrations and measures of aggression have been reported, although not entirely consistently in nonpsychiatric subjects (see refs. 89 and 91 for review). Correlations have also been reported in volunteers between reports of relatives and spouses of their global aggressive behavior and both testosterone and NE (92), and basal testosterone levels have been reported to be higher in subjects with high-normal than those with low-normal aggressiveness (93).

Plasma testosterone levels have also been reported to be higher in psychiatric and criminal populations characterized by high aggression. For example, male criminals with personality disorders had significantly higher levels of circulating testosterone than criminal patients with schizophrenia (94), and high free testosterone concentrations were associated with increased aggression in Finnish violent offenders with alcoholism (8). Plasma concentrations of testosterone appear to be higher in persons with alcoholism with a history of repeated episodes of domestic violence than in comparison groups (95). In criminal offenders, higher CSF testosterone concentrations were found in antisocial impulsive violent offenders, but they were not found in nonantisocial impulsive or nonimpulsive violent offenders, in comparison with a healthy volunteer control group.

There are some reports from in prospective, blinded studies suggesting that administration of exogenous testosterone may result in aggressive behaviors (96), but the percentage experiencing severe mental disturbances is likely to be small (97 ,98 and 99). Anabolic steroid administration may not be uncommon among prisoners (100), and it may induce abnormal personality traits in body builders (101). Naturalistic studies of testosterone concentrations are limited in their interpretation because of the pulsatile nature of testosterone release, so particularly plasma concentrations may be quite variable, whereas CSF may be more reflective of average values. Studies of exogenous steroid administration are complicated by, in most cases, the uncontrolled nature of this steroid use.

Cortisol

In general, cortisol concentrations are reported to be relatively low in aggressive individuals. This finding is consistent with the negative correlation between cortisol and testosterone concentrations (102) in volunteer subjects under controlled conditions. Correlations have been found among plasma PRL, testosterone, and aggression, but not with cortisol, although cortisol and PRL concentrations were correlated with each other in their day-to-day changes in one study (103). In children, increases in cortisol during the day were correlated with increased aggression (104). A low concentration of salivary cortisol was associated with persistent aggression in boys referred because of disruptive behavior (105). In criminal offender or antisocial populations, reduced urinary-free cortisol and CSF adrenocorticotropic hormone (ACTH) concentrations are found compared with healthy volunteers (8). Plasma cortisol has also been reported to be reduced in persons with alcoholism and a history of repeated domestic violence (95) but increased after cessation of drinking in incarcerated persons with alcoholism (106). However, the data in both animals and man are inconsistent because a positive correlation has been observed between plasma ACTH and aggression in nonhuman primates (19), and cortisol rises during competition for dominance in vervet monkey males (107). The relationship between cortisol and aggression is thus likely to be complex and dependent on social context and stress.

Neuropeptides

Among peptides neurotransmitters and modulators, limited data suggest a positive relationship between aggression and central vasopressin and central opioid activity.

Vasopressin

Whereas Virkkunen et al. reported no difference in CSF vasopressin concentrations among impulsive and nonimpulsive violent offenders (8), Coccaro et al. reported a positive correlation between CSF vasopressin concentration and life history of aggression in 26 personality-disordered persons, particularly men (108). Despite a significant inverse correlation between CSF vasopressin and PRL[$_d$-FEN] response, the positive relationship between aggression and CSF vasopressin remained even after the influence of PRL[$_d$-FEN] response on the aggression score was taken into account. These data are consistent with those from animal studies in which vasopressin antagonists reduced aggression in golden hamsters, whereas 5-HT uptake inhibitors increased central 5-HT activity and reduced central vasopressin concentration and levels of aggressive behavior in the same species (109). Differences between the two human studies may be accounted for by significant differences in the sample population (e.g., criminally versus noncriminally violent subjects). It is possible that agents that dampen central vasopressinergic activity could have antiaggressive efficacy.

Opiates

Although opiate withdrawal may precipitate aggressive behavior, there has been little study of the relationship of aggression with endogenous opiates. In one study, a CSF opioid-binding protein was positively correlated with “assaultiveness” in healthy male volunteers (110). Circulating levels of metenkephalins have been associated with self-injurious behaviors in a limited number of studies (111). Postmortem brain studies of violent suicide victims have also found greater numbers of μ-opioid receptors as well (112). Male undergraduates receiving 45 mg of oral morphine tablets or placebo displayed more aggression in the morphine-induced than the placebo conditions (113), whereas naltrexone, an opiate blocker, attenuated self-injurious behavior (114). These studies suggested that increased opioid activity may increase the likelihood of aggressive behavior.

Cholesterol and Fatty Acids

Some studies, both from persons who attempted suicide and from healthy persons, suggested that reduced serum cholesterol may be associated with aggressive behavior (115 ,116). Male monkeys randomized to a low-fat and low-cholesterol diet displayed more aggressive behavior and less prosocial behaviors than those randomized to a high-cholesterol diet (117 ,118). In humans, pharmacologic reduction of cholesterol may increase the risk of non–illness-related mortality such as death by suicide or trauma related to aggression (119). Naturally occurring reduced cholesterol may also be associated with non–illness-related mortality (116 ,120 ,121 ,122 and 123), largely attributable to suicide (116 ,122). Low serum cholesterol has been reported in psychiatric inpatients associated with suicide attempts (124 ,125 ,126 and 127). Reduced serum cholesterol has also been related to the severity of borderline personality disorder traits (115), as well as to antisocial personality disorder and the violence associated with it (128 ,129). It is also found in male forensic patients (130), aggressive conduct-disordered children and adolescents with attention deficit disorder (131), and suicidal adolescents (132).

MOLECULAR GENETICS

Part of "119 - Pathophysiology and Treatment of Aggression "

An understanding of the molecular genetics of impulsive aggression is currently emerging with the rise of association studies involving various DNA polymorphisms of candidate genes. One of the first notable studies in this area was that of Brunner et al. (133), who reported an association between a point mutation for the monoamine oxidase A (MAO A)

gene in an extended family pedigree and impulsive violence in males of low intelligence. The presence of this mutation was associated with evidence of altered catecholamine metabolism (i.e., reduced brain catecholamine breakdown and increased activity at central catecholamine receptors). Although no other families with this specific MAO A point mutation have been reported, this report highlighted the potential of the candidate gene approach to the molecular genetics of aggression. At about the same time, Nielson et al. reported that the presence of the L allele for the (intronic) biallelic tryptophan hydroxylase (TPH) polymorphism was associated with a reduction of CSF 5-HIAA concentration in impulsive violent offenders (nearly all with DSM-III IED) (134). In the same study, the presence of the L allele was also associated with history of suicide attempts in all violent offenders. Although this finding was not replicated by Abbar et al. (135), using a different TPH polymorphism, Nielson et al. did replicate this finding in a second group of violent offenders (136), specifically in impulsive violent offenders and more specifically for severe suicide attempts. Linkage for this TPH polymorphism was also noted in a sib-pair analysis in this same report. New et al. reported a linear association between the TPH genotypes and dimensional measures of impulsive aggression (137). In this brief report of only 21 personality-disordered subjects, those with the LL genotype had significantly higher aggression scores than subjects with the UU genotype. However, an association with the U allele of the TPH polymorphism in patients with a history of suicide attempts has also been reported (138), as has an association with the U allele and aggression scores in community-recruited healthy volunteers (139). It may be that the TPH polymorphism is in linkage disequilibrium with different genes in different populations. Lappalainen et al. reported an association between “antisocial alcoholism” (i.e., alcoholism with antisocial personality disorder or DSM-III IED) and the C allele biallelic polymorphism for the 5-$HT_{1D\beta}$ receptor (140). A study of a 5-HT_6 receptor allelic variant in patients with schizophrenia and in controls was negative for an association with aggressive behavior (141). Finally, Lachman et al. (142) reported replication of a report by Strous et al. (143) of an association between the “low activity” allele of a biallelic polymorphism for catechol-O-methyltransferase (an enzyme that degrades NE) and violence in patients with schizophrenia.

NEUROPSYCHOLOGY OF AGGRESSION

Part of "119 - Pathophysiology and Treatment of Aggression "

The relationship between aggression and neuropsychology is in part dependent on the syndrome in which aggression is observed. For example, the cognitive impairment of dementia may be associated with aggressive behavior. In adolescents with conduct disorder, verbal processing deficits are associated with greater aggressiveness and antisocial behavior (144). Low executive cognitive function is also related to aggressive antisocial behavior (145 ,146). Low executive function also contributed to failures to inhibit responses to stimuli associated with punishment on a “go/no-go” learning task, and poor “self-awareness” possibly related to right cerebral dysfunction has also been related to increased hostility. Verbal signal decoding and P300 amplitudes in an evoked potential paradigm predicted impulsiveness and anger in prison inmates (147). In terms of regional localization, neuropsychological tasks sensitive to frontal and temporal dysfunction have best characterized aggressive antisocial subjects (148 ,149). Thus, neuropsychological and cognitive studies do suggest that abnormalities of higher integrative functions, consistent with reduced cortical inhibitory influences on aggression, result in more disinhibition of aggressive behaviors. As described earlier, certain laboratory paradigms may discriminate aggressive individuals from comparison groups including the PSAP (150) and a “go/no go” version of the Continuous Performance Task (151). The PSAP has been externally validated in violent and nonviolent male parolees, in that violent parolees emit more aggressive responses than nonviolent parolees; furthermore, the number of aggressive responses correlated with other psychometric measures of aggression (76). However, the heritability of these laboratory measures has not been systematically assessed in studies of families or sibs of impulsive or aggressive probands, a logical prerequisite to an endophenotypic approach to borderline personality disorder.

NEUROIMAGING OF AGGRESSION

Part of "119 - Pathophysiology and Treatment of Aggression "

Neuroanatomy of Aggression

Prefrontal cortex, particularly prefrontal orbital cortex and adjacent ventral medial cortex, appears to play a central role in the regulation of aggressive behavior, but temporal cortex, cingulate cortex, and amygdala may also play important roles in the generation of aggression as well. The critical role of prefrontal orbital cortex is exemplified by the case of Phineas Gage, a solid, upstanding railroad worker, who, after a penetrating injury to his orbital frontal cortex, became irritable, hostile, and displayed poor social judgment. Careful reexamination of his skull led to a reconstruction of the location of the lesion at the anterior and mesial aspects of the orbital cortex as well as anterior cingulate and anterior mesial aspects of frontal cortex superior to orbital cortex, with more marked damage in the left hemisphere (152). Other clinical cases support the central role of orbital prefrontal cortex in regulation of aggression (153 ,154 ,155 ,156 and 157). Irritability and angry outbursts have also been associated with damaged orbital frontal cortex in neurologic patients (158), and frontal and temporal hypoperfusion has been noted in frontotemporal dementias (159). Lesions of prefrontal cortex, particularly orbital frontal cortex, early in childhood can result in antisocial disinhibited, aggressive behavior later in life (160).

Temporal lobe lesions have also been associated with a susceptibility to violent behavior, as suggested by multiple case reports of patients with temporal lobe tumors. In one study of violent patients, many anterior inferior temporal lobe tumors were reported (161 ,162), and aggressive behavior has been associated with temporal lesions (163). Although temporal disease may express itself in a variety of ways, there does appear to be a clear association between temporal pathology and aggressive behavior.

The amygdala is also implicated in the regulation of aggression both in electrical stimulation studies, associated with rage attacks, and studies of patients who have undergone amygdalectomy (164), although destructive behaviors have also been observed in the context of coagulation of the amygdala (165). Patients with bilateral amygdala damage judged unfamiliar persons to be more trustworthy than controls, a finding consonant with the role of the amygdala in social judgments of potential threat (166).

The association of violent behavior with aggressive behavior with localized seizure activity provides a further guide to brain regions implicated in the modulation of aggression. Aggressive behavior has been found to be associated with frontal lobe seizure activity (167 ,168) and temporal lobe seizures (169). However, only a few patients with temporal lobe epilepsy engage in aggressive behaviors in the interictal or periictal periods (170 ,171 and 172). These clinical correlations, although pointing to regions of interest for imaging studies, cannot directly address the circuitry involved in impulsive aggression in the absence of specific neurologic disease. These data do support the importance of prefrontal, temporal, and limbic cortex in the regulation of aggression.

Structural Imaging and Aggression

Reduced prefrontal gray matter has been associated with autonomic deficits in patients with antisocial personality disorders characterized by aggressive behaviors (173). Although these deficits are not visually perceptible, they reach statistical significance and are consistent with the neurologic literature (described earlier) and functional imaging data (described in the next section).

Functional Imaging and Aggression

One technique used to identify brain activity in individuals displaying aggressive behavior is the assessment of *in vivo* cerebral glucose metabolism through positron emission tomography. Studies of this type tend to implicate brain hypometabolism in a variety of regions but particularly frontal and temporal cortex. In psychiatric patients with a history of repetitive violent behavior, decreased blood flow consistently has been found in temporal cortex and, to some extent, in frontal cortex (174 ,175). In a study of homicide offenders, bilateral diminution of glucose metabolism was observed in both medial frontal cortex and at a trend level in orbital frontal cortex (176). These deficits were more pronounced in persons without psychosocial deprivation (177). In a study of patients with personality disorders, an inverse relationship was found between life history of aggressive impulsive behavior and regional glucose metabolism in orbital frontal cortex and right temporal lobe. Patients meeting criteria for borderline personality disorder had decreased metabolism in frontal regions corresponding to Brodmann's areas 46 and 6 and increased metabolism in superior and inferior frontal gyrus (Brodmann's areas 9 and 45) (178). Single photon emission computed tomography studies have also suggested reduced perfusion in prefrontal cortex, as well as focal abnormalities in left temporal lobe and increased activity in anteromedial frontal cortex in limbic system in aggressive persons with reduced prefrontal perfusion in antisocial personality-disordered alcoholism (179), and hypoperfusion in the left frontoparietal region associated with attacks of bizarre, impulsive behaviors (180). Cingulate cortex has also been implicated especially in posterior regions in aggressive borderline patients (178), a finding consistent with the putative role of cingulate cortex in the control of affective evaluation of incoming stimuli (134).

Extensive connections between amygdala and prefrontal cortex have been described, suggesting an inhibitory influence of frontal cortex on the amygdala (181). Amygdalotomy has been associated with reduced aggressive outbursts in patients with intractable aggression (182), but there have been no direct imaging studies to date reporting the relationship of amygdala activity and aggression.

Imaging Neurotransmitter Systems in Aggression

Serotonin

Ascending serotonergic neurons from the raphe nuclei project widely throughout the brain, including projections to dorsolateral prefrontal cortex and medial temporal lobe. Dorsal raphe-median forebrain bundle also directly innervates the amygdala, where dorsal raphe tracts outside the medial forebrain bundle project to parietotemporal cortex. Diffuse tracts extend from dorsal and medial raphe project to frontal lobe. Both 5-HT_{2A} and 5-HT_{1A} receptors are found in high concentrations in human prefrontal cortex, as are 5-HT transporter sites (183), and patients with localized frontotemporal contusions show significantly lower 5-HT metabolites in CSF than patients with diffuse cerebral contusions (184). Greater ß-CIT binding to 5-HT transporters has also been reported in nonhuman primates with a higher ß-CIT binding associated with greater aggressiveness (185). 5-HT_{2A} receptor number has been inversely related to aggressive behavior in posterior orbital frontal cortex and medial frontal cortex in the amygdala, whereas increased 5-HT_{2A} number in orbital frontal cortex, posterior temporal

cortex, and amygdala have been correlated with prosocial behavior in primates (186). Thus, serotonergic modulation of frontal and temporal cortical activity by 5-HT receptors, possibly of the 5-HT_{2A} type, may be particularly important in aggression.

The administration of FEN has been shown to increase cortical metabolism in frontal, temporal, and parietal cortex (187 ,188 and 189). In a study of depressed patients that included patients with a comorbid diagnosis of borderline personality disorder and a history of suicide attempts, activation of cortex including orbital and cingulate cortex was significantly blunted in the depressed patients, particularly in those who attempted suicide, compared with the control subjects. The depressed patients showed no significant changes in their glucose metabolic response to FEN compared with placebo, in contrast to the controls (189). In another study, intravenous administration of m-CPP in patients with alcoholism resulted in blunted glucose metabolic responses in right orbital frontal cortex, left anterolateral prefrontal cortex, posterior cingulate cortex, and thalamus compared with controls (190). In the first study directly comparing glucose metabolism after FEN and placebo in personality-disordered patients with impulsive aggression, neurologically normal subjects showed increased metabolism in orbital frontal and adjacent ventral medial frontal cortex as well as cingulate and inferior parietal cortex after FEN compared with placebo, whereas impulsive-aggressive patients appeared to show significant increases only in the inferior parietal lobe. Between-group comparisons demonstrated blunted responses of glucose metabolism in orbital frontal, ventral medial frontal, and cingulate cortex in the impulsive personality-disordered patients compared with the neurologically normal subjects. This study's results were replicated in a study of patients with borderline personality disorder (191), who displayed reduced regional uptake of fluorodeoxyglucose (relative to placebo) compared with control subjects in right medial and orbital frontal cortex, left middle and superior temporal gyri, left parietal lobe, and left caudate. In more recent pilot data from a study of patients with impulsive-aggressive personality disorders and controls that evaluated glucose metabolism after the administration of the 5-HT_2 agonist m-CPP, reduced metabolic responses were found in the aggressive patients, particularly in orbital frontal cortex, compared with controls (192), a finding inviting more direct assessment of components of serotonergic activity such as 5-HT_{2A} receptor number, transporter site number, and 5-HT_{2A} receptors.

In summary, imaging studies of the 5-HT system in impulsive-aggressive patients suggest reduced activation by ascending serotonergic projections on critical cortical inhibitory regions such as orbital frontal and related medial frontal cortex (137). Reduced serotonergic activation of these inhibitory regions mediated in part through 5-HT_{2A} receptors, but probably by other serotonergic mediators as well, may have a disinhibiting effect on the generation of aggression by amygdala and related structures.

Dopamine

In animal studies, 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine–induced unilateral striatal dopamine deficiency in vervet monkeys was associated with increased frequency of aggressive behaviors toward other members of the group in the monkey colony (193). Greater heterogeneity was also found in striatal dopamine transporter density, as assessed by ^{123}I (β-CIT distribution) of impulsive violent offenders than controls (88), a finding possibly consistent with hypotheses that aggressive behavior is associated with increased dopaminergic transmission

PHARMACOLOGIC TREATMENT OF AGGRESSION

Part of "119 - Pathophysiology and Treatment of Aggression "

The rational clinical psychopharmacology of aggressive behavior began in the mid-1970s with the first placebo-controlled, double-blind, study of lithium carbonate in prison inmates (9). In this study, impulsive, but not premeditated (or other antisocial behavior), aggression was reduced to extremely low levels during a 3-month course of treatment with lithium carbonate; levels of aggression remained unchanged in inmates treated with placebo. Notably, all gains were lost within a month after a switch to placebo. An antiaggressive effect of lithium was replicated in subsequent studies including a blinded placebo-controlled trial in hospitalized aggressive children with conduct disorder (194) and a blinded, placebo-controlled trial of 42 mentally disabled patients (195). The mechanism of action for lithium in this regard is unknown, but it likely includes an enhancement of 5-HT function and a dampening of catecholaminergic function.

Other agents that may have antiaggressive efficacy include 5-HT–enhancing agents (i.e., 5-HT selective uptake inhibitors and 5-HT_{1A} agonists), anticonvulsants, typical and atypical neuroleptics, ß-blockers, and antiandrogenic agents, among others.

The rationale to treat impulsive-aggressive patients with 5-HT–enhancing agents is rooted in the consistent findings of an inverse relationship between 5-HT and impulsive aggression, reviewed earlier. Since the early 1990s, numerous open and blinded, placebo-controlled, studies have documented the efficacy of these agents, specifically with respect to the 5-HT selective uptake inhibitors (SSUIs). Among the controlled trials, SSUIs have been shown to reduce verbal and nonassaultive physical aggression in personality-disordered patients selected for a history of recurrent, problematic, impulsive-aggressive behavior (82), to reduce nonassaultive physical aggression in patients with borderline personality disorder who were recruited from the community (196),

to reduce anger attacks in depressed patients undergoing a clinical trial for the treatment of aggression (197), and to reduce impulsive aggression in adults with autistic disorder (198).

Although enhancement of central 5-HT function by the SSUI is presumed to underlie the antiaggressive effect in these subjects, the one study that examined 5-HT function before treatment actually found a positive relationship between pretreatment 5-HT function, assessed by PRL[$_d$-FEN] response, and improvement in aggression scores at end of trial (199). These data suggest that SSUIs may work best in patients whose postsynaptic 5-HT receptors are normal, or only moderately impaired, in function. If so, other agents that do not work primarily on presynaptic neurons may be necessary in patients with severe impairment of postsynaptic neurons. Such agents could include 5-HT receptor agonists or anticonvulsants, for example. Although evidence for the antiaggressive efficacy for 5-HT$_{1A}$ agonists is limited, buspirone, at doses of 20 to 50 mg per day, was shown to reduce aggression in two studies of mentally retarded patients (200 ,201). More data, however, are available to support the antiaggressive efficacy of anticonvulsants.

Carbamazepine has been shown in blinded, placebo-controlled, trials to reduce episodes of behavioral dyscontrol markedly in borderline personality disorder (202) and to reduce agitation and aggression in nursing home patients with dementia (203), although not in children with conduct disorder (204). Phenytoin was also shown to reduce impulsive, but not premeditated, aggressive behavior in a blind, placebo-controlled, study of prison inmates (10). Divalproex sodium, another anticonvulsant, showed promise in open (205 ,206 ,207 and 208) and early blinded, placebo-controlled, trials (50) of personality-disordered patients and conduct-disordered adolescents. Open trials also suggested that this agent may reduce behavioral agitation and mood lability in elderly demented patients (209 ,210) and in patients with aggression and mood lability secondary to brain trauma (211).

The use of typical neuroleptics in the treatment of aggression has largely been aimed at sedating psychotic patients or to reduce the psychotic thinking that may trigger aggressive behavior. In addition, these agents have also been used in nonpsychotic patients to treat aggression and agitation, with mixed results. Thioridazine was reported to diminish impulsive behavior in an open-label study of patients with borderline personality disorder (212), and low-dose haloperidol was reported to reduce hostility and impulsivity compared with amitriptyline and placebo in patients with borderline or schizotypal personality disorder (213). These findings were not reproduced in two other double-blind, placebo-controlled studies of borderline and schizotypal personality-disordered patients treated with thiothixine (214) or trifluoperazine (202). Treatment with the newer, atypical neuroleptics may prove to be more effective than that with the earlier typical antipsychotic agents. Antagonism of 5-HT$_2$ receptors appears to decrease aggression in animal models, and this effect may explain the ability of newer antipsychotic agents (which, unlike the older medications, block 5-HT$_2$ receptors) to produce a reduction in aggression and agitation independent of effects on psychotic symptoms (215 ,216). Studies suggest that the overall frequency of assaults, use of seclusion, mechanical restraint, and chemical restraint in patients with schizophrenia who are treated with clozapine are reduced over traditional neuroleptics (217). In a double-blind study, risperidone had a greater selective effect on hostility than haloperidol or placebo in patients with schizophrenia (218). Finally, an open-label study of olanzapine in 11 patients with borderline personality disorder reported significant reductions in anger (219), a finding suggesting that the potential benefit of atypical neuroleptics in treating aggression may extend to nonpsychotic patients as well.

Given the potential facilitory role of the central noradrenergic system, agents that dampen the function of this system could be expected to have antiaggressive efficacy. Notably, ß-noradrenergic blockers have been found effective in reducing aggressive behavior in patients with organic brain syndromes or chronic psychosis. Propranolol has been shown to reduce aggressive behavior in patients with traumatic brain injuries (220 ,221) or in patients with dementia (222). Both propranolol and nadolol have been shown to be effective in reducing aggressive behavior in chronic psychiatric inpatients, independent of psychotic symptoms (223 ,224). The use of these medications is limited, however, by hypotension and bradycardia, which can be side effects at the higher doses that are often used in these cases.

The use of antiandrogens in the treatment of some types of aggression, specifically sexual aggression, has undergone limited study. Antiandrogens such as medroxyprogesterone acetate and cyproterone acetate appear to lower both deviant and nondeviant sexual drive and activity in men with paraphilias, and this behavioral improvement is associated with decreases in testosterone level (225). However, although no data support the routine use of antiandrogens in nonsexual aggressive behavior, patients with sexual aggression do appear to respond, in some cases, to antiandrogen treatment.

SUMMARY

Part of "119 - Pathophysiology and Treatment of Aggression "

The study of the pathophysiology and pharmacologic treatment of aggression has undergone much progress since the 1980s. Extensive evidence supports an important role for central 5-HT function in the regulation impulsive aggressive behavior. In addition, more is known about potential regulatory roles of other central neurotransmitters and modulators, as well as their possible interaction with 5-HT. This knowledge has led to the development of a more rational approach to the psychopharmacologic treatment of impulsive aggression (e.g., treatment with 5-HT uptake inhibitors). Recent work examining the relationship of DNA polymorphisms

and aggression and work examining the relationship between brain structure and function in aggressive persons could yield critical data that will bring our understanding of the pathophysiology of aggression to a new level.

REFERENCES

1. National Center for Health Statistics 1996–97, U.S. Government, Hyattsville, MD, 1997.

2. Robins LN, Regier DA. *Psychiatric disorders in America*. New York: Free Press, 1991.

3. Zimmerman M, Mattia J, Younken S, et al. *The prevalence of DSM-IV impulse control disorders in psychiatric outpatients*. APA new research abstracts no. 265. Washington, DC: American Psychiatric Association, 1998.

4. Coccaro EF. Intermittent explosive disorder. *Curr Psychiatry Rep* 2000;2:67–71.

5. Bergeman CS, Seroczynski AD. Genetic and environmental influences on aggression and impulsivity In: Maes M, Coccaro E, eds. *Neurobiology and clinical views on aggression and impulsivity*. New York: John Wiley, 1998:63–80.

6. Miles, Carey G. Genetic and environmental architecture of human aggression. *J Pers Social Psychol* 1997;72:207–217.

7. Linnoila M, Virkkunen M, Scheinin M, et al. Low cerebrospinal fluid 5-hydroxylindolacetic acid concentration differentiates impulsive from nonimpulsive violent behavior. *Life Sci* 1983;33:2609–2614.

8. Virkkunen M, Rawlings R, Tokola R, et al. CSF biochemistries, glucose metabolism, and diurnal activity rhythms in alcoholic, violent offenders, fire setters, and healthy volunteers. *Arch Gen Psychiatry* 1994;51:20–27.

9. Sheard M, Manini J, Bridges C, et al. The effect of lithium on impulsive aggressive behavior in man. *Am J Psychiatry* 1976;133:1409–1413.

10. Barratt ES, Stanford MS, Felthous AR, et al. The effects of phenytoin on impulsive and premeditated aggression: a controlled study. *J Clin Psychopharmacol* 1997;17:341–349.

11. Asberg M, Traksman L, Thoren P. 5-HIAA in the cerebrospinal fluid: a biochemical suicide predictor? *Arch Gen Psychiatry* 1976;33:1193–1197.

12. Brown GL, Goodwin FK, Ballenger JC, et al. Aggression in human correlates with cerebrospinal fluid amine metabolites. *Psychiatry Res* 1979;1:131–139.

13. Brown GL, Ebert MH, Goyer PF, et al. Aggression, suicide, and serotonin: relationships to CSF amine metabolites. *Am J Psychiatry* 1982;139:741–746.

14. Lidberg L, Tuck JR, Asberg M, et al. Homicide, suicide and CSF 5-HIAA. *Acta Psychol Scand* 1985;71:230–236.

15. Virkkunen M, Nuutila A, Goodwin FK, et al. Cerebrospinal fluid monoamine metabolite levels in male arsonists. *Arch Gen Psychiatry* 1987;44:241–247.

16. Virkkunen M, DeJong J, Bartko J, et al. Relationship of psychobiological variables to recidivism in violent offenders and impulsive fire setters. *Arch Gen Psychiatry* 1989;46:600–603.

17. Limson R, Goldman D, Roy A, et al. Personality and cerebrospinal fluid monoamine metabolites in alcoholics and controls. *Arch Gen Psychiatry* 1991;48:437–441.

18. Kruesi MJP, Rapoport JL, Hamberger S, et al. Cerebrospinal fluid metabolites, aggression, and impulsivity in disruptive behavior disorders of children and adolescents. *Arch Gen Psychiatry* 1990;47:419–462.

19. Higley, JD, Mehlman PT, Taub DM, et al. Cerebrospinal fluid monoamine and adrenal correlates of aggression in free-ranging rhesus monkeys. *Arch Gen Psychiatry* 1992;49:436–441.

20. Westergaard GC, Suomi SJ, Higley JD, et al. CSF 5-HIAA and aggression in female macaque monkeys: species and interindividual differences. *Psychopharmacology (Berl)* 1999;146:440–444.

21. Gardner DL, Lucas PB, Cowdry RW. CSF metabolites in borderline personality disorder compared with normal controls. *Biol Psychiatry* 1990;28:247–254.

22. Castellanos FX, Elia J, Kreusi MJP, et al. Cerebrospinal fluid monoamine metabolites in boys with attention deficit hyperactivity disorder. *Psychiatry Res* 1994;52:305–316.

23. Moller SE, Mortensen EL, Breum L, et al. Aggression and personality: association with amino acids and monoamine metabolites. *Psychol Med* 1996;26:323–331.

24. Coccaro EF, Kavoussi RJ, Trestman RL, et al. Serotonin function in personality and mood disorder: intercorrelations among central indices and aggressiveness. *Psychiatry Res* 1997;73:1–14.

25. Coccaro EF, Kavoussi RJ, Cooper TB, et al. Central serotonin and aggression: inverse relationship with prolactin response to D-fenfluramine, but not with CSF 5-HIAA concentration in human subjects. *Am Journal Psychiatry* 1997;154:1430–1435.

26. Coccaro EF, Siever LJ, Klar HM, et al. Serotonergic studies in affective and personality disorder: correlates with suicidal and impulsive aggressive behavior. *Arch Gen Psychiatry* 1989;46:587–599.

27. Coccaro EF, Kavoussi RJ, Oakes M, et al. Amesergide completely blocks the prolactin response to D-fenfluramine in healthy male subjects. *Psychopharmacology (Berl)* 1996;126:24–30.

28. Hodge GK, Butcher LL. Catecholamine correlates of isolation-induced aggression in mice. *Eu J Pharmacol* 1975;31:81–93.

29. Siever LJ, Trestman RL, Coccaro EF, et al. The growth hormone response to clonidine in acute and remitted depressed male patients. *Neuropsychopharmacology* 1992;6:165–177.

30. Coccaro EF, Lawrence T, Trestman R, et al. Growth hormone responses to intravenous clonidine challenge correlates with behavioral irritability in psychiatric patients and in healthy volunteers. *Psychiatry Res* 1991;39:129–139.

31. Park SBG, Cowen PJ. Effect of pindolol on the prolactin response to D-fenfluramine. *Psychopharmacology (Berl)* 1995;118:471–474.

32. Coccaro, EF, Kavoussi RJ, Cooper TB, et al. 5-HT-3 receptor antagonism by ondansetron does not attenuate the prolactin response to D-fenfluramine challenge in healthy male subjects. *Psychopharmacology (Berl)* 1996;127:108–112.

33. Coccaro EF, Gabriel S, Siever LJ. Buspirone challenge: preliminary evidence for a role for 5-HT-1a receptors in impulsive aggressive behavior in humans. *Psychopharm Bull* 1990;26:393–405.

34. Cherek DR, Moeller FG, Khan-Dawwod F, et al. Prolactin response to buspirone was reduced in violent compared to non-violent parolees. *Psychopharmacology (Berl)* 1999;142:144–148.

35. Coccaro EF, Kavoussi RJ, Hauger RL. Physiologic responses to D-fenfluramine and ipsapirone challenge correlate with indices of aggression in males with personality disorder. *Int Clin Psychopharmacol* 1995;10:177–180.

36. Moeller FG, Allen T, Cherek DR, et al. Ipsapirone neuroendocrine challenge: relationship to aggression as measured in the human laboratory. *Psychiatry Res* 1998;18:31–38.

37. New AS, Trestman RL, Mitroupoulou V, et al. Serotonergic function and self-injurious behavior in personality disorder patients. *Psychiatry Res* 1997;69:17–26.

38. Moss HB Yao JK, Panzak GL. Serotonergic responsivity and behavioral dimensions in antisocial personality disorder with substance abuse. *Biol Psychiatry* 1990;28:325–338.

39. Handlesman L, Holloway K, Kahn RS, et al. Hostility is associated with a low prolactin response to meta-chlorophenylpiperazine in abstinent alcoholics. *Alcohol Clin Exp Res* 1996;20:824–829.

40. Lopez-Ibor JJ, Lana F, Saiz Ruiz J. Conductas autoliticas impulsivas y serotonina. *Actas Luso Esp Neurol Psiquiatr* 1990;18:316–325.

41. O'Keane V, Moloney E, O'Neill H. Blunted prolactin responses to D-fenfluramine in sociopahty: evidence of subsensitivity of central serotonergic function. *Br J Psychiatry* 1992;160:643–646.

42. Manuck SB, Flory JD, McCaffery JM, et al. Aggression, impulsivity, and central nervous system serootnergic responsivity in a nonpatient sample. *Neurpsychopharmacology* 1998;19:287–299.

43. Botchin MB, Kaplan JR, Manuck SB, et al. Low versus high prolactin responders to fenfluramine challenge: marker of behavioral differences in adult male cynomolgus macaques. *Neuropsychopharmacology* 1993;9:93–99.

44. Fishbein DH, Lozovsky D, Jaffe JH. Impulsivity, aggression, and neuroendocrine responses to serotonergic stimulation in substance abusers. *Biol Psychiatry* 1989;25:1049–1066.

45. Bernstein DP, Handlesman L. The neurobiology of substance abuse and personality disorders. In: Ratey JJ, ed. *Neuropsychiatry of behavior disorders*. Oxford: Blackwell Scientific Publications, 1995:120–148.

46. Halperin JM, Sharma V, Siever LJ, et al. Serotonergic function in aggressive and nonaggressive boys with attention deficit hyperactivity disorder. *Am J Psychiatry* 1994;151:243–248.

47. Pine DS, Coplan JD, Wasserman GA, et al. Neuroendocrine response to *d,l*-fenfluraminje challenge in boys: associations with aggressive behavior and adverse rearing. *Arch Gen Psychiatry* 1997;54:839–846.

48. Stoff DM, Pastiempo AP, Yeung JH, et al. Neuroendocrine responses to challenge with *d,l*-fenfluramine and aggression in disruptive behavior disorders of children and adolescents. *Psychiatry Res* 1992;43:263–276.

49. Halperin JM, Newcorn JH, Schwartz ST, et al. Age-related changes in the association between serotonergic function ansd aggression in boys with ADHD. *Biol Psychiatry* 1997;41:682–689.

50. Donovan AM, Halperin JM, Newcorn, et al. Thermal response to serotonergic challenge and aggression in attention-deficit hyperactivity disorder children. *J Child Adolesc Psychopharmacol* 1999;9:85–91.

52. Whitaker-Azmitia PM, Zhang X, Clarke C. Effects of gestational exposure to monoamine oxidase inhibitors in rats: preliminary behavioral and neurochemical studies *Neuropsychopharmacology* 1994;11:125–132.

53. Stoff DM, Pollock L, Vitiello B. Reduction of 3-H-imipramine binding sites on platelets of conduct disordered children. *Neuropsychopharmacology* 1987;1:55–62.

54. Birmaher B, Stanley M, Greenhill L, et al. Platelet imipramine binding in children and adolescents with impulsive behavior *J Am Acad Child Adolesc Psychiatry* 1990;29:914–918.

55. Simeon D, Stanley B, Frances A, et al. Self-mutiliation in personality disorders: psychological and biological correlates. *Am J Psychiatry* 1992;149:221–226.

56. Coccaro EF, Kavoussi RJ, Sheline YI, et al. Impulsive aggression in personality disorder: correlates with 3-H-aroxetine binding in the platelet. *Arch Gen Psychiatry* 1996;53:531–536.

57. Marazziti D, Rotondo A, Presta S, et al. Role of serotonin in human aggressive behavior. *Aggress Behav* 1993;9:347–353.

58. Sarne Y, Mandel J, Goncalves MH, et al. Imipramine binding to blood platelets and aggressive behavior in offenders, schizophrenics and normal volunteers. *Neuropsychobiology* 1995;31:120–124.

59. Brown CS, Kent TA, Bryant SG, et al. Blood platelet uptake of serotonin in episodic aggression. *Psychiatry Res* 1989;27:5–12.

60. Stoff DM, Ieni J, Friedman E, et al. Platelet 3H-imipramine binding, serotonin uptake, and plasma α1 acid glycoprotein in disruptive behavior disorders. *Biol Psychiatry* 1991;29:494–498.

61. McBride PA, Brown RP, DeMeo M, et al. The relationship of platelet 5-HT-2 receptor indices to major depressive disorder, personality traits, and suicidal behavior. *Biol Psychiatry* 1994;35:295–308.

62. Blumensohn R, Ratzoni G, Weizman A, et al. Reduction in serotonin 5HT2 receptor binding on platelets of delinquent adolescents. *Psychopharmacology (Berl)* 1995;118:354–356.

63. Coccaro EF, Kavoussi RJ, Sheline YI, et al. Impulsive aggression in personality disorder: Correlates with ^{125}I-LSD binding in the platelet. *Neuropsychopharmacology* 1997;16:211–216.

64. Pliszka S, Rogeness G, Renner P, et al. Plasma neurocehmistry in juvenile offenders. *J Am Acad Child Adolesc Psychiatry* 1988;27:588–594.

65. Unis AS, Cook EH, Vincent JG, et al. Platelet serotonin measures in adolescents with conduct disorder. *Biol Psychiatry* 1997;42:553–559.

66. Mann JJ, McBride PA, Anderson GM, et al. Platelet and whole blood serotonin content in depressed inpatients: correlations with acute and life-time psychopathology. *Biol Psychiatry* 1992;32:243–257.

67. Greenberg AS, Coleman M. Depressed 5-hydroxyindole levels associated with hyperactive and aggressive behavior. *Arch Gen Psychiatry* 1976;33:331–336.

68. Cook E, Stein MA, Ellison T, et al. Attention deficit hyperactivity disorder and whole blood serotonin levels: effects of comorbidity. *Psychiatry Res* 1995;57:13–20.

69. Branchey L, Branchey M, Shaw S, et al. Depression, suicide, and aggression in alcoholics and their relationship to plasma amino acids. *Psychiatry Res* 1984;12:219–226.

70. Buydens-Branchey L, Branchey MH, Noumair D, et al. Age of alcoholism onset. II. Relationship to susceptability to serotonin precursor availability. *Arch Gen Psychiatry* 1989;46:231–236.

71. Virkkunen M, Narvanen S. Plasma insulin, tryptophan and serotonin levels during the glucose tolerance test among habitually violent and impulsive offenders. *Neuropsychobiology* 1987;17:19–23.

72. Eriksson T, Lidberg L. Increased plasma concentrations of the 5-HT precursor amino acid tryptophan and other large neutral amino acids in violent criminals. *Psychol Med* 1997;27:477–481.

73. Wingrove J, Bond AJ, Cleare AJ, et al. Plasma tryptophan and trait aggression. *J Psychopharmacol* 1999;13:235–237.

74. Taylor SP. Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *J Pers* 1967;35:297–310.

75. Cherek DR, Moeller G, Schnapp W, et al. Studies of violent and non-violent male parolees. I. Laboratory and psychometric measurements of aggression. *Biol Psychiatry* 1997;41:514–522.

76. Pihl R-O, Young S, Harden P, et al. Acute effect of altered tryptophan levels and alcohol on aggression in normal human males. *Psychopharmacology (Berl)* 1995;119:353–360.

77. Cleare AJ, Bond AJ. The effect of tryptophan depletion and enhancement on subjective and behavioural aggression in normal male subjects. *Psychopharmacology (Berl)* 1995;118:72–81.

78. Moeller FG, Dougherty DM, Swann AC, et al. Tryptophan depletion and aggressive responding in healthy males. *Psychopharmacology (Berl)* 1996;126:97–103.

79. Dougherty DM, Bjork JM, Marsh DM, et al. Influence of trait hostility on tryptophan depletion-induced laboratory aggression. *Psychiatry Res* 1999;88:227–232.

80. Smith SE, Pihl R-O, Young SN, et al. Elevation and reduction of plasma tryptophan and their effects on aggression and perceptual sensitivity in normal males, *Aggress Behav* 1986;12:393–407.

81. Cherek DR, Spiga R, Creson DL. Acute effects of eltoprazine on aggressive and point-maintained responding of normal male participants: phase I study. *Exp Clin Psychopharmacol* 1995;3:287–293.

82. Coccaro EF, Kavoussi RJ. Fluoxetine and impulsive aggressive behavior in personality disordered subjects. *Arch Gen Psychiatry* 1997;54:1081–1088.

83. Matsumoto K, Ojima K, Watanabe H. Noradrenergic denervation attenuates desipramine enhancement of aggressive behavior in isolated mice. *Pharmacol Biochem Behav* 1995;50:481–484.

84. Coccaro EF. Neurotransmitter correlates of impulsive aggression in humans. *Ann NY Acad Sci* 1996;794:82–89.

85. Siever L, Trestman RL. The serotonin system and aggressive personality disorder. *Int Clin Psychopharmacol* 1993;8[Suppl 2]:33–39.

86. Barrett JA, Edinger H, Siegel A. Intrahypothalamic injections of norepinephrine facilitate feline affective aggression via alpha-2 adrenoceptors. *Brain Res* 1990;525:285–293.

87. Agren H, Mefford IN, Rudorfer MV, et al. Interacting neurotransmitter systems: a non-experimental approach to the 5-HIAA-HVA correlation in human CSF. *J Psychiatr Res* 1986;20:175–193.

88. Kuikka JT, Tiihonen J, Bergstrom KA, et al. Abnormal structure of human striatal dopamine re-uptake sites in habitually violent alcoholic offenders: a fractal analysis. *Neurosci Lett* 1998;253:195–197.

89. Archer J. The influence of testosterone on human aggression. *Br J Psychol* 1991;82:1–28.

90. Rubinow DR, Schmidt PJ. Androgens, brain, and behavior. *Am J Psychiatry* 1996;153:974–984.

91. Salvadora A, Suay F, Martinez-Sanchis S, et al. Correlating testosterone and fighting in male participants in judo contests. *Physiol Behav* 1999;68:205–209.

92. Gerra G, Avanzini P, Zaimovic A, et al. Neurotransmitter and endocrine modulation of aggressive behavior and its components in normal humans. *Behav Brain Res* 1996;81:19–24.

93. Gerra G, Avanzini P, Zaimovic A, et al. Neurotransmitter-neuroendocrine responses to experimentally induced aggression in humans: influence of personality variable. *Psychiatry Res* 1997;66:33–43.

94. Rasanen P, Hakko H, Visuri S, et al. Serum testosterone levels, mental disorders and criminal behaviour. *Acta Psychiatr Scand* 1999;99:348–352.

95. Bergman B, Brismar B. Characteristics of imprisoned wife-beaters. *Forensic Sci Int* 1994;65:157–167.

96. Pope HG Jr, Kouri EM, Hudson JI. Effects of supraphysiologic doses of testosterone on mood and aggression in normal men: a randomized controlled trial. *Arch Gen Psychiatry* 2000;57:133–140; discussion 155–156.

97. Anderson RA, Bancroft J, Wu FC. The effects of exogenous testosterone on sexuality and mood of normal men. *J Clin Endocrinol Metab* 1992;75:1503–1507.

98. Tricker R, Casaburi R, Storer TW, et al. The effects of supraphysiological does of testosterone on angry behavior in healthy eugonadal men: clinical research center study. *J Clin Endocrinol Metab* 1996;81:3754–3758.

99. Bahrke MS, Yesalis CE 3rd, Wright JE. Psychological and behavioural effects of endogenous testosterone and anabolicandrogenic steroids: an update. *Sports Med* 1996;22:367–390.

100. Pope HG Jr, Kouri EM, Powell KF, et al. Anabolic-androgenic steroid use among 133 prisoners. *Compr Psychiatry* 1996;37:322–327.

101. Cooper CJ, Noakes TD, Dunne T, et al. A high prevalence of abnormal personality traits in chronic users of anabolic-androgenic steroids. *Br J Sports Med* 1996;3:246–250.

102. Essman WB. Drug effects upon aggressive behavior. In: Valzelli I, Morgese I, eds. *Aggression and violence: a psychobiological and clinical approach*. Milan: Edizioni Saint Vincent, 1981:150–175.

103. Jeffcoate WJ, Lincoln NB, Selby C, et al. Correlation between anxiety and serum prolactin in humans. *J Psychosom Res* 1986;30:217–222.

104. Dettling AC, Gunnar MR, Donzella B. Cortisol levels of young children in full-day childcare centers: relations with age and temperament. *Psychoneuroendocrinology* 1999;24:19–36.

105. McBurnett K, Lahey BB, Rathouz PJ, et al. Low salivary cortisol and persistent aggression in boys referred for disruptive behavior. *Arch Gen Psychiatry* 2000;57:38–43.

106. Buydens-Branchcy L, Branchey MH. Cortisol in alcoholics with a disordered aggression control. *Psychoneuroendocrinology* 1992;17:45–54.

107. McGuire MT, Brammer GL, Raleigh MJ. Resting cortisol levels and the emergence of dominant status among male vervet monkeys. *Horm Behav* 1986;20:106–117.

108. Coccaro EF, Kavoussi RJ, Hauger RL, et al. Cerebrospinal fluid vasopressin: correlates with aggression and serotonin function in personality disordered subjects. *Arch Gen Psychiatry* 1998;55:708–714.

109. Ferris CF, Delville Y. Vasopressin and serotonin interactions in the control of agonistic behavior. *Psychoneuroendocrinology* 1994;19:593–601.

110. Post RM, Pickar D, Ballenger JC, et al. Endogenous opiates in cerebrospinal fluid: Relationship to mood and anxiety. In: Post RM, Ballenger JC, eds. *Neurobiology of mood disorders*. Baltimore: Williams & Wilkins, 1984:356–368.

111. Coid J, Allolio B, Rees LH. Raised plasma metenkephalin in patients who habitually mutilate themselves. *Lancet* 1983;2:545–546.

112. Gross-Isseroff R, Dillon KA, Israeli M, et al. Regionally selective increases in mu opioid receptor density in the brains of suicide victims. *Brain Res* 1990;530:312–316.

113. Berman M, Taylor S, Marged B. Morphine and human aggression. *Addict Behav* 1993;18:263–268.

114. Sandman CA, Barron JL, Colman H. An orally administered opiate blocker, naltrexone, attenuates self-injurious behavior. *Am J Ment Retard* 1990;95:93–102.

115. New AS, Sevin EM, Mitropoulou V, et al. Serum cholesterol and impulsivity in personality disorders. *Psychiatry Res* 1999;85:145–150.

116. Muldoon MF, Rossouw JE, Manuck SB, et al. Low or lowered cholesterol and risk of death from suicide and trauma. *Metabolism* 1993;42[Suppl 1]:45–56.

117. Kaplan JR, Manuck SB, Shively C. The effects of fat and cholesterol on social behavior in monkeys. *Psychosom Med* 1991;53:634–642.

118. Kaplan JR, Shively CA, Fontenot MB, et al. Demonstration of an association among dietary cholesterol, central serotonergic activity, and social behavior in monkeys. *Psychom Med* 1994;56:479–484.

119. Frick, MH, Elo O, Haapa K, et al. Helsinki heart study: primary prevention trial with gemfibrozil in middle-aged men with syslipidemia. Safety of treatment, changes in risk factors, and incidence of coronary heart disease. *N Engl J Med* 1987;317:1237–1245.

120. Muldoon, MF, Manuck SB, Matthews KA. Lowering cholesterol concentration and mortality: a quantitative review of primary prevention trials. *BMJ* 1990;301:309–314.

121. Neaton, JD, Blackburn H, Jacobs D, et al. Serum cholesterol level and mortality findings for men screened in the Multiple Risk Factor Intervention Trial Research Group. *Arch Intern Med* 1992;152:1490–1500.

122. Maes M, Smith R, Christophe A, et al. Lower serum high density lipoprotein cholesterol (HDL-C) in major depression and in depressed men with serious suicidal attempts: relationship with immune-inflammatory markers. *Acta Psychiatr Scand* 1997;95:212–221.

123. Markovitz JH, Smith D, Raczynski JM, et al. Lack of relations of hostility, negative affect and high risk behavior with low plasma lipid levels in the Coronary Artery Risk Development in Young Adults Study. *Arch Intern Med* 1997;157:1953–1959.

124. Goier, JA, Marzuk PM, Leon AC, et al. Low serum cholesterol and attempted suicide. *Am J Psychiatry* 1996;152:419–423.

125. Zureik M, Courbon D, Ducimetiere P. Serum cholesterol concentration and death from suicide in men: Paris prospective study I. *BMJ* 1996;313:649–651.

126. Maes M, Sharpe P, D'Hondt P, et al. Biochemical metabolic and immune correlates of seasonal variation in violent suicide: a chronoepidemiologic study. *Eur Psychiatry* 1996;11:21–33.

127. Kunugi H, Takei N, Aoki H, et al. Low serum cholesterol in suicide attempters. *Biol Psychiatry* 1997;41:196–200.

128. Virkkunen M. Serum cholesterol in antisocial personality disorder. *Neuropsychobiology* 1979;5:27–32.

129. Virkkunen M. Serum cholesterol in levels in homicidal offenders: a low cholesterol level is connected with a habitually violent tendency under the influence of alcohol. *Neuropsychobiology* 1983;10:65–69.

130. Hillbrand M, Spitz RT, Foster HG. Serum cholesterol and aggression in hospitalized male forensic patients *J Behav Med* 1995;18:33–43.

131. Virkkunen M, Pentitinen H. Serum cholesterol in aggressive conduct disorder: a preliminary study. *Biol Psychiatry* 1984;19:435–439.

132. Apter A, Laufer N, Bar-Sever M, et al. Serum cholesterol suicidal tendencies, impulsivity, aggression, and depression in adolescent psychiatric inpatients. *Biol Psychiatry* 1999;46:532–541.

133. Brunner HG, Nelen M, Breakefeld XO, et al. Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science* 1995;262:578–580.

134. Nielson DA, Goldman D, Virkkunen M, et al. Suicidality and 5-hydroxyindolacetic acid concentration associated with a tryptophan hydoxylase polymorphism. *Arch Gen Psychiatry* 1994;51:34–38.

135. Abbar M, Courtet P, Amadeo S, et al. Suicidal behaviors and the tryptophan hydroxylase gene. *Arch Gen Psychiatry* 1995;52:846–849.

136. Nielson DA, Virkkunen M, Lappalainen J, et al. A tryptophan hydroxylase gene marker for suicidality and alcoholism. *Arch Gen Psychiatry* 1998;55:593–602.

137. New AS, Gelernter J, Yovell Y, et al. Tryptophan hydroxylase genotype is associated with impulsive aggression measures. *Am J Medical Genet* 1998;81:3–17.

138. Mann JJ, Malone KM, Nielsen DA, et al. Possible association of a polymorphism of the tryptophan hydroxylase gene with suicidal behavior in depressed patients. *Am J Psychiatry* 1997;154:1451–1453.

139. Manuck SB, Flory JD, Ferrell RE, et al. Aggression and anger-related traits associated with a polymorphism of the tryptophan hydroxylase gene. *Biol Psychiatry* 1999;45:603–614.

140. Lappalainen J, Long JC, Eggert M, et al. Linkage of antisocial alcoholism to the serotonin 5-HT-1B receptor gene in two populations. *Arch Gen Psychiatry* 1998;55:989–994.

141. Tsai S-J, Chiu H-J, Wang Y-C, et al. Association study of serotonin-6 receptor variant (C267T) with schizophrenia and aggressive behavior. *Neurosci Letters* 1999;271:135–137.

142. Lachman HM, Nolan KA, Mohr P, et al. Association between catechol-O-methyltransferase genotype and violence in schizophrenia and schizoaffective disorder. *Am J Psychiatry* 1998;155:835–837.

143. Strous RD, Bark N, Parsia SS, et al. Analysis of a functional catechol-O-methyltransferase gene polymorphism in schizophrenia: evidence for association with aggressive and antisocial behavior. *Psychiatry Res* 1997;69:71–77.

144. Dery M, Toupin J, Pauze R, et al. Neuropsychological characteristics of adolescents with conduct disorder: association with attention-deficit-hyperactivity and aggression. *J Abnorm Child Psychol* 1999;27:225–236.

145. Giancola PR, Zeichner A. Neuropsychological performance on tests of frontal-lobe functioning and aggressive behavior in men. *J Abnorm Psychol* 1994;103:832–835.

146. Giancola PR, Moss H, Martin C, et al. Executive cognitive functioning predicts reactive aggression in boys at high risk for substance abuse: a prospective study. *Alcohol Clin Exp Res* 1996;20:740–744.

147. Barratt ES, Stanford MS, Kent TA, et al. Neuropsychological and cognitive psychophysiological substrates of impulsive aggression. *Biol Psychiatry* 1997;41:1045–1061.

148. Yeudall LT. Neuropsychological assessment of forensic disorders. *Can Ment Health* 1977;25:7–15.

149. Dolan M. Psychopathy: a neurobiological perspective. *Br J Psychiatry* 1994;165:151–159.

150. Cherek DR. Effects of smoking different doses of nicotine on human aggressive behavior. *Psychopharmacology (Berl)* 1981;75:339–349.

151. LeMarquand DG, Benkelfat C, Pihl RO, et al. Behavioral disinhibition induced by tryptophan depletion in nonalcoholic young men with multigenerational family histories of paternal alcoholism. *Am J Psychiatry* 1999;156:1771–1779.

152. Damasio H, Grabowski T, Frank R, et al. The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science* 1994;264:1102–1105.

153. Brickner RM. An interpretation of frontal lobe function based upon the study of a case of partial bilateral lobectomy. *Res Nerv Ment Dis* 1934;13:259–351.

154. Damasio AR, Van Hoesen GW. Emotional disturbances associated with focal lesions of the limbic frontal lobe. In: Heilman KM, Satz P, eds. *Neuropsychology of human emotion*. New York: Guilford Press, 1983.

155. Grafman J, Vance S, Wingartner H, et al. The effects of lateralized frontal lesions on mood regulation. *Brain* 1986;109:1127–1148.

156. Grafman J, Schwab K, Warden D, et al. Frontal lobe injuries, violence, and aggression: a report of the Vietnam Head Injury Study. *Neurology* 1996;46:1231–1238.

157. Volvaka J. Neurological, neuropsychological, and electrophysiological correlates of violent behavior. *Neurobiology of violence*. Washington, DC: American Psychiatry Press, 1995:77–122.

158. Weiger WA, Bear DM. An approach to the neurology of aggression. *J Psychiatr Press* 1988;22:85–98.

159. Miller BL, Darby A, Benson DF, et al. Aggressive, socially disruptive and antisocial behaviour associated with fronto-temporal dementia. *Br J Psychiatry* 1997;170:150–154.

160. Anderson SW, Bechara A, Damasio H, et al. Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nat Neurosci* 1999;2:1032–1037.

161. Tonkonogy JM. Violence and temporal lobe lesion: head CT and MRI data. *J Neuropsychiatry Clin Neurosci* 1991;3:189–196.

162. Tonkonogy JM, Geller JL. Hypothalamic lesions and intermittent explosive disorder. *J Neuropsychiatry Clin Neurosci* 1992;4:45–50.

163. Martinius J. Homicide of an aggressive adoloscent boy with right temporal lesion: a case report. *Neurosci Biobehav Rev* 1983;7:419–422.

164. Mark VH, Ervin FR. *Violence in the brain*. New York: Harper and Row, 1970.

165. Hitchcock E, Cairns V. Amygdalotomy. *Postgrad Med J* 1973;49:894–904.

166. Adolphs R, Tranel D, Damasio AR. The human amygdala in social judgment. *Nature* 1998;393:470–474.

167. Gedye A. Episodic rage and aggression attributed to frontal lobe seizures. *J Ment Defic Res* 1989;33:369–379.

168. Williams D. Neural factors related to habitual aggression. *Brain Res* 1969;92:503–520.

169. DeVinsky O, Bear D. Varieties of aggressive behavior in temporal lobe epilepsy. *Am J Psychiatry* 1984;141:651–656.

170. Rodin EA. Psychomotor epilepsy and aggressive behavior. *Arch Gen Psychiatry* 1973;28:210–213.

171. Currie S, Heathfield KW, Henson RA, et al. Clinical course and progress of temporal lobe epilepsy: a survey of 666 patients. *Brain* 1971;94:173–190.

172. Delgado-Escueta AV, Mattson RH, King L, et al. The nature of aggression during epileptic seizures. *N Engl J Med* 1981;305:711–716.

173. Raine A, Lencz T, Bihrle S, et al. Reduced prefrontal gray volume and autonomic deficits in antisocial personality disorder. *Arch Gen Psychiatry* 2000;57:119–127.

174. Volkow ND, Tancredi LR. Neural substrates of violent behavior: a preliminary study with positron emission tomography *Br J Psychiatry* 1987;151:668–673.

175. Volkow ND, Tancredi LR, Grant C, et al. Brain glucose metabolism in violent psychiatric patients: a preliminary study. *Psychiatry Res* 1995;61:243–253.

176. Raine A, Buchsbaum M, Stanley J, et al. Selective reductions in prefrontal glucose metabolism in murderers. *Biol Psychiatry* 1994;36:365–373.

177. Raine A, Stoddard J, Bihrle S, et al. Prefrontal glucose deficits in murderers lacking psychosocial deprivation. *Neuropsychiatry Neuropsychol Behav Neurol* 1998;11:1–7.

178. Goyer PF, Andreason PJ, Semple WE, et al. Positron-emission-tomography and personality disorders. *Neuropsychopharmacology* 1994;10:21–28.

179. Amen DG, Stubblefield M, Carmicheal B. Brain SPECT findings and aggressiveness. *Ann Clin Psychiatry* 1996;3:129–137.

180. Finkelstein Y, Yardi J, Hod I. Impulsive artistic creativity as a presentation of transient cognitive alterations. *Behav Med* 1991;17:91–94.

181. Porrino LJ, Goldman-Rakic PS. Brainstem innervation of prefontal and anterior cingulate cortex in the rheus monkey revealed by retrograde transport of HRP. *J Comp Neurol* 1982;205:63–76.

182. Lee GP, Bechara A, Adolphs R, et al. Clinical and physiological effects of sterotaxic bilateral amygdalotomy for intractable aggression. *J Neuropsychiatry Clin Neurosci* 1998;10:413–420.

183. Arango V, Ernsberger P, Marzuk PM, et al. Autoradiographic demonstration of increased serotonin 5-HT2 receptors and beta-adrenergic receptor binding sites in the brain of suicide victims. *Arch Gen Psychiatry* 1990;47:1038–1047.

184. VanWoerken TCAM, Teelken AW, Minderhoud JM. Difference in neurotransmitter metabolism in fronto-temporal-lobe contusion and diffuse cerebral confusion. *Lancet* 1977;1:812–813.

185. Heinz A, Higley JD, Gorey JG, et al. *In vivo* association between alcohol intoxication, aggression, and serotonin transporter availability in nonhuman primates. *Am J Psychiatry* 1998;155:1023–1028.

186. Raleigh MJ, Brammer GL. Individuals differences in serotonin-2 receptors and social behavior in monkeys. *Soc Neurosci Abstr* 1993;19:592.

187. Kapur S, Myer J, Wilson AA, et al. Activation of specific cortical regions by apomorphine A [^{15}O] H_2O PET study in humans. *Neurosci Lett* 1994;176:21–24.

188. Mann J, Malone K, Diehl D, et al. Positron emission tomographic imaging of serotonin activation effects on prefrontal cortex in healthy volunteers. *J Cereb Blood Flow Metab* 1996;16:418–426.

189. Mann J, Malone KM, Diehl DJ, et al. Demonstration *in vivo* of reduced serotonin responsivity in the brain of untreated depressed patients. *Am J Psychiatry* 1996;153:174–182.

190. Hommer D, Andreason P, Rio D, et al. Effects of m-chlorophenylpiperazine on regional brain glucose utilization: a positron emission tomography comparison of alcoholic and control subjects. *J Neurosci* 1997;17:2796–2802.

191. Soloff PH, Meltzer CC, Greer PJ, et al. A fenfluramine-activated FDG-PET study of borderline personality disorder. *Biol Psychiatry* 2000;47:540–547.

192. Siever LJ, Buchsbaum MS, New AS, et al. Serotonergic responsiveness in impulsive/aggressive personality disorders. *Biol Psychiatry* 2000;47:1S-173S.

193. Melega WP, Raleigh MJ, Stout DB, et al. Longitudinal behavioral and 6-[^{18}F] fluoro-L-DOPA-PET assessment in MPTP-hemiparkinsonian monkeys. *Exp Neurol* 1996;141:318–329.

194. Campbell M, Adams PB, Small AM, et al. Lithium in hospitalized aggressive children with conduct disorder: a double-blind and placebo-controlled study. *J Am Acad Child Adolesc Psychiatry* 1995;34:445–453.

195. Craft M, Ismail IA, Krishnamurti D, et al. Lithium in the treatment of aggression in mentally handicapped patients: a double-blind trial. *Br J Psychiatry* 1987;150:685–689.

196. Salzman C, Wolfson AN, Schatzberg A, et al. Effect of fluoxetine on anger in symptomatic volunteers with borderline personality disorder. *J Clin Psychopharmacol* 1995;15:23–29.

197. Fava M, Rosenbaum JF, Pava JA, et al. Anger attacks in unipolar depression. I. Clinical correlates and response to fluoxetine treatment. *Am J Psychiatry* 1993;150:1158–1163.

198. McDougle CJ, Naylor ST, Cohen DJ, et al. A double-blind, placebo-controlled study of fluvoxamine in adults with autistic disorder. *Arch Gen Psychiatry* 1996;53:1001–1008.

199. Coccaro EF, Kavoussi RJ, Hauger RL. Serotonin function and antiaggressive responses to fluoxetine: a pilot study. *Biol Psychiatry* 1997;42:546–552.

200. Ratey JJ, Sovner R, Parks A, et al. Buspirone treatment of aggression and anxiety in mentally retarded patients: a multiple baseline, placebo lead-in study. *J Clin Psychiatry* 1991;52:159–161.

201. Verhoeven WM, Tuinier S. The effect of buspirone on challenging behaviour in mentally retarded patients: an open prospective multiple-case study. *J Intellect Disabil Res* 1996;40:502–508.

202. Cowdry RW, Gardner DL. Pharmacotherapy of borderline personality disorder: alprazolam, carbamazepine, trifluroperazine, and trancypromine. *Arch Gen Psychiatry* 1998;45:111–119.

203. Tariot PN, Erb R, Podgorski CA, et al. Efficacy and tolerability of carbamazepine for agitation and aggression in dementia. *Am J Psychiatry* 1998;155:54–61.

204. Cueva JE, Overall JE, Small AM, et al. Carbamazepine in aggressive children with conduct disorder: a double-blind and placebo-controlled study. *J Am Acad Child Adolesc Psychiatry* 1996;35:480–490.

205. Stein DJ, Simeon D, Frenkel M, et al. An open trial of valproate in borderline personality disorder. *J Clin Psychiatry* 1995;56:506–510.

206. Wilcox JA. Divalproex sodium as a treatment for borderline personality disorder. *Ann Clin Psychiatry* 1995;7:33–37.

207. Donovan SJ, Susser ES, Nunes EV, et al. Divalproex treatment of disruptive adolescents: a report of 10 cases. *J Clin Psychiatry* 1997;58:12–15.

208. Kavoussi RJ, Coccaro EF. Divalproex for impulsive aggressive behavior in personality disordered patients. *J Clin Psychiatry* 1998;59:676–680.

209. Lott AD, McElroy SL, Keys MA. Valproate in the treatment of behavioral agitation in elderly patients with dementia. *J Neuropsychiatry Clin Neurosci* 1995;7:314–319.

210. Mellow AM, Solano-Lopez C, Davis S. Sodium valproate in the treatment of behavioral disturbance in dementia. *J Geriatr Psychiatry Neurol* 1993;6:205–209.

211. Wroblewski BA, Joseph AB, Kupfer J, et al. Effectiveness of valproic acid on destructive and aggressive behaviours in patients with acquired brain injury. *Brain Injury* 1997;11:37–47.

212. Teicher MH, Glod CA, Aaronson ST, et al. Open assessment of the safety and efficacy of thioridazine in the treatment of patients with borderline personality disorder. *Psychopharmacol Bull* 1989;25:535–549.

213. Soloff PH, George A, Nathan S, et al. Amitriptyline versus haloperidol in borderlines: final outcomes and predictors of response. *J Clin Psychopharmacol* 1989;9:238–246.

214. Goldberg SC, Schulz SC, Schulz PM, et al. Borderline and schizotypal personality disorders treated with low-dose thiothixene versus placebo. *Arch Gen Psychiatry* 1986;43:680–686.

215. Buckley P, Bartell J, Donenwirth K, et al. Violence and schizophrenia: clozapine as a specific antiaggressive agent. *Bulle Am Acad Psychiatry Law* 1995;23:607–611.

216. Rabinowitz J, Avnon M, Rosenberg V. Effect of clozapine on physical and verbal aggression. *Schizophr Res* 1996;22:249–255.

217. Ratey JJ, Leveroni C, Kilmer D, et al. The effects of clozapine on severely aggressive psychiatric inpatients in a state hospital. *J Clin Psychiatry* 1993;54:219–223.

218. Czobor P, Volavka J, Meibach RC. Effect of risperidone on hostility in schizophrenia. *J Clin Psychopharmacol* 1995;15:243–249.

219. Schulz SC, Camlin KL, Beryy SA, et al. Olanzapine safety and efficacy in patients with borderline personality disorder and comorbid dysthymia. *Biol Psychiatry* 1999;46:1429–1435.

220. Yudofsky S, Williams D, Gorman J. Propranolol in the treatment of rage and violent behavior in patients with chronic brain syndromes. *Am J Psychiatry* 1981;138:218–220.

221. Brooke MM, Patterson DR, Questad KA. The treatment of agitation during initial hospitalization after traumatic brain injury. *Arch Phys Med Rehabil* 1992;73:917–921.

222. Shankle WR, Nielson KA, Cotman CW. Low-dose propranolol reduces aggression and agitation resembling that associated with orbitofrontal dysfunction in elderly demented patients. *Alzheimer Dis Assoc Disord* 1995;9:233–237.

223. Sorgi PJ, Ratey JJ, Polakoff S. Beta adrenergic blockers for the control of aggressive behaviors in patients with chronic schizophrenia. *Am J Psychiatry* 1986;143:775–776.

224. Ratey JJ, Sorgi P, O'Driscoll GA, et al. Nadolol to treat aggression and psychiatric symptomatology in chronic psychiatric inpatients: a double-blind, placebo-controlled study, *J Clin Psychiatry* 1992;53:41–46.

225. Kravitz HM, Haywood TW, Kelly J, et al. Medroxyprogesterone treatment for paraphiliacs. *Bull Am Acad Psychiatry Law* 1995;23:19–33.

120

Pathologic Gambling and Impulse Control Disorders

Mark N. Potenza

Eric Hollander

Mark N. Potenza: Director, Problem Gambling Clinic; Department of Psychiatry, Yale University School of Medicine and Connecticut Mental Health Center, New Haven, Connecticut.

Eric Hollander: Professor of Psychiatry, Seaver Autism Research Center, Department of Psychiatry, Mount Sinai School of Medicine, New York, New York.

Impaired regulation of impulse is a central feature in several psychiatric disorders and behaviors, including drug use disorders, cluster B personality disorders such as borderline personality disorder, bipolar disorders, and suicide attempts. The DSM-IV's "impulse control disorders (ICDs) not specified elsewhere" (1) have historically received less attention than other psychiatric conditions. This heterogeneous group of illnesses includes intermittent explosive disorder, kleptomania, pyromania, pathologic gambling (PG), trichotillomania, and ICDs not otherwise specified. We review the neurobiology and treatment of one of these ICDs, PG, and describe the nature and treatments of several other potentially related conditions that have recently received increased attention: (a) compulsive buying (CB); (b) compulsive sexual behavior (CSB); and (c) compulsive computer use (CCU). These disorders, linked by a failure to resist urges to engage in ultimately self-destructive behaviors, appear to be relatively common, frequently go unrecognized for considerable periods, and may constitute greater threats to personal health than is often appreciated.

PATHOLOGIC GAMBLING

Part of "120 - Pathologic Gambling and Impulse Control Disorders "

Descriptions of gambling and gambling disorders are found in some of the earliest human records (2 ,3). Historically, gambling has been viewed as a sin and later as a vice (2 ,3 ,4 and 5). More recently, disordered gambling has been seen as an illness determined by genetic and environmental factors and individual decision making. The most extreme form of disordered gambling, PG, was first included in the DSM in 1980 (1). Since that time, there has been increasing research into the clinical features and neurobiological causes of PG.

Theoretic Conceptualizations

PG has been described as sharing features with several groups of disorders (Fig. 120.1). Some authors have hypothesized that PG lies along a compulsive-impulsive spectrum (6 ,7), and PG represents an obsessive-compulsive (OC) spectrum disorder (8 ,9 and 10). Consistent with the classification of PG as an OC-spectrum disorder, individuals with PG engage in repetitive (gambling-related) behaviors, often in response to overwhelming thoughts to engage in the behavior (11). Studies of comorbidity between OC disorder (OCD) and PG have yielded mixed results, with some studies finding rates of OCD in individuals with PG higher than in the general population (12 ,13), whereas others have not found elevated rates of comorbidity (14 ,15 ,16 and 17) or elevated rates of positive family histories for PG in patients with OCD (18). A direct investigation into OC characteristics of individuals with PG found that those with PG scored significantly higher than those without on the Padua Inventory (19). The differences clustered within two factors corresponding to obsessive qualities of impaired control over mental activity and worries of losing control over motor behavior, respectively (19). Although the findings support the notion that PG lies toward the impulsive end of a compulsive-impulsive spectrum, the authors cite a central difference between gambling in PG and repetitive behaviors in OCD (19). Namely, gambling and actions in other ICDs are often related as pleasurable or egosyntonic, whereas performance of repetitive activities in OCD is generally described as egodystonic. Although one study reported the possibility of the association of PG and OCD with the Huntington disease mutation in a family with PG, OCD, and Huntington disease (20), the neurobiological similarities and differences between PG and OCD remain to be defined more clearly, to explore their relatedness further.

FIGURE 120.1. Proposed conceptual model for relationships between pathologic gambling (PG) and other psychiatric conditions. ADHD, attention-deficit/hyperactivity disorder.

Researchers and clinicians have also described PG as an addiction and have cited similarities to substance use disorders (21 ,22 and 23). In fact, the diagnostic criteria for PG were modeled after those for substance dependence (1) and include

aspects of tolerance, withdrawal, and failed attempts to control the destructive behavior. High rates of comorbidity are observed between PG and substance use disorders. Individuals with PG have high rates of substance use disorders, with rates of nicotine dependence approaching 70% (24), alcohol abuse or dependence in the range of 45% to 55% (12 ,25), and other drug use problems nearing 40% (26). Conversely, individuals with substance use disorders are four- to tenfold more likely to have PG (27): 9% of opiate addicts in methadone maintenance (28), 17% of alcohol abusers (29), and 15% of cocaine addicts (30) have PG. The high rates of comorbidity have implications with regard not only to potential similarities in the underlying neurobiological bases of PG and substance use disorders, but also to the clinical needs of individuals with PG. Specifically, individuals dually diagnosed with a substance use disorder and PG were found to require more psychiatric admissions and detoxifications than individuals with a substance use disorder without PG (31). A separate study found that individuals with comorbid substance use disorders and PG were at greater risk for contemplated and attempted suicide than individuals with either diagnosis alone (32). These and other findings (33 ,34) indicate that dually diagnosed individuals with PG appear to be more severely ill than those with either illness alone. Taken together with emerging data suggesting neurobiological similarities between substance use disorders and PG (see the later discussions of genetics and neuroimaging), there is mounting evidence supporting the notion of substance use disorders and PG lying along an addiction spectrum.

High rates of other psychiatric disorders, particularly mood, attention-deficit, and antisocial personality disorders, have also been described in individuals with PG (24 ,35 ,36 and 37). Some data suggest that individuals with features of some of these disorders (e.g., cycling mood disorders) could benefit from different treatment interventions (see the later discussion on pharmacotherapy). Further studies are warranted to investigate the precise relationships between these disorders and PG.

Biochemistry

Multiple factors, including behavioral initiation, arousal, reward and reinforcement, and behavioral disinhibition, have been described as contributing to or disordered gambling behavior in PG (38). Unique roles for specific neurotransmitters have been hypothesized as mediating aspects of PG and other ICDs (Table 120.1). Specifically, serotonin (5-HT) has been described as important in behavioral regulation (behavioral initiation and inhibition, including control of aggressive and other impulses) (38 ,39 ,40 and 41). Data support a central role for norepinephrine (NE) in the control of levels of arousal and detection of novel or aversive stimuli (42). Multiple lines of evidence from studies of human and other organisms cite dopamine (DA) function, particularly within the mesocorticolimbic (MCL) pathways, as critical in processing and modulating rewarding and reinforcing stimuli and behaviors (43 ,44 and 45). Abnormalities in these neurotransmitter systems as they relate to PG are explored in the following sections.

Neurotransmitter	Proposed Role
Norepinephrine	Arousal, excitement
Serotonin	Behavioral initiation and cessation
Dopamine	Reward, reinforcment
Opioids	Pleasure, urges

TABLE 120.1. PROPOSED ROLES FOR NEUROTRANSMITTER SYSTEMS IMPLICATED IN THE PATHOPHYSIOLOGY OF PATHOLOGIC GAMBLING

Serotonin

A role for 5-HT system dysfunction in the neurobiology of PG has come from results of pharmacologic challenge studies (38 ,46). The 5-HT and NE reuptake inhibitor clomipramine (CMI) has been used to investigate neurochemical responses in individuals with PG as compared with those without PG (46). Eight men and women with PG and eight age- and gender-matched controls received a relatively low intravenous dose of clomipramine (12.5 mg), one that the authors argued targeted primarily the 5-HT transporter (46). The persons with PG in comparison with controls were found to have at baseline lower prolactin levels and exhibited significantly blunted prolactin increases 60 minutes after clomipramine administration (46). The blunted prolactin response suggests the possibility of diminished 5-HT transporter binding in individuals with PG.

An independent challenge study investigating 5-HT function in individuals with PG was undertaken by DeCaria and colleagues (38). The investigators administered meta-chlorophenylpiperazine (m-CPP) to 10 men with PG and 10 healthy male control subjects. m-CPP, a metabolite of the antidepressant trazodone and a partial 5-HT_1 and 5-HT_2 receptor agonist, binds with high affinity to 5-HT_{1A}, 5-HT_{1D}, 5-HT_{2A}, 5-HT_{2C}, and 5-HT_3 receptors, with particularly high affinity for 5-HT_{2C} receptors (47 ,48 and 49). 5-HT_{2C} receptors, localized to brain regions including the cortex and caudate, have been implicated in mediating aspects of mood, anxiety, appetite, behavior (including sexual activity), and neuroendocrine function (50 ,51). The investigators found that individuals with PG reported a euphoric response or "high" after m-CPP administration, a finding similar to those reported for other disorders such as antisocial personality disorder (52), borderline personality disorder (53), trichotrillomania (54), and alcohol abuse or dependence (55), in which impulsive or compulsive behaviors are prominent. In addition to differences in behavioral reactions, PG subjects demonstrated altered biochemical responses to the m-CPP challenges. Specifically, increases in prolactin levels were observed in persons with PG, with greater prolactin responses correlating with increased gambling severity.

Additional support for 5-HT dysfunction in individuals with PG has been obtained from investigations of cerebrospinal fluid (CSF) (56). Initial studies into the chemical composition of CSF from men with PG found no significant differences in levels of 5-HT or its metabolite 5-hydroxyindolacetic acid (5-HIAA) as compared to levels in healthy men (57 ,58 and 59). Given that multiple factors can complicate evaluation of CSF data (56 ,60 ,61) and the finding that men with PG have significantly longer CSF tapping times than healthy male controls (62), the authors calculated the concentrations of monoamine metabolites per minute of tapping to obtain an estimate of mass flow through the lumbar puncture needle. When taking tapping time into account, levels of 5-HIAA, as well as those of the NE metabolite 4-hydroxy-3-methoxyphenyl glycol (HMPG), were found to be significantly lower in the group of men with PG. These findings, particularly in light of reports of low CSF levels of 5-HIAA in individuals with impulsive characteristics, such as those attempting suicide (39 ,40 and 41), lend further support to a central role for 5-HT in the underlying pathology of PG. Additional data emerging from pharmacotherapy trials, neuroimaging studies, and investigations into monoamine oxidase (MAO) function (see later) are also consistent with 5-HT dysfunction in PG. Further studies are needed to define the nature and extent of 5-HT perturbations in PG more precisely, particularly as they relate to specific aspects of PG (e.g., behavioral initiation or cessation) and subgroups of individuals with PG (e.g., as distinguished by such characteristics as gender or comorbid diagnosis).

Dopamine

Data support positing a role for DA in reinforcing and rewarding aspects of gambling in PG. Multiple lines of evidence from studies investigating the neurochemical bases of drug use disorders have implicated the MCL DA system in the mediation of rewarding and reinforcing behaviors (43 ,44 and 45). Studies in humans with cocaine dependence have found MCL regional brain activations after a cocaine-induced rush (63) or viewing of cocaine-related videotapes (64), and occupancy of the DA transporter has been correlated with cocaine's euphorigenic effects (65). A role for DA in the rewarding and reinforcing aspects of gambling has been proposed (38 ,66). To explore this hypothesis, Bergh et al. analyzed CSF from ten men with PG and seven matched male controls (59). Decreased levels of DA and increased levels of the DA metabolites 3,4-dihydroxyphenylacetic acid (DOPAC) and homovanillic acid (HVA) were observed in the PG subjects (59). The authors concluded these findings to be consistent with an increased rate of DA neurotransmission, although more recently the same group did not find decreased HVA levels when correcting for CSF flow rate (56).

A separate study investigated peripheral levels of DA under gaming conditions (67). Plasma levels of DA were measured when subjects played Pachinko machines, described as a popular form of recreation in Japan and one that combines elements of pinball and slot machines. The authors reported that after a winning streak described as a "fever," six men who were regular Pachinko players were found to have elevated levels of DA. The authors suggested the DA changes may be related to the motivational processes underlying repeated Pachinko playing. Alterations in measures of arousal, NE, endogenous opiates, and immune system function were also described. Although the results of the foregoing studies and additional evidence from neuroimaging and molecular genetic studies suggest DA dysfunction in PG (see later), more studies are needed to clarify the involvement of DA pathology in PG.

Norepinephrine

NE has been hypothesized as mediating aspects of arousal, attention, and sensation seeking in individuals with PG (38 ,57 ,58 ,68). To investigate, Roy et al. measured urinary, peripheral, and central levels of NE or the NE metabolites MHPG and vanillylmandelic acid (VMA) and found the PG subjects to have higher CSF levels of MHPG and higher urinary measures of NE (57). In a subsequent report, the investigators found in the same group of 17 men with PG originally studied, scores of extraversion on the Eysenck Personality Questionnaire were found to correlate positively and significantly with CSF and plasma levels of MHPG, urinary measures of VMA, and the sum of urinary levels of NE and NE metabolites (58). More recently, increased CSF

levels of NE and MHPG were found in a second group of men with PG (59), although a subsequent report from the same research team reported findings of decreased MHPG in men with PG when correcting for CSF flow.

As mentioned in the section on DA, a study of male Pachinko players found NE system changes under gaming conditions (67). Specifically, blood levels of NE were found to increase from baseline over time during Pachinko play, with statistically significant changes noted at the onset and end of Pachinko "fever." Levels of NE decreased but remained significantly elevated 30 minutes after the end of "fever." Alterations in heart rate, a physiologic measurement associated with arousal, was also observed, with peak heart rate measured at the start of "fever."

Monoamine Oxidase Activity

The MAOs, subtypes MAO A and MAO B, are enzymes that metabolize NE, 5-HT, and DA (69). Peripheral MAO derived from platelets is of the MAO B subtype and has been suggested to be an indicator of 5-HT function (70 , 71), although MAO B also binds with high affinity to and catabolizes DA (69). Decreased platelet MAO activity has been reported in association with impulsive behaviors (72 ,73), high levels of sensation-seeking (74 ,75 and 76), and ICDs, including eating disorders (77). Individuals with PG have also been reported to exhibit decreased platelet MAO activities (78 ,79). In one study, 15 men with PG were found to have MAO activities 26% lower than those in a group of 25 male controls (78). A separate study involving 27 men with PG found MAO activity levels 41% lower than in matched male controls (79). Each group investigated personality and sensation-seeking characteristics of the PG and control groups and found statistically significant differences. However, no clear picture emerged regarding correlation of the characteristics with MAO levels, with no associations maintained after Bonferroni correction application in one study (79) and a positive correlation between MAO and several measures of sensation-seeking in the other (78).

Stress Response Pathways

Immune system and cortisol changes have been related to gaming behaviors. The foregoing study involving Pachinko players (67) found alterations in T cells (decreased number of T cells and decreased CD3-56 activity) and natural killer cells (increased number of natural killer cells without change in activity). A separate study of male and female Kimberley aborigines found significantly higher cortisol levels on days in which gambling behavior was concentrated (Thursdays and Fridays) as compared with temporally distinct days (Mondays and Tuesdays) (80). Significantly higher epinephrine levels were also found on the gambling-concentrated days, a finding further implicating adrenergic function in gambling behaviors. Although higher on gambling-concentrated days, differences in blood pressure measurements did not reach statistical significance. A separate study found no evidence of abnormal cortisol responsivity on a dexamethasone suppression test in 21 men with PG (81). Independently, CSF levels of corticotropin-releasing hormone and corticotropin were found not to be significantly different between healthy and PG-affected men (57). Additional studies investigating immune and stress hormone function in gambling and PG situations are warranted, in light of the foregoing data and reports of increased rates of physical health problems in individuals with PG (82).

Opioidergic Pathways

Given (a) the role of endogenous opioids in mediating levels of pleasure and (b) the μ-opioid receptor (mOR) in modulating reward and reinforcement DA pathways by disinhibition of γ-aminobutyric acid (GABA) input to DA neurons in the ventral tegmental area (83 ,84), studies exploring ß-endorphin function in gaming behaviors has been explored. In the previously detailed Pachinko study (67), blood levels of ß-endorphins were found to be elevated during Pachinko play, peaking during the start of "fever." These findings, in conjunction with the results from preliminary studies of the mOR antagonist naltrexone in the treatment of PG (see later), suggest further investigation into opioid function in PG are warranted.

Other Neurotransmitter Systems

Studies of the possible dysregulation of other neurotransmitter systems as related to PG have been undertaken. results of most of these studies to date have been negative, with statistically nonsignificant differences observed between PG and healthy male subjects with regard to CSF levels of neuropeptide Y (57 ,85), galanin (86), GABA (87), "diazepam-inhibitor binding" (57 ,88), neurotensin (57), somatostatin (57), or growth hormone–releasing hormone (57). One study did find decreased levels of the inhibitory neurotransmitter taurine in the CSF of men with PG as compared with healthy controls (89).

Neuroimaging Studies

Few neuroimaging studies directly investigating PG have been performed to date. One study investigated the potential role of the MCL DA system in a study in which participants were paid increasing amounts of money depending on the skill level reached while playing a video game (90 , 98). Positron emission tomography (PET) studies using ^{11}C-labeled raclopride, a ligand with high affinity for D2-like DA receptors (D2Rs), found decreased levels of striatal binding in eight male study subjects playing a tank video game as compared with when they viewed a gray screen image (90). The authors concluded that the observed 13%

reduction in [su>11C]raclopride signal during the gaming condition is consistent with at least a twofold increase in levels of extracellular DA. Because the game involved increasing monetary reward associated with each skill level reached during the video game, the paradigm is similar but not identical to actual gambling.

An independent study investigated for specific DA and 5-HT abnormalities in individuals with PG (91). Using PET, the researchers found decreased striatal binding in PG subjects of [^{11}C]*N*-methylspiperone, a ligand with high affinity for D2Rs and 5-HT_{2A} and 5-HT_{2C} receptors (92). The striatal signal, corresponding to D2R-receptor occupancy, could be explained by multiple, non–mutually exclusive possibilities including decreased numbers of available D2Rs, decreased affinity of D2Rs for the tracer, or increased synaptic concentrations of DA. PG subjects were also found to have impaired performance on multiple neurocognitive tests, including the Halstead-Reitan, Wisconsin Card Sort, Shipley, and California Verbal Learning. Regional cerebral blood flow to the frontal cortex and anterior cingulate was found to be significantly lower in PG as compared with healthy subjects during performance of an auditory continuous performance attention task. These findings are consistent with prior studies implicating involvement of the frontal cortex and anterior cingulate in attention (93), among other processes, and suggest a role for these brain regions in mediating attentional deficits in individuals with PG (35).

Multiple neuroimaging studies into the neural bases of drug use disorders have been performed. Studies of drug craving, a central component in relapsing behavior, have repeatedly identified the involvement of anterior cingulate activation (63 ,64 ,94 ,95), among other brain regions. Studies investigating whether similar brain regions may be involved in PG are under way, with preliminary data suggesting the involvement of parallel neural activities (96 ,97).

Decision Making

An instrument called the Iowa Gambling Task has been developed and used in investigations into decision making (98 ,99 ,100 and 101). The tool involves four piles of cards, each associated with predetermined patterns of rewards and punishments. Selection from two of the piles, each associated with lower rewards and lower punishments, will ultimately result in long-term gains, and selection from the other two piles will result in long-term losses. Without prior instruction into the reward and punishment profile of each pile, individuals are instructed to select from the piles and to maximize gains. Interestingly, individuals with stroke lesions in either the ventromedial prefrontal cortex (VM) or amygdala perform worse than healthy subjects on the task (100 ,101). Additionally, those with VM lesions not only do not improve over time with repeat performance, but also fail to exhibit changes in skin conductance associated with the decision-making processes (102). Individuals with substance use disorders also have demonstrated impaired performance on the Iowa Gambling Task (103 ,104 and 105), and poor performance has been shown to correlate with decreased blood flow measurements to the VM in cocaine-dependent subjects (105 ,106). The extent to which dysfunction of the VM, amygdala, or other brain regions involved in regulation of emotion and decision making may be involved in the pathophysiology of PG remains to be explored more completely.

Genetics

Twin studies investigating disordered gambling behaviors have been published (107 ,108). One study observed significantly greater rates of similarities in male monozygotic as compared with male dizygotic twins with regard to participation in past-year high-action forms of gambling (e.g., casino card, lottery, or gambling machine) (108). No differences were observed in the two groups of males with regard to measures of low action forms of gambling or in female monozygotic versus dizygotic groups with regard to past-year participation in either high- or low-action forms gambling. A larger study used the monozygotic (n = 1,869 pairs) and dizygotic (n = 1490 pairs) twins who served in the military in the Vietnam War era on whom questions pertaining to PG from the Diagnostic Interview Schedule Version III-R were available (107). The authors found inherited factors to contribute between 35% and 54% of the liability for each of the five individual PG-related factors. Higher degrees of familial contribution were estimated for the reporting of three (56%) or four (62%) or more of the individual PG-related factors. The results are comparable to findings derived from the same sample for the heritability of drug use disorders, with 34% and 28% of the variance accounted for by genetic and shared environmental factors, respectively (109). The findings reported by Eisen et al. suggest that familial factors explain a substantial portion of the risk for experiencing symptoms or behaviors consistent with PG. The findings are consistent with the notion that genetic influence significantly affects the risk of developing PG.

Molecular genetic investigations into the origin of PG have been performed (110 ,111 ,112 ,113 ,114 and 115). Investigations described to date have been association studies exploring the involvement of genes related to NE, 5-HT, and DA systems. The first of these studies investigated a role for the *D2A1* allele of the D2R, an allele previously reported by the same research group to be implicated in such compulsive-addictive behaviors as drug abuse, cocaine abuse, and compulsive eating and smoking (116 ,117). In a group of 171 whites with PG, 50.9% carried the *D2A1* allele as compared with 25.9% of the control group (odds ratio (OR) = 2.96; *p* .0000001) (110). Additionally, gambling severity was found to correlate with an increased likelihood of carrying the *D2A1* allele,

and the group of individuals without compared with those with a history of a major depressive episode were more likely to carry the *D2A1* allele (110). This latter finding suggests that differences in underlying motivations for gambling and comorbidity may be important factors in relation to the genetics of PG.

Comings and colleagues also investigated associations of PG with polymorphic variants of the D1, D3, and D4 receptors (111 ,112 ,115). The authors found the frequency of the *Dde I* allele of the D1 receptor (D1R) to be significantly higher as compared with controls in each of three groups: pathologic gamblers, tobacco smokers, and Tourette syndrome probands (111). A negative association for heterozygosity at the *Dde I* polymorphism was observed for all three disease groups. Given the findings of this (111) and their prior study (110) and the roles of the D1R and D2R in modulating rewarding and reinforcing behaviors (118), the authors proposed for the genetic variants at the D2R and D1R with regard to PG the possible existence of heterosis. *Heterosis*, with regard to populations, refers to a situation in which the progeny (hybrid) has a significantly greater effect on phenotype than either parental strain (e.g., certain hybrid strains of corn exhibiting increased vigor) (115). More investigations are needed to replicate the findings in other populations of individuals with PG and to determine the functional significance of the findings.

Allelic variants of the D4 receptor (D4R) differing in the number of 48-base pair nucleotide repeats have been implicated in some studies of novelty-seeking behavior (119 ,120), but not others (121 ,122 ,123 and 124). Moreover, the corresponding proteins derived from the allelic variants demonstrate functional differences (125), a finding lending further support to the concept that differences in the genetic composition of the groups at the D4R site may be directly related to differences in D4R function. Two groups have independently investigated a role for the D4R in PG (113 ,115), with each group finding differences in groups of individuals with PG. Perez de Castro and colleagues reported a significant positive association between PG and the longest allele of the D4R (D7), with a stronger association observed in the female group and a nonsignificant relationship seen in the male group (113). Conversely, Comings and colleagues reported no significant association between the D7 allele carriers and PG, although a significant positive association was found with regard to the number of individuals with a high number of 48-base pair repeats (five to eight) and PG (115). The authors also reported an increase in heterozygosity at the D4R allele in association with PG, invoking the notion of heterosis. Discrepancies in the findings of the two groups may be explained by genetic heterogeneity, or the genetic contributions from these loci may be additive or modest. Taken together, the findings support a potential role for the D4R in PG, and further studies are warranted to clarify the relationship. Comings and colleagues also found a decrease in heterozygosity of the *Msc I* allele of the D3 receptor individually in groups with PG or Tourette syndrome (112). The multiple genetic findings implicating dopaminergic genes in PG lend further support for a role for DA in PG.

Molecular studies have been performed with regard to genes involved in modulating 5-HT function (112 ,114). One report did not find a statistically significant association between PG and allelic variants of the tryptophan 2,3-dioxygenase gene, whose gene product regulates 5-HT metabolism (112). A variant of the 5-HT transporter (5-HTT) promoter region (126), associated with altered protein expression (127), was previously implicated in anxiety (127) and depression (128). Specifically, individuals with at least one copy of the short variant, associated with decreased protein levels, were found to have higher measures of anxiety or depression (127 ,128). Ibanez and Perez de Castro and their colleagues reported an increased association between the short (less functional) variant and PG in the group of males but not females studied, with increasing association observed with increasing severity of PG (129). These findings (a) further support a role for 5-HT dysregulation in PG and (b) suggest the direct target of 5-HT reuptake inhibitors (SRIs), drugs with apparent efficacy in the treatment of PG (see later), may be differentially regulated in certain groups with PG. Further studies are warranted to replicate and extend these findings.

Treatment

Multiple interventions, including imaginal desensitization (130) and aversion therapies (131), have been examined for the treatment of PG (132). Many treatments explored to date have often been inadequate. Arguably the most long-standing form of treatment, Gambler's Anonymous (GA), is associated with an 8% 1-year retention rate, with most participants leaving after one or two meetings (133). Preliminary results of investigations into the efficacy of cognitive behavioral therapy appear promising (134 ,135), although additional studies, particularly with regard to and in conjunction with pharmacotherapies, are warranted.

Psychopharmacology

Relatively few investigations have been performed into the tolerability and efficacy of drug treatments for PG (Table 120.2). Although most studies involve case reports or series, larger-scale, placebo-controlled trials are emerging.

Catagery Reference	Drug	Sample	Design		Outcome
Mood stabilizers					
	Moskowitz, 1980 (137)	Lithium	3 males with comorbid cycling mood disorders	Open-label	Improved control of gambling, cycling mood, risk taking, and mania/hypomania
	Haller & Hinterhuber, 1994 (138)	Carbemazepine	1 male	Placebo-controlled, double-blind, crossover	Decreased gambling behavior maintained at 30 wk
Serotonin reuptake inhibitors					
	Hollander, et al., 1992 (140)	Clomipramine	1 female with comorbid social phobia	Placebo-controlled, double-blind, crossover	Gambling behaviors discontinued persisting through 38 wk
	Hollander, et al., 1998 (141)	Fluvoxamine	16 subjects entered, 10 completed (4 female, 6 male)	Placebo-controlled, single-blind, 16-wk trial (8-wk placebo, 8-wk active)	Seven of 10 completers determined to be responders by PG-CGI and PG-YBOCS scores
	Hollander, et al., 2000 (142)	Fluvoxamine	15 subjects enrolled, 10 completed (10 male)	Placebo-controlled, double-blind, crossover (1-wk placebo Lead-in, 8-wk active/ placebo, 8-wk crossover)	Seven of 10 completers determined to be responders by PG-CGI and PG-YBOCS scores; fluvoxamine superior to placebo, particularly at end of 16 wk of treatment
	Blanco-Jerez, et al., 1999 (143)	Fluvoxamine	34 subjects enrolled	Placebo-controlled trial of 6 mo	No statistically significant difference in response rates for placebo as compared with active drug; high rates of discontinuation were seen
	De La Gandera, et al., 1999 (144)	Fluoxetine	20 subjects enrolled (11 receiving drug and psychotherapy, 9 psychotherapy only)	Open-label trial of 6 mo	Fluoxetine plus psychotherapy better than psychotherapy alone at 6 mo as measured by CGI scores and other measures
	Kim, 2000 (145)	Paroxetine	41 subjects (20 receiving paroxetine, 21 placebo)	Placebo-controlled, double-blind, parallel group (1-wk placebo lead-in, 8 wk of active medication of placebo)	Paroxetine group significantly improved as compared with placebo as determined by CGI; no statistically significant difference on other outcome measures
Opioid antagonists	Kim, 1998 (152)	Naltrexone	1 male with comorbid compulsive shopping behavior	Open-label	Gambling and excessive shopping behaviors eliminated at 100 mg/d with gains maintained through 9 mo
	Crockford & El-Guebaly, 1998 (153)	Naltrexone	1 male with comorbid alcohol dependence and depression	Open-label	Cessation in gambling and alcohol cravings observed through 4 wk following addition of naltrexone to fluoxetine

CGI, clinical global impression; PG, pathologic gambling; YBOCS, Yale-Brown Obsessive-Compulsive Scale.

TABLE 120.2. PSYCHOPHARMACOLOGIC TRIALS IN PATHOLOGIC GAMBLING

Mood Stabilizers

Lithium, a salt with mood stabilizing properties believed to modulate 5-HT systems (136), has been examined in the treatment of PG (137). In three men with PG and comorbid cycling mood disorders, lithium, at daily doses reported up to 1,800 mg per day, was found to be at least partially effective in controlling gambling, cycling mood, hypomania

and mania, and risk-taking behaviors. Durations of treatment were not clearly specified, although at least one patient was maintained for up to 1½ years. No adverse effects were described. The author concluded that lithium may target "an affective component with excitability and impulsiveness-explosiveness." Larger controlled studies seem warranted and are currently ongoing to determine the tolerability and efficacy of lithium, particularly in groups of individuals with PG and cycling mood disorders.

A case report involving the use of carbamazepine in the treatment of a 37-year-old man with PG has been described (138). The gambler had a 16-year history of significant gambling, with periods of abstinence lasting only 2 to 3 months apiece despite participation in GA, behavior therapy, and psychoanalysis. A placebo-controlled, double-blind trial of carbamazepine was undertaken, with no improvement noted in gambling behavior over the 12-week placebo phase. Carbamazepine was introduced at 200 mg per day was increased to 600 mg per day, with blood levels of 4.8 to 9.5 µg/mL achieved. Gambling behaviors decreased 2 weeks into treatment, and gains were maintained at 30 months.

Serotonin Reuptake Inhibitors

Given the efficacy of SRIs in targeting OC behaviors in OCD (139) and the data supporting 5-HT dysregulation in PG, trials of SRIs have been performed. The first of these studies involved a 31-year-old woman with PG and comorbid social phobia and OC personality traits who had been gambling persistently despite multiple prior treatments (140). Clomipramine was administered in double-blind, placebo-controlled fashion in a crossover design. Minimal improvement was seen after 10 weeks of placebo treatment. After initiation of active drug at 25 mg per day with an increase up to 175 mg per day, gambling behavior was discontinued at week 3, with absence of gambling remaining at 38 weeks. The adverse effect of increased irritability was effectively treated with a temporary decrease in dose.

More recently, a single-blind crossover study of the selective SRI (SSRI) fluvoxamine was performed (141). Sixteen subjects entered the 16-week trial (8-week placebo lead-in, 8-week active), with seven of ten completers judged to be responders by (a) a score of "much improved" or "very much improved" on the Clinical Global Impression score for gambling severity (PG-CGI) and (b) greater than 25% reduction in scores on the PG modification of the Yale-Brown Obsessive-Compulsive Scale (PG-YBOCS). Of the completers, four were female and six were male. The medication was well-tolerated, and the average dose for completers was 220 mg per day at endpoint, with responders tending to be treated with a slightly lower dose (207 mg per day on average). Noncompleters left the study during the placebo phase (four for noncompliance, two for lack of response). Of the three nonresponders, two were the only completers with histories of cyclothymia, a finding raising the possibility that individuals with a comorbid cycling mood disorder may respond better to an alternate pharmacotherapy. Further studies are warranted to investigate this possibility.

Hollander and colleagues performed a randomized double-blind, placebo-controlled crossover study of fluvoxamine in the treatment of PG (142). The trial lasted 16 weeks after a 1-week placebo lead-in phase, with subjects randomized to receive either active medication or placebo during the first 8-week phase followed by the alternate treatment during the second 8-week phase. Fifteen subjects meeting the criteria for PG but not for active substance use disorders or past or present major axis I disorders were enrolled, and 10 individuals (all male) completed the study. Two of the five noncompleters left during the placebo lead-in phase, one during placebo treatment in phase I, and two during phase I treatment with fluvoxamine (one for noncompliance, one for interaction with an as-needed medication). Study drug dosing was initiated at 50 mg per day, with fixed increases to 100 and 150 mg per day during the second and third weeks, respectively. Thereafter, the dose was adjusted in 50-mg increments each week with a maximum of 250 mg per day and a minimum of 100 mg per day, based on clinical response and drug tolerance. Mean endpoint dose of fluvoxamine was 195±50 mg per day (range, 100 to 250 mg per day). Adverse effects documented during fluvoxamine treatment were of only mild intensity and were consistent with SSRI treatment, and they were not associated with early withdrawal from the study. Outcome measures included scores from the PG-CGI and PG-YBOCS, as earlier. Data from the investigation demonstrated active drug to be superior to placebo in targeting gambling behavior over time. Both the groups receiving active medication and placebo showed improvement in control of gambling behaviors during the first 8 weeks, and the most significant difference in response was observed at the end of the second 8-week block (Fig. 120.2). In other words, during phase II, improvements seen during the course of the 16-week trial in the placebo-fluvoxamine treatment group were more likely to persist over time, whereas initial gains observed in the fluvoxamine-placebo treatment group declined. These findings are consistent with a high initial rate of placebo responders and suggest that acute trials of longer duration may be important in better distinguishing response to placebo and active drug.

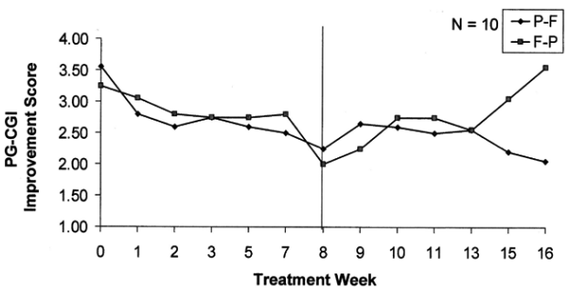


FIGURE 120.2. Clinical Global Impression (CGI) scores are shown for subjects completing a 16-week placebo-controlled, double-blind study of fluvoxamine for the treatment of PG. Measures are shown for individuals receiving placebo in phase I followed by fluvoxamine in phase II *(diamonds)* or fluvoxamine in phase I followed by placebo in phase II *(squares)*. ANOVAR: [F = 14.8 (df = 1,8), p = .005 (drug effect)]; ANOVAR: [F = 6.0 (df = 1,8), p = .040 (phase order × drug interaction)]; *post hoc* ANOVA: phase I: [F = 0.113 (df = 1,7), p = .747]; phase II:[F = 12.45 (df = 1,7), p = .010].

A longer-term placebo-controlled trial of fluvoxamine in the treatment of PG was reported by an independent group (143). In their study, 34 patients were treated for 6 months with placebo or fluvoxamine at 200 mg per day. Outcome was measured by quantification of time and money spent on gambling. The authors found no statistically significant differences in response rates to placebo as compared with active drug for the overall sample. The authors reported observing a statistically significant superiority of fluvoxamine as compared with placebo in the male and younger-aged subgroups of individuals with PG in the study. Strikingly,

a large proportion of individuals did not complete the 6-month study, and this complicated interpretation of the results and suggested that long-term compliance with drug treatment will be a significant consideration for individuals with PG.

A separate longer-term, open-label trial of a different SSRI, fluoxetine, was performed (144). The study compared the results of treatment with fluoxetine at 20 mg per day with support psychotherapy (n = 11) as compared with psychotherapy alone (n = 9). Measures of outcome included scores on the CGI and Ludo-Cage test. The treatment group receiving fluoxetine showed significantly improved outcomes as measured by CGI scores (fluoxetine plus psychotherapy: 1.5±0.8; psychotherapy alone: 3.2±0.7; $p < .001$) and Ludo-Cage mean scores ($p = .004$) at the 6-month assessments. Individuals in the combined fluoxetine and psychotherapy treatment group also demonstrated better adherence to treatment guidelines.

An independent double-blind, placebo-controlled trial of a third SSRI, paroxetine, was performed (145). The study used a parallel group design with each group receiving a 1-week placebo lead-in followed by 8 weeks of either placebo or active medication. Dosing was initiated at 20 mg per day with increases up to 60 mg per day as clinically indicated. Forty-one patients meeting the criteria for PG and no other axis I diagnosis participated in the study (20 paroxetine, 21 placebo). Adverse effects were observed with greater frequency in the paroxetine-treated group (2.3 treatment-emergent symptoms per patient in the paroxetine group as compared with 1.2 in the placebo group). The treatment-emergent symptoms were consistent with SSRI treatment, most frequently involving reports of headaches, fatigue, and dry mouth. Outcome was measured by scores on the patient- and clinician-rated CGI and the Gambling Symptom Assessment Scale (G-SAS). The paroxetine treatment as compared with placebo resulted in statistically significant improvement as determined by the clinician-rated CGI (random regression analysis: $z = -1.99$, $p < .05$), with no statistically significant differences observed between groups on other measures.

Taken together, findings from these initial studies suggest that SSRIs are well-tolerated, efficacious drugs for the treatment of PG. Larger scale (e.g., multicenter), placebo-controlled trials of SRIs are warranted to extend these initial promising results and to define better the short- and long-term efficacies and tolerabilities of specific SRIs in groups of individuals with PG.

Opioid-Receptor Antagonists

The mOR, involved in regulation of DA reward- and reinforcement-related pathways, has been the target for pharmacotherapies in the treatment of addictive disorders. Naltrexone, an mOR antagonist, has been shown to reduce alcohol intake and alcohol cravings in the treatment of alcohol dependence (146 ,147 and 148), as well as to target impulsive, self-injurious behaviors in multiple other patient populations (149 ,150 and 151). Two case reports described a potential role for naltrexone in the treatment of individuals with PG (152 ,153). In an open-label case series of individuals with ICDs, Kim described a 55-year-old man with PG and CB (152). Naltrexone at 50 mg per day was initiated with no clinical change observed at 2 weeks. Several days after an increase to 100 mg per day, a significant decrease in gambling urge intensity was reported by the patient. This decrease was followed by elimination of gambling and excessive buying, with gains maintained for at least 9 months.

An independent case report described the open-label treatment of a 49-year-old man with comorbid PG, depression, and alcohol dependence (153). The patient was initially treated with fluoxetine (dose and duration not specified), with improvements in mood and persistence in urges to drink and gamble. Naltrexone at 50 mg per day was added to the fluoxetine with a cessation in gambling and alcohol cravings observed over a 4-week period. Given the data supporting gambling-induced opioidergic changes (67), the role of opioid function in modulating DA reward

and reinforcement pathways, and the results of the present study, larger-scale, placebo-controlled trials of mOR antagonists seem warranted.

COMPULSIVE BUYING

Part of "120 - Pathologic Gambling and Impulse Control Disorders "

Although recognized by Kraeplin and Bleuler a century ago, CB, then termed oniomania and more recently compulsive shopping or impulsive or addictive buying, has been relatively understudied in psychiatry (154 ,155 and 156). Although not formally listed in the DSM-IV (1), CB has a set of proposed diagnostic criteria (157), which include maladaptive preoccupation with or engagement in buying and the preoccupations or actual buying leading to significant distress or impairment. Additionally, the behavior cannot be better accounted for by a manic episode. Prevalence estimates have been made at 1% to 8% of the general population (157 ,158 and 159). Initial reports describe individuals with CB as generally in their thirties and predominantly female, with a 4:1 or greater female-to-male ratio (156 ,159 ,160). Individuals with CB are reported to have elevated rates of psychiatric comorbidity, particularly anxiety disorders, mood disorders, substance abuse or dependence, eating disorders, ICDs, and personality disorders (157 ,159 ,161).

Pharmacotherapy

Thymoleptic Treatment

Some authors have proposed depression as a significant underlying motivational factor related to engagement in CB (158 ,162 ,163). An early description of pharmacotherapeutic interventions in PG described the use of three antidepressant medications, fluoxetine, bupropion, and nortriptyline, in CB (164). Each of the three patients receiving the medications reported a partial or complete reduction in CB symptoms. In a larger study of 20 individuals with CB, nine of 13 patients who had received thymoleptic pharmacotherapy while they were symptomatic (69%) reported their CB to be in full (n = 5) or partial (n = 4) remission (157). The 20 study participants were predominantly female (n = 16) and often carried active comorbid diagnoses, most frequently mood (n = 18) or anxiety (n = 12) disorders. The effective drugs used varied widely and included bupropion, lithium, valproate, nortriptyline, desipramine, fluoxetine, sertraline, trazodone, clonazepam, diazepam, levothyroxine, and methylphenidate, often used in combination of two or more drugs simultaneously (157). Doses and durations of pharmacotherapy were not clearly defined in the report. The authors described full remissions up to only 7 months and partial remissions up to 13 months and noted that several of the drug trials were terminated after only a short period secondary to intolerable adverse effects (n = 2) or hypomania (n = 1). Although the relationship between pharmacotherapy and CB symptoms cannot be precisely determined from the study, the findings support the need for further systematic investigations into drug treatments for CB.

Selective Serotonin Reuptake Inhibitors

Given the repetitive, ritualistic buying behaviors and the intrusive preoccupations with buying associated with CB, the efficacy of SSRIs in the treatment of OCD, and the initial findings with SSRIs described earlier, an open-label trial of fluvoxamine in CB was undertaken (165). The ten participants in the study met the criteria for CB, as proposed by McElroy et al. (157), and not for an active mood or substance use disorder. Nine participants were female, and the average age of the group was 41.4±9.2 years. The study design included a 1-week placebo lead-in followed by an 8-week period of treatment with fluvoxamine and a subsequent drug taper and discontinuation (over 3 to 4 days) and reassessment off medication at the end of week 13. Responses were measured with the YBOCS modified for CB (YBOCS-SV), the CGI, patient self-rating, and other standardized scales for depression, disability, and OC symptoms. Nine of ten individuals were deemed responders, having a more than 50% reduction in YBOCS-SV scores at week 9 as compared with baseline. Highly significant improvements were observed at week 9 as compared with baseline in scores on both the obsession and compulsion subscales of the YBOCS-SV, the National Institute of Mental Health OC scale, patient self-rating reports, subscales of the Sheehan Disability Scale, and the CGI severity and improvement scales. Symptoms appeared to worsen but often remained improved from baseline during the 4-week discontinuation phase. The adverse effects reported were consistent with fluvoxamine's use in other patient populations, with sedation, headache, dry mouth, and gastrointestinal disturbances reported most frequently. The appearance of adverse effects did not result in discontinuation of the drug for any of the participants. The results from this initial study of fluvoxamine in the treatment of CB suggest it to be efficacious and well-tolerated and support the need for larger scale, placebo-controlled, double-blind studies of SSRIs in CB.

Opioid Antagonists

Given data supporting efficacy of the μ-opioid antagonist naltrexone in urge regulation and the role of μ-opioid function in modulating MCL DA pathways, a trial of naltrexone in the treatment of ICDs (including CB) was reported (152). Two patients with CB treated with naltrexone were described in detail in a series of 15 individuals with ICDs, with an additional three responders with CB mentioned in the report. One of the two responders had comorbid PG and CB, and this response is described earlier (in the PG

section). A second individual, a 46-year-old woman with comorbid CB and bulimia nervosa, was started on naltrexone at 50 mg per day. She initially developed diarrhea, which later resolved without discontinuation of the drug. After not experiencing improvement in target symptoms, her dose was increased to 100 mg per day at week 2. At this dose, she reported a significant decrease in thoughts and behaviors related to excessive shopping and disordered eating. She maintained her gains at 7 months and tolerated the medication with normal liver function tests and without adverse effects. The results from this initial report of open-label, high-dose naltrexone administration suggest that the drug may be effective in targeting symptoms of CB. Larger-scale, placebo-controlled, double-blind studies are warranted to define better the efficacy and tolerability of the drug in the short- and long-term treatment of individuals with CB.

COMPULSIVE SEXUAL BEHAVIOR

Part of "120 - Pathologic Gambling and Impulse Control Disorders "

Traditionally, the majority of attention given to disordered sexual behaviors has arguably been focused on the paraphilias. These disorders involve sexual arousal from inappropriate objects or partners and include fetishism, exhibitionism, voyeurism, sadomasochism, pedophilia, and zoophilia. Nonparaphilic excessive sexual behavior, currently classified as an “ICD not otherwise specified” in the DSM, involves repetitive, interfering sexual behavior without the use of inappropriate objects or partners (166). The term CSB has been used to encompass both paraphilic and nonparaphilic sexual disorders (167). CSB has been estimated to affect 3% to 6% of individuals in the United States (167 ,168 and 169), with most of those with the disorder thought to be male (167 ,170 ,171). Given the relatively high estimated prevalence rates and the clinical or social impairment often experienced with CSB, there exists a need for further well-defined studies into the epidemiology and treatment of CSB.

Pharmacotherapy

Thymoleptics

High rates of mood disorders have been reported in individuals with CSB (167 ,172). Case reports have been described supporting the efficacy of multiple thymoleptics in the treatment of CSB. Specifically, the following have been reported: electroconvulsive therapy (173) and treatments with lithium (174 ,175 and 176), buspirone (177), imipramine (178 ,179), desipramine (180), clomipramine (180), and the SSRIs (172 ,178 ,181 ,182 ,183 ,184 ,185 ,186 and 187), particularly fluoxetine and sertraline. In the following section, we describe one of the larger, systematic investigations performed to date.

Selective Serotonin Reuptake Inhibitors

Initial studies into the efficacy and tolerability of SSRIs in the treatment of paraphilic and nonparaphilic CSBs have been performed (187). In one study, 20 men with CSB were entered into a 12-week open-label trial of fluoxetine (172). Ten of the men had solely nonparaphilic CSBs, and the other ten had both paraphilic and nonparaphilic CSBs. Nineteen met the criteria for comorbid dysthymia and 11 for current major depression. Outcome measures included the Inventory to Diagnose Depression (IDD) and the Sexual Outlet Inventory (SOI). IDD scores were obtained at baseline and weeks 4, 8, and 12. Of the 20 entered participants, four discontinued (three nonparaphilic and one paraphilic, one each for alcohol abuse, no change in CSB, increase in CSB after initial remission, and increased anxiety and CSB). The mean dose of fluoxetine at week 12 was 39.37±14.81 mg per day. Significant reductions in both depressive and CSB symptoms were observed, with improvement in sexual symptoms independent of baseline depression scores. Sexual symptoms showing significant improvement included total sexual outlet and unconventional forms of masturbation, sexual activity, desire intensity, and sexual interests. Conventional sexual symptoms were not adversely effected. The promising results of this open-label study warrant larger, placebo-controlled, double-blind studies of specific subgroups of individuals with CSB to determine further the efficacy and tolerability of fluoxetine and other SSRIs.

Dopamine Augmentation

In individuals who respond incompletely to SSRIs, trials of augmentation with the DA-enhancing drugs methylphenidate or bupropion have been described. The rationale for use of these drugs has been described as related to multiple findings, including the efficacy of similar augmentation strategies in depressive disorders, improvement of SRIinduced adverse effects with these DA “agonists,” and comorbidity and similarities with attention-deficit/hyperactivity disorders (172). One investigator reports having treated more than 30 patients with the combination of an SRI and a DA drug (172). Further studies are needed to both explore possible DA dysfunction in CSB and to determine the efficacies and tolerabilities of DA drugs in CSB.

Hormone System Treatments

Several classes of drugs modulating hormonal systems, including antiandrogens, estrogens, and gonadotropin-releasing hormone (GnRH) analogues, have been investigated in the treatment of CSBs (187 ,188 ,189 ,190 ,191 ,192 and 193). The group of antiandrogens includes medroxyprogesterone acetate (MPA) and cyproterone acetate (CPA). MPA, a potent progestogen lacking antiandrogen effects at the androgen receptor level, has been tested in targeting CSB (reviewed in refs. 187 and 193).

CPA, also a potent progestogen but also with testosterone antagonist activity at the receptor level, has also been studied in the treatment of CSB (190 ,194). The results of placebo-controlled trials of MPA, CPA, or both (190 ,194 ,195), as well as data from a large number of open-label trials and case reports (reviewed in refs. 187 and 193), suggest a role for the drugs in the management of groups of individuals with CSBs, particularly sex offenders and elderly individuals with aggressive sexual behaviors. Although these agents are not effective for all patients (196), accumulating data suggest this family of drugs may be effective in subgroups of individuals with CSB, particularly those with repetitive deviant sexual behaviors (193). The drugs are limited by the emergence of adverse effects, including commonly weight gain, fatigue, hypertension, headaches, hyperglycemia, leg cramps, and diminished spermatogenesis (187). More rarely, feminizing effects may be seen, and thromboembolic phenomena may be seen more frequently with use of the drugs (187). Additional studies are needed to determine the long-term efficacy and tolerability of the antiandrogen drugs MPA and CPA in nonparaphilic and paraphilic CSBs.

Fewer studies have been performed to date to test the efficacy and tolerability in CSB of estrogens such as diethylstilbestrol (DES) (197) or transdermal estrogen (198) and GnRH analogues such as triptorelin (191) and leuprolide (196). Estrogen treatment works in a similar fashion to MPA and CPA in terms of decreasing luteinizing hormone (LH) and follicle-stimulating hormone (FSH) secretion and thereby decreasing testosterone production. In contrast, GnRH analogues suppress testosterone production by stimulation of LH and FSH secretion leading to increased testosterone and estrogen levels. Continued administration results in insensitivity of the pituitary to GnRH, significantly reducing LH and FSH levels. One of the larger studies includes an open-label trial of triptorelin in the treatment of 30 men with paraphilic CSB. Treatment of 8 to 42 months' duration with injectable triptorelin at 3.75 mg per month resulted in significant decreases in deviant sexual fantasies and behaviors along with significant decreases in serum testosterone levels (191). Adverse effects included erectile failure, hot flashes, and decrease in bone mineral densities. As with antiandrogens, adverse effect profiles of estrogens and GnRH analogues may preclude widespread use of these classes of drugs in the treatment of CSB. However, a role for these drugs may exist in the treatment of specific subgroups of individuals with CSBs.

COMPULSIVE COMPUTER USE

Part of "120 - Pathologic Gambling and Impulse Control Disorders "

With the increasing availability of personal computers and the rapid expansion in use of the Internet, the emergence of disordered computer use has been described (199 ,200 ,201 ,202 and 203). The Internet provides immediate access to a broad range of impulsive behaviors, including gambling (both traditional forms as well as stock trading), shopping, and sexual behaviors, and it may be associated with an increase in the prevalence of related ICDs. Individuals with excessive and interfering computer use have been termed "webaholics" or "cyberholics," and their computer-related behaviors have been termed computer or Internet addiction or dependency, Internet addictive disorder, cyberaddiction, or CCU (204 ,205 and 206). Given the recent emergence of CCU, formal diagnostic criteria have not been developed or endorsed, pharmacologic treatment studies have not been reported, and its relationship with other disorders (e.g., OCD, substance use, mood and ICDs) has not been adequately investigated. Nonetheless, initial investigations into the characteristics of individuals with CCU have been performed.

One study using a 94-item questionnaire described four factors relating to CCU in a sample of college students (205). From 341 completed questionnaires, four factors (two major and two minor) were identified and were found to explain 31% of the variance. Factor 1 focused on problematic use of the Internet and included significant contributions from questions relating to staying on line too long, having a restricted repertoire of interests, and interference with sleep, diet, exercise, and attending meetings. Within this factor, lower levels of correlation but each above 0.35 were noted for computer hacking and gambling and use of the Internet to relieve sadness or loneliness. The second major factor focused on the usefulness and general purpose of computers or the Internet. Positive correlations with extensive use of the Internet and finding information loaded onto this factor. Interestingly, questions related to on-line shopping and downloading of nude images loaded onto factor 2, as compared with on-line gambling, which loaded onto factor 1. Given the estimated high rates of Internet use for sexually related activities (1% of a group of on-line computer users spending more than 11 hours per week) (207), definition of the characteristics of ICD-related online behaviors seems particularly important. A positive correlation with campus and negative correlation with home computer use also was observed within factor 2. Factor 3 focused on questions relating to a combination of shyness and introversion and using the Internet for sexual gratification (high correlation with physical arousal and downloading of nude images), whereas factor 4 focused on the absence of Internet problems and a mild aversion to or lack of interest in the technology. These initial findings suggest that some similar motivations that underlie other ICDs may lead to CCU and provide data that could be helpful in defining diagnostic characteristics for CCU.

A second investigation recruited 21 individuals (16 men, five women) with self-reported excessive computer use that interfered with social or occupational functioning or caused personal distress (206). Participants were mainly between 20 and 50 years of age, with 43% of the total sample falling between the ages of 21 and 29 years, inclusively. Observed were high rates of attempting to cut back on computer usage (62%) and comorbid psychiatric illness, including mood,

substance use, anxiety, and personality disorders. The findings of the studies by Black et al. and Praterelli et al. suggest the need for additional studies into the epidemiology and neurobiology of CCU, as well as pharmacologic and behavioral treatments for individuals with CCU.

CONCLUSIONS AND FUTURE DIRECTIONS

Part of "120 - Pathologic Gambling and Impulse Control Disorders "

Although understudied for a significant period of time by the psychiatric community, PG and other ICDs appear to be receiving an increasing amount of clinical attention. Significantly more research is needed to diminish the gap in our knowledge of ICDs as compared with other psychiatric illnesses and to optimize treatment strategies for the large number of individuals suffering from these disorders.

ACKNOWLEDGMENTS

Part of "120 - Pathologic Gambling and Impulse Control Disorders "

Dr. Hollander has received research support and/or served as a consultant or on a speaker's bureau for the following companies: Solvay, Abbott, SmithKline Beecham, Lilly, Wyeth-Ayerst and Bristol Myers Squibb.

REFERENCES

1. American Psychiatric Association Committee on Nomenclature and Statistics. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington, DC: American Psychiatric Association, 1994.

2. Gabriel K. Gambling and spirituality: a new anthropological perspective. www.nmia.com/˜kgabriel/ 1998.

3. Anonymous. *Mahabharata: adiparva*. Ganguli KM, trans. Calcutta and Evanston, IL: Bharata Press and American Theological Library, 1884:647.

4. Weems ML. *God's revenge against gambling exemplified in the miserable lives and untimely deaths of a number of persons from both sexes, who had sacrificed their health, wealth, and honor at the gaming tables*. Philadelphia: 1812.

5. Cosby AG, May DC, Frese W, et al. Legalization of crimes against the moral order: results from the 1995 United States survey of gaming and gambling. *Deviant Behav Interdisc J* 1996;17:369–389.

6. McElroy SL, Hudson JI, Pope HG Jr, et al. The DSM-III-R impulse control disorders not elsewhere classified: clinical characteristics and relationship to other psychiatric disorders. *Am J Psychiatry* 1992;149:318–327.

7. Hollander E. Treatment of obsessive-compulsive spectrum disorders with SSRIs. *Br J Psychiatry* 1998;173[Suppl 35]:7–12.

8. Hollander E, Wong CM. Obsessive-compulsive spectrum disorders. *J Clin Psychiatry* 1995;56[Suppl 4]:3–6.

9. Hollander E, Wong CM. Body dysmorphic disorder, pathological gambling, and sexual compulsions. *J Clin Psychiatry* 1995;56[Suppl 4]:7–12; discussion, 13.

10. Hollander E, Benzaquin SD. Is there a distinct OCD spectrum? *CNS Spectrums* 1996;1:17–26.

11. Siever LJ. Relationship between impulsivity and compulsivity: a synthesis. In: Oldham J, Hollander E, Skodol AE, et al., eds. *Impulsivity and compulsivity*. Washington, DC: American Psychiatric Association, 1996:261–272.

12. Linden RD, Pope HG Jr, Jonas JM. Pathological gambling and major affective disorder: preliminary findings. *J Clin Psychiatry* 1986;47:201–203.

13. Simeon D, Hollander E, Cohen L. Obsessive compulsive related disorders. In: Hollander E, Zohar J, Marazziti D, eds. *Current insights in obsessive compulsive disorder*. Chichester, UK: Wiley, 1994:53–66.

14. Rasmussen SA, Eisen JL. *Epidemiology and clinical features of obsessive-compulsive disorder*. Littleton, MA: Medical Publishers, 1990:10–27.

15. Hollander E, Stein DJ, Kwon JH, et al. Psychosocial function and economic costs of obsessive-compulsive disorder. *CNS Spectrums* 1997;2:16–25.

16. Kagan DM. Addictive personality factors. *J Psychol* 1987;121:533–538.

17. Specker SM, Carlson GA, Edmonson KM, et al. Psychopathology in pathological gamblers seeking treatment. *J Gambling Stud* 1996;12:67–81.

18. Black DW, Goldstein RB, Noyes R, et al. Compulsive behaviors and obsessive-compulsive disorder (OCD): lack of a relationship between OCD, eating disorders, and gambling. *Compr Psychiatry* 1994;35:145–148.

19. Blaszczynski A. Pathological gambling and obsessive-compulsive spectrum disorders. *Psychol Rep* 1999;84:107–113.

20. De Marchi N, Morris M, Mennella R, et al. Association of obsessive-compulsive disorder and pathological gambling with Huntington's disease in an Italian pedigree: possible association with Huntington's disease mutation. *Acta Psychiatr Scand* 1998;97:62–65.

21. Lesieur HR, Blume SB, Zoppa RM. Alcoholism, drug abuse, and gambling. *Alcohol Clin Exp Res* 1986;10:33–38.

22. Marks I. Behavioural (non-chemical) addictions. *Br J Addict* 1990;85:1389–1394.

23. DeCaria C, Hollander E. Pathological gambling. In: Hollander E, ed. *Obsessive-compulsive related disorders*. Washington, DC: American Psychiatric Association, 1993:155–178.

24. Crockford DN, el-Guebaly N. Psychiatric comorbidity in pathological gambling: a critical review. *Can J Psychiatry* 1998;43:43–50.

25. McCormick RA, Russo AM, Ramirez LF, et al. Affective disorders among pathological gamblers seeking treatment. *Am J Psychiatry* 1984;141:215–218.

26. Cunningham-Williams RM, Cottler LB, Compton WM 3rd, et al. Taking chances: problem gamblers and mental health disorders—results from the St. Louis Epidemiologic Catchment Area Study. *Am J Public Health* 1998;88:1093–1096.

27. Spunt B, Lesieur H, Hunt D, et al. Gambling among methadone patients. *Int J Addict* 1995;30:929–962.

28. Feigelman W, Kleinman PH, Lesieur HR, et al. Pathological gambling among methadone patients. *Drug Alcohol Depend* 1995;39:75–81.

29. Haberman PW. Drinking and other self-indulgences: complements or counter-attractions? *Int J Addict* 1969;4:157–167.

30. Steinberg MA, Kosten TA, Rounsaville BJ. Cocaine abuse and pathological gambling. *Am J Addict* 1992;1:121–132.

31. Harvard Medical School Division on Addictions. Screening veterans for pathologial gambling. *The WAGER* 1998;3(14):1.

32. Kaplan G, Davis B. *Gambling, alcohol, and other drugs: prevalence and implications of dual problem clients*. Winnipeg, Canada: Addicitions Foundation of Manitoba, 1997.

33. Smart RG, Ferris J. Alcohol, drugs and gambling in the Ontario adult population, 1994. *Can J Psychiatry* 1996;41:36–45.

34. Harvard Medical School Division on Addictions. Screening veterans for pathological gambling. *The WAGER* 1996;1(43):1.

35. Rugle L, Melamed L. Neuropsychological assessment of attention problems in pathological gamblers. *J Nerv Ment Disease* 1993;181:107–112.

36. Blaszczynski A, McConaghy N, Frankova A. Crime, antisocial personality and pathological gambling. *J Gambling Behav* 1989;5:137–152.

37. Blaszczynski AP, McConaghy N. Antisocial personality disorder and pathological gambling. *J Gambling Stud* 1994;10:129–145.

38. DeCaria C, Begaz T, Hollander E. Serotonergic and noradrenergic function in pathological gambling. *CNS Spectrums* 1998;3:38–47.

39. Asberg M, Traskman L, Thoren P. 5–HIAA in the cerebrospinal fluid: a biochemical predictor. *Arch Gen Psychiatry* 1976;33:1193–1197.

40. Linnoila M, Virkunnen M, Scheinen M, et al. Low cerebrospinal fluid 5–hydroxy indolacetic acid concentrations differentiates impulsive from non impulsive violent behavior. *Life Sci* 1983;33:2609–2614.

41. Coccaro EF, Siever LJ, Klar HM, et al. Serotonergic studies in patients with affective and personality disorders: correlates with suicidal and impulsive aggressive behavior. *Arch Gen Psychiatry* 1989;46:587–599.

42. Usdin E, Snyder S, eds. *Frontiers in catecholamine research*. Elmsford, NY: Pergamon, 1973.

43. Koob GF. Drugs of abuse: anatomy, pharmacology and functions of reward pathways. *Trends Pharmacol Sci* 1992;13:177–184.

44. Self DW, Nestler EJ. Molecular mechanisms of drug reinforcement. *Annu Rev Neurosci* 1995;18:463–495.

45. Koob GF, Nestler EJ. The neurobiology of drug addiction. *J Neuropsychiatry CIin Neurosci* 1997;9:482–497.

46. Moreno I, Saiz-Ruiz J, Lopez-Ibor JJ. Serotonin and gambling dependence. *Hum Psychopharmacol* 1991;6[Suppl]:9–12.

47. Caccia S, Ballabio M, Saminin R, et al. M-CPP, a central 5–HT agonist, is a metabolite of trazodone. *J Pharm Pharmacol* 1981;33:477–478.

48. Hamik A, Peroutka SJ. 1–(m-Chlorophenyl) piperazine (mCPP) interactions with neurotransmitter receptors in the human brain. *Biol Psychiatry* 1989;25:569–575.

49. Hoyer D. Functional correlates of serotonin 5–HT1 recognition sites. *J Recept Res* 1988;8:59–81.

50. Kennett GA, Curzon G. Evidence that mCPP may have behavioural effects mediated by the 5HT1C receptor. *Br J Pharmacol* 1988;94:137–147.

51. Kennett GA, Curzon G. Evidence that hypophagia induced by mCPP and TFMPP requires 5HT1C and 5HT1B receptors: hypophagia induced by RU 24969 requires only 5HT1B receptors. *Psychopharmacology (Berl)* 1998;96:93–100.

52. Moss HB, Yao JK, Panzak GL. Serotonergic responsivity and behavioral dimensions in antisocial personality associated with substance abuse. *Biol Psychiatry* 1990;28:325–338.

53. Hollander E, Stein D, DeCaria CM, et al. Serotonergic sensitivity in borderline personailty disorder: preliminary findings. *Am J Psychiatry* 1994;151:277–280.

54. Stein DJ, Hollander E, DeCaria C, et al. Behavioral responses to m-chlorophenyl-poperazine and clonidine in trichotrillomania. *J Serotonin Res* 1997;4:11–15.

55. Benkelfat C, Murphy DL, Hill JL, et al. Ethanol like properties of the serotonergic partial agonist m-chlorophenylpiperazine in chronic alcoholic patients. *Arch Gen Psychiatry* 1991;48:383.

56. Nordin C, Eklundh T. Altered CSF 5–HIAA disposition in pathologic male gamblers. *CNS Spectrums* 1999;4:25–33.

57. Roy A, Adinoff B, Roehrich L, et al. Pathological gambling: a psychobiological study. *Arch Gen Psychiatry* 1988;45:369–373.

58. Roy A, De Jong J, Linnoila M. Extraversion in pathological gamblers: correlates with indexes of noradrenergic function. *Arch Gen Psychiatry* 1989;46:679–681.

59. Bergh C, Eklund T, Sodersten P, et al. Altered dopamine function in pathological gambling. *Psychol Med* 1997;27:473–475.

60. Bertilsson L, Asberg M, Lantto O, et al. Gradients of monoamine metabolites and cortisol in cerebrospinal fluid of psychiatric patients and healthy controls. *Psychiatry Res* 1982;6:77–83.

61. Nordin C, Swedin A, Zachau A. Tapping time influences concentrations of 5–HIAA in the CSF. *J Psychiatr Res* 1993;27:409–414.

62. Nordin C, Eklundh T. Tapping-time is longer in pathological male gamblers than in healthy male controls. *J Psychiatr Res* 1998;32:421–422.

63. Breiter HC, Gollub RL, Weisskopf RM, et al. Acute effects of cocaine on human brain activity and emotion. *Neuron* 1997;19:591–611.

64. Maas LC, Lukas SE, Kaufman MJ, et al. Functional magnetic resonance imaging of human brain activation during cueinduced cocaine craving. *Am J Psychiatry* 1998;155:124–126.

65. Volkow ND, Wang GJ, Fischman MW, et al. Relationship between subjective effects of cocaine and dopamine transporter occupancy. *Nature* 1997;386:827–830.

66. Bergh C, Kuhlhorn E. Social, psychological and physical consequences of pathological gambling in Sweden. *J Gambling Stud* 1994;10:275–285.

67. Shinohara K, Yanagisawa A, Kagota Y, et al. Physiological changes in Pachinko players: beta-endorphin, catecholamines, immune system substances and heart rate. *Appl Hum Sci* 1999;18:37–42.

68. Zuckerman M. *Sensation seeking: beyond the optimal level of arousal*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.

69. Deutch AY, Roth RH. Neurochemical systems in the central nervous system. In: Charney DS, Nestler EJ, Bunney BS, eds. *Neurobiology of mental illness*. New York: Oxford University Press, 1999:10–25.

70. Oreland L, Wiberg A, Asberg M, et al. Platelet MAO activity and monoamine metabolites in cerebrospinal fluid in depressed and suicidal patients and in healthy controls. *Psychiatry Res* 1981;4:21–29.

71. Levitt P, Pintar JE, Braekefield XO. Immunocytochemical demonstration of monoamine oxidase B in brain astrocytes and serotonin neurons. *Proc Natl Acad Sci USA* 1982;79:6385–6389.

72. Buchsbaum MS, Haier RJ, Murphy DL. Suicide attempts, platelet monoamine oxidase and the average evoked response. *Acta Psychiatr Scand* 1977;56:69–79.

73. von Knorring L, Oreland L, von Knorring AL. Personality traits and platelet MAO activity in alcohol and drug abusing teenage boys. *Acta Psychiatr Scand* 1987;75:307–314.

74. Fowler CJ, Von Knorring L, Oreland L. Platelet monoamine activity in sensation seekers. *Psychiatry Res* 1980;3:273–279.

75. Ward PB, Catts SV, Norman TR, et al. Low platelet monoamine oxidase and sensation seeking in males: an established relationship? *Acta Psychiatr Scand* 1987;75:86–90.

76. Carrasco JL, Saiz-Ruiz J, Diaz-Marsa M, et al. Low platelet monoamine oxidase activity in sensation-seeking bullfighters. *CNS Spectrums* 1999;4:21–24.

77. Hallman J, Sakurai E, Oreland L. Blood platelet monoamine oxidase activity, serotonin uptake, and release rates in anorexia and bulimia patients and in healthy controls. *Acta Psychiatr Scand* 1990;81:73–77.

78. Carrasco JL, Saiz-Ruiz J, Hollander E, et al. Low platelet monoamine oxidase activity in pathological gambling. *Acta Psychiatr Scand* 1994;90:427–431.

79. Blanco C, Orensanz-Munoz L, Blanco-Jerez C, et al. Pathological gambling and platelet MAO activity: a psychobiological study. *Am J Psychiatry* 1996;153:119–121.

80. Schmitt LH, Harrison GA, Spargo RM. Variation in epinephrine and cortisol excretion rates associated with behavior in an Australian aboriginal community. *Am J Phys Anthropol* 1998;106:249–253.

81. Ramirez LF, McCormick RA, Lowy MT. Plasma cortisol and depression in pathological gamblers. *Br J Psychiatry* 1988;153:684–686.

82. James KC, Bible WA, Dobson JC, et al. National Gambling Impact Study Commission: final report to Congress. http://www.ngisc.gov/reports/fullrpt.html 1999.

83. Phillips AG, LePiane FG. Reinforcing effects of morphine microinjection on to the ventral tegmental area. *Pharmacol Biochem Behav* 1980;12:965–968.

84. von Wolfswinkel L, van Ree JM. Effects of morphine and naloxone on thresholds of ventral tegmental electrical self-stimulation. *Naunyn Schmiedebergs Arch Pharmacol* 1985;330:84–92.

85. Roy A, Berrettini W, DeJong J, et al. CSF neuropeptide Y in alcoholics and normal controls. *Psychiatry Res* 1990;33:215–219.

86. Roy A, Berrettini W, Adinoff B, et al. CSF galanin in alcoholics, pathological gamblers, and normal controls: a negative report. *Biol Psychiatry* 1990;27:923–926.

87. Roy A, DeJong J, Ferraro T, et al. CSF GABA and neuropeptides in pathological gamblers and normal controls. *Psychiatry Res* 1989;30:137–144.

88. Roy A, Pickar D, Gold P, et al. Diazepam-binding inhibitor and corticotropin-releasing hormone in cerebrospinal fluid. *Acta Psychiatr Scand* 1989;80:287–291.

89. Nordin C, Eklundh T. Lower CSF taurine levels in male pathological gamblers than in healthy controls. *Hum Psychopharmacol* 1996;11:401–403.

90. Koepp MJ, Gunn RN, Lawrence AD, et al. Evidence for striatal dopamine release during a video game. *Nature* 1998;393:266–268.

91. Goyer PF, Semple WE, Rugle L, et al. Brain blood flow and dopamine receptor PET imaging in pathological gamblers. National Conference on Problem Gambling. Detroit: 1999.

92. Nyberg S, Eriksson B, Oxenstierna G, et al. Suggested minimal effective dose of risperidone based on PET-measured D2 and 5-HT2A receptor occupancy in schizophrenic patients. *Am J Psychiatry* 1999;156:869–875.

93. Botvinick M, Nystrom LE, Fissell K, et al. Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* 1999;402:179–181.

94. Childress AR, Mozely PD, McElgin W, et al. Limbic activation during cue-induced cocaine craving. *Am J Psychiatry* 1999;156:11–18.

95. Volkow ND, Wang G-J, Fowler JS, et al. Association of methylphenidate-induced craving with changes in right striato-orbitofrontal metabolism in cocaine abusers. *Am J Psychiatry* 1999;156:19–26.

96. Lambert C. Deep cravings. *Harvard Mag* 2000;102:60–68.

97. Potenza MN, Armentano CJ, Steinberg MA, et al. An fMRI study of gambling urges in individuals with pathological gambling. World Congress of Biological Psychiatry Convention. Berlin: 2001.

98. Bechara A, Damasio AR, Damasio H, et al. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 1994;50:7–15.

99. Bechara A, Damasio H, Tranel D, et al. Deciding advantageously before knowing the advantageous strategy [see Comments]. *Science* 1997;275:1293–1295.

100. Bechara A, Damasio H, Tranel D, et al. Dissociation of working memory from decision making within the human prefrontal cortex. *J Neurosci* 1998;18:428–437.

101. Bechara A, Damasio H, Damasio HR, et al. Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J Neurosci* 1999;19:5473–5481.

102. Bechara A, Damasio H, Damasio HR. Emotion, decision making and the orbitofrontal cortex. *Cereb Cortex* 2000;10:295–307.

103. Kaiser E, Johnson D, Hommer D. Alcoholics show impaired decision making on the Iowa Gambling Task. In: *Society for Neuroscience Annual Convention*. Miami Beach, FL: Society for Neuroscience, 1999.

104. Bechara A, Dolan S, Hindes A, et al. Decision-making deficits, linked to a dysfunctional orbitofrontal cortex, revealed in alcohol and stimulant abusers. *Society for Neuroscience Annual Convention*. Miami Beach, FL: Society for Neuroscience, 1999.

105. Grant SJ, Contoreggi CC, London ED. Drug abusers show impaired performance in a laboratory test of decision making. *Neuropsychologia* 2000;38(8):1180–1187.

106. London ED, Ernst M, Grant S, et al. Orbitofrontal cortex and human drug abuse: functional imaging. *Cereb Cortex* 2000;10:334–342.

107. Eisen SA, Lin N, Lyons MJ, et al. Familial influences on gambling behavior: an analysis of 3359 twin pairs. *Addiction* 1998;93:1375–1384.

108. Winters KC, Rich T. A twin study of adult gambling behavior. *J Gambling Stud* 1998;14:213–225.

109. Tsuang M, Lyons MJ, Eisen SA, et al. Genetic influences on DSM-III-R drug abuse and dependence: a study of 3,372 twin pairs. *Am J Med Genet* 1996;67:473–477.

110. Comings DE, Rosenthal RJ, Lesieur HR, et al. A study of the dopamine D2 receptor gene in pathological gambling. *Pharmacogenetics* 1996;6:223–234.

111. Comings DE, Gade R, Wu S, et al. Studies of the potential role of the dopamine D1 receptor gene in addictive behaviors. *Mol Psychiatry* 1997;2:44–56.

112. Comings DE. The molecular genetics of pathological gambling. *CNS Spectrums* 1998;3:20–37.

113. Perez de Castro I, Ibanez A, Torres P, et al. Genetic association study between pathological gambling and a functional DNA polymorphism at the D4 receptor gene. *Pharmacogenetics* 1997;7:345–348.

114. Perez de Castro I, Ibanez A, Saiz-Ruiz J, et al. Genetic contribution to pathological gambling: possible association between a DNA polymorphism at the serotonin transpoter gene (5HTT) and affected men. *Pharmacogenetics* 1999;9:397–400.

115. Comings DE, Gonzalez N, Wu S, et al. Studies of the 48 bp repeat polymorphism of the DRD4 gene in impulsive, compulsive, addictive behaviors: Tourette syndrome, ADHD, pathological gambling, and substance abuse. *Am J Hum Genet* 1999;88:358–368.

116. Blum K, Sheridan PJ, Wood RC, et al. Dopamine D2 receptor gene variants: association and linkage studies studies in impulsive-addictive-compulsive behavior. *Pharmacogenetics* 1995;5:121–141.

117. Blum K, Cull JG, Braverman ER, et al. Reward deficiency syndrome. *Am Scientist* 1996;84:132–145.

118. Self DW, Barnhart WJ, Lehman DA, et al. Opposite modulation of cocaine-seeking behavior by D1–and D2-like dopamine receptor agonists. *Science* 1996;271:1586–1589.

119. Benjamin J, Patterson C, Greenber BD, et al. Population and familial association between the D4 dopamine receptor and measures of novelty seeking. *Nat Genet* 1996;12:81–84.

120. Ebstein RP, Novick O, Umansky R, et al. Dopamine D4 receptor (DRD4) exon III polymorphism associated with the human personality trait of novelty seeking. *Nat Genet* 1996;12:78–80.

121. Malhotra AK, Virkunnen M, Rooney W, et al. The association between the dopamine D4 (D4DR) 16 amino acid repeat and novelty seeking. *Mol Psychiatry* 1996;1:388–391.

122. Gelernter J, Kranzler H, Coccaro E, et al. D4 dopamine-receptor (DRD4) alleles and novelty-seeking in substance-dependent, personality-disorder and control subjects. *Am J Hum Genet* 1997;61:1144–1152.

123. Jonssen EG, Nothen MM, Gustavson JP, et al. Lack of evidence for allelic association between personality traits and the dopamine D4 receptor gene polymorphisms. *Am J Psychiatry* 1997;154:697–699.

124. Sullivan PF, Fifield WJ, Kennedy MA, et al. No association between novelty seeking and the type 4 dopamine receptor gene (DRD4) in two New Zealand samples. *Am J Psychiatry* 1998;155:98–101.

125. Van Tol HH, Wu CM, Guan HC, et al. Multiple dopamine D4 receptor variants in the human population. *Nature* 1992;358:149–152.

126. Heils A, Teufel A, Petri S, et al. Allelic variation of human serotonin transporter gene expression. *J Neurochem* 1996;66:2621–2624.

127. Lesch KP, Bengal D, Heils A, et al. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 1996;274:1527–1531.

128. Collier DA, Stober G, Li T, et al. A novel functional polymorphism within the promoter of the serotonin transporter gene: possible role in susceptibility to affective disorders. *Mol Psychiatry* 1996;1:453–460.

129. Ibanez A, Saiz-Ruiz J, Perez De Castro I, et al. Is serotonin transporter gene associated with pathological gambling? American Psychiatric Association Annual Convention. Toronto: American Psychiatric Association, 1998.

130. McConaghy N, Blaszczynski A, Frankova A. Comparison of imaginal desensitisation with other behavioural treatments of pathological gambling: a two- to nine-year follow-up. *Br J Psychiatry* 1991;159:390–393.

131. McConaghy N, Armstrong MS, Blaszczynski A, et al. Controlled comparison of aversive therapy and imaginal desensitization in compulsive gambling. *Br J Psychiatry* 1983;142:366–372.

132. Petry NM, Armentano C. Prevalence, assessment, and treatment of pathological gambling: a review. *Psychiatr Serv* 1999;50:1021–1027.

133. Stewart RM, Brown RI. An outcome study of Gamblers Anonymous. *Br J Psychiatry* 1988;152:284–288.

134. Bujold A, Ladouceur R, Sylvain C, et al. Treatment of pathological gamblers: an experimental study. *J Behav Ther Exp Psychiatry* 1994;25:275–282.

135. Sylvain C, Ladouceur R, Boisvert JM. Cognitive and behavioral treatment of pathological gambling: a controlled study. *J Consult Clin Psychol* 1997;65:727–732.

136. Price LH, Charney DS, Delgado PL, et al. Lithium and serotonin function: implications for the serotonin hypothesis of depression. *Psychopharmacology (Berl)* 1990;100:3–12.

137. Moskowitz JA. Lithium and lady luck: use of lithium carbonate in compulsive gambling. *NY State J Med* 1980;80:785–788.

138. Haller R, Hinterhuber H. Treatment of pathological gambling with carbamazepine. *Pharmacopsychiatry* 1994;27:129.

139. Goodman WK, Price LH, Rasmussen SA, et al. Efficacy of fluvoxamine in obsessive-compulsive disorder: a double-blind comparison with placebo. *Arch Gen Psychiatry* 1989;46:36–43.

140. Hollander E, Frenkel M, Decaria C, et al. Treatment of pathological gambling with clomipramine [Letter]. *Am J Psychiatry* 1992;149:710–711.

141. Hollander E, DeCaria C, Mari E, et al. Short-term single-blind fluvoxamine treatment of pathological gambling. *Am J Psychiatry* 1998;155:1781–1783.

142. Hollander E, DeCaria CM, Finkell JN, et al. A randomized double-blind fluvoxamine/placebo crossover trial in pathological gambling. *Biol Psychiatry* 2000;47:813–817.

143. Blanco-Jerez C, Petkova E, Ibanez A, et al. A long-term, double-blind, placebo-controlled study of fluvoxamine for pathological gambling. In: *American Psychiatric Association Annual Convention*. Washington, DC: American Psychiatric Association, 1999.

144. De La Gandara JJ, Sanz O, Gilaberte I. Fluoxetine: open-trial in pathological gambling. In: *American Psychiatric Association Annual Convention*. Washington, DC: American Psychiatric Association, 1999.

145. Kim SW, unpublished data, 2000.

146. O'Malley SS, Jaffe AJ, Chang G, et al. Naltrexone and coping skills therapy for alcohol dependence: a controlled study. *Arch Gen Psychiatry* 1992;49:881–887.

147. Volpicelli JR, Alterman AI, Hayashida M, et al. Naltrexone in the treatment of alcohol dependence [see Comments]. *Arch Gen Psychiatry* 1992;49:876–880.

148. Volpicelli JR, Watson NT, King AC, et al. Effect of naltrexone on alcohol "high" in alcoholics. *Am J Psychiatry* 1995;152:613–615.

149. Barrett RP, Feinstein C, Hole WT. Effects of naloxone and naltrexone on self-injury: a double-blind, placebo-controlled analysis. *Am J Ment Retard* 1989;93:644–651.

150. Walters AS, Barrett RP, Feinstein C, et al. A case report of naltrexone treatment of self-injury and social withdrawal in autism. *J Autism Dev Disord* 1990;20:169–176.

151. Roth AS, Ostroff RB, Hoffman RE. Naltrexone as a treatment for repetitive self-injurious behaviour: an open-label trial. *J Clin Psychiatry* 1996;57:233–237.

152. Kim SW. Opioid antagonists in the treatment of impulse-control disorders. *J Clin Psychiatry* 1998;59:159–164.

153. Crockford DN, el-Guebaly N. Naltrexone in the treatment of pathological gambling and alcohol dependence [Letter]. *Can J Psychiatry* 1998;43:86.

154. Kraeplin E. *Psychiatrie*. Leipzig, Germany: Verlag Von Johann Ambrosius Barth, 1915:408–409.

155. Bleuler E. *Textbook of psychiatry*. 1924, New York: Macmillan, 1924:540.

156. Christenson GA, Faber RJ, de Zwaan M, et al. Compulsive buying: descriptive characteristics and psychiatric comorbidity. *J Clin Psychiatry* 1994;55:5–11.

157. McElroy SL, Keck PE, Pope HG Jr, et al. Compulsive buying: a report of 20 cases. *J Clin Psychiatry* 1994;55:242–248.

158. Lejoyeaux M, Ades J, Tassain V, et al. Phenomenology and psychopathology of uncontrolled buying. *Am J Psychiatry* 1996;153:1524–1529.

159. Black DW, Repertinger S, Gaffney GR, et al. Family history and psychiatric comorbidity in persons with compulsive buying: preliminary findings. *Am J Psychiatry* 1998;155:960–963.

160. Faber RJ, O'Guinn TC. A clinical screener for compulsive buying. *J Consumer Res* 1992;19:459–469.

161. Schlosser S, Black DW, Repertinger S, et al. Compulsive buying: demography, phenomenology and comorbidity in 46 subjects. *Gen Hosp Psychiatry* 1994;16:205–212.

162. Lejoyeaux M, Hourtane M, Ades J. Compulsive buying and depression. *J Clin Psychiatry* 1995;56:38.

163. Lejoyeaux M, Tassain V, Solomon J, et al. Study of compulsive buying in depressed patients. *J Clin Psychiatry* 1997;58:169–173.

164. McElroy SL, Satlin A, Pope HG Jr, et al. Treatment of compulsive shopping with antidepressants: a report of three cases. *Ann Clin Psychiatry* 1991;3:199–204.

165. Black DW, Monahan P, Gabel J. Fluvoxamine in the treatment of compulsive buying. *J Clin Psychiatry* 1997;58:159–163.

166. Coleman E. Sexual compulsivity: definition, etiology, and treatment considerations. In: Coleman E, ed. *Chemical dependency and intimacy dysfunction*. New York: Hawarth, 1987.

167. Black DW. The epidemiology and phenomenology of compulsive sexual behavior. *CNS Spectrums* 2000;5:26–35.

168. Coleman E. Is your patient suffering from compulsive sexual behavior? *Psychiatr Ann* 1992;22:320–325.

169. Carnes P. *Don't call it love: recovery for sexual addiction*. New York: Bantam Books, 1991.

170. Carnes PJ, Delmonico DL. Childhood abuse and multiple addictions: research findings in identified sexual addicts. *Sex Addict Compuls* 1996;3:258–268.

171. Schneider JP, Schneider BH. Couple recovery from sexual addiction/coaddiction: results of a survey of 88 marriages. *Sex Addict Compuls* 1996;3:111–126.

172. Kafka MP, Prentky R. Fluoxetine treatment of nonparaphilic sexual addictions and paraphilias in men. *J Clin Psychiatry* 1992;53:351–358.

173. Eyres A. Transvestism: employment of somatic therapy with subsequent improvement. *Dis Nerv Syst* 1960;21(1):52–53.

174. Bartova D, Nahunek K, Svetske J. Pharmacological treatment of deviant sexual behavior. *Activ Nerv Sup (Praha)* 1978;20:72–74.

175. Ward NG. Successful lithium treatment of transvestitism associated with manic-depression. *J Nerv Ment Dis* 1975;161:204–206.

176. Cesnik JA, Coleman E. Use of lithium carbonate in the treatment of autoerotic asphyxia. *Am J Psychother* 1989;63:277–286.

177. Federoff JP. Buspirone hydrochloride in the treatment of transvestic fetishism. *J Clin Psychiatry* 1988;49:408–409.

178. Kafka MP. Successful antidepressant treatment of nonparaphilic sexual addictions and paraphilias in men. *J Clin Psychiatry* 1991;52:60–65.

179. Snaith RP. Five exhibitionists and a method of treatment. *Br J Psychiatry* 1981;132:126–130.

180. Kruesi MJP, Fine S, Valladares L, et al. Paraphilias: a double-blind crossover comparison of clomipramine vs. desipramine. *Arch Sex Behav* 1993;21:587–593.

181. Federoff JP. Serotonergic drug treatment of deviant sexual interests. *Ann Sex Res* 1993;6:105–121.

182. Greenberg DM, Bradford JMW. Treatment of the paraphilic disorders: a review of the role of the selective serotonin reuptake inhibitors. *Sex Abuse J Treat Res* 1997;9:349–360.

183. Kafka MP. Sertraline pharmacotherapy for paraphilias and paraphilia-related disorders: an open trial. *Ann Clin Psychiatry* 1994;6:189–195.

184. Perilstein RD, Lipper S, Friedman LJ. Three cases of paraphilias responsive to fluoxetine treatment. *J Clin Psychiatry* 1991;52:169–170.

185. Bradford JMW. An open pilot study of sertraline in the treatment of outpatients with pedophilia. In: *American Psychiatric Association Annual Convention*. Washington, DC: American Psychiatric Association, 1995.

186. Greenberg DM, Bradford JMW, Curry S, et al. A comparison of treatment of paraphilias with three serotonin reuptake inhibitors: a retrospective study. *Bull Am Acad Psychiatry Law* 1996;24:525–532.

187. Kafka MP. Psychopharmacologic treatments for nonparaphilic compulsive sexual behaviors. *CNS Spectrums* 2000;5:49–59.

188. Meyer WI, Cole C, Emory E. Depo provera treatment for sex offending behavior: an evaluation of outcome. *Bull Am Acad Psychiatry Law* 1992;20:249–259.

189. Gottesman HG, Schubert DSP. Low-dose oral medroxyprogesterone acetate in the management of the paraphilias. *J Clin Psychiatry* 1993;54:182–188.

190. Bradford JMW, Pawlak A. Double-blind placebo crossover study of cyproterone acetate in the treatment of the paraphilias. *Arch Sex Behav* 1993;22:383–403.

191. Rosler A, Witzum E. Treatment of men with paraphilia with a long-acting analogue of gonadotropin-releasing hormone. *N Engl J Med* 1998;338:416–422.

192. Levitsky AM, Owens NJ. Pharmacologic treatment of hypersexuality and paraphilias in nursing home residents. *J Am Geriatr Soc* 1999;47:231–234.

193. Grossman LS, Martis B, Fichtner CG. Are sex offenders treatable? A research overview. *Psychiatr Serv* 1999;50:349–361.

194. Cooper AJ, Sandhu S, Losztyn S, et al. A double-blind placebo-controlled trial of medroxyprogesterone acetate and cyproterone acetate with seven pedophiles. *Can J Psychiatry* 1992;37:687–693.

195. Kiersch T. Treatment of sex offenders with Depo-provera. *Bull Am Acad Psychiatry Law* 1990;18:179–187.

196. Dickey R. The management of a case of treatment-resistant paraphilia with a long-acting LHRH agonist. *J Clin Psychiatry* 1992;37:567–569.

197. Kyomen HH, Satlin A, Hennen J, et al. Estrogen therapy and aggressive behavior in elderly patients with moderate-to-severe dementia: results from a short-term, randomized, double-blind trial. *Am J Geriatr Psychiatry* 1999;7:339–348.

198. Lothstein LM, Fogg-Waberski, J, Reynolds P. Risk management of sexual disinhibition in geriatric patients. *Conn Med* 1997;61:609–618.

199. Young KS. Pathological Internet use: the emergence of a new clinical disorder. American Psychological Association Annual Convention. Toronto: American Psychological Association, 1996.

200. Young KS. Psychology of computer use. XL. Addictive use of the Internet: a case that breaks the stereotype. *Psychol Rep* 1996;79:899–902.

201. Belsare TJ, Gaffney GR, Black DW. Compulsive computer use. *Am J Psychiatry* 1997;154:289.

202. Young KS. Internet addiction: what makes computer-mediated communication habit-forming? American Psychological Association Annual Convention. Chicago: American Psychological Association, 1997.

203. Griffiths MD. Internet addiction: does it really exist? In: Gackenbach J, ed. *Psychology and the Internet: intrapersonal, interpersonal, and transpersonal implications*. New York: Academic, 1998:61–75.

204. Fearing JD. Computer addiction: "hooked on the Net." www.nationalcounseling.com/cmpadict.html 2000.

205. Pratarelli ME, Browne BE, Johnson K. The bits and bytes of computer/Internet addiction: a factor analytic approach. *Behav Res Methods Instr Comput* 1999;31:305–314.

206. Black DW, Belsare G, Schlosser S. Clinical features, psychiatric comorbidity, and health-related quality of life in persons reporting compulsive computer use behavior. *J Clin Psychiatry* 1999;60:839–844.

207. Cooper A, Delmonico DL, Burg R. Cybersex users, abusers, and compulsives: new findings and implications. Sexual Addiction and Compulsion 2000;7(1 and 2):5–29.

121

Compulsive and Impulsive Aspects of Self-Injurious Behavior

Dan J. Stein

Joseph Zohar

Daphne Simeon

Dan J. Stein: University of Stellenbosch, Cape Town, South Africa and University of Florida, Gainesville, Florida.

Joseph Zohar: Sackler Medical School, Tel Aviv, Israel.

Daphne Simeon: Mt. Sinai School of Medicine, New York, New York.

Self-injurious behavior (SIB) is seen in a range of different psychiatric disorders, including stereotypic movement disorder, Tourette's disorder, borderline personality disorder, and psychotic disorders. An immediate question is to what extent similar psychobiological mechanisms mediate SIB across so diverse a spectrum of conditions. It has been argued that specific neurobiological mechanisms may be responsible for various symptoms across diagnostic categories; could the same hold true for self-injurious behaviors?

A closer exploration of the phenomenology of SIB suggests that there are a number of distinct types of self-mutilation (1). "Compulsive" self-injurious behavior is arguably exemplified by the stereotypic self-injurious behavior of stereotypic movement disorder (SMD). By definition, such behavior is composed of repetitive, seemingly driven, but nonfunctional motor behavior. Stereotypical SIB is also seen in a range of specific syndromes, including Lesch-Nyhan syndrome, Cornelia de Lange syndrome, and Prader-Willi syndrome. Although SMD is commonly encountered in patients with developmental disabilities, there is also evidence of it in intellectually normal adult samples (2), in the form of behaviors such as skin picking and nail biting.

The *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV) specifically excludes from the diagnosis of SMD the repetitive self-injurious symptoms of a range of other disorders such as pervasive developmental disorder, Tourette's disorder, and trichotillomania. Some of these behaviors arguably have an "impulsive" component in that they are preceded by increasing tension, and followed by pleasure, gratification, or relief. Similarly, certain self-injurious behaviors, perhaps predominantly in personality-disordered patients, can plausibly be described as "impulsive" in nature. This phenomenologic distinction between "compulsive" and "impulsive" is intended only to be heuristic; clinically, there may be significant overlap, with the impulsive wrist-cutter, for example, going on to develop apparently compulsive repetitive SIB.

This chapter briefly reviews work from animal studies of SIB, including animal stereotypies and veterinary behavioral disorders, before going on to consider the neuropsychopharmacology of a range of clinical disorders characterized by SIB. These include SMD in developmentally disabled patients, specific syndromes characterized by stereotypic SIB, SMD in intellectually normal patients (e.g., skin picking, nail biting, etc.), Tourette's disorder, autistic disorder, trichotillomania, and impulsive SIB in predominantly personality-disordered patients. Throughout the chapter we follow a schema of reviewing phenomenology, neurochemistry, and neuroanatomy in turn.

ANIMAL STEREOTYPIES

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

Phenomenology

Stereotypy may be defined as the excessive production of one type of motor act (3). Stereotypic behavior is species specific, with common behaviors including grooming, gnawing, and pacing. A subset of these stereotypies are self-injurious; these include excessive grooming and self-biting. Animal stereotypies can be induced by confinement (e.g., to a small enclosure) and by deprivation (e.g., being reared alone). Deprivation stereotypies are more likely to result in SIB (4), particularly in higher mammals (3).

Neurochemistry

Animal research on stereotypic self-injurious behavior has highlighted the role of the dopaminergic, serotoninergic,

and opioid systems. These systems also have important interactions with one another. Nevertheless, for the sake of simplicity, we list studies on each system in turn.

Early work with rodents demonstrated that amphetamines act to produce an increase in stereotypic behavior (5), including automutilation (6), and such findings have been replicated with different dopamine agonists in different species. These responses are prevented by 6-hydroxy-dopamine–induced lesions or by dopamine antagonists (7 ,8), perhaps particularly by D1 antagonists (9 ,10). Furthermore, early destruction of dopaminergic neurons may lead to hypersensitivity of D1 receptors, with increased self-biting behavior in response to later administration of dopamine agonists (11).

Traditional serotonin models include the administration of a monoamine oxidase inhibitor with L-tryptophan to induce a hyperactivity syndrome in rats, and the use of 5-hydroxytryptophan to induce the head-twitch syndrome in mice and the wet dog shake in rats. Reviewing this literature, Jacobs and Fornal (12) conclude that serotonin facilitates gross motor output and inhibits sensory information processing. Certainly, isolation may be associated with decreased serotonin turnover (13). Furthermore, in confining or depriving environments, the animal may activate dorsal raphe neurons by performance of stereotypies (12).

In animal work, opioid agonists may induce autoaggression, and opioid antagonists may be particularly effective in reducing self-injurious stereotypies in younger animals (14). Indeed, it has been suggested that excessive opioid activity is responsible for SIB (15). However, an alternative hypothesis emphasizes that pain associated with SIB results in the release of brain endorphins and draws a parallel between such endogenous release of endorphins and addiction to an exogenous substance (16).

From an integrated brain-mind perspective, interactions and overlaps between environmental and pharmacologic inducers of stereotypy would be predicted. Indeed, compared to normally reared rodents, isolation-reared subjects show increased stereotypic responses after administration of amphetamine (17) or tail-pinch (18).

Neuroanatomy

There is evidence that corticostriatal circuits are important in mediating stereotypic behavior (3). First, infusion of the dopamine into the caudate results in stereotyped orofacial behaviors (grooming, gnawing). Conversely, infusion of dopamine blockers into the same areas reduces amphetamine-induced stereotypy. Second, frontal lesions and dopamine agonists produce similar behavioral and cognitive effects, and frontal lesions exacerbate the effects of these agents. Furthermore, striatal lesions may also lead to stereotypic behavior. Reviewing this literature and his own work, Ridley (3) concludes that loss of inhibition induced by frontal lesions results in a release of previously stored response sequences. This view is consistent with a view of striatal function that emphasizes the development, maintenance, and selection of motoric and cognitive procedural strategies. Different terms given to allude to this group of functions have included habit system, response set, and procedural mobilization (19).

VETERINARY BEHAVIORAL DISORDERS

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

Phenomenology

In addition to animal laboratory studies, there is an interesting literature on veterinary behavioral disorders characterized by SIB (20). Acral lick dermatitis, for example, is a condition characterized by excessive paw licking and scratching in dogs. This results in the characteristic dermatitis; in severe cases sequelae include osteomyelitis. The disorder is seen in certain breeds of large dog, and within breeds may be more common in particular families.

Different species may manifest a range of different kinds of SIB (20). Psychogenic alopecia is found in cats, with excessive depilation leading to bare patches. Feather-picking in birds is seen in a range of avian species, and can be complicated by severe hemorrhage. Although stereotypies may appear to arise spontaneously in companion and domestic animals, as in laboratory animals, confinement and deprivation are robust elicitors of stereotypic and self-injurious behavior.

Neurochemistry

The pharmacotherapeutic profile of acral lick dermatitis overlaps remarkably neatly with that of obsessive-compulsive disorder (OCD) (21). Thus, the disorder responds to selective serotonin reuptake inhibitors (SSRIs), but fails to respond to desipramine or fenfluramine. However, the disorder may also respond to opioid agents. Although not well studied pharmacologically, there are a number of reports indicating that psychogenic alopecia responds to treatment with SSRIs. In addition, administration of a dopamine blocker has been noted to lead to a decrease in symptoms. Similarly, feather-picking in birds and SIB in horses may respond to treatment with SSRIs (20).

We noted earlier the cross-sensitization between pharmacologic and environmental inducers of stereotypies (22). Conversely, it can be emphasized that environmentally induced SIB may respond to psychopharmacologic intervention. For example, in a placebo-controlled study of fluoxetine in isolation-reared primates, this SSRI was effective in reducing such symptoms (23).

Neuroanatomy

In the previous section, the possible involvement of frontostriatal circuits in animal stereotypic behavior was noted.

With regard to stereotypic SIB in higher mammals, a particularly interesting finding is that socially isolated primates develop striatal cellular disorganization together with stereotypic and self-injurious behaviors (24). The neuroanatomy of early social isolation and other developmental stressors would seem to be a promising area for further investigation, one that may have direct relevance to some of the clinical disorders considered in the next part of this chapter.

SMD IN DEVELOPMENTAL DISABILITY

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

Phenomenology

SIB in patients with developmental disabilities may fall within the diagnostic category of SMD. This disorder is characterized by repetitive, seemingly driven, but nonfunctional motor behavior. Examples include body rocking, hand waving, head banging, and skin picking. The DSM-IV provides a subtype “with self-injurious behavior,” to be used when bodily damage requires medical treatment. We have reviewed this disorder elsewhere (25 ,26), and draw extensively on those reviews here.

A number of the DSM-IV criteria for SMD address the severity of the behavior. It must “markedly interfere with normal activities” or result in “bodily injury that requires medical treatment.” In patients with developmental disability the behavior must be sufficiently severe to be a focus of treatment, and the behavior must persist for at least 4 weeks. The DSM-IV also states that stereotypic behaviors in SMD should not be better accounted for by the compulsions of obsessive-compulsive disorder, the stereotypies of pervasive developmental disorder, or the tics of Tourette’s disorder.

SIB in patients with developmental disability may not always meet the rather strict DSM-IV criteria for SMD. Nevertheless, there is evidence that stereotyped, self-injurious, and compulsive behaviors appear to be correlated in such patients (27). Certainly, the relatively large body of work on the pathogenesis and treatment of stereotypic behaviors in this population may be useful in understanding and managing SIB.

Reports of the incidence of SIB in patients with developmental disability range from 3% to 46% (27 ,28). Head banging, head and body hitting, eye gouging, biting, and scratching are the most common of these behaviors (29). Behaviors may cause permanent and disabling tissue damage and may sometimes be life threatening. For example, severe head banging or hitting may lead to cuts, bleeding, infection, retinal detachment, and blindness. Incidence of SIB in patients with developmental disability is dependent on a number of factors that vary within this heterogeneous population, including extent of cognitive impairment (29) and institutionalization status (28).

Neurochemistry

It has been argued that there is phenomenologic evidence of similarities between SIB in patients with developmental disability and symptoms in patients with OCD (28). Furthermore, as noted earlier, in patients with developmental disability, there are significant positive associations between the occurrence of self-injury, stereotypy, and compulsions (27). Nevertheless, relatively few studies have directly explored the role of serotonin in mediating SIB in such patients.

There is, however, some evidence of the value of SSRIs in the treatment of SIB in developmental disabled patients. A retrospective review suggested that response of SIB in such patients was higher for SSRIs and was not predicted by comorbid depressive symptoms (30). Indeed, both open and controlled (31) studies have confirmed the efficacy of SSRIs in the treatment of SIB in developmental disabilities. Although this research is consistent with a role for serotonin in the mediation of SIB in mental retardation, questions remain about whether the effects of these agents are not “downstream” of their primary actions and about their specificity. Several methodologies are available for delineating different aspects of serotonin dysfunction in psychiatric patients, and the “pharmacologic dissection” strategy of comparing responses to clomipramine and desipramine has been useful in showing that serotonin plays a specific role in several disorders characterized by unwanted repetitive and/or self-injurious behaviors (32 ,33). A similarly designed study in mental retardation patients with SIB would be of interest.

It may be noted here that other agents with serotoninergic effects have also been studied for the treatment of SIB in developmental disability (28 ,34). 5-Hydroxytryptophan (5-HT) has been shown useful in only a minority of open studies. Buspirone (15 to 45 mg/day), a 5-HT_{1A} agonist, has been somewhat effective in small groups of adults with developmental disability and SIB. Eltoprazine, a selective 5-HT_{1A} and 5-HT_{1B} agonist, has yielded conflicting evidence of efficacy.

A number of agents, such as lithium and beta-blockers, have multiple neurotransmitter effects including serotoninergic effects. Although early studies in this area suffered from methodologic flaws, lithium has long been used with some apparent success in the treatment of SIB and aggressive behaviors in patients with mental retardation (28 ,34). Propranolol (90 to 410 mg/day) was reported to reduce SIB and aggression in a small case series of patients with mental retardation. Also, in a controlled study, pindolol (40 mg/day) was significantly more effective than placebo in 14 patients with developmental disability and SIB (35).

There has been relatively little direct work on the dopamine system in patients with developmental disability and SIB. Dopamine blockers, however, are often successfully used to manage SIB in such patients (28 ,34). Preliminary evidence suggests that the atypical neuroleptics, which have both dopaminergic and serotoninergic effects, may also be useful in SIB and other target symptoms in this patient population (36 ,37). Given their apparently favorable side-effect

profile, controlled trials with such agents are warranted.

Increased plasma enkephalin levels in patients with developmental disability compared with normal controls have been reported (38). Although this may support the excessive opioid hypothesis, it is also possible that decreased endogenous brain opioid levels ultimately lead to compensatory overproduction (39). It has also been argued, however, that opioid effects on self-injury may be primarily mediated via the dopamine or serotonin system (28).

There is some evidence that the opiate antagonists naloxone and naltrexone lead to a reduction in frequency of self-injury in different patient populations, including those with developmental disability (28 ,34). However, the total number of patients in such studies is relatively small, and the study designs have been criticized (39 ,40). Indeed, in a placebo-controlled study of 32 subjects with mental retardation and SIB and/or autism, naltrexone (50 mg/day) failed to have an effect on SIB and increased the incidence of stereotypic behavior (40). Although this finding does not entirely rule out a role for the opioid system, it further emphasizes the need for caution in drawing conclusions from open trials of treatment for SIB.

SIB may vary in women with mental retardation according to the stage of menstrual cycle. In one study, fluctuations in SIB were associated with early and late follicular phases. Although there are currently insufficient data for a specific association between self-injurious behavior and hormonal factors, such an association warrants further attention.

Further attention should perhaps also be paid to the role of the γ-aminobutyric acid (GABA)ergic system in SIB in developmental disability. The use of benzodiazepines in this population has not been well studied. The anticonvulsant valproic acid, however, was effective in reducing SIB and aggression in 12 of 18 patients with mental retardation and affective symptoms in a 2-year open trial (41). Positive response to valproate was associated with a past history of seizure disorder.

SYNDROMES WITH SELF-INJURIOUS BEHAVIOR

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

Phenomenology

Lesch-Nyhan syndrome (LNS) is an X-linked recessive disorder of purine synthesis. Patients present with hyperuricemia and neuropsychiatric symptoms including spasticity, choreoathetosis, dystonia, mental retardation, aggression, and self-injurious behavior. SIB in LNS is dramatic, and most commonly consists of biting of fingers and lips, although head banging, tongue biting, eye/nose poking, and self-scratching also occur (42).

Cornelia de Lange syndrome (CLS) is a rare congenital disorder characterized by a distinctive appearance and mental retardation. Patients with CLS manifest excessive grooming behavior (hand licking and hair stroking) and SIB including head slapping and self-scratching (43).

Prader-Willi syndrome (PWS) is a congenital disorder that affects approximately 1 in 10,000 newborns and is one of the five commonest abnormalities seen in birth defect clinics (53a ,53b). PWS is associated with marked hyperphagia, and the disorder is the most common dysmorphic form of obesity. In addition, PWS is characterized by behavioral disturbances, mental retardation, sleep disturbances, neonatal hypotonia, and hypogonadism. Behavioral disturbances include compulsive self-mutilation, impulsive temper outbursts, and classic obsessive-compulsive behaviors (44 ,45). SIB is common and not necessarily associated with cognitive impairment (45). It includes skin and nose picking, nail biting, lip biting, and hair pulling (45). Patients frequently have chronic skin sores.

Neurochemistry

Both the biochemical abnormality [virtual absence of hypoxanthine-guanine phosphoribosyltransferase (HPRT)] and the underlying genetic defect (a mutation of the HPRT locus located at q26-28 of the X-chromosome) in LNS are well identified. However, the mechanisms underlying neuropsychiatric symptoms are less clear. Nevertheless, the biochemical systems implicated do include the dopaminergic and serotoninergic systems.

Early postmortem findings in three patients with LNS demonstrated dramatically reduced levels of dopamine, homovanillic acid, and dopa decarboxylase in the basal ganglia (46), placing particular emphasis on the dopamine system. Animal studies with HPRT-deficient models and clinical studies of neurotransmitters and metabolite levels in patients with the disorder support the importance of dopaminergic mediation of symptoms (47).

It is interesting to note that although patients with both LNS and Parkinson’s disorder demonstrate decreased dopamine neurons and motoric symptoms, there are important differences. Parkinson’s disorder, for example, is characterized by diminished motor output, whereas LNS is characterized by uncontrolled and exaggerated motor activity. These differences may reflect the importance of the developmental stage at which dopaminergic deficits occur (11). Although dopamine supersensitivity has also been hypothesized to play a role in SIB in LNS (9), and dopamine blockers have been used for their treatment, long-term treatment with these agents is not clearly of benefit in LNS. Further understanding of the developmental neurobiology underlying this syndrome is therefore needed.

Phenomenologic similarities have been noted between excessive grooming in CLS and animal dopamine agonist-induced stereotypies. Nevertheless, there is little direct research to support a dopamine hypothesis of CLS. Indeed, the underlying neurobiology of this interesting syndrome remains relatively poorly understood.

Research has also focused on the serotonin system in LNS. Studies of cerebrospinal fluid (CSF) 5-hydroxyindoleacetic acid (5-HIAA) have been inconsistent in LNS, but there may be slightly increased putamen serotonin and 5-HIAA (46). In addition, early reports suggested 5-HT (1 to 8 mg/kg) was useful in the treatment of self-injurious behaviors in LNS. However, only a minority of subsequent studies have confirmed this finding (28).

In an early study, CLS patients were found to have reduced whole blood serotonin levels (48). Several possible mechanisms underlying this finding were considered, including a dysfunction in serotonin metabolism, failure to bind to platelets, and transporter abnormalities. Again, however, putative serotonin dysfunction may simply reflect dysfunction in other systems.

Serotoninergic mediation of PWS is raised by the role of serotonin and the efficacy of serotoninergic agents in appetite control and eating disorders, compulsive skin picking, impulsive aggression, and obsessive-compulsive–related disorders. A double-blind trial of fenfluramine found that this agent was useful for weight loss and other-directed aggressive behavior in PWS patients, but did not affect SIB (49). However, the SSRI fluoxetine has been described as useful for SIB in a number of cases of PWS. Similarly, a survey of caregivers suggested that SSRIs may be helpful for both impulsive-aggressive and compulsive symptoms in some PWS patients (45).

There is relatively little work on other systems in these disorders. GABAergic systems have been postulated to play a role in LNS, but again this hypothesis has not translated into successful pharmacotherapeutic strategies. Opioid antagonists have been reported to decrease appetite in some PWS patients, but controlled work has not supported their efficacy.

Abnormalities of chromosome 15 have been implicated in the etiology of PWS, and recent research using cytogenetic and molecular techniques suggests that identification of a specific genetic basis is possible in most patients. In about 70% of patients a cytogenetically visible deletion can be detected in the paternally derived chromosome 15 (15q11q13), whereas in about 20% of patients both copies of chromosome 15 are inherited from the mother (maternal uniparental disomy). Genotype-phenotype correlations in 167 patients with PWS found no significant difference in skin picking between patients with or without a chromosomal deletion (50). Nevertheless, the abnormal gene product in PWS may ultimately provide crucial information on the neurochemistry of the SIB characteristic of these patients.

Neuroanatomy

Most recently, volumetric magnetic resonance imaging (MRI) and positron-emission tomography (PET) techniques have documented reduced caudate volume and reduced dopamine transporters in caudate and putamen (51) as well as decreased dopa decarboxylase activity and dopamine storage throughout the dopaminergic system in LNS (52). It is possible that in LNS there is reorganization of the cortical–basal ganglia–thalamic pathways during development (52).

A number of authors have suggested hypothalamic-pituitary dysfunction in PWS. Again, however, further work is needed to extend these preliminary findings and to determine the relationship with behavioral symptoms. Indeed, at present little is ultimately understood about the underlying neurobiology of self-injury and other behavioral symptoms in this fascinating disorder.

SMD IN INTELLECTUALLY NORMAL INDIVIDUALS

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

Phenomenology

Skin picking and scratching appear to be not uncommon symptoms (53). The incidence of so-called neurotic excoriations in dermatology clinics has been estimated to be around 2%. Medical complications of skin picking include infection and scarring. Furthermore, skin picking may be associated with significant distress and dysfunction. At times, patients with these behaviors may meet criteria for OCD. However, in many other patients, this is not the case. Skin picking may also be seen in patients with the putative OCD spectrum disorders trichotillomania and body dysmorphic disorder.

Nail biting (onychophagia) is a common behavior that is not, however, necessarily benign (33). Nail biting may be associated with serious infection, nail bed damage and scarring, craniomandibular dysfunction, and dental disorders. The apparent ubiquity of mild nail biting should not discourage clinical and research attention to patients with more severe forms of the behavior.

A range of other common self-injurious stereotypies may be seen in intellectually normal adults including lip biting and eye rubbing (1). Certain kinds of stereotypies, such as thumb sucking and head banging, appear more common in children, although on occasion these behaviors may also be seen in intellectually normal adults (2). In a college population, the total number of stereotypic behaviors was significantly associated with increased scores of obsessivecompulsive symptoms, of perfectionism, and of impulsive-aggressive traits (54). Importantly, stereotypic behaviors may be associated with significant medical complications, and they may also lead to distressing feelings of shame and lowered self-esteem, as well as to social avoidance and occupational impairment.

Neurochemistry

Case reports suggested that the SSRIs may have a role in the treatment of skin picking. Indeed, in a series of 30

patients with skin picking, an open trial of sertraline demonstrated efficacy (55). Similarly, a retrospective treatment review of body dysmorphic disorder patients with skin picking indicated that SSRIs were often effective, whereas other agents were not (56). Finally, in their controlled study Simeon and colleagues (53) found that fluoxetine was significantly superior to placebo in decreasing compulsive skin picking in intellectually normal patients.

In nail biting, clomipramine appeared more effective than desipramine, although results were not perhaps as robust as those seen in classic OCD (33). The authors emphasized that there was a high dropout rate at every stage of the study, which appeared in sharp contrast to that seen in other psychiatric populations. They did, however, suggest that their data were consistent with the hypothesis that similar biological systems mediate a spectrum of grooming disorders, including OCD and trichotillomania.

Castellanos and colleagues (2) compared clomipramine and desipramine in a crossover trial of SMD patients. Although clomipramine appeared promising in a number of cases, too few patients completed the trial to demonstrate a clear benefit of clomipramine over desipramine. Nevertheless, several case reports suggest that SSRIs may be useful in patients with skin picking, head banging, and other self-injurious stereotypic behaviors (26). Given that dopamine agonists may result in SIB (57), a possible role for dopamine blockers, and the new atypical neuroleptics in particular, also warrants further consideration. Ultimately, controlled and long-term studies are needed to formulate rational approaches to the pharmacotherapy of SMD.

Neuroanatomy

To our knowledge, there have been no studies on the neuroanatomy of stereotypic movement disorder in normal controls. Given the ubiquity of these behaviors, and the presumptive role of cortico-striatal-thalamic-cortical (CSTC) circuits, this seems an area that may be worth investigating in more detail.

AUTISM

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

Phenomenology

Autistic disorder (or classic autism) is a pervasive developmental disorder characterized by impairment in social interactions, communication deficits, and restrictive and stereotyped behaviors. Stereotyped SIBs are common in patients with this disorder and may also be seen in other pervasive developmental disorders that do not meet the narrower criteria for autistic disorder (58). Common forms of SIB in autism include hand/wrist biting, head banging, self-scratching, self-hitting, self-pinching, and hair pulling.

It has been argued that repetitive behaviors in autism cannot simply be subsumed under the banner of OCD. Indeed, compared to patients with OCD, adults with autism were found to have a different range of repetitive symptoms. They were more likely to demonstrate repetitive ordering, hoarding, touching, tapping, or rubbing, and self-injurious behaviors (59). Nevertheless, it may be postulated that there are at least some similarities in the underlying neurobiological mediation of autism and OCD.

Neurochemistry

There does seem to be evidence of serotoninergic dysfunction in autism (60). Several studies have found elevated platelet serotonin levels in autism. Neuroendocrine challenge studies with serotoninergic agents have indicated reduced serotoninergic responsivity in autism. Furthermore, in a tryptophan depletion study, autism resulted in increased SIB, motor stereotypies, and anxiety.

Despite early reports of the efficacy of fenfluramine in open trials in autism, subsequent controlled trials were disappointing. However, both open and placebo-controlled (61) trials with SSRIs have demonstrated efficacy in reducing symptoms such as SIB in autism. Furthermore, the SSRI clomipramine was more effective than the noradrenergic reuptake inhibitor desipramine in autism (62). Nevertheless, not all studies of these agents have been positive (63).

Other neurochemical systems may also play a role in the mediation of self-injurious behaviors in autism. A PET study demonstrated reduced dopaminergic activity in the anterior medial prefrontal cortex (64). Controlled trials have demonstrated that dopamine blockers (like SSRIs) are effective in about 50% of patients with autism for target symptoms including SIBs (65 ,66). Clinical experience indicates that where a medication is ineffective in autism, an agent from a different class of medication may be useful (60). The atypical neuroleptics, with their combined dopaminergic and serotoninergic effects, also warrant further study.

Various authors have suggested a role for the opioid system in autism (40). However, studies of opioid levels in autism have been inconsistent. Furthermore, despite promising open trials, in controlled studies the effect of opioid blockers on target symptoms including SIB in autism has been disappointing.

There is promise for delineating the specific albeit multiple genetic factors underlying autism (60). Interestingly, there is preliminary evidence of a familial link with Tourette’s disorder. Most recently, a possible link to the serotonin-transporter gene has been suggested. Such work may ultimately lead to a clearer understanding of the neurochemistry of autism and self-injury and to specific therapeutic interventions.

Neuroanatomy

The neuroanatomy of autism has also received increasing attention in recent years (67). Preliminary postmortem studies have found abnormalities in the cerebellum and limbic

system, including the hippocampus and amygdala. Neurophysiologic research has demonstrated various abnormalities including aberrant processing in frontal association cortex. Early work with pneumoencephalography suggested left temporal horn dilatation, and an early MRI study found hypoplasia of the posterior cerebellar vermis, but later studies have been inconsistent. Functional brain imaging studies are also so far inconsistent, although perhaps suggestive of dysfunction in association cortex. Clearly, much remains to be done to understand the neuroanatomy of SIB, and indeed to integrate behavioral and biological findings in this disorder.

TOURETTE'S SYNDROME

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

Phenomenology

Compulsive self-injurious behavior is only rarely seen in OCD. In contrast, SIB is seen in 13% to 53% of Tourette's syndrome (TS) patients (68). A wide range of behaviors may be seen, particularly head banging and self-punching or slapping, but also including lip biting and tongue biting, eye poking, skin picking, and self-punching or -slapping. Medical complications have included subdural hematoma and vision impairment.

In a large study, SIB in TS was not correlated with intellectual function, but was significantly associated with severity of motor tics and with scores of hostility and obsessionality (68). Furthermore, SIB has been described as one of the compulsions that are more common in patients with TS than in OCD. Thus, although the neurobiology of SIB per se in TS has not been well studied, it is possible that this overlaps with that underlying tics and compulsions.

Neurochemistry

Several neurochemical systems have been implicated in TS, most notably the dopamine system, but including also the serotoninergic, noradrenergic, opiatergic, hormonal, and immunologic systems (see Chapter 117). However, to our knowledge little of this work has focused specifically on the neurochemistry of SIB in TS.

Neuroanatomy

From a neuroanatomic perspective, there is strong evidence that prefrontal–basal ganglia–thalamic circuits are involved in OCD. There is also increasing evidence that these circuits are among those that mediate TS (see Chapter 117). Of note, increased metabolism in the orbitofrontal cortex and putamen correlated with complex behavioral and cognitive features such as self-injurious behavior (69). As in OCD, further work is needed to determine whether this reflects a primary deficit or functional compensation.

TRICHOTILLOMANIA

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

Phenomenology

The term *trichotillomania* was coined over a century ago to describe patients with hair pulling. Hair pulling most frequently occurs from the scalp, although it can occur from a wide range of body areas, including the eyebrows, eyelashes, beard, axillae, and pubis (70). Plucking may be confined to a single patch, may involve different areas, or may cover the entire scalp. Some patients also report pulling hair from a child, significant other, or pet.

Patients with hair pulling may demonstrate a range of other stereotypic and self-injurious behaviors (70 ,71). Although hair pulling may lead to significant medical complications, including trichobezoar after ingestion of pulled hair, it is perhaps more commonly associated with significant feelings of shame and lowered self-esteem. Indeed, both the personal and the economic costs of this disorder may be significant.

Neurochemistry

Research on the neurobiology of hair pulling was boosted by a seminal trial comparing clomipramine and desipramine in trichotillomania (32). As in OCD, trichotillomania responded selectively to the SSRI. Nevertheless, although the SSRIs have seemed effective for trichotillomania in a number of open trials, these agents have proved disappointing in placebo-controlled trials (72).

Furthermore, although Swedo and colleagues found that trichotillomania response to clomipramine may be sustained over time, there are also reports that initial response to SSRIs in patients with hair pulling may be lost during continued treatment (72). Taken together, this work indicates that it may be premature to overly emphasize the specific role of serotonin in trichotillomania.

Indeed, few studies of trichotillomania patients have directly assessed monoamine concentrations. Ninan and colleagues (73) obtained CSF from a small group of patients with trichotillomania and found that CSF 5-HIAA levels did not differ from normal controls. However, baseline CSF 5-HIAA did correlate significantly with degree of response to SSRIs. This finding is redolent of some work on OCD and suggests that in both disorders response to SSRIs may be accompanied by a fall in CSF 5-HIAA levels.

There are also few studies of serotoninergic pharmacologic challenges in trichotillomania. Stein and colleagues (74) found that the 5-HT agonist m-chlorophenylpiperazine (mCPP), which has exacerbated OCD symptoms in some studies, did not lead to an increase in hair pulling in women with trichotillomania (74). The interpretation of these data is not straightforward; for example, whereas OCD symptoms may be present throughout the day, hair pulling is often triggered only in particular settings. Of interest,

however, trichotillomania subjects described an increase in feeling “high,” a phenomenon previously documented in patients with borderline personality disorder. It might be speculated that hair pulling in trichotillomania and self-injurious stereotyped behaviors in impulsive personality patients have some overlapping characteristics (71).

There is increasing evidence that dopamine plays a role in OCD and related disorders, perhaps particularly in those with a marked motoric component. There is some preliminary data that dopamine also plays a role in hair pulling. One report noted exacerbation of hair pulling by methylphenidate in a series of children (75). A similar phenomenon can be seen in adults with trichotillomania. Furthermore, preliminary open data suggests that augmentation of SSRIs with dopamine blockers may be useful in the treatment of hair pulling (76 ,77). The atypical neuroleptics, which have dopamine and serotonin antagonist effects, may also be effective augmenting agents in OCD and trichotillomania (78).

Christenson and colleagues found no significant differences in either pain detection or pain tolerance thresholds between trichotillomania patients and controls. On the other hand, this group have suggested that the opioid blocker naltrexone may be effective in the treatment of trichotillomania, indicating that further research on the opioid system in hair pulling may be useful (74).

Trichotillomania is predominantly a disorder of women in the clinical setting. It frequently begins around the time of the menarche, and in some women there is premenstrual exacerbation of symptoms (70). Nevertheless, to our knowledge there are no studies that directly explore hormonal mechanisms and hair pulling. Although there is therefore currently insufficient evidence to indicate a specific link between hair pulling and hormonal mechanisms, further work in this area seems warranted.

Genetic studies might shed light on the particular neurochemical mediators of trichotillomania. Nevertheless, there have been few such studies. Christenson et al. (79) reported that 8% of 161 trichotillomania patients had first-degree relatives with hair pulling. Another study failed to show elevated rates of trichotillomania in first-degree relatives of trichotillomania probands, but did find elevated rates of OCD (80). Furthermore, elevated rates of trichotillomania have been found in a cohort of patients with both OCD and TS as compared to those with TS or OCD alone (Miguel et al., unpublished data).

Neuroanatomy

The neuroanatomy of trichotillomania is comparatively poorly researched. There are only occasional reports of hair pulling in association with neurologic disorders, although once again basal ganglia lesions have been implicated (74). Neuropsychological and neurologic soft sign studies have also been partly consistent with involvement of the CSTC system in trichotillomania (74). Furthermore, a few studies of brain imaging in trichotillomania have now been undertaken.

Stein and colleagues (81) employed brain MRI and found no differences in caudate volume in female patients with trichotillomania and normal controls. O’Sullivan and colleagues (82) similarly found no difference in caudate volumes in trichotillomania and controls on MRI, but did find that patients with trichotillomania had reduced left putamen volumes. This finding is of particular interest given work demonstrating reduced left putamen volumes in Tourette’s syndrome.

Swedo et al. (83) found increased right and left cerebellar and right superior parietal glucose metabolic rates in trichotillomania patients compared with normal controls. This finding does not seem to support the hypothesis that orbitofrontal–basal ganglia circuits are key to this disorder, and differs from findings obtained in OCD and Tourette’s. However, patients were scanned at rest, rather than during hair pulling or during the performance of a neuropsychological test that might have activated these structures. Swedo and colleagues also found that anterior cingulate and orbital-frontal metabolism correlated negatively with clomipramine response, a result they had previously found in OCD. They concluded that increased orbital-frontal metabolism may comprise a compensatory response to basal ganglia pathology in both these disorders.

Stein et al. (74) studied single photon emission computed tomography (SPECT) scans in patients with trichotillomania before and after pharmacotherapy with the SSRI citalopram (74). During treatment there was a reduction in activity in left and right inferior-posterior and other frontal areas. In nonresponders there was an increase in baseline left and right superior-lateral frontal areas. These data are again to some extent consistent with work suggesting that trichotillomania, like OCD, is mediated by corticostriatal circuits.

The neuroimmunology of OCD and TS has recently been an important focus of study. Of particular interest to research on trichotillomania, Swedo et al. (84) reported that like OCD symptoms, hair pulling may relapse after streptococcal infections. In addition, a case report of a patient in whom hair-pulling symptoms appeared closely linked with Sydenham’s chorea has been published (85). Both choreiform symptoms and hair pulling remitted in response to penicillin treatment.

Nevertheless, no data have yet been published that establish a causal connection between *Streptococcus* or Sydenham’s chorea and hair pulling. Furthermore, Niehaus et al. (86) recently found that D8/17, a marker of susceptibility to developing sequelae after *Streptococcus* infection, was not more frequent in patients with trichotillomania than in controls. It remains possible, however, that particular subtypes

of trichotillomania have a specific neuroimmunologic etiology. Further research in this area is warranted.

IMPULSIVE SIB

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

Phenomenology

Common behaviors in this category include skin cutting, skin burning, and self-hitting. These behaviors frequently permit those who engage in them to obtain rapid but short-lived relief from a variety of intolerable states, in this sense serving a morbid and pathologic but life-sustaining function. Descriptive and systematic data indicate that SIBs of this type are typically impulsive, and hence their proposed classification in our schema.

Bennum (87), for example, reported that 70% of self-mutilators feel they have no control over the act. Favazza and Conterio (88) found that 78% of individuals in their sample decided to self-mutilate on the spur of the moment, and another 15% made the decision within an hour of the act. The act was then always (30%) or almost always (51%) carried out. Other studies by this group have also emphasized an association between impulsivity and self-mutilation (89). In another sample, less than 15% of self-mutilators reported any inner struggle to resist the behavior (90). Simeon et al. (91) found a significant correlation between the degree of self-mutilation and an independent measure of impulsivity.

An aggressive component has also been descriptively identified in impulsive SIB. Studies of impulsive self-mutilators show that 18% to 45% of individuals report anger toward themselves and 10% to 32% report anger toward others leading up to the acts of self-injury (87 ,90). Bennum (87) found that self-mutilators had greater outwardly directed hostility than nonmutilating depressives, while not differing in inwardly directed hostility. Simeon et al. (91) reported that, compared to nonmutilating controls matched for personality disorder diagnoses, self-mutilators had lifetime histories of greater aggression and sociopathy, and the degree of self-mutilation correlated significantly with chronic anger.

Some individuals engage in these self-injurious behaviors only a limited number of times in their lifetimes, whereas others do so quite frequently and habitually (1). As with other conditions on the compulsive-impulsive spectrum, the distinction between compulsive and impulsive repetitive self-injury may not always be sharp and clear. Impulsive repetitive self-injury can at times become so habitual as to occur on a daily or weekly basis and without clearly identifiable precipitating external events or affective states, as if it were a compulsion. It could be speculated that individuals who have an obsessive-compulsive predisposition may be more prone to become fixated on SIBs that were initially more episodic and impulsive and over time become habitual and dystonic. Interestingly, the few studies that have examined obsessionality in impulsive self-injury may indeed support such a notion. In one study comparing 22 female inpatient nonpsychotic “repetitive self-cutters” with 22 demographically matched inpatient controls, the patients with self-injury were significantly more obsessional than the controls while not differing in depression, anxiety, phobia, or hysteria (90). In another study comparing mutilating and nonmutilating antisocial women confined to a criminal ward, the self-injurers were found to be significantly more obsessional (92). On another note, impulsive and compulsive self-injurious behaviors can be comorbidly encountered in a certain proportion of individuals.

Although epidemiologic studies are lacking, it has been indirectly estimated that the incidence of impulsive self-injury may be at least 1/1,000 people annually (1). It is more common in females, and typically begins in adolescence or early adulthood, although it has been described as early as in the latency or even the preschool years. It is more commonly associated with certain disorders such as borderline personality disorder, antisocial personality disorder, posttraumatic stress disorder, dissociative disorders, and eating disorders. Of all these diagnoses, the one that appears most commonly associated with SIB is borderline personality disorder, but this is neither a necessary nor a sufficient condition, and the assumption of such may lead to premature diagnostic formulation and closure.

The relationship between impulsive self-injurious behaviors and suicide attempts warrants brief mention here. The coexistence of both behaviors in individuals with severe personality disorders may be more the rule than the exception. In a large, thoroughly studied series of chronically hospitalized highly disturbed patients, of 141 borderline females ten had histories of self-injury alone, whereas 20 had histories of both self-injury and suicide attempts (93). However, a 15-year follow-up of these patients suggested that history of self-injury alone was not a predictor of future suicide whereas a history of suicide attempts was; this observation indeed supports the distinct conceptualization of the two behaviors.

Traumatic experiences commonly predate and appear to contribute to the development of impulsive self-injury, and these traumatic experiences are more typically childhood ones. Although speculative, comparisons can arguably be made between the stereotypic and aggressive behavior of isolation reared primates, and symptoms of impulsive aggression and autoaggression in patients with impulsive personality disorders. In one large study of habitual female self-mutilators, childhood abuse was noted in 62% of the subjects. Of these, 29% reported both sexual and physical abuse, 17% reported only sexual abuse, and 16% reported only physical abuse. The onset of the abuse was typically early, reported as early latency, and it often involved family members (88). Indeed, abused and neglected children can begin to exhibit SIB from a disturbingly early age.

One comprehensive study attempting to tease out the contribution of various factors to the genesis and perpetuation of self-destructive behaviors followed 74 personality-disordered individuals over a 4-year period (94). A portion of this group had a history of self-mutilation, and within the self-mutilating group there was an 89% incidence of major disruptions in parental care and a 79% incidence of childhood trauma such as physical abuse, sexual abuse, or witnessing domestic violence. Sexual abuse most strongly predicted self-injury, which was also associated with younger age at the time of the abuse, as well as with childhood chaos, separations, and neglect. Dissociative experiences also correlated with self-mutilation. Interestingly, neglect and separation from caregivers predicted continuation of the self-injury in the face of treatment efforts, leading the authors to postulate that the latter traumas impaired the capacity to form trusting stable bonds with others that could then facilitate treatment change. Another study of borderline personality disorder inpatients compared to personality-disordered controls similarly found that both parental sexual abuse and emotional neglect were significantly related to self-mutilation (95).

Neurochemistry

Research over the last quarter century, has highlighted the role of the serotoninergic system in impulsive aggression. Initially focusing on serotoninergic dysregulation in attempted or completed suicide, studies later demonstrated similar neurochemical derangement in outwardly directed acts of impulsive aggression, as well as in inwardly directed aggression without suicidal intent.

As a background we briefly review the studies establishing a connection between serotoninergic dysregulation and impulsive aggression. An important series of studies by Brown et al. (96 ,97) demonstrated a decrease in CSF 5-HIAA in patients with personality disorder, and found that this decrease correlated with scores on a lifetime aggression scale. A series of subsequent studies have confirmed a relationship between CSF 5-HIAA and impulsive or aggressive behaviors. Linnoila et al.'s work (98) is of particular interest insofar as it specifically divided aggressive behaviors into impulsive and nonimpulsive forms, and CSF 5-HIAA correlated only with impulsive aggression.

A range of other static measures of serotonin function is consistent with serotonin hypofunction in impulsive aggression (99). In addition, neuroendocrine challenge studies, which provide a dynamic measure of serotoninergic function, confirm this relationship. Coccaro et al. (100), for example, administered fenfluramine to personality-disordered patients and found that prolactin response, a measure of net serotonin function, correlated inversely with impulsive aggression. Similarly, abnormal neuroendocrine responses after mCPP challenge have been observed in borderline personality disorder.

The literature on serotonin and suicide brings another dimension to the relationship between serotonin and impulsive aggression. Early postmortem studies found that brainstem levels of serotonin or 5-HIAA are decreased in suicide victims, and subsequent studies confirm that decreased serotonin and/or 5-HIAA in brainstem (raphe nuclei) and/or subcortical nuclei (hypothalamus) have been consistent postmortem changes in suicide completers (101). Recent genetic findings suggest that particular polymorphisms in the serotonin system may account for some of the variance in symptoms in impulsive personality disorders; such work may ultimately prove of significant value in understanding the neurobiology of SIB.

We previously presented descriptive data conceptualizing impulsive SIB as a form of impulsive aggression. Direct neurobiological data in impulsive SIB are very limited, yet consistent with serotoninergic dysregulation, and we summarize them below. Lopez-Ibor et al. (102) found that inpatient depressives with histories of self-injury without suicidal intent had lower CSF 5-HIAA than those without self-injury histories. Coccaro et al. (100) reported a significant correlation between self-damaging acts and a blunted prolactin response to the serotonin agonist fenfluramine in patients with personality disorders. Simeon et al. (91) reported a significant negative correlation between the frequency of self-mutilation and platelet imipramine binding, a peripheral serotoninergic index. In contrast to these studies, Gardner et al. (103) compared CSF 5-HIAA levels in borderline personality-disordered patients with and without self-mutilation, and found no difference. However, it cannot be ruled out that the different incidence of suicide attempt histories in the two groups may have concealed an association between CSF 5-HIAA and self-mutilation. In this regard, Simeon et al. (91) found a 44% reduction in CSF 5-HIAA in a small subsample of self-mutilators compared to nonmutilators who had never made suicide attempts. In a study by Siever and Trestman (104), self-directed aggression in personality-disordered subjects, as measured by a composite score of suicide attempts and SIB, was inversely correlated to the prolactin response to fenfluramine challenge. In a smaller sample of female personality-disordered patients, the relationship did not reach statistical significance (105).

More indirect evidence for underlying serotoninergic dysfunction in self-mutilation comes from open-treatment treatment studies utilizing SSRIs. Fluoxetine has been reported useful in decreasing SIB in a number of studies of borderline personality disorder. For example, in an open study of fluoxetine in 22 borderline and schizotypal patients, a 97% decrease in self-mutilating episodes was found (106). Clomipramine was apparently useful for SIB in OCD patients, many of whom had a history of sexual abuse (107). In an open trial of venlafaxine in subjects with borderline personality disorder (108), five out of seven subjects who engaged in self-injury ceased to do so after 12 weeks of treatment; of note, venlafaxine is both a serotonin and norepinephrine

reuptake inhibitor. Further pharmacologic trials, conducted in a double-blind fashion, and assessing larger numbers of patients, are needed to conclusively determine whether SSRIs are efficacious in treating self-injury, whether the response is selective in comparison to other pharmacologic agents, how the response compares to the change in other target symptoms, as well as to clarify the time course and maintenance of the therapeutic response.

Other neurotransmitter systems may also be involved in the neurobiology of impulsive SIB but have been even less studied. Siever and Trestman (104) have suggested that, in addition to serotoninergic dysfunction, noradrenergic dysregulation may be implicated in the expression of impulsive aggression. Specifically, noradrenergic hyperreactivity may mediate increased arousal and irritability and trigger acts of impulsive aggression, in conjunction with behavioral disinhibition mediated by serotoninergic dysfunction. A relationship between CSF MHPG and impulsive aggression has not been consistently replicated, but there is some evidence that personality-disordered patients may have increased growth hormone responses to clonidine challenge. We are not, however, aware of any studies of the noradrenergic system specifically in self-injuring individuals. A treatment study describing some success with venlafaxine in self-injury was mentioned above.

Another system implicated in impulsive SIB is the opioid system. Indeed, habitual self-mutilators commonly report the relief of depersonalization and other dissociative states as the motivation to injure themselves, and relative analgesia to the self-inflicted injury is often present in the emotional state surrounding such acts. The role of the endogenous opioid system in stress-induced analgesia seems established.

Actual data examining the opioid system in SIB or SIB-prone groups are limited. Russ and colleagues (109) examined the pain response to a cold pressor test in individuals with borderline personality disorder. Compared to subjects who experienced pain during self-injury, those who experienced analgesia reported less pain during the experiment. However, naloxone pretreatment did not increase discomfort induced by the test (110), casting some doubt on an endogenous opioid hypothesis. In one study of the opioid system, Coid et al. (111) found that habitually self-mutilating individuals had higher plasma metenkephalin than normal comparison subjects, but the finding might have been secondary to recent self-injury itself. A recent open study of naltrexone in female patients with SIB accompanied by analgesia and dysphoria reduction, and typically accompanied by a history of abuse, found that SIB symptoms ceased in six of seven subjects (112). However, these results need to be replicated in controlled studies.

Neuroanatomy

There has been little study of the neuroanatomy of SIB per se in borderline personality disorder (BPD). However, it would again seem relevant to this chapter to emphasize the increasingly large literature correlating frontal hypofunction with impulsive aggression. The classic case of Phineas Gage is an excellent example of the marked dysfunction, with increased impulsivity and perseveration, that results from frontal lesions. Subsequent studies of head injury and of frontal lobe surgery have led to multiple descriptions of such dysfunction.

A body of neuropsychological literature, although open to different interpretations, also points to a relationship between frontal dysfunction and impulsive aggression. Some of the strongest evidence of involvement of frontal lobe dysfunction, however, in impulsive aggression emerges from recent brain imaging studies. Preliminary findings include the demonstration of a significant association between decreased metabolic rates in prefrontal cortex and aggression in personality-disordered and aggressive patients. Similarly, there is increasing imaging evidence for neuronal dysfunction in antisocial personality disorder (113). Also, BPD subjects had diminished response to fenfluramine challenge in a number of areas of prefrontal cortex (114).

It is perhaps important to emphasize the way in which neurobiological factors in BPD may intersect with psychosocial ones. An abusive background, for example, may result in neurobiological changes that then further promote risk-seeking behavior in adult life. Research on the neurobiology of posttraumatic stress disorder similarly may provide a useful framework for developing hypotheses about childhood abuse and subsequent changes in personality. Such work may also shed light on the multiple roots of “impulsive” SIB.

CONCLUSION

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

In the literature on the obsessive-compulsive spectrum, it has been suggested that compulsive and impulsive symptoms and disorders reflect different kinds of psychobiological mechanisms. Compulsive symptoms, for example, may be mediated by serotonin and frontal hyperfunction, whereas impulsive symptoms may be mediated by serotonin and frontal hypofunction. In reality the complex neurobiology of compulsivity and impulsivity cannot be captured by so simplistic a contrast. Nevertheless, such contrasts arguably have some heuristic value.

Does this contrast apply also to compulsive and impulsive SIB? Perhaps patients with SMD have frontal/serotoninergic hyperfunction, whereas patients with BPD have frontal/serotoninergic hypofunction? Currently, there are insufficient data to make so bold a claim. Furthermore, in clinical and biological reality, the situation may be much more complex than this, with patients demonstrating both compulsive and impulsive features, and with disorders such as OCD, SMD, and BPD exhibiting brain areas of both serotonin hyperfunction and serotonin hypofunction.

Indeed, an alternative hypothesis is that stereotypic SIB can be seen after a number of different situations. First, in patients with overactivation of the basal ganglia, there may be excessive release of various motoric sequences, including SIB. Second, in patients with hypofrontal function, there may be an inability to control such programmed sequences. Environmental conditions such as deprivation may also disrupt the balance in corticostriatal circuits (3), in such a way as to produce symptoms.

Fortunately there a now a range of different approaches to understanding the neurobiology of unwanted repetitive symptoms. These include brain imaging, genetics, neuroimmunology, and pharmacological dissection. Although such approaches have shed light on some relationships between the OCD spectrum disorders, much further work is required before it is possible to make phenomenologic distinctions (such as compulsive vs. impulsive SIB) on the basis of differential psychobiological mechanisms.

In tackling this area of research, it may also be possible to integrate neurobiological with psychosocial data; compulsive SIB, for example, may be seen not only in response to a pharmacologic challenge but also after isolation, whereas impulsive behavior that is induced by an environmental stressor may nevertheless be mediated by specific neurobiological mechanisms. Future research should see an expansion in our understanding of the neuropsychopharmacology of compulsive and impulsive SIB, and thus in our ability to intervene effectively with patients who demonstrate these distressing and disabling symptoms.

ACKNOWLEDGMENT

Part of "121 - Compulsive and Impulsive Aspects of Self-Injurious Behavior "

Dr. Stein is supported by a grant from the Medical Research Council (MRC).

REFERENCES

1. Favazza A, Simeon D. Self-mutilation. In: Hollander E, Stein DJ, eds. *Impulsivity and aggression*. Chichester: Wiley, 1995.

2. Castellanos FX, Ritchie FG, Marsh WL, et al. DSM-IV stereotypic movement disorder: persistence of stereotypies of infancy in intellectually normal adolescents and adults. *J Clin Psychiatry* 1996;57:116–122.

3. Ridley RM. The psychology of perseverative and stereotyped behavior. *Prog Neurobiol* 1994;44:221–231.

4. Cross HA, Harlow HF. Prolonged and progressive effects of partial isolation on the behavior of macaques monkeys. *J Exp Res Person* 1965;1:39–44.

5. Randrup A, Munkvad I, Usden P. Adrenergic mechanisms and amphetamine induced abnormal behavior. *Acta Pharmacol Toxicol* 1963;20:145–157.

6. Mueller K, Nyhan W. Pharmacologic control of pemoline-induced self-injurious behavior in rats. *Pharmacol Biochem Behav* 1982;16:957–963.

7. Phillips AG, Fibiger HC. Deficits in stimulation-induced feeding after intraventricular administration of 6-hydroxydopamine in rats. *Behav Biol* 1973;9:749–754.

8. Phillips AG, Nikaido RS. Disruption of brain stimulation-induced feeding by dopamine receptor blockade. *Nature* 1975;258:750–751.

9. Goldstein M, Kuga S, Kusano N, et al. Dopamine agonist induced self-mutilative biting behavior in monkeys with unilatreal ventromedial tegmental lesions of the brainstem: Possible pharmacological model for Lesch-Nyhan syndrome. *Brain Res Bull* 1986;367:114–119.

10. Criswell HE, Mueller RA, Breese GR. Clozapine antagonism of D-1 and D-2 dopamine receptor-mediated behavior. *Eur J Pharmacol* 1989;159:141–147.

11. Breese GR, Baumeister AA, McGown T, et al. Behavioral differences between neonatal and adult 6-hydroxydopamine-treated rats to dopamine agonists: relevance to neurological symptoms in clinical syndromes with reduced brain dopamine. *J Pharmacol Exp Ther* 1984;231:343–354.

12. Jacobs BJ, Fornal CA. Serotonin and behavior: a general hypothesis. In: Blood FE, Kupfer DJ, eds. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995.

13. Valzelli L, Bernasconi S. Aggressiveness by isolation and brain serotonin turnover changes in different strains of mice. *Neuropsychobiology* 1979;5:129–135.

14. Cronin GM, Wiepkema PR, Van Ree JM. Endorphins implicated in stereotypies of tethered sows. *Experientia* 1986;42:198–199.

15. Herman BH, Hammock MK, Egan J, et al. A role of opioid peptides in self-injurious behavior: dissociation from autonomic nervous system functioning. *Dev Pharmacol Ther* 1989;12:81–89.

16. Cataldo MF, Harris J. The biological basis for self-injury in the mentally retarded. *Anal Intervent Dev Disabil* 1982;2:21–39.

17. Sahakian BJ, Robbins TW, Morgan JJ, et al. The effects of psychomotor stimulants on stereotypy and locomotor activity in socially deprived and control rats. *Brain Res* 1975;84:195–205.

18. Sahakian BJ, Robbins TW. Isolation-rearing enhances tail-pinch induced oral behaviour in rats. *Physiol Behav* 1977;18:53–58.

19. Saint-Cyr JA, Taylor AE, Nicholson K. Behavior and the basal ganglia. In: Weiner WJ, Lang AE, eds. *Behavioral neurology of movement disorders*. New York: Raven Press, 1995.

20. Dodman N, Moon A, Stein DJ. Animal models of obsessive-compulsive disorder. In: Hollander E, Stein DJ, eds. *Obsessive-compulsive disorders: etiology, diagnosis, treatment*. New York: Marcel Decker, 1997.

21. Rapoport JL, Ryland DH, Kriete M. Drug treatment of canine acral lick. An animal model of obsessive-compulsive disorder. *Arch Gen Psychiatry* 1992;48:517–521.

22. Antelman SM, Eichler AJ, Black CA, et al. Interchangeability of stress and amphetamine in sensitization. *Science* 1980;207:329–331.

23. Wessels CJ, Seier J, Mdhluli C, et al. Fluoxetine decreases stereotyped behaviour in primates. Presented at the 9th Biennial Meeting of the International Society for Comparative Psychology, Cape Town, September 1998.

24. Martin LJ, Spicer DM, Lewis MH, et al. Social deprivation of infant monkeys alters the chemoarchitecture of the brain: I. Subcortical regions. *J Neurosci* 1991;11:3344–3358.

25. Stein DJ, Niehaus DJN. Self-injurious behaviors in developmental disorders: biology and Psychopharmacology. In: Simeon D, Hollander E, eds. *Self-injurious behavior*. Washington, DC: American Psychiatric Press, 2001.

26. Stein DJ, Simeon D. Compulsive self-injurious behavior. In: Simeon D, Hollander E, eds. *Self-injurious behavior*. Washington, DC: American Psychiatric Press, 2001.

27. Bodfish JM, Crawford TM, Powell SB, et al. Compulsions in adults with mental retardation: prevalence, phenomenology, and comorbidity with stereotypy and self-injury. *Am J Ment Retard* 1995;100:183–192.

28. Winchel RM, Stanley M. Self-injurious behavior: a review of the behavior and biology of self-mutilation. *Am J Psychiatry* 1991;148:306–317.

29. King BH. Self-injury by people with mental retardation: a compulsive behavior hypothesis. *Am J Ment Retard* 1993;98:93–112.

30. Mikkelsen EJ, Albert LG, Emens M, et al. The efficacy of antidepressant medication for individuals with mental retardation. *Psych Ann* 1997;27:198–206.

31. Lewis MH, Bodfish JW, Powell SB, et al. Clomipramine treatment for self-injurious behavior in mental retardation: a double-blind comparison with placebo. *Am J Retard* 1996;100:654–665.

32. Swedo SE, Leonard HL, Rapoport JL, et al. A double-blind comparison of clomipramine and desipramine in the treatment of trichotillomania (hair pulling). *N Engl J Med* 1989;321:497–501.

33. Leonard HL, Lenane MC, Swedo SE, et al. A double-blind comparison of clomipramine and desipramine treatment of severe onychophagia (nail biting). *Arch Gen Psychiatry* 1991;48:821–827.

34. Osman OT, Loschen EL. Self-injurious behavior in the developmentally disabled: pharmacologic treatment. *Psychopharmacol Bull* 1992;28:439–449.

35. Ratey JJ, Lindem KJ. B-Blockers as primary treatment for aggression and self-injury in the developmentally disabled. In: Ratey JJ, ed. *Mental retardation: developing pharmacotherapies*. Washington, DC: American Psychiatric Press, 1991:51–81.

36. Cohen SA, Underwood MT. The use of clozapine in a mentally retarded and aggressive population. *J Clin Psychiatry* 1994;55:440–444.

37. Rubin M. Use of atypical antipsychotics in children with mental retardation, autism, and other developmental disabilities. *Psych Ann* 1997;27:219–221.

38. Sandman CA. B-Endorphin dysregulation in autistic and self-injurious behavior: a neurodevelopmental hypothesis. *Synapse* 1988;2:193–199.

39. Pies RW, Popli AP. Self-injurious behavior: pathophysiology and implications for treatment. *J Clin Psychiatry* 1995;56:580–588.

40. Willemsen-Swinkels SH, Buitelaar JK, Nijhof GJ, et al. Failure of naltrexone hydrochloride to reduce self-injurious and autistic behavior in mentally retarded adults. Double-blind placebo-controlled studies. *Arch Gen Psychiatry* 1995;52:766–773.

41. Kastner T, Finesmith R, Walsh K. Long-term administration of valproic acid in the treatment of affective symptoms in people with mental retardation. *J Clin Psychopharmacol* 1993;13:448–451.

42. Nyhan WL. Behavior in the Lesch-Nyhan syndrome. *J Autism Child Schizophr* 1976;19:79–95.

43. Shear CS, Nyhan WL, Kirman BH, et al. Self-mutilative behaviour as a feature of the de Lange syndrome. *J Pediatr* 1971;78:506–509.

44. Clarke DJ, Waters J, Corbett JA. Adults with Prader-Willi syndrome: abnormalities of sleep and behavior. *J R Soc Med* 1989;82:21–24.

45. Stein DJ, Keating J, Zar HJ, et al. A survey of the phenomenology and pharmacotherapy of compulsive and impulsive-agressive symptoms in Prader-Willi syndrome. *J Neuropsychiatry Clin Neurosci* 1994;6:23–29.

46. Lloyd KG, Hornykiewics O, Davidson L, et al. Biochemical evidence of dysfunction of brain neurotransmitters in the Lesch-Nyhan syndrome. *N Engl J Med* 1981;305:1106–1111.

47. Jinnah HA, Wojcik BE, Hunt M, et al. Dopamine deficiency in a genetic mouse model of Lesch-Nyhan disease. *J Neurosci* 14:1164–1175, 1994

48. Greenberg A, Coleman M. Depressed whole blood serotonin levels associated with behavioral abnormalities in the de Lange syndrome. *Pediatrics* 1973;52:720–724.

49. Selikowitz M, Sunman J, Pendergast A, et al. Fenfluramine in Prader-Willi syndrome: a double-blind, placebo controlled trial. *Arch Dis Child* 1990;65:112–114.

50. Gillessen-Kaesbach G, Robinson W, Lohmann D, et al. Genotype-phenotype correlation in a series of 167 deletion and non-deletion patients with Prader-Willi syndrome. *Hum Genet* 1995;96:638–643.

51. Ernst M, Zametkin AJ, Matochik JA, et al. Presynaptic dopaminergic deficits in Lesch-Nyhan disease. *N Engl J Med* 1996;334:1568–1572.

52. Wong DF, Harris JC, Naidu S, et al. Dopamine transporters are markedly reduced in Lesch-Nyhan disease in vivo. *Proc Natl Acad Sci USA* 1996;93;5539–5543.

53. Simeon D, Stein DJ, Gross S, et al. A double-blind trial of fluoxetine in pathologic skin picking. *J Clin Psychiatry* 1997;58:341–347.

53a. Prader A, Labhart A, Willi H. Ein syndrom von adipositas, kleinwuchs, kryptochismus und oligophrenie nach myatonieartigem zustand im neugeborenenalter. *Schweizerische Medizinische Wochenschrift* 1956;86:1260–1261.

53b. Holm VA. The diagnosis of Prader-Willi syndrome. In Holm VA, Sulzbacher SJ, Pipes P, eds. *Prader-Willi syndrome*. Baltimore: University Park Press, 1981:27–44.

54. Niehaus DJ, Emsley RA, Brink PA, et al. Stereotypies: prevalence and association with compulsive and impulsive symptoms in college students. *Psychopathology* 2000;33:31–35.

55. Kalivas J, Kalivas L, Gilman D, et al. Sertraline in the treatment of neurotic excoriations and related disorders. *Arch Dermatol* 1996;132:589–590

56. Phillips KA, Taub SL. Skin picking as a symptom of body dysmorphic disorder. *Psychopharmacol Bull* 1995;31:279–288.

57. Sokol MS, Campbell M, Goldstein M, et al. Attention deficit disorder with hyperactivity and the dopamine hypothesis. *J Child Adolesc Psychiatry* 1987;26:428–433.

58. Rothenberger A. Psychopharmacological treatment of self-injurious behavior in autism. *Acta Paedopsychiatr* 1993;56:99–104.

59. McDougle CJ, Kresch LE, Goodman WK, et al. A case-controlled study of repetitive thoughts and behavior with autistic disorder and obsessive-compulsive disorder. *Am J Psychiatry* 1995;152:772–777.

60. McBride PA, Anderson GM, Shapiro T. Autism research: bringing together approaches to pull apart the disorder. *Arch Gen Psychiatry* 1996;53:980–983.

61. McDougle CJ, Naylor ST, Cohen DJ, et al. A double-blind, placebo-controlled study of fluvoxamine in adults with autistic disorder. *Arch Gen Psychiatry* 1996;53:1001–1008.

62. Gordon CT, State RC, Nelson JE, et al. A double-blind comparison of clomipramine, desipramine, and placebo in the treatment of autistic disorder. *Arch Gen Psychiatry* 1993;50:441–447.

63. Sanchez LE, Campbell M, Small AM, et al. A pilot study of clomipramine in young autistic children. *J Am Acad Child Adolesc Psychiatry* 1997;35:537–544.

64. Ernst M, Zametkin AJ, Matochik JA, et al. Low medial prefrontal dopaminergic activity in autistic children. *Lancet* 1997;350:638.

65. Anderson LT, Campbell M, Adams P, et al. The effects of haloperidol on discrimination learning and behavioral symptoms in autistic children. *J Autism Dev Disord* 1989;19:227–239.

66. Perry R, Campbell M, Adams P, et al. Long-term efficacy of haloperidol in autistic children: continuous versus discontinuous administration. *J Am Acad Child Adolesc Psychiatry* 1989;28:87–92.

67. Rapin I. Autism. *N Engl J Med* 1997;337:97–104.

68. Robertson MM, Trimble MR, Lees A. Self-injurious behavior and the Gilles de la Tourette syndrome: a clinical study and review of the literature. *Psychol Med* 1989;19:611–625.

69. Braun AR, Randolph C, Stoetter B, et al. The functional neuroanatomy of Tourette's syndrome: an FDG-PET study. II. Relationships between regional cerebral metabolism and associated behavioral and cognitive features of the illness. *Neuropsychopharmaoclogy* 1995;13:151–168.

70. Christenson GA, Mackenzie TB, Mitchell JE. Characteristics of 60 adult chronic hair pullers. *Am J Psychiatry* 1991;148:365–370.

71. Simeon D, Cohen LJ, Stein DJ, et al. Comorbid self-injurious behaviors in 71 female hair-pullers: a survey study. *J Nerv Ment Disord* 1997;185:117–119.

72. O'Sullivan RA, Christenson GA, Stein DJ. Pharmacotherapy of trichotillomania. In: Stein DJ, Christenson GA, Hollander E, eds. *Trichotillomania*. Washington, DC: American Psychiatric Press, 1999.

73. Ninan PT, Rothbaum BO, Stipetic M, et al. CSF 5HIAA as a predictor of treatment response in trichotillomania. *Psychopharmacol Bull* 1992;28:451–455.

74. Stein DJ, O' Sullivan R, Hollander E. Neurobiology of trichotillomania. In: Stein DJ, Christenson GA, Hollander E, eds. *Trichotillomania*. Washington, DC: American Psychiatric Press, 1999.

75. Martin A, Scahill L, Vitulani L, et al. Stimulant use and trichotillomania. *J Am Acad Child Adolesc Psychiatry* 1998;37:349–350.

76. Stein DJ, Hollander E. Low-dose pimozide augmentation of serotonin reuptake blockers in the treatment of trichotillomania. *J Clin Psychiatry* 1992;53:123–126.

77. van Ameringen M, Mancini C. Treatment of trichotillomania with haloperidol. *J Affect Disord* 1999;56:219–226.

78. Stein DJ, Bouwer C, Hawkridge S, et al. Risperidone augmentation of serotonin reuptake inhibitors in obsessive-compulsive and related disorders. *J Clin Psychiatry* 1997;58:119–122.

79. Christenson GA, Mackenzie TB, Reeve EA. Familial trichotillomania. *Am J Psychiatry* 1992;149:283

80. Lenane MC, Swedo SE, Rapoport JL, et al. Rates of obsessive compulsive disorder in first degree relatives of patients with trichotillomania: a research note. *J Child Psychol Psychiatry* 1992;33:925–933.

81. Stein DJ, Coetzer R, Lee M, et al. Magnetic resonance brain imaging in women with obsessive-compulsive disorder and trichotillomania. *Psychiatry Res* 1997;74:177–182.

82. O'Sullivan RL, Rauch SL, Breiter HC, et al. Reduced basal ganglia volumes in trichotillomania measured via morphometric MRI. *Biol Psychiatry* 1997;42:39–45.

83. Swedo SE, Rapoport JL, Leonard HL, et al. Regional cerebral glucose metabolism of women with trichotillomania. *Arch Gen Psychiatry* 1991;48:828–833.

84. Swedo SE, Leonard HL, Lenane MC, et al. Trichotillomania: a profile of the disorder from infancy through adulthood. *Int Pediatr* 1992;7:144–150.

85. Stein DJ, Wessels C, Carr J, et al. Hair-pulling in a patient with Sydenham's chorea (letter). *Am J Psychiatry* 1997;154:1320.

86. Niehaus DJH, Knowles JA, van Kradenberg J, et al. D8/17 in obsessive-compulsive disorder and trichotillomania (letter). *S Afr Med J* 1999;89:755–756.

87. Bennum I. Depression and hostility in self-mutilation. *Suicide Life Threat Behav* 1983;13:71–84.

88. Favazza AR, Conterio K. Female habitual self-mutilators. *Acta Psychiatr Scand* 1989;79:283–289.

89. Herpertz S, Sass S, Favazza A. Impulsivity in self-mutilative behaviour: psychometric and biological findings. *Psychiatry Res* 1997;241:451–465

90. Gardner AR, Gardner AJ. Self-mutilation, obsessionality and narcissism. *Br J Psychiatry* 1975;127:127–132.

91. Simeon D, Stanley B, Frances A, et al. Self-mutilation in personality disorders: psychological and biological correlates. *Am J Psychiatry* 1992;149:221–226.

92. McKerracher DW, Loughnane T, Watson RA. Self-mutilation in female psychopaths. *Br J Psychiatry* 1968;114:829–832.

93. Stone MH. *The fate of borderline patients. Successful outcome and psychiatric practice*. New York: Guilford, 1990:61–63.

94. van der Kolk BA, Perry JC, Herman JL. Childhood origins of self-destructive behavior. *Am J Psychiatry* 1991;148:1665–1671.

95. Dubo ED, Zanarini MC, Lewis RE, et al. Childhood antecedents of self-destructiveness in borderline personality disorder. *Can J Psychiatry* 1997;42:63–69.

96. Brown G, Goodwin F, Ballenger J, et al. Aggression in humans correlates with cerebrospinal fluid amine metabolites. *Psychiatry Res* 1979;1:131–139.

97. Brown G, Ebert M, Goyer P, et al. Aggression, suicide, and serotonin. *Am J Psychiatry* 1982;139:741–746.

98. Linnoila M, Virkkunen M, Scheinin M, et al. Low cerebrospinal fluid 5-hydroxyindoleacetic acid concentrations differentiates impulsive from non-impulsive violent behavior. *Life Sci* 1983;33:2609–2614.

99. Stein DJ, Hollander E, Liebowitz MR: Neurobiology of impulsivity and impulse control disorders. *J Neuropsychiatry Clin Neurosci* 1993;5:9–17.

100. Coccaro EF, Siever LJ, Klar HM, et al. Serotonergic studies in patients with affective and personality disorders: correlates with suicidal and impulsive aggressive behavior. *Arch Gen Psychiatry* 1989;46:587–599.

101. Stanley M, Mann JJ. Biological factors associated with suicide, In: Frances AJ, Hales RE, eds. *Review of psychiatry*. Washington, DC: American Psychiatric Press, 1988.

102. Lopez-Ibor JJ, Saiz-Ruiz J, Perez de los Cobos JC. Biological correlations of suicide and aggressivity in major depressions (with melancholia): 5-hydroxyindoleacetic acid and cortisol in cerebral spinal fluid, dexamethasone suppression test and therapeutic response to 5-hydroxytryptophan. *Neuropsychobiology* 1985;14:67–74.

103. Gardner DL, Lucas PB, Cowdry RW. CSF metabolites in borderline personality disorder compared with normal controls. *Biol Psychiatry* 1990;28:247–254.

104. Siever LJ, Trestman RL. The serotonin system and aggressive personality disorder. *Int Clin Psychopharmacol* 1993;8S:33–39.

105. New AS, Trestman RL, Mitropoulou V, et al. Serotonergic function and self-injurious behavior in personality disorder patients. *Psychiatry Res* 1997;69:17–26.

106. Markowitz PJ, Calabrese JR, Schulz SC. Fluoxetine treatment of borderline and schizotypal personality disorder. *Am J Psychiatry* 1991;148:1064–1067.

107. Yaryura-Tobias JA, Neziroglu FA, Kaplan S. Self-mutilation, anorexia, and dysmenorrhea in obsessive compulsive disorder. *Int J Eat Disord* 1995;17:33–38.

108. Markowitz PJ, Wagner SC. Venlafaxine in the treatment of borderline personality disorder. *Psychopharmacol Bull* 1995;31:773–777.

109. Russ MJ, Roth SD, Kakuma T, et al. Pain perception in self-injurious borderline patients: naloxone effects. *Biol Psychiatry* 1994;35:207–209.

110. Russ MJ, Shearin EN, Clarkin JF, et al. Subtypes of self-injurious patients with borderline personality disorder. *Am J Psychiatry* 1993;150:1869–1871.

111. Coid J, Allolio B, Rees LH. Raised plasma metenkephalin in patients who habitually mutilate themselves. *Lancet* 1983;545–546.

112. Roth AS, Ostroff RB, Hoffman RE. Naltrexone as a treatment for repetitive self-injurious behavior: an open-label trial. *J Clin Psychiatry* 1996;57:233–237.

113. Damasio AR. A neural basis for psychopathy. *Arch Gen Psychiatry* 2000;57:128–129.

114. Soloff PH, Meltzer CC, Greer PJ, et al. A fenfluramine-activated FDG-PET study of borderline personality disorder. *Biol Psychiatry* 2000;47:540–547.

Section XII
Movement Disorders and Proxysmal Disorders

Carol A. Tamminga

Carol A. Tamminga: University of Maryland School of Medicine; Deputy Director, Maryland Psychiatric Research Center, Baltimore, Maryland.

Movement Disorders and Proxysmal Disorders - Introduction

Diseases of the brain include human motor disorders and epilepsy, which are distinguished from primary behavioral disorders by the nature of their illness manifestations. These diseases have observable symptoms and, in several cases, cerebral pathology associated with genetic determinants. Especially the features of cerebral pathology and genetic association make it easier to model the critical disease elements in the animal and to move more directly to the neural underpinnings of pathology in human brain. Identified cerebral pathology provides an opportunity for targeting the involved central nervous system (CNS) regions and known brain networks for modeling and focused study. These kinds of CNS illnesses not only serve as models for behavioral disorders of a successful experimental approach, but also generate a valuable knowledge base for understanding human brain physiology in other human CNS illnesses. The known pathophysiology of Parkinson's disease (PD), not information on etiology, has allowed investigators in this field to focus on a relevant region of the brain to explain the neural substrate of Parkinsonian symptoms.

The chapter by Zigmond and Burke addresses the clinical features and critical pathophysiology of PD, and illustrates the value of critical genetic information in the understanding of PD, as well as clear descriptions of tissue pathology. The hypotheses of pathophysiologic mechanisms in PD, whether free-radical–mediated injury, programmed cell death, or effects of protein aggregation in tissue (or some combination of these processes) are identified. Moreover, the authors observe that this disease has been know for two centuries, and its critical pathology described for a half century, but it has only been in recent years that neurobiologic understanding of PD is emerging; this presents a gauge for the difficulty of the work and insight involved in understanding the mechanisms of brain diseases.

The chapter by Wichmann and DeLong does a masterful job of presenting the molecular anatomy and physiology of the basal ganglia thalamocortical network as it relates to PD. It is this feature of network pathology, advanced for several decades by these investigators, that continues to inform not only motor disorder research but aspects of cognition and affective research as well. Moreover, identification of the importance of the structures in this network in these brain diseases has allowed many different investigators to contribute data to produce a rich anatomic and electrophysiologic database of broad relevance for neural mechanisms.

Gracies and Olanow's chapter complements the first two chapters, providing an exhaustive and up-to-date review of therapeutic approaches and their pharmacologic mechanisms. They weave together clinical experience and practical clinical findings with the known neuropharmacology of the illness. L-dopa, dopamine agonists, anticholinergics, and other drug treatments are critically reviewed in detail as are the newer surgical and transplantation therapies now widely discussed. The broad knowledge base available for understanding therapeutic approaches in PD is illustrated in the chapter. After encountering the significant advances made in PD, discoveries in other CNS diseases appear more modest.

Ross and Margolis articulate progress in the understanding of Huntington's disease (HD), the known autosomal-dominant disorder where the gene has been identified for nearly a decade. They describe the characteristic motor, cognitive, and psychiatric symptoms and the usual symptomatic treatments. The known genetic etiology of HD uniquely provides a range of diagnostic and research opportunities

not available in other diseases, including the availability of a diagnostic test and a precise animal model. The genetic mechanism in HD has introduced the consideration of polyglutamine mechanisms in diseases of the brain and rational prophylactic therapies.

The chapter by Tamminga and Woerner presents the clinical course, risk factors, and pathophysiologic considerations for the iatrogenic hyperkinetic motor disorder tardive dyskinesia (TD). Since the etiology of this disorder is known, namely, chronic blockade of the dopamine receptor with antipsychotic drugs, it can be modeled in the experimental animal. This body of research has generated information relevant to TD pathophysiology and treatment, and has relied on the well-described experimental data characterizing basal ganglia structure, neurochemistry, and function.

Schwarcz, Scharfman, and Bertram complete this section with a chapter noting the contribution of extrahippocampal brain regions to the mechanisms of temporal lobe epilepsy. Although of considerable etiologic complexity, temporal lobe epilepsy demonstrates several common aspects of clinical presentation, molecular and cellular changes, and pathologic sequelae. These features are broadly developed in this chapter. The discussion of suspected cellular changes, synaptic reorganization, and neurogenesis in the manifestations of epilepsy makes the chapter an extremely timely contribution.

This section contains a diversity of cutting-edge presentations on human brain diseases that advance the boundaries of not only clinical phenomena but also neuroscience research.

122

Neurocircuitry of Parkinson’s Disease

Thomas Wichmann

Mahlon R. Delong

Thomas Wichmann and Mahlon R. Delong: Department of Neurology, Emory University School of Medicine, Atlanta, Georgia.

Recent progress in neuroscience research has led to major insights into the structure and function of the basal ganglia and into the pathophysiologic basis of Parkinson’s disease (PD) and other movement disorders of basal ganglia origin (3 ,4 ,116 ,313 ,314). The availability of suitable animal models, such as primates rendered parkinsonian by treatment with 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), has been crucial in this progress (17 ,57 ,173). In addition, the renaissance of stereotactic surgery for PD and other movement disorders has provided valuable neuronal recording and imaging data from human subjects. Newer genetic models, for instance mice that overexpress αsynuclein, should provide further insights into the genetic backdrop upon which PD develops. This chapter summarizes from a systems perspective the pathophysiologic concepts that have arisen from the animal models and from work in patients with PD.

ETIOLOGY AND PATHOLOGY IN PARKINSON'S DISEASE

Part of "122 - Neurocircuitry of Parkinson’s Disease "

Idiopathic PD is a disorder characterized by the cardinal signs of akinesia (impaired movement initiation and poverty of movement), bradykinesia (slowness of movement), muscular rigidity, and tremor at rest. The etiology of the disease is uncertain and likely multifactorial, with both genetic and environmental/toxic factors playing a role (see below; see refs. 279 and 287 for review). Idiopathic PD must be distinguished from a large number of other disorders (“atypical” parkinsonism, or “Parkinson-plus” syndromes) which share some of the features of PD but exhibit additional signs (for instance, signs indicative of upper motor neuron, cerebellar or oculomotor involvement). Among these disorders are, for instance, the multiple systems atrophies, progressive supranuclear palsy, and corticobasal ganglionic degeneration. These “atypical” forms of parkinsonism are associated with different and more widespread pathologic abnormalities than those seen in PD proper, and will not be dealt with further in this chapter.

The salient pathologic feature of idiopathic PD is relatively selective degeneration of dopaminergic neurons in the substantia nigra pars compacta (SNc) that project to the striatum (99 ,137), and, to a lesser extent, to other basal ganglia nuclei such as the external and internal segments of the globus pallidus (GPe, GPi, respectively), the subthalamic nucleus (STN), and the substantia nigra pars reticulata (SNr) (99 ,137). Consistent with the early manifestations of motor dysfunction, in the early stages of PD, dopamine depletion is greatest in the sensorimotor territory of the striatum, the postcommissural portion of the putamen (157).

Although it appears that only a small minority of patients suffer from purely inherited forms of PD, investigations into the genetic mechanism that may underlie these cases are being very actively pursued in hopes of discovering pathogenetic mechanism for parkinsonism in general. Inherited forms of parkinsonism in fact have been known for many years (11 ,24 ,113 ,270 ,272), and it has been shown that specific forms of parkinsonism may be caused by different genetic mechanisms. For instance, in a large kindred with *autosomal-dominant* parkinsonism, the disorder was linked to genetic markers on chromosome 4 (PARK1) (233), and has subsequently been shown to be due to a mutation in the α-synuclein gene (215 ,234). α-Synuclein is one of the major components of Lewy bodies, i.e., eosinophilic inclusions in degenerating neurons in the SNc that have long been accepted as one of the pathologic hallmarks of PD (273). A form of *autosomal-recessive* juvenile parkinsonism is caused by a mutation in a gene on chromosome 6, called parkin (PARK 2) (160 ,192 ,286). Finally, *mutations in the mitochondrial DNA*, particularly those affecting complex I function, may also cause or contribute to PD (81 ,120 ,165 ,203 ,248). An involvement of mitochondrial dysfunction in

the development of some forms of parkinsonism is also suggested by findings indicating that the toxicity of MPTP may be due to its inhibition of the mitochondrial complex I enzyme reduced nicotinamide adenine dinucleotide (NADH) coenzyme Q1 reductase (58 ,120 ,197 ,247), and the recent discovery that systemic administration of the pesticide rotenone, a mitochondrial complex I inhibitor, induces striatal dopamine depletion in rats (33).

Overall, however, a *genetic predisposition* for environmental insults that lead to parkinsonism may be far more common than gene mutations that directly result in the disease (87 ,102 ,171 ,288 ,289). Epidemiologic studies have shown an association with rural living, well-water drinking, pesticide exposure, and wood-preservative use (70 ,163 ,236 ,237 ,253). An interesting inverse relationship has been reported between PD and smoking (112 ,128 ,134 ,153 ,240). Among specific toxins that may contribute to PD are MPTP and other isoqinoline derivatives (172 ,197), organophosphate pesticides (38), and perhaps mitochondrial toxins such as rotenone (33).

ANATOMIC SUBSTRATE FOR CIRCUIT DYSFUNCTION IN PARKINSONISM

Part of "122 - Neurocircuitry of Parkinson's Disease "

To understand how the relatively selective loss of dopamine in the basal ganglia leads to parkinsonism, it is necessary to consider in some detail the circuitry, molecular anatomy, and physiology of the basal ganglia and related structures.

The basal ganglia are a group of functionally related subcortical nuclei that include the neostriatum (composed of the caudate nucleus and the putamen), ventral striatum, GPe, STN, GPi, SNr, and SNc. These structures are anatomically related to large portions of the cerebral cortex, thalamus, and brainstem. The striatum, and, to a lesser extent, the STN, are the main entries for cortical and thalamic inputs into the basal ganglia. From these input nuclei, this information is conveyed to the basal ganglia output nuclei—GPi and SNr. Basal ganglia outflow is directed at a variety of targets, among them frontal areas of the cerebral cortex (via the ventrolateral and intralaminar thalamic nuclei), various brainstem structures (superior colliculus, pedunculopontine nucleus, parvicellular reticular formation), and the lateral habenular nucleus.

Input to the Basal Ganglia

The most abundant inputs to the basal ganglia are the topographically segregated corticostriatal projections (7 ,8 ,225). In primates, projections from the somatosensory, motor, and premotor cortices terminate in the postcommissural putamen, the motor portion of the striatum (97 ,98 ,167 ,168). Similarly, associative cortical areas project to the caudate nucleus and the precommissural putamen (110 ,254 ,320 ,321 and 322) and projections from limbic cortices, amygdala, and hippocampus terminate preferentially in the ventral striatum, which includes the nucleus accumbens and the olfactory tubercle (9 ,121 ,166 ,242).

Cortical inputs also terminate in the STN (1 ,126 ,210). The corticosubthalamic projection is derived from the primary motor, prefrontal, and premotor cortices (1 ,126 ,167 ,210). The segregation of cortical projections found in the striatum is also present in the STN. Thus, afferents from the primary motor cortex reach the dorsolateral part of the STN (126 ,210), whereas afferents from premotor and supplementary motor areas innervate mainly the medial third of the nucleus (126 ,168 ,210 ,212). The prefrontal-limbic cortices project to the ventral portion and the medialmost tip of the STN (1 ,27 ,126 ,193).

A second major group of inputs to striatum and STN arises from the intralaminar thalamic nuclei, the centromedian and parafascicular nucleus (CM/Pf). These nuclei have long been identified as major source of excitatory afferents to the basal ganglia (86 ,92 ,154 ,245 ,281 ,317). The projections to striatum and STN arise largely from different neurons in the parafascicular nucleus of the thalamus in rats (92 , but see ref. 79). In primates, CM projects to the motor portions of putamen and STN, whereas Pf projects largely to the associative and limbic territories (208 ,244 ,245 ,265).

Other thalamostriatal inputs arise from the ventral anterior (VA), ventrolateral nucleus (VL), and possibly even the cerebellar-receiving areas of the thalamus (VPLo) (196 ,265). These thalamostriatal projections are less well documented, and their functional significance is unclear. The available evidence indicates that these projections are much less prominent than the projections from the intralaminar nuclei.

Intrinsic Basal Ganglia Connections

The topographically segregated cortical information is conveyed from the striatum to the output nuclei of the basal ganglia (GPi and SNr). Striatofugal projections maintain the striatal organization into motor, limbic, associative, and oculomotor territories (8). The connections between the striatum and the output nuclei of the basal ganglia are thought to be organized into two distinct pathways, the so-called direct and indirect pathways (3 ,6 ,29). The direct pathway arises from a set of neurons that projects monosynaptically to neurons in GPi and SNr, whereas the indirect arises from a different set of neurons that projects to GPe (see ref. 106 for review). In deviation from this strict scheme, some striatofugal neurons may collateralize more extensively, reaching GPe, GPi, and SNr (226). GPe conveys the information it receives either directly or via the STN to GPi and SNr [and, as was recently shown, back to the striatum (159 , 274)].

Several studies have demonstrated highly ordered and specific relationships between the neurons in GPe, STN, and GPi that constitute the indirect pathway (256 ,262 ,266).

Thus, populations of neurons within sensorimotor, cognitive, and limbic territories in GPe are reciprocally connected with populations of neurons in the same functional territories of STN, and neurons in each of these regions, in turn, innervate the same functional territory of GPi (256 ,262), although additional, more divergent circuits may also exist (149 ,256 ,262 ,266).

The STN also provides a dense feedback projection to the GPe (35 ,52 ,205 ,216 ,256 ,258 ,264) and projections to the striatum (22 ,230 ,265), the SNc (158 ,261 ,264), the pedunculopontine nucleus (124 ,158 ,230), and the spinal cord (285). STN output is highly collateralized in the rat (77 ,297), but is more specific in primates (27 ,122 ,230 ,256 ,297) (but see refs. 228 and 229).

The subpopulation of striatal neurons that gives rise to the direct pathway can be further characterized by the presence of the neuropeptides substance P and dynorphin, by the preferential expression of the dopamine D1 receptors, and by the fact that these neurons (as well as most striatal interneurons) appear to be the targets of thalamic inputs from the centromedian nucleus (231 ,260). The subpopulation that gives rise to the indirect pathway expresses preferentially enkephalin and dopamine D2 receptors (105 ,176 ,283), and may be the principal target of cortical inputs (231 ,260).

Although the segregation of D1 and D2 receptors between the direct and indirect pathways is probably not as strict as initially proposed (2 ,282 ,283), it may still serve to explain the apparent dual action of dopamine, released from the nigrostriatal pathway arising in the substantia nigra pars compacta, on striatal output. Dopamine appears to modulate the activity of the basal ganglia output neurons in GPi and SNr by *facilitation* of transmission over the direct pathway and *inhibition* of transmission over the indirect pathway (104). The net effect of striatal dopamine release appears to be to reduce basal ganglia output to the thalamus and other targets (see below). This implies that a reduction of dopamine release as is seen in PD results in a net increase in basal ganglia output.

Output Projections of the Basal Ganglia

Basal ganglia output arises from both GPi and SNr. The segregation of GPi into a caudoventral “motor” portion and rostromedial associative and limbic areas (225) is maintained in the pallidothalamic projections (259). The motor territory of GPi projects almost exclusively to the posterior part of the ventrolateral nucleus (VLo in macaques), which in turn sends projections toward the supplementary motor area (SMA) (143 ,249 ,280), the primary motor cortex (MI) (135 ,136 ,143 ,148 ,152 ,213 ,241), and premotor (PM) cortical areas (135). The outflow from pallidal motor areas directed at cortical areas MI, PM, and SMA appears to arise from separate populations of pallidothalamic neurons (135), indicating that the motor circuit itself can be subdivided into subcircuits, each centered on specific cortical motor and premotor areas. Associative and limbic areas project preferentially to the parvocellular part of the VA and the dorsal VL nucleus (80 ,155 ,259), and may be transmitted in turn to prefrontal cortical areas (111 ,198), as well as motor and supplementary motor regions (68 ,143).

Other output projections from GPi arise mostly as collaterals from the pallidothalamic projection. Thus, prominent axon collaterals are sent in a segregated manner to the CM/Pf complex, which project to the striatum (see above), constituting one of the many feedback circuits in the basal ganglia–thalamocortical circuitry (259). Additional axon collaterals reach the noncholinergic portion of the pedunculopontine nucleus (PPN) (125 ,227 ,243 ,257 ,277), which gives rise to descending projections to pons, medulla, and spinal cord, and ascending projections to basal ganglia, thalamus, and basal forebrain (see ref. 144 for review).

Although the overlap between motor and nonmotor areas is probably greater in the SNr than in GPi (127), the SNr can be broadly subdivided into a dorsolateral sensorimotor and a ventromedial associative territory (78). By and large, projections from these areas target the same nuclei that also receive GPi output, but tend to terminate in different regions of these nuclei. Projections from the medial SNr to the thalamus terminate mostly in the medial magnocellular division of the ventral anterior nucleus (VAmc) and the mediodorsal nucleus (MDmc), which, in turn, innervate anterior regions of the frontal lobe including the principal sulcus (Walker’s area 46) and the orbital cortex (Walker’s area 11) in monkeys (140). Neurons in the lateral SNr project preferentially to the lateral posterior region of VAmc and to different parts of the MD. These areas of the thalamus are predominately related to posterior regions of the frontal lobe including the frontal eye field and areas of the premotor cortex, respectively (140). As is the case with GPi, SNr also sends projections to the noncholinergic neurons in the medial two-thirds of the PPN (117 ,243 ,271 ,277). Additional projections reach the parvicellular reticular formation, a region whose neurons are directly connected with orofacial motor nuclei (55 ,204 ,304), and the superior colliculus, which may play a critical role in the control of saccades (319). The latter projection is far more prominent in phylogenetically old animal species (amphibians) than in primates (189).

ROLE OF THE BASAL GANGLIA–THALAMOCORTICAL CIRCUITRY IN THE CONTROL OF MOVEMENT

Part of "122 - Neurocircuitry of Parkinson’s Disease "

At the most basic level, voluntary movements appear to be initiated at the cortical level of the motor circuit with output to brainstem and spinal cord, and to multiple subcortical targets, including the thalamus, putamen, and the STN. The exact nature of the information reaching either striatum

or the STN is not clear. Thus, studies of the electrophysiologic properties of corticostriatal projection neurons have shown that these neurons are different from corticospinal projection neurons (20 ,295) and tend to have slower conduction velocities and lower spontaneous rates, and are usually not responding to somatosensory input.

According to the current model of the functions of the basal ganglia–thalamocortical circuitry, activation of an ensemble of striatal neurons that give rise to the direct pathway leads to a reduction of inhibitory basal ganglia output from targeted neurons with subsequent disinhibition of related thalamocortical neurons (142). The net effect is increased activity in appropriate cortical neurons, resulting in a *facilitation* of the movement. In contrast, activation of the striatal neurons that give rise to the indirect pathway will lead to increased basal ganglia output and, presumably, to *suppression* of movement. Because the majority of neurons in GPi increase their firing rate with movement (103 ,202), the presumed increased suppression of unintended competing movements may be a particularly important role of the basal ganglia. Depending on the precise timing and anatomic connectivity, this dual action on movement could result in limiting the spatial or temporal extent of movements.

Clinical and experimental studies suggest that the basal ganglia play a role in specifying the amplitude or velocity of movement (14 ,46 ,59 ,74 ,142 ,296) or in maintaining postural stability during arm movements (142). The combination of information traveling via the direct and the indirect pathways of the motor circuit has been proposed to serve to either scale or focus movements (7 ,200 ,211). *Scaling* would be achieved by a temporal sequence of activity changes in the basal ganglia. Striatal output would first inhibit specific neuronal populations in GPi/SNr via the direct pathway, thus facilitating movement, followed by disinhibition of the *same* GPi/SNr neuron via inputs over the indirect pathway, leading to inhibition (“braking”) of the ongoing movement. In the *focusing* model, by contrast, inhibition of relevant pallidal/nigral neurons via the direct pathway would allow intended movements to proceed, whereas unintended movements would be suppressed by concomitant increased excitatory input via the indirect pathway in *other* GPi/SNr neurons (see discussions in refs. 145 and 309). Overall, the effect exerted by the two pathways in this case would be to further shape or sculpt the movement.

Both models are not entirely compatible with the available data. The focusing model, however, is difficult to reconcile with the fact that basal ganglia neurons become active after changes in cortex and thalamus are manifest (13 ,63 ,73 ,75 ,103 ,202 ,293 ,294 ,309). Both models are at odds with the fact that although STN lesions (thus an interference with the indirect pathway) result in spontaneous dyskinesias, they do not directly disrupt or alter voluntary movements.

The view that the basal ganglia are involved in the direct control of ongoing movements is too simplistic. A multitude of other motor functions of the basal ganglia are strong candidates, such as a role in self-initiated (internally generated) movements, in motor (procedural) learning, and in movement sequencing (115 ,250 ,318). These can only be mentioned in passing here, but will probably gain greater prominence in future models of basal ganglia function.

CHANGES IN BASAL GANGLIA CIRCUIT ACTIVITY IN PARKINSONISM

Part of "122 - Neurocircuitry of Parkinson’s Disease "

Regardless of the precise causation of the disease, all of the proposed proparkinsonian mechanisms have in common interference with the synthesis, and release or action of dopamine in the basal ganglia as well as cortex and thalamus. The study of pathophysiologic changes in the basal ganglia that result from loss of dopaminergic transmission in the basal ganglia has been greatly facilitated by the discovery that primates treated with MPTP develop behavioral and anatomic changes that closely mimic the features of PD in humans (17 ,47 ,100 ,170).

Changes in the activity over striatopallidal pathways were first suggested by studies in MPTP-induced parkinsonism in primates that indicated that the metabolic activity (as measured with the 2-deoxyglucose technique) is increased in both pallidal segments (60 ,201 ,222 ,252). This was interpreted as evidence of increased activity of the striatum-GPe connection and the STN-GPi pathway, or, alternatively, as evidence of increased activity via the projections from the STN to both pallidal segments. It was then shown directly with microelectrode recordings of neuronal activity that MPTP-induced parkinsonism in primates is associated with reduced tonic neuronal discharge in GPe, and increased discharge in the STN and GPi, as compared to normal controls (see example recordings in Fig. 122.1) (31 ,39 ,94 ,95 ,199).

In parkinsonian patients undergoing pallidotomy it has also been shown that the discharge rates in GPe are significantly lower than those in GPi (83 ,182 ,284 ,302), as had previously been shown in the MPTP-primate model. Recently, we have shown that treatment with MPTP results also in changes of neuronal activity in the second output nucleus of the basal ganglia, the SNr (Fig. 122.1). These changes in activity are qualitatively similar to those occurring in GPi (312). In addition, loss of dopamine in the striatum should also lead to reduced activity via the inhibitory direct pathway. To date, this has not been directly demonstrated, however.

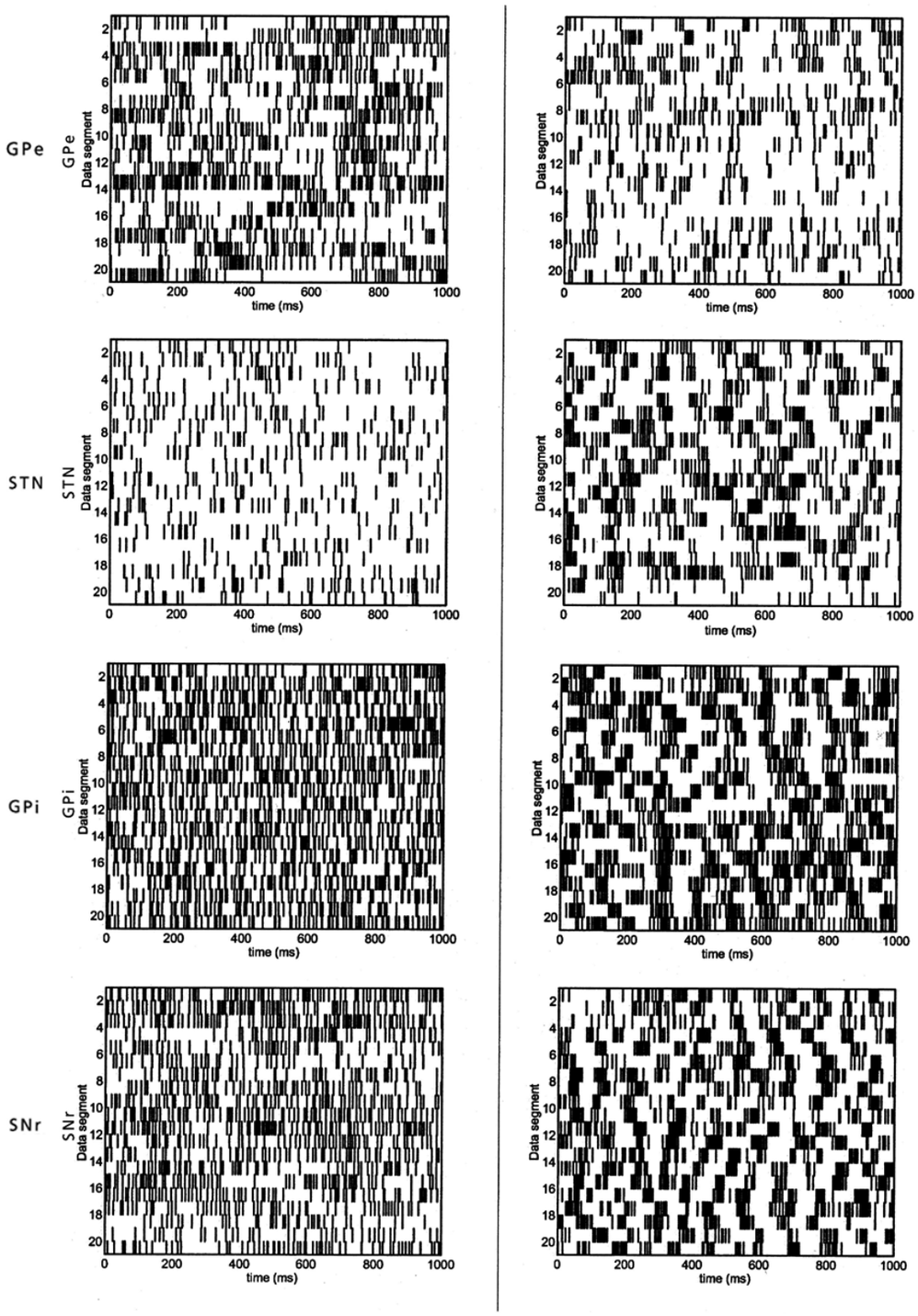


FIGURE 122.1. Raster displays of spontaneous neuronal activity recorded in different basal ganglia structures within the basal ganglia circuitry in normal and parkinsonian primates. Shown are ten consecutive 1000-msec segments of data from the external and internal segments of the globus pallidus (GPe, GPi, respectively), the subthalamic nucleus (STN), and the substantia nigra pars reticulata (SNr). The neuronal activity is reduced in GPe, and increased in STN, GPi, and SNr. In addition to the rate changes, there are also obvious changes in the firing patterns of neurons in all four structures, with a marked prominence of burstiness and oscillatory discharge patterns in the parkinsonian state. For further explanations, see text.

The changes in discharge rates in the subnuclei of the basal ganglia have been interpreted as indicating that striatal dopamine depletion leads to increased activity of striatal neurons of the indirect pathway, resulting in inhibition of GPe, and subsequent disinhibition of STN and GPi/SNr. The proposed pathophysiologic model of changes in the level of activity in the basal ganglia–thalamocortical motor circuit is summarized in Fig. 122.2 .

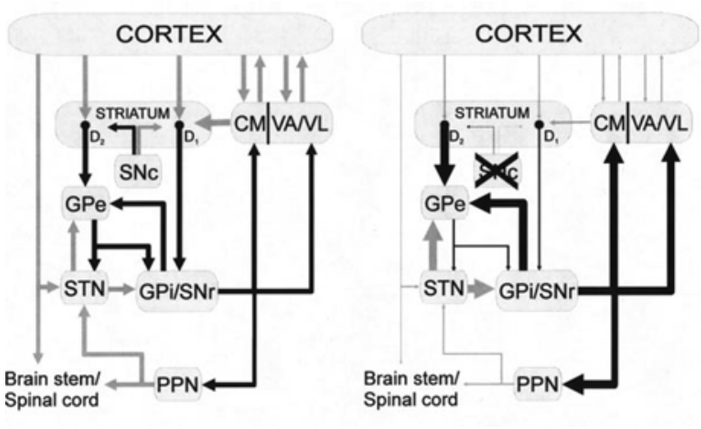


FIGURE 122.2. Model of the proposed rate changes in the basal ganglia–thalamocortical circuitry under normal (**left**) and parkinsonian conditions (**right**). In parkinsonism, dopaminergic neurons in the the substantia nigra pars compacta (SNc) degenerate, which results, via a cascade of changes in the other basal ganglia nuclei, in increased basal ganglia output from GPi and SNr. This, in turn, is thought to lead to inhibition of related thalamic and cortical neurons. In addition to the changes shown here, there are prominent alterations in discharge patterns (see text).

The basal ganglia circuitry incorporates multiple negative and positive feedback loops that may play a prominent role in the development and maintenance of abnormal discharge in the basal ganglia output structures. Some of the primary feedback loops that may directly affect GPi activity involve intrinsic basal ganglia structures such as GPe and STN (the two pathways labeled *3* in Fig. 122.3), or structures outside of the basal ganglia, such as the thalamic nucleus CM (labeled *1* in Fig. 122.3), the PPN (labeled *2* in Fig, 122.3) (101 ,117 ,161 ,263), and the habenula (e.g., GPi → lateral habenula → raphe nuclei → SNc → striatum → direct, indirect pathway → GPi; not shown in Fig. 122.3). Positive feedback loops, such as the one involving PPN and the STN (labeled *2*) and the pathway through CM and the putamen (labeled *1*) will tend to aggravate or enhance the abnormalities of discharge in the basal ganglia output nuclei associated with movement disorders, such as PD, whereas negative feedback circuits, such as a feedback involving CM and STN (not shown) will act to normalize neuronal discharge in the basal ganglia output nuclei. It is worth noting that via the CM nucleus, activity changes in the indirect pathway may influence the activity along the direct pathway. Thus, increased STN output in parkinsonism, by an action via GPi and CM, may result in a reduction of activity along the direct pathway.

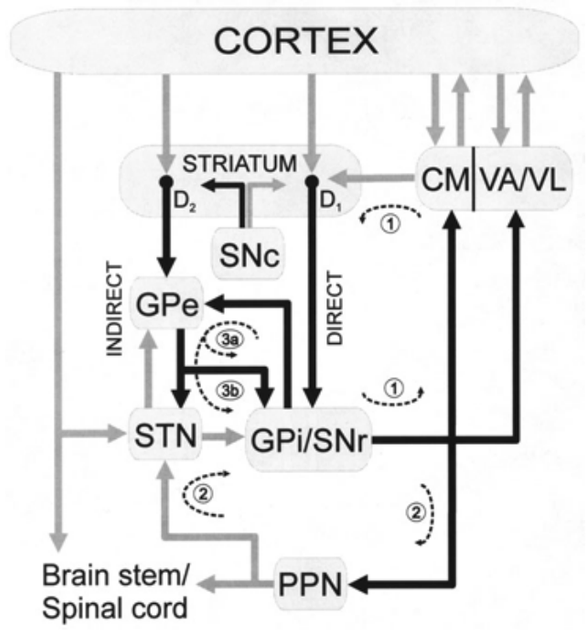


FIGURE 122.3. Simplified schematic diagram of the basal ganglia–thalamocortical circuitry under normal conditions. Inhibitory connections are shown as *filled arrows*, excitatory connections as *open arrows*. The principal input nuclei of the basal ganglia, the striatum, and the STN are connected to the output nuclei—GPi and SNr. Basal ganglia output is directed at several thalamic nuclei [ventral anterior/ventrolateral (VA/VL) and centromedian (CM)] and at brainstem nuclei [pedunculopontine nucleus (PPN) and others]. Some of the many important feedback connections are shown by the *dashed lines*. For further explanation of the model, see text.

The pathophysiology of early parkinsonism may differ from that of late parkinsonism in several aspect. For instance, increased STN output in early parkinsonism may have a compensatory function by increasing glutamatergic drive on SNc neurons. Thus, it has been shown that local injections of glutamate receptor blockers into the SNc significantly worsen motor signs in early stages of MPTP-induced parkinsonism (36 ,37), whereas such worsening is

no longer seen in later stages of the disease, probably reflecting loss of the majority of dopamine neurons in the SNc. At the same time, increased glutamatergic drive onto surviving SNc neurons may also be (excito-) toxic (239).

The reciprocal changes in activity in the indirect and direct pathways following dopamine depletion should both result in increased activity in GPi/SNr, and, subsequently, increased basal ganglia output to the thalamus and increased inhibition of thalamocortical neurons. The 2-deoxyglucose studies mentioned above demonstrated increased (synaptic) activity in the VA and VL nucleus of thalamus (60 ,201 ,252), presumably reflecting increased inhibitory basal ganglia output to these nuclei. Consistent with this are positron emission tomography (PET) studies in parkinsonian patients that have consistently shown reduced activation of motor and premotor areas in such patients (42 ,48 ,54 ,88 ,90), although no changes have been seen in the thalamus. Alterations of cortical activity in motor cortex and supplementary motor areas have also been demonstrated with single-cell recording in hemiparkinsonian primates (306).

The finding that SNr activity is also abnormal in parkinsonism is potentially important, because output from this nucleus reaches different cortical targets than output from GPi. For instance, the movement-related output from the SNr appears to reach predominately premotor areas, and could conceivably play a role in some aspects of akinesia (141). In addition, the SNr carries a substantial portion of the nonmotor circuitry of the basal ganglia. Abnormal SNr discharge may therefore be associated with some of the non-(limb)-motor abnormalities in parkinsonism, including oculomotor disturbances as well as cognitive, behavioral, and emotional disturbances.

Brainstem areas such as the PPN may also be involved in the development of parkinsonian signs. It has been shown that lesions of this nucleus in normal monkeys can lead to hemiakinesia, possibly by reducing stimulation of SNc neurons by input from the PPN, or by a direct influence on descending pathways (51 ,146 ,162 ,206). It remains unclear, however, whether the motor abnormalities seen after PPN inactivation in fact are related to parkinsonism or represent changes in the behavioral state or other disturbances that have no direct relation to PD. It is noteworthy that these animals do not manifest rigidity or tremor, which appear to be critically dependent on thalamic circuitry (see below).

It is important to realize that parkinsonism is a *network* disease. Changes that arise in any portion of the complex basal ganglia–thalamocortical circuitry will have significant consequences in all other areas of the network. This implies that the search for a parkinsonism-inducing “source” of abnormalities in the neuronal activity within the network may be futile, but suggests also that surgical or pharmacologic interventions at a variety of targets within the network could be successful. This can indeed be appreciated when considering the results of lesion studies in parkinsonian primates. One of the most important and dramatic in this regard was the demonstration that lesions of the STN in MPTP-treated primates reverse all of the cardinal signs of parkinsonism, presumably by reducing GPi activity (16 ,30 ,119). Similarly, GPi and SNr inactivation have been shown to be effective against at least some parkinsonian signs in MPTP-treated primates (179 ,181 ,308 ,315).

Over the last decade, these results from animal studies have rekindled interest in functional neurosurgical approaches to the treatment of medically intractable PD. This was first employed in the form of GPi lesions (pallidotomy) (19 ,85 ,169 ,183 ,276 ,301) and, more recently, with STN lesions (108). In addition, high-frequency deep brain stimulation (DBS) of both the STN and GPi have been shown to reverse parkinsonian signs. The mechanism of action of DBS remains controversial. It appears most likely, however, that DBS and lesions act similarly in that both result in an overall reduction of basal ganglia output.

PET studies in pallidotomy patients performing a motor task have shown that frontal motor areas whose metabolic activity was reduced in the parkinsonian state became again active after the procedure (53 ,85), providing support for the concept of excessive pallidal inhibition of thalamocortical systems in PD, which, when eliminated, reverses the major parkinsonian signs. DBS of the STN and GPi have revealed similar changes with PET, further supporting this concept as well as the belief that DBS appears to act functionally like ablation.

The experience with inactivation or deep brain stimulation of the SNr is very limited at this point. There are no studies of the effects of (exclusive) lesioning of the SNr available, and one case report on the effects of (inadvertent) stimulation in the ventral STN/dorsal SNr area reported the appearance of psychiatric depression during episodes of stimulation (23). This clearly needs further study, but it seems that the SNr may not be a feasible target for surgical interventions, because of its prominent involvement in nonmotor functions, and possibly also because of the greater degree of overlap between the different functional territories in this nucleus (123 ,127 ,186).

Controversial Issues

It has long been clear that the aforementioned models of the pathophysiology of parkinsonism are too simplistic, and that they cannot explain many of the clinical and experimental features of the disease. Thus, although the results of lesions in parkinsonism seem at first glance easily explained by the above-mentioned rate-based model of parkinsonism, more detailed studies of the results of lesions in patients with parkinsonism have brought to light several important findings that are not compatible with the models. For instance, in contrast to the prediction of simple rate-based models, lesions of the “basal ganglia–receiving” areas of the thalamus (VA/VL) do not lead to parkinsonism and in fact are beneficial in the treatment of both tremor and

rigidity (45 ,109 ,220 ,290)). Similarly, lesions of GPi in the setting of parkinsonism lead to improvement in all aspects of PD without any obvious detrimental effects. Furthermore, they are, often in the same patient, effective against both parkinsonism and drug-induced dyskinesias (217 ,235). In contrast to the abnormalities seen in parkinsonism, such dyskinesias are thought to arise from pathologic *reduction* in basal ganglia outflow (223 ,313), and thus should not respond positively to further reduction of pallidal outflow (190).

The assumption that parkinsonism may at least in part result from altered processing of proprioceptive input, abnormal timing, patterning, and synchronization of discharge that introduces errors and nonspecific noise into the thalamocortical signal may help to explain these seemingly paradoxical findings. Alterations in discharge patterns and synchronization between neighboring neurons have been extensively documented in parkinsonian monkeys and patients. For instance, neuronal responses to passive limb manipulations in STN, GPi and thalamus (31 ,95 ,199 ,299) have been shown to occur more often, to be more pronounced, and to have widened receptive fields after treatment with MPTP. There is also a marked change in the synchronization of discharge between neurons in the basal ganglia. Cross-correlation studies have revealed that a substantial proportion of neighboring neurons in the globus pallidus and STN discharge in unison in MPTP-treated primates (31). This is in contrast to the virtual absence of synchronized discharge of such neurons in normal monkeys (309). Finally, the proportion of cells in STN, GPi, and SNr that discharge in oscillatory or nonoscillatory bursts is greatly increased in the parkinsonian state (31 ,94 ,199 ,300 ,302 ,311). Oscillatory burst discharge patterns are often seen in conjunction with tremor. The question of whether this is simply a reflection of tremor-related proprioceptive input or of active participation of basal ganglia in the generation of tremor is still unsettled (see below).

Conceivably, altered neuronal activity patterns in the basal ganglia may play an important role in parkinsonism. Thus, increased phasic activity in the basal ganglia may erroneously signal excessive movement or velocity to precentral motor areas, leading to a slowing or premature arrest of ongoing movements and to greater reliance on external clues during movement. Alternatively, phasic alteration of discharge in the basal ganglia may simply introduce noise into thalamic output to the cortex that is detrimental to cortical operations. The polarity and exact nature of the abnormal patterning and overall activity in the basal ganglia–thalamocortical pathways may determine the nature of the resulting movement disorder.

The foregoing discussion indicates that in patients with movement disorders it is not only the loss of basal ganglia contribution to movement that must be compensated for, but also the disruptive influence of the inappropriate basal ganglia output. The therapeutic benefits of GPi and STN lesions suggest that in patients with PD and other movement disorders the absence of basal ganglia input to the still intact portions of the basal ganglia–thalamocortical network is more tolerable than abnormal input. Near-normal motor function is still possible in these disorders once the abnormal basal ganglia–thalamocortical input is removed. It needs to be emphasized, however, that the surgical interventions do not necessarily normalize cortical motor mechanisms in parkinsonian subjects, but rather may allow the intact portions of the thalamocortical and brainstem system to more effectively compensate for the loss of the basal ganglia contribution to movement.

Another recent further challenge to the proposed pathophysiologic model of parkinsonism has arisen from histochemical studies on the amount of messenger RNA (mRNA) for GAD_{67}, one of the enzymes synthesizing γ-aminobutyric acid (GABA) in basal ganglia neurons. In contrast to GAD itself, which is found in neuronal cell bodies or terminals, the mRNA for the enzyme is thought to be contained exclusively in cell bodies. In these studies the GAD mRNA activity in a given nucleus is therefore taken as a parameter for the level of activity of GABAergic neurons in the nucleus under study. Experiments in parkinsonian primates have shown that, as expected from the above-mentioned model, GAD_{67} mRNA activity is increased in GPi neurons (131 ,132 ,269), and is reversed with levodopa administration. GAD_{67} mRNA activity in the GPi of humans with parkinsonism, however, was found to be similar to that in controls, possibly because these patients were chronically treated with levodopa (131 ,132). Some of the findings regarding GAD mRNA in GPe, however, are at odds with the above-mentioned model in which the activity of GABAergic neurons in GPe is decreased. In rats, primates and humans, GAD_{67} mRNA in GPe was either unchanged in the parkinsonian state or even increased (56 ,72 ,131 ,132 ,268). These results have been interpreted as evidence that GPe and GPi function may not be as tightly linked via the indirect pathway as proposed by the model outlined above, and that the observed activity changes in the basal ganglia may primarily be due to altered activity via the corticosubthalamic projection or dopaminergic inputs to STN itself, which, by changing STN activity, may cause the neuronal activity in both nuclei to increase, possibly due to a greater tendency of neurons to discharge in bursts (178). However, the consistent finding of significantly decreased GPe discharge in MPTP-treated animals (93 ,94 ,199) and patients with PD (83 ,182 ,284 ,302) is difficult to reconcile with the lack of change in GAD_{67} mRNA in GPe. Conceivably, GAD_{67}-mRNA levels may reflect something other than neuronal discharge rates (191 ,238), or may be greatly influenced by the emergence of burst discharges. In a recent study it was shown that GABA levels in the STN, which are at least in part reflective of GABA release from terminals of GPe axons, were reduced in MPTP-treated primates, as predicted by the above-mentioned model (267). This finding casts further

doubt on the assumption that GAD_{67}-mRNA levels are a reliable predictor of the activity along the GPe outflow pathways.

PATHOPHYSIOLOGY OF INDIVIDUAL PARKINSONIAN MOTOR SIGNS

Part of "122 - Neurocircuitry of Parkinson's Disease "

Although the cardinal parkinsonian signs of tremor, rigidity, akinesia, and bradykinesia are generally all present in a given patient, they can occur independently of each other. For instance, patients with severe akinesia/bradykinesia do not necessarily exhibit tremor or rigidity, and severely akinetic patients may not experience significant bradykinesia or rigidity. This suggests that the different signs may depend on different pathophysiologic mechanisms, possibly involving different subcircuits of the larger motor circuit. The physiologic basis of the cardinal parkinsonian motor signs will be briefly considered in the following subsections.

Akinesia

Akinesia, the hallmark of PD, is characterized by a global impairment of movement initiation, affecting gross and fine movements as well as gait. In extreme cases, akinesia is experienced as freezing episodes, i.e., periods of complete motor block (107). Although there is some evidence that certain aspects of akinesia may be related to abnormal activity along the brainstem projections of the basal ganglia output nuclei (162), most authors attribute akinesia to changes in cortical processing, due to altered basal ganglia output to the thalamus. Freezing episodes may be the manifestation of temporary near-complete failure of compensatory mechanisms. This happens more often in late than in earlier stages of the disease, suggesting that the compensatory reserve of remaining intact thalamic, cortical, and brainstem circuits becomes smaller as the disease progresses.

As mentioned above, as a first approximation, overall discharge rates in the basal ganglia output nuclei have an impact on movement. GPi/SNr rates are determined by the amount of striatal dopamine, which in turn determines the balance between overall discharge in the direct and indirect pathways. There are many possible ways in which increased basal ganglia output could lead to akinesia. For instance, increased tonic inhibition of thalamocortical neurons by excessive output from GPi/SNr may reduce the responsiveness of cortical mechanisms involved in motor control. Increased tonic inhibition of thalamocortical neurons by increased basal ganglia output in parkinsonism may also render precentral motor areas less responsive to other inputs normally involved in initiating movements or may interfere with "set" functions that have been shown to be highly dependent on the integrity of basal ganglia pathways (6).

Akinesia may be a good example of a parkinsonian sign whose development appears to depend on discharge abnormalities in specific subcircuits of the motor loop. PET studies of cortical activation in akinesia-predominant parkinsonism suggest that the supplementary (SMA) and dorsal premotor areas are hypoactive in such patients (44 ,147). Moreover, pallidotomy results in increased metabolism in these areas in association with improvement in akinesia and bradykinesia (43 ,88 ,91 ,114 ,129 ,246). Further evidence for abnormal activity in these nuclei comes from studies of the Bereitschaftspotential (readiness potential), a slow negative cortical potential that precedes self-paced movements and is thought to reflect the neural activity in SMA (71). The early portion of the Bereitschaftspotential is smaller in parkinsonian patients than in age-matched controls (82 ,218), suggesting a deficit in the normal function of the SMA in the early stages of preparation for self-initiated movements. Akinesia may be thus related to abnormal discharge in a subcircuit whose activity may be to a large degree "preparatory" (5 ,12 ,40 ,61 ,148 ,251), interfering with the planning and early execution stages of movement. A disorganization of preparatory activity in SMA neurons was indeed identified with electrophysiologic methods in hemiparkinsonian primates (306).

One of the inconsistencies with the concept of akinesia as a consequence of increased inhibition of thalamocortical neurons is the finding that thalamic lesions per se do not appear to result in akinesia, as predicted by the model (but see ref. 49), although VA/VL lesions are effective in reducing rigidity and tremor. These findings argue against the view that increased tonic pallidal output and resulting inhibition of the neurons in the VA/VL nuclei is the sole or even the major reason for the development of akinesia. Alternatively, the fact that ventral thalamic lesions do not appear to influence akinesia may indicate that akinesia develops as a consequence of abnormally reduced activity in the intralaminar thalamic nuclei, or in the PPN with its descending brainstem projections. As discussed earlier, CM/Pf involvement may be the reason for the finding of prominent changes in cortical activity associated with akinesia, whereas involvement of the PPN is suggested by the finding that lesions of this structure result in poverty of movement (162 ,206).

Bradykinesia

Although bradykinesia is usually associated with akinesia, as mentioned earlier, these two signs can be strikingly dissociated in some patients. The pathophysiology of bradykinesia may be closely associated with the postulated scaling function of basal ganglia output (see above) (25 ,305) and is probably also dependent on abnormal processing in prefrontal cortical areas that are strongly influenced by increased basal ganglia output. In normal monkeys, neurophysiologic studies and, more recently, PET studies investigating cerebral blood flow have described an influence

of velocity/amplitude on the discharge of neurons in these premotor cortical areas (21 ,62 ,130 ,296).

Conceivably, abnormally increased phasic GPi/SNr output during movement may signal excessive speed and/or amplitude of ongoing movement, leading to a corrective reduction in cortical motor output (as mentioned above). PET studies, measuring cerebral blood flow in human parkinsonian patients investigated before and during deep brain stimulation of GPi, have revealed that stimulation that improved bradykinesia led to an increase in blood flow in the ipsilateral premotor cortical areas (69). A PET study has shown a significant correlation between movement speed and basal ganglia activation (296), and the loss of this in PD (Turner et al., personal communication). Thus there are several independent lines of evidence for the role of the basal ganglia motor circuitry in the scaling of movement and the disruption of this in diseases such as PD.

Rigidity

Parkinsonian rigidity is characterized by a uniform (“plastic”) increase in resistance to passive movements about individual joints. A “cogwheel” feature may result from superimposed, and usually subclinical, tremor (96). The pathophysiology of rigidity is elusive, but it has been suggested that altered basal ganglia output, mediated via the PPN and its output to the pontine nucleus gigantocellularis and the dorsal longitudinal fasciculus of the reticulospinal projection, may lead to increased inhibition of spinal Ib interneurons, which in turn may disinhibit α-motoneurons (50 ,76 ,139 ,175). Abnormalities of long-latency reflexes (LLRs) may also play a role in abnormal α-motoneuron excitability (26 ,174 ,291 ,292 ,316), although the velocityindependence of rigidity suggests that it is not a reflex phenomenon per se. The finding that rigidity can be abolished by interruption of the basal ganglia–thalamocortical circuit at multiple levels (STN, GPi, and thalamus) suggests that pallidal output leads to rigidity via the thalamocortical route rather than via brainstem projections.

Tremor

Parkinsonian tremor is typically a 4- to 5-Hz tremor at rest that is suppressed by voluntary movement. Parkinsonian tremor has been shown to be critically dependent on the integrity of the thalamic nucleus ventralis intermedius (Vim), which contains neurons that exhibit oscillatory discharge at the tremor frequency (177 ,214 ,219), although a tight correlation between oscillatory discharge and tremor is often not observed (28 ,221). Lesions of Vim have also been shown to abolish tremor (207). It has been proposed that thalamic oscillatory discharge may be induced by hyperpolarization of these cells induced by increased inhibitory basal ganglia output (224). This increases the likelihood that these cells will discharge in bursts (156 ,180 ,278). On the other hand, tremor may also arise from oscillatory discharge originating within the basal ganglia, based on the finding of oscillatory discharge patterns in the STN and GPi in parkinsonian patients and animals (31 ,84 ,150 ,284 ,311). These oscillatory discharge patterns may arise from local pacemaker networks, such as a feedback circuit involving GPe and STN (232). It has also been speculated that intrinsic membrane properties of basal ganglia neurons are conducive to the development of oscillatory discharge (15 ,34 ,209) in basal ganglia neurons themselves, or that they may contribute to the generation of oscillatory discharge in the thalamus (303 ,310 ,313), which may then be transmitted to the cortex. Finally, and perhaps most likely, oscillations throughout the entire basal ganglia–thalamocortical network may be tightly related to each other, so that no one oscillator can be identified as their sole source (187 ,188).

Support for the concept that altered basal ganglia discharge either alone or as part of the basal ganglia–thalamocortical network is important in the pathogenesis of tremor comes from lesion studies showing that parkinsonian tremor in MPTP-treated African green monkeys and in patients with parkinsonism is significantly ameliorated by lesions of the STN and GPi (19 ,29 ,284 ,310). It has been suggested that loss of *extrastriatal* dopamine may contribute to the development of tremor, because primates in which MPTP treatment affects the dopamine supply to GPi (African green monkeys) tend to develop tremor, whereas species in which the dopamine supply to GPi is not as severely affected (Rhesus monkeys) rarely develop parkinsonian tremor (28). Furthermore, in a postmortem study it was found that the degree of dopamine loss in the striatum did not correlate with the extent of tremor in parkinsonian patients, whereas the degree of dopamine loss in the pallidum did (32). More direct evidence for a role of extrastriatal dopamine in the development of tremor is lacking, however.

It remains unclear which oscillation frequency within the basal ganglia in fact is related to tremor. Oscillatory discharge in the basal ganglia of MPTP-treated primates has been shown to occur in at least two different frequency bands—the 3- to 5-Hz range, and the 8- to 15-Hz range (28 ,31 ,310). Although, intuitively, discharge in the lower frequency range would be expected to be more directly related to tremor at the typical parkinsonian frequency range, there is some evidence that oscillations in the higher frequency range in fact may be an important determinant of tremor. Thus, in MPTP-treated animals in which tremor had been eliminated with STN lesions, oscillations in the 8- to 10-Hz range in GPi were also greatly reduced, whereas oscillatory discharge in the lower frequency band persisted (310). In addition, basal ganglia neurons in tremulous animals show considerable coherence in the 8- to 15-Hz range, but not in the 3- to 5-Hz range (28). A process by which 10-Hz oscillations could be transformed in the thalamus into 5-Hz oscillatory discharge has been proposed by Pare

et al. (224). However, even primate species that generally do not show tremor after MPTP treatment (i.e., Rhesus monkeys) develop 8- to 13-Hz oscillations in the basal ganglia output nuclei (28 ,31 ,199 ,312). The difference between monkeys with and those without tremor may therefore lie not in the presence or absence of oscillations, but rather in the degree of synchronization between neighboring neurons. For tremor to occur, significant synchrony with little phase difference in large neuronal assemblies may be required. Thus, in one study the phase shift distribution of oscillatory cross-correlograms of neighboring pallidal cells in MPTP-treated vervet monkey were tightly clustered around a 0-degree phase shift, whereas the oscillatory correlograms in the MPTP-treated rhesus monkey were more widely scattered between 0 and 180 degrees (28). It is tempting to implicate the motor subcircuit that is centered on the motor cortex in the pathogenesis of tremor, because tremor-related neurons are focused primarily in the most ventral portions of the sensorimotor GPi (118), a region that in the monkey has been shown to project (via the thalamus) to the motor cortex.

Nonmotor Symptoms of Parkinson's Disease

Besides the cardinal (and early) skeletomotor abnormalities, parkinsonism is also associated with oculomotor abnormalities, such as hypometric and slow saccades (41 ,138 ,184 ,185 ,255 ,298 ,307), autonomic dysfunction, depression, anxiety, sleep disturbances, impaired visuospatial orientation and cognitive abnormalities (18 ,64 ,67 ,194 ,275). It is likely that at least some of these abnormalities rely on abnormal discharge in nonmotor circuits of the basal ganglia, which may be affected by dopamine loss in much the same way as the motor circuit. This is particularly true for oculomotor abnormalities that may directly result from dopamine depletion in the caudate nucleus (151 ,164). Similarly, some (20) of the cognitive and psychiatric disturbances seen in parkinsonian patients are reminiscent of syndromes seen after lesions of the dorsolateral prefrontal cortex (problems with executive functions) or of the anterior cingulate (apathy, personality changes), and may be the result of loss of dopamine in the dorsolateral or ventral caudate nucleus, respectively (65). Besides abnormalities in dopaminergic transmission in the striatum, several studies indicate that concomitant noradrenergic and serotoninergic deficiencies could contribute to the mood alteration in PD. For instance, cerebrospinal fluid levels of the serotonin metabolite 5-hydroxyindolacetic acid are reduced in depressed parkinsonian patients compared with nondepressed parkinsonian patients, and selective serotonin reuptake inhibitors are effective in treating parkinsonian depression (10 ,67 ,195). Finally, disturbance of the normal function of cortical–basal ganglia–thalamocortical circuits has also been implicated in the occurrence of obsessive-compulsive disease, as well as in Tourette's syndrome (66 ,89 ,133).

CONCLUSION

Part of "122 - Neurocircuitry of Parkinson's Disease "

From the considerations above, a complex model of parkinsonism emerges in which relatively selective dopamine depletion in the striatum and other basal ganglia nuclei results in increased and disordered discharge and synchronization in motor areas of the basal ganglia–thalamocortical motor loops. Abnormal activity in one or more of the basal ganglia feedback loops may contribute to the development of parkinsonism. Individual parkinsonian motor signs appear to be caused by distinct abnormalities in basal ganglia discharge, and by involvement of specific subcircuits related to distinct cortical targets. It is probable that progressive loss of dopamine in nonmotor areas of the striatum and other basal ganglia nuclei may underlie the nonmotor abnormalities of PD. By contrast, drug-induced dyskinesias are characterized by decreased pallidal output. Differences in the balance between direct and indirect pathways, and in the degree of synchronization of discharge in the basal ganglia output nuclei must be invoked to explain the striking clinical differences between PD and the drug-induced dyskinesias and dystonia. A critical analysis of the effects of pallidal and thalamic lesions in hypo- and hyperkinetic disorders strongly suggests that the main features accounting for the different signs of movement disorders are the appearance of not only changes in discharge rate, but also altered discharge patterns, changes in the degree of synchronization of discharge, altered proprioceptive feedback, and "noise" in the basal ganglia output signal. It is proposed that both ablation and deep brain stimulation are effective in treating both hypo- and hyperkinetic disorders because they both remove the abnormal signals directed to the thalamus and brainstem, thus allowing these intact systems to compensate more effectively.

The current models of basal ganglia pathophysiology are incomplete and should be taken as a first draft of basal ganglia dysfunction in the different disease states. Most pertinently, changes in phasic discharge patterns, and new anatomic connections need to be better incorporated into any new concept of basal ganglia function and a greater emphasis placed on the manner in which thalamic, brainstem, and cortical neurons utilize basal ganglia output.

REFERENCES

1. Afsharpour S. Topographical projections of the cerebral cortex to the subthalamic nucleus. *J Comp Neurol* 1985;236:14–28.

2. Aizman O, Brismar H, Uhlen P, et al. Anatomical and physiological evidence for D1 and D2 dopamine receptor colocalization in neostriatal neurons. *Nature Neurosci* 2000;3:226–230.

3. Albin RL, Young AB, Penney JB. The functional anatomy of basal ganglia disorders. *Trends Neurosci* 1989;12:366–375.

4. Albin RL, Young AB, Penney JB. The functional anatomy of disorders of the basal ganglia. *Trends Neurosci* 1995;18:63–64.

5. Alexander GE, Crutcher MD. Coding in spatial rather than joint coordinates of putamen and motor cortex preparatory activity preceding planned limb movements. In: Sambrook MA, Crossman AR, eds. *Neural mechanisms in disorders of movement*. London: Blackwell, 1989:55–62.

6. Alexander GE, Crutcher MD. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci* 1990;13:266–271.

7. Alexander GE, Crutcher MD, DeLong MR. Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, “prefrontal” and “limbic” functions. *Prog Brain Res* 1990;85:119–146.

8. Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 1986;9:357–381.

9. Alheid GF, Heimer L. New perspectives in basal forebrain organization of special relevance for neuropsychiatric disorders: the striatopallidal, amygdaloid, and corticopetal components of substantia innominata. *Neuroscience* 1988;27:1–39.

10. Allain H, Schuck S, Mauduit N. Depression in Parkinson’s disease [editorial]. *BMJ* 2000;320:1287–1288.

11. Allan W. 1937. Inheritance of shaking palsy. *Arch Intern Med* 1937;60:424–436.

12. Anderson M, Inase M, Buford J, et al. Movement and preparatory activity of neurons in pallidal-receiving areas of the monkey thalamus. In: Mano N, Hamada I, Delong MR, eds. *Role of cerebellum and basal ganglia in voluntary movement*. Amsterdam: Elsevier, 1992:39.

13. Anderson ME, Horak FB. Changes in firing of pallidal neurons during arm reaching movements in a reaction time task. *Soc Neurosci Abstr* 1981;7:241.

14. Anderson ME, Turner RS. A quantitative analysis of pallidal discharge during targeted reaching movement in the monkey. *Exp Brain Res* 1991;86:623–632.

15. Awad H, Hubert GW, Smith Y, et al. Activation of metabotropic glutamate receptor 5 has direct excitatory effects and potentiates NMDA receptor currents in neurons of the subthalamic nucleus. *J Neurosci* 2000;20:7871–7879.

16. Aziz TZ, Peggs D, Sambrook MA, et al. Lesion of the subthalamic nucleus for the alleviation of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-induced parkinsonism in the primate. *Mov Disord* 1991;6:288–292.

17. Bankiewicz KS, Oldfield EH, Chiueh CC, et al. Hemiparkinsonism in monkeys after unilateral internal carotid artery infusion of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP). *Life Sci* 1986;39:7–16.

18. Barbosa ER, Limongi JC, Cummings JL. Parkinson’s disease. *Psychiatr Clin North Am* 1997;20:769–790.

19. Baron MS, Vitek JL, Bakay RAE, et al. Treatment of advanced Parkinson’s disease by GPi pallidotomy: 1 year pilot-study results. *Ann Neurol* 1996;40:355–366.

20. Bauswein E, Fromm C, Preuss A. Corticostriatal cells in comparison with pyramidal tract neurons: contrasting properties in the behaving monkey. *Brain Res* 1989;493:198–203.

21. Bauswein E, Fromm C, Werner W, et al. Phasic and tonic responses of premotor and primary motor cortex neurons to torque changes. *Exp Brain Res* 1991;86:303–310.

22. Beckstead RM. A reciprocal axonal connection between the subthalamic nucleus and the neostriatum in the cat. *Brain Res* 1983;275:137–142.

23. Bejjani BP, Damier P, Arnulf I, et al. Transient acute depression induced by high-frequency deep-brain stimulation. *N Engl J Med* 1999;340:1476–1480.

24. Bell J, Clark AJ. A pedigree of paralysis agitans. *Ann Eugen* 1926;1.

25. Berardelli A, Hallett M, Rothwell JC, et al. Single-joint rapid arm movements in normal subjects and in patients with motor disorders. *Brain* 1996;119:661–674.

26. Berardelli A, Sabra A, Hallett M. Physiological mechanisms of rigidity in Parkinson’s disease. *J Neurol Neurosurg Psychiatr* 1983;46:45–53.

27. Berendse HW, Groenewegen HJ. The connections of the medial part of the subthalamic nucleus in the rat. Evidence for a parallel organization. In: Bernardi G, Carpenter MB, Di Chiara G, et al., eds. *The basal ganglia II*. New York: Plenum Press, 1989:89–98.

28. Bergman H, Raz A, Feingold A, et al. Physiology of MPTP tremor. *Mov Disord* 1998;13(suppl 3):29–34.

29. Bergman H, Wichmann T, DeLong MR. Amelioration of parkinsonian symptoms by inactivation of the subthalamic nucleus (STN) in MPTP treated green monkeys. *Mov Disord* 1990;5(suppl1):79.

30. Bergman H, Wichmann T, DeLong MR. Reversal of experimental parkinsonism by lesions of the subthalamic nucleus. *Science* 1990;249:1436–1438.

31. Bergman H, Wichmann T, Karmon B, et al. The primate subthalamic nucleus. II. Neuronal activity in the MPTP model of parkinsonism. *J Neurophysiol* 1994;72:507–520.

32. Bernheimer H, Birkmayer W, Hornykiewicz O, et al. Brain dopamine and the syndromes of Parkinson and Huntington. *J Neurol Sci* 1973;20:415–455.

33. Betarbet R, Sherer TB, MacKenzie G, et al. Chronic systemic pesticide exposure reproduces features of Parkinson’s disease. *Nature Neurosci* 2000;3:1301–1306.

34. Beurrier C, Congar P, Bioulac B, et al. Subthalamic nucleus neurons switch from single-spike activity to burst-firing mode. *J Neurosci* 1999;19:599–609.

35. Bevan MD, Clarke NP, Bolam JP. Synaptic integration of functionally diverse pallidal information in the entopeduncular nucleus and subthalamic nucleus in the rat. *J Neurosci* 1997;17:308–324.

36. Bezard E, Boraud T, Bioulac B, et al. Compensatory effects of glutamatergic inputs to the substantia nigra pars compacta in experimental parkinsonism. *Neuroscience* 1997;81:399–404.

37. Bezard E, Boraud T, Bioulac B, et al. Involvement of the subthalamic nucleus in glutamatergic compensatory mechanisms. *Eur J Neurosci* 1999;11:2167–2170.

38. Bhatt MH, Elias MA, Mankodi AK. Acute and reversible parkinsonism due to organophosphate pesticide intoxication: five cases. *Neurology* 1999;52:1467–1471.

39. Boraud T, Bezard E, Guehl D, et al. Effects of L-DOPA on neuronal activity of the globus pallidus externalis (GPe) and globus pallidus internalis (GPi) in the MPTP-treated monkey. *Brain Res* 1998;787:157–160.

40. Boussaoud D, Kermadi I. The primate striatum: neuronal activity in relation to spatial attention versus motor preparation. *Eur J Neurosci* 1997;9:2152–2168.

41. Brooks BA, Fuchs AF, Finocchio D. Saccadic eye movement deficits in the MPTP monkey model of Parkinson’s disease. *Brain Res* 1986;383:402–407.

42. Brooks DJ. Detection of preclinical Parkinson’s disease with PET. *Neurology* 1991;41(suppl 2):24–27.

43. Brooks DJ. Motor disturbance and brain functional imaging in Parkinson’s disease. *Eur Neurol* 1997;38(suppl 2):26–32.

44. Brooks DJ. PET and SPECT studies in Parkinson’s disease. *Baillieres Clin Neurol* 1997;6:69–87.

45. Brophy BP, Kimber TJ, Thompson PD. Thalamotomy for parkinsonian tremor. *Stereotact Funct Neurosurg* 1997;69:1–4.

46. Brotchie P, Iansek R, Horne MK. Motor function of the monkey globus pallidus. 1. Neuronal discharge and parameters of movement. *Brain* 1991;114:1667–1683.

47. Burns RS, Chiueh CC, Markey SP, et al. A primate model of parkinsonism: selective destruction of dopaminergic neurons in the pars compacta of the substantia nigra by N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine. *Proc Natl Acad Sci USA* 1983;80:4546–4550.

48. Calne D, Snow BJ. PET imaging in Parkinsonism. *Adv Neurol* 1993;60:484.

49. Canavan AG, Nixon PD, Passingham RE. Motor learning in monkeys (Macaca fascicularis) with lesions in motor thalamus. *Exp Brain Res* 1989;77:113–126.

50. Cantello R, Gianelli M, Civardi C, et al. Pathophysiology of Parkinson's disease rigidity. Role of corticospinal motor projections. *Adv Neurol* 1996;69:129–133.

51. Carlson JD, Pearlstein RD, Buchholz J, et al. Regional metabolic changes in the pedunculopontine nucleus of unilateral 6-hydroxydopamine Parkinson's model rats. *Brain Res* 1999;828:12–19.

52. Carpenter MB, Batton RRI, Carleton SC, et al. Interconnections and organization of pallidal and subthalamic nucleus neurons in the monkey. *J Comp Neurol* 1981;197:579–603.

53. Ceballos-Bauman AO, Obeso JA, Vitek JL, et al. Restoration of thalamocortical activity after posteroventrolateral pallidotomy in Parkinson's disease. *Lancet* 1994;344:814.

54. Ceballos-Baumann AO, Brooks DJ. Basal ganglia function and dysfunction revealed by PET activation studies. *Adv Neurol* 1997;74:127–139.

55. Chandler SH, Turman J Jr, Salem L, et al. The effects of nanoliter ejections of lidocaine into the pontomedullary reticular formation on cortically induced rhythmical jaw movements in the guinea pig. *Brain Res* 1990;526:54–64.

56. Chesselet MF, Delfs JM. Basal ganglia and movement disorders: an update. *Trends Neurosci* 1996;19:417–422.

57. Chiueh CC, Burns RS, Markey SP, et al. Primate model of parkinsonism: selective lesion of nigrostriatal neurons by 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine produces an extrapyramidal syndrome in rhesus monkeys. *Life Sci* 1985;36:213–218.

58. Cochiolo JA, Ehsanian R, Bruck DK. Acute ultrastructural effects of MPTP on the nigrostriatal pathway of the C57BL/6 adult mouse: evidence of compensatory plasticity in nigrostriatal neurons. *J Neurosci Res* 2000;59:126–135.

59. Contreras-Vidal JL, Stelmach GE. A neural model of basal ganglia-thalamocortical relations in normal and parkinsonian movement. *Biol Cybernet* 1995;73:467–476.

60. Crossman AR, Mitchell IJ, Sambrook MA. Regional brain uptake of 2-deoxyglucose in N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-induced parkinsonism in the macaque monkey. *Neuropharmacology* 1985;24:587–591.

61. Crutcher MD, Alexander GE. Supplementary motor area (SMA): coding of both preparatory and movement-related neural activity in spatial rather than joint coordinates. *Soc Neurosci Abstr* 1988;14:342.

62. Crutcher MD, Alexander GE. Movement-related neuronal activity selectively coding either direction or muscle pattern in three motor areas of the monkey. *J Neurophysiol* 1990;64:151–163.

63. Crutcher MD, DeLong MR. Relation of putamen neuronal discharge to direction of movement or pattern of muscular activity. *Soc Neurosci Abstr* 1981;7:778.

64. Cummings JL. Depression and Parkinson's disease: a review [see comments]. *Am J Psychiatry* 1992;149:443–454.

65. Cummings JL. Frontal-subcortical circuits and human behavior. *Arch Neurol* 1993;50:873–880.

66. Cummings JL, Cunningham K. Obsessive-compulsive disorder in Huntington's disease. *Biol Psychiatry* 1992;31:263–270.

67. Cummings JL, Masterman DL. Depression in patients with Parkinson's disease. *Int J Geriatr Psychiatry* 1999;14:711–718.

68. Darian-Smith C, Darian-Smith I, Cheema SS. Thalamic projections to sensorimotor cortex in the macaque monkey: use of multiple retrograde fluorescent tracers. *J Comp Neurol* 1990;299:17–46.

69. Davis KD, Taub E, Houle S, et al. Globus pallidus stimulation activates the cortical motor system during alleviation of parkinsonian symptoms. *Nature Med* 1997;3:671–674.

70. De Michele G, Filla A, Volpe G, et al. Environmental and genetic risk factors in Parkinson's disease: a case-control study in southern Italy. *Mov Disord* 1996;11:17–23.

71. Deecke L. Cerebral potentials related to voluntary actions: parkinsonism and normal subjects. In: Delwaide PJ, Agnoli A, eds. *Clinical neurophysiology in parkinsonism*. Amsterdam: Elsevier, 1985:91–105.

72. Delfs JM, Anegawa NJ, Chesselet MF. Glutamate decarboxylase messenger RNA in rat pallidum: comparison of the effects of haloperidol, clozapine and combined haloperidol-scopolamine treatments. *Neuroscience* 1995;66:67–80.

73. DeLong MR. Activity of basal ganglia neurons during movement. *Brain Res* 1972;40:127–135.

74. DeLong MR. The neurophysiologic basis of abnormal movements in basal ganglia disorders. *Neurobehav Toxicol Teratol* 1983;5:611–616.

75. DeLong MR, Crutcher MD, Georgopoulos AP. Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. *J Neurosci* 1983;3:1599–1606.

76. Delwaide PJ, Pepin JL, Maertens de Noordhout A. Contribution of reticular nuclei to the pathophysiology of parkinsonian rigidity. *Adv Neurol* 1993;60:381–385.

77. Deniau JM, Hammond C, Riszk A, et al. Electrophysiological properties of identified output neurons of the rat substantia nigra (pars compacta and pars reticulata): evidences for the existence of branched neurons. *Exp Brain Res* 1978;32:409–422.

78. Deniau JM, Thierry AM. Anatomical segregation of information processing in the rat substantia nigra pars reticulata. *Adv Neurol* 1997;74:83–96.

79. Deschenes M, Bourassa J, Doan VD, et al. A single-cell study of the axonal projections arising from the posterior intralaminar thalamic nuclei in the rat. *Eur J Neurosci* 1996;8:329–343.

80. DeVito JL, Anderson ME. An autoradiographic study of efferent connections of the globus pallidus in Macaca mulatta. *Exp Brain Res* 1982;46:107–117.

81. Di Monte DA. Mitochondrial DNA and Parkinson's disease. *Neurology* 1991;41(suppl 2):38–42.

82. Dick JPR, Rothwell JC, Day BL, et al. The Bereitschaftspotential is abnormal in Parkinson's disease. *Brain* 1989;112:233–244.

83. Dogali M, Beric A, Sterio D, et al. et al. Anatomic and physiological considerations in pallidotomy for Parkinson's disease. *Stereotact Funct Neurosurg* 1994;62:53–60.

84. Dogali M, Beric A, Sterio D, et al. Anatomic and physiological considerations in pallidotomy for Parkinson's disease. *Acta Neurochir Suppl* 1995;64:9–12.

85. Dogali M, Fazzini E, Kolodny E, et al. Stereotactic ventral pallidotomy for Parkinson's disease. *Neurology* 1995;45:753–761.

86. Dube L, Smith AD, Bolam JP. Identification of synaptic terminals of thalamic or cortical origin in contact with distinct medium-size spiny neurons in the rat neostriatum. *J Comp Neurol* 1988;267:455–471.

87. Duvoisin RC. Role of genetics in the cause of Parkinson's disease. *Mov Disord* 1998;13(suppl 1):7–12.

88. Eidelberg D, Edwards C. Functional brain imaging of movement disorders. *Neurol Res* 2000;22:305–312.

89. Eidelberg D, Moeller JR, Antonini A, et al. The metabolic anatomy of Tourette's syndrome. *Neurology* 1997;48:927–934.

90. Eidelberg D, Moeller JR, Dhawan V, et al. The metabolic topography of parkinsonism. *J Cereb Blood Flow Metab* 1994;14:783–801.

91. Eidelberg D, Moeller JR, Kazumata K, et al. Metabolic correlates of pallidal neuronal activity in Parkinson's disease. *Brain* 1997;120:1315–1334.

92. Feger J, Bevan M, Crossman AR. The projections from the parafascicular thalamic nucleus to the subthalamic nucleus and the striatum arise from separate neuronal populations: a comparison with the corticostriatal and corticosubthalamic efferents in a retrograde double-labelling study. *Neuroscience* 1994;60:125–132.

93. Filion M, Boucher R, Bedard P. Globus pallidus unit activity in the monkey during the induction of parkinsonism by 1-methyl-4-phenyl-1,2,3,6,-tetrahydropyridine (MPTP). *Soc Neurosci Abstr* 1985;11:1160.

94. Filion M, Tremblay L. Abnormal spontaneous activity of globus pallidus neurons in monkeys with MPTP-induced parkinsonism. *Brain Res* 1991;547:142–151.

95. Filion M, Tremblay L, Bedard PJ. Abnormal influences of passive limb movement on the activity of globus pallidus neurons in parkinsonian monkeys. *Brain Res* 1988;444:165–176.

96. Findley LJ, Gresty MA, Halmagyi GM. Tremor, the cogwheel phenomenon and clonus in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1981;44:534–546.

97. Flaherty AW, Graybiel AM. Corticostriatal transformations in the primate somatosensory system. Projections from physiologically mapped body-part representations. *J Neurophysiol* 1991;66:1249–1263.

98. Flaherty AW, Graybiel AM. Two input systems for body representations in the primate striatal matrix: experimental evidence in the squirrel monkey. *J Neurosci* 1993;13:1120–1137.

99. Forno LS. Neuropathology of Parkinson's disease. *J Neuropathol Exp Neurol* 1996;55:259–272.

100. Forno LS, DeLanney LE, Irwin I, et al. Similarities and differences between MPTP-induced parkinsonism and Parkinson's disease. *Adv Neurol* 1993;60:600–608.

101. Futami T, Takakusaki K, Kitai ST. Glutamatergic and cholinergic inputs from the pedunculopontine tegmental nucleus to dopamine neurons in the substantia nigra pars compacta. *Neurosci Res* 1995;21:331–342.

102. Gasser T. Genetics of Parkinson's disease. *Ann Neurol* 1998;44:S53–57.

103. Georgopoulos AP, DeLong MR, Crutcher MD. Relations between parameters of step-tracking movements and single cell discharge in the globus pallidus and subthalamic nucleus of the behaving monkey. *J Neurosci* 1983;3:1586–1598.

104. Gerfen CR. Dopamine receptor function in the basal ganglia. *Clin Neuropharmacol* 1995;18:S162–S177.

105. Gerfen CR, Engber TM, Mahan LC, et al. D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* 1990;250:1429–1432.

106. Gerfen CR, Wilson CJ. The basal ganglia. In: Björklund A, Hökfeld T, Swanson L, eds. *Handbook of chemical neuroanatomy, integrated systems of the CNS, part III*. Amsterdam: Elsevier, 1996:369.

107. Giladi N, McMahon D, Przedborski S, et al. Motor blocks in Parkinson's disease. *Neurology* 1992;42:333–339.

108. Gill SS, Heywood P. Bilateral subthalamic nucleotomy can be accomplished safely. *Mov Disord* 1998;13:201.

109. Giller CA, Dewey RB, Ginsburg MI, et al. Stereotactic pallidotomy and thalamotomy using individual variations of anatomic landmarks for localization. *Neurosurgery* 1998;42:56–62.

110. Goldman PS, Nauta WJH. An intricately patterned prefronto-caudate projection in the rhesus monkey. *J Comp Neurol* 1977;171:369–386.

111. Goldman-Rakic PS, Porrino LJ. The primate mediodorsal (MD) nucleus and its projection to the frontal lobe. *J Comp Neurol* 1985;242:535–560.

112. Gorell JM, Rybicki BA, Johnson CC, et al. Smoking and Parkinson's disease: a dose-response relationship [see comments]. *Neurology* 1999;52:115–119.

113. Gowers WR. *Diseases of the nerves and spinal cord*, third ed. Philadelphia: P. Blakiston, 1900.

114. Grafton ST, Waters C, Sutton J, et al. Pallidotomy increases activity of motor association cortex in Parkinson's disease: a positron emission tomographic study. *Ann Neurol* 1995;37:776–783.

115. Graybiel AM. Building action repertoires: memory and learning functions of the basal ganglia. *Curr Opinion Neurobiol* 1995;5:733–741.

116. Graybiel AM. Basal ganglia: new therapeutic approaches to Parkinson's disease. *Curr Biol* 1996;6:368–371.

117. Grofova I, Zhou M. Nigral innervation of cholinergic and glutamatergic cells in the rat mesopontine tegmentum: light and electron microscopic anterograde tracing and immunohistochemical studies. *J Comp Neurol* 1998;395:359–379.

118. Gross RE, Lombardi WD, Hutchison WD, et al. Lesion location and outcome following pallidotomy: support for multiple output channels in the human pallidum. *Mov Disord* 1998;13(suppl 2):262.

119. Guridi J, Herrero MT, Luquin R, et al. Subthalamotomy improves MPTP-induced parkinsonism in monkeys. *Stereotact Funct Neurosurg* 1994;62:98–102.

120. Haas RH, Nasirian F, Nakano K, et al. Low platelet mitochondrial complex I and complex II/III activity in early untreated Parkinson's disease. *Ann Neurol* 1995;37:714–722.

121. Haber SN, Kunishio K, Mizobuchi M, et al. The orbital and medial prefrontal circuit through the primate basal ganglia. *J Neurosci* 1995;15:4851–4867.

122. Haber SN, Lynd-Balta E, Mitchell SJ. The organization of the descending ventral pallidal projections in the monkey. *J Comp Neurol* 1993;329:111–128.

123. Haber SN, Lynd-Balta E, Spooren WPJM. Integrative aspects of basal ganglia circuitry. In: Percheron G, McKenzie JS, Feger J, eds. *The basal ganglia IV. New ideas and data on structure and function*. New York: Plenum Press, 1994:71–80.

124. Hammond C, Rouzaire-Dubois B, Feger J, et al. Anatomical and electrophysiological studies on the reciprocal projections between the subthalamic nucleus and nucleus tegmenti pedunculopontinus in the rat. *Neuroscience* 1983;9:41–52.

125. Harnois C, Filion M. Pallidofugal projections to thalamus and midbrain: a quantitative antidromic activation study in monkeys and cats. *Exp Brain Res* 1982;47:277–285.

126. Hartmann-von Monakow K, Akert K, Kunzle H. Projections of the precentral motor cortex and other cortical areas of the frontal lobe to the subthalamic nucleus in the monkey. *Exp Brain Res* 1978;33:395–403.

127. Hedreen JC, DeLong MR. Organization of striatopallidal, striatonigral and nigrostriatal projections in the macaque. *J Comp Neurol* 1991;304:569–595.

128. Hellenbrand W, Seidler A, Robra BP, et al. Smoking and Parkinson's disease: a case-control study in Germany. *Int J Epidemiol* 1997;26:328–339.

129. Henselmans JM, de Jong BM, Pruim J, et al. Acute effects of thalamotomy and pallidotomy on regional cerebral metabolism, evaluated by PET. *Clin Neurol Neurosurg* 2000;102:84–90.

130. Hepp-Reymond MC, Husler EJ, Maier MA, et al. Force-related neuronal activity in two regions of the primate ventral premotor cortex. *Can J Physiol Pharmacol* 1994;72:571–579.

131. Herrero MT, Levy R, Ruberg M, et al. Glutamic acid decarboxylase mRNA expression in medial and lateral pallidal neurons in the MPTP-treated monkeys and patients with Parkinson's disease. *Adv Neurol* 1996;69:209–216.

132. Herrero MT, Levy R, Ruberg M, et al. Consequence of nigrostriatal denervation and L-dopa therapy on the expression of glutamic acid decarboxylase messenger RNA in the pallidum. *Neurology* 1996;47:219–224.

133. Hollander E, Cohen L, Richards M, et al. A pilot study of the neuropsychology of obsessive-compulsive disorder and Parkinson's disease: basal ganglia disorders. *J Neuropsychiatry Clin Neurosci* 1993;5:104–107.

134. Honig LS. Smoking and Parkinson's disease [letter; comment]. *Neurology* 1999;53:1158.

135. Hoover JE, Strick PL. Multiple output channels in the basal ganglia. *Science* 1993;259:819–821.

136. Hoover JE, Strick PL. The organization of cerebellar and basal ganglia outputs to primary motor cortex as revealed by retrograde transneuronal transport of herpes simplex virus type 1. *J Neurosci* 1999;19:1446–1463.

137. Hornykiewicz O, Kish SJ. Biochemical pathophysiology of Parkinson's disease. *Adv Neurol* 1987;45:19–34.

138. Hotson JR, Langston EB, Langston JW. Saccade responses to dopamine in human MPTP-treated parkinsonism. *Ann Neurol* 1986;20:456–463.

139. Hyland B, Chen DF, Maier V, et al. What is the role of the supplementary motor area in movement initiation? *Prog Brain Res* 1989;80:431–436; discussion 427–430.

140. Ilinsky IA, Jouandet ML, Goldman-Rakic PS. Organization of the nigrothalamocortical system in the rhesus monkey. *J Comp Neurol* 1985;236:315–330.

141. Ilinsky IA, Jouandet ML, Goldman-Rakic PS. Organization of the nigrothalamocortical system in the rhesus monkey. *J Comp Neurol* 1985;236:315–330.

142. Inase M, Buford JA, Anderson ME. Changes in the control of arm position, movement, and thalamic discharge during local inactivation in the globus pallidus of the monkey. *J Neurophysiol* 1996;75:1087–1104.

143. Inase M, Tanji J. Thalamic distribution of projection neurons to the primary motor cortex relative to afferent terminal fields from the globus pallidus in the macaque monkey. *J Comp Neurol* 1995;353:415–426.

144. Inglis WL, Winn P. The pedunculopontine tegmental nucleus: where the striatum meets the reticular formation. *Prog Neurobiol* 1995;47:1–29.

145. Jaeger D, Gilman S, Aldridge JW. Neuronal activity in the striatum and pallidum of primates related to the execution of externally cued reaching movements. *Brain Res* 1995;694:111–127.

146. Jaffer A, van der Spuy GD, Russell VA, et al. Activation of the subthalamic nucleus and pedunculopontine tegmentum: does it affect dopamine levels in the substantia nigra, nucleus accumbens and striatum? *Neurodegeneration* 1995;4:139–145.

147. Jenkins IH, Fernandez W, Playford ED, et al. Impaired activation of the supplementary motor area in Parkinson's disease is reversed when akinesia is treated with apomorphine. *Ann Neurol* 1992;32:749–757.

148. Jinnai K, Nambu A, Yoshida S, et al. The two separate neuron circuits through the basal ganglia concerning the preparatory or execution processes of motor control. In: Mamo N, Hamada I, DeLong MR, eds. *Role of the cerebellum and basal ganglia in voluntary movement*. Amsterdam: Elsevier Science, 1993:153–161.

149. Joel D, Weiner I. The connections of the primate subthalamic nucleus: indirect pathways and the open-interconnected scheme of basal ganglia-thalamocortical circuitry. *Brain Res Brain Res Rev* 1997;23:62–78.

150. Karmon B, Bergman H. Detection of neuronal periodic oscillations in the basal ganglia of normal and parkinsonian monkeys. *Isr J Med Sci* 1993;29:570–579.

151. Kato M, Miyashita N, Hikosaka O, et al. Eye movements in monkeys with local dopamine depletion in the caudate nucleus. I. Deficits in spontaneous saccades. *J Neurosci* 1995;15:912–927.

152. Kayahara T, Nakano K. Pallido-thalamo-motor cortical connections: an electron microscopic study in the macaque monkey. *Brain Res* 1996;706:337–342.

153. Kelton MC, Kahn HJ, Conrath CL, et al. The effects of nicotine on Parkinson's disease. *Brain Cogn* 2000;43:274–282.

154. Kemp JM, Powell TPS. 1971. The connections of the striatum and globus pallidus: synthesis and speculation. *Philos Trans R Soc Lond* 1971;262:441–457.

155. Kim R, Nakano K, Jayaraman A, et al. Projections of the globus pallidus and adjacent structures: an autoradiographic study in the monkey. *J Comp Neurol* 1976;169:263–290.

156. Kim U, Sanchez-Vives MV, McCormick DA. Functional dynamics of GABAergic inhibition in the thalamus. *Science* 1997;278:130–134.

157. Kish SJ, Shannak K, Hornykiewicz O. Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. *N Engl J Med* 1988;318:876–880.

158. Kita H, Kitai ST. Efferent projections of the subthalamic nucleus in the rat: light and electron microscope analysis with the PHA-L method. *J Comp Neurol* 1987;260:435–452.

159. Kita H, Tokuno H, Nambu A. Monkey globus pallidus external segment neurons projecting to the neostriatum. *Neuroreport* 1999;10:1467–1472.

160. Kitada T, Asakawa S, Hattori N, et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* 1998;392:605–608.

161. Kitai ST. Afferent control of substantia nigra compacta dopamine neurons: anatomical perspective and role of glutamatergic and cholinergic inputs. *Adv Pharmacol* 1998;42:700–702.

162. Kojima J, Yamaji Y, Matsumura M, et al. Excitotoxic lesions of the pedunculopontine tegmental nucleus produce contralateral hemiparkinsonism in the monkey. *Neurosci Lett* 1997;226:111–114.

163. Koller W, Vetere-Overfield B, Gray C, et al. Environmental risk factors in Parkinson's disease. *Neurology* 1990;40:1218–1221.

164. Kori A, Miyashita N, Kato M, et al. Eye movements in monkeys with local dopamine depletion in the caudate nucleus. II. Deficits in voluntary saccades. *J Neurosci* 1995;15:928–941.

165. Kosel S, Hofhaus G, Maassen A, et al. Role of mitochondria in Parkinson disease. *Biol Chem* 1999;380:865–870.

166. Kunishio K, Haber SN. Primate cingulostriatal projection: limbic striatal versus sensorimotor striatal input. *J Comp Neurol* 1994;350:337–356.

167. Kunzle H. Bilateral projections from precentral motor cortex to the putamen and other parts of the basal ganglia. An autoradiographic study in macaca fascicularis. *Brain Res* 1975;88:195–209.

168. Kunzle H. An autoradiographic analysis of the efferent connections from premotor and adjacent prefrontal regions (areas 6 and 9) in macaca fascicularis. *Brain Behav Evol* 1978;15:185–234.

169. Laitinen LV, Bergenheim AT, Hariz MI. Leksell's posteroventral pallidotomy in the treatment of Parkinson's disease. *J Neurosurg* 1992;76:53–61.

170. Langston JW. MPTP: the promise of a new neurotoxin. In: Marsden CD, Fahn S, eds. *Movement disorders*, vol 2. London: Butterworths, 1987:73–90.

171. Langston JW. Epidemiology versus genetics in Parkinson's disease: progress in resolving an age-old debate. *Ann Neurol* 1998;44:S45–52.

172. Langston JW, Ballard P. Chronic parkinsonism in humans due to a product of meperidine-analog synthesis. *Science* 1983;219:979–980.

173. Langston JW, Forno LS, Rebert CS, et al. Selective nigral toxicity after systemic administration of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyrine (MPTP) in the squirrel monkey. *Brain Res* 1984;292:390–394.

174. Lee RG, Murphy JT, Tatton WG. Long-latency myotatic reflexes in man: mechanisms, functional significance, and changes in patients with Parkinson's disease or hemiplegia. In: Desmedt JE, ed. *Motor control mechanisms in health and disease*. New York: Raven Press, 1983:489–507.

175. Lelli S, Panizza M, Hallett M. Spinal cord inhibitory mechanisms in Parkinson's disease. *Neurology* 1991;41:553–556.

176. LeMoine C, Normand E, Guitteny AF, et al. Dopamine receptor gene expression by enkephalin neurons in rat forebrain. *Proc Natl Acad Sci USA* 1990;182:611–612.

177. Lenz FA, Tasker RR, Kwan HC, et al. Single unit analysis of the human ventral thalamic nuclear group: correlation of thalamic "tremor cells" with the 3–6 Hz component of parkinsonian tremor. *J Neurosci* 1988;8:754–764.

178. Levy R Hazrati LN, Herrero MT, et al. Re-evaluation of the functional anatomy of the basal ganglia in normal and Parkinsonian states. *Neuroscience* 1997;76:335–343.

179. Lieberman DM, Corthesy ME, Cummins A, et al. Reversal of experimental parkinsonism by using selective chemical ablation of the medial globus pallidus. *J Neurosurg* 1999;90:928–934.

180. Llinas R, Jahnsen H. Electrophysiology of mammalian thalamic neurones in vitro. *Nature* 1982;29:406–408.

181. Lonser RR, Corthesy ME, Morrison PF, et al. Convection-enhanced selective excitotoxic ablation of the neurons of the globus pallidus internus for treatment of parkinsonism in nonhuman primates [see comments]. *J Neurosurg* 1999;91:294–302.

182. Lozano A, Hutchison W, Kiss Z, et al. Methods for microelectrode-guided posteroventral pallidotomy. *J Neurosurg* 1996;84:194–202.

183. Lozano AM, Lang AE, Galvez-Jimenez N, et al. Effect of GPi pallidotomy on motor function in Parkinson's disease. *Lancet* 1995;346:1383–1387.

184. Lueck CJ, Crawford TJ, Henderson L, et al. Saccadic eye movements in Parkinson's disease: II. Remembered saccades—towards a unified hypothesis? *Q J Exp Psychol [A] Hum Exp Psychol* 1992;45:211–233.

185. Lueck CJ, Tanyeri S, Crawford TJ, et al. Saccadic eye movements in Parkinson's disease: I. Delayed saccades. *Q J Exp Psychol [A] Hum Exp Psychol* 1992;45:193–210.

186. Lynd-Balta E, Haber SN. Primate striatonigral projections: a comparison of the sensorimotor-related striatum and the ventral striatum. *J Comp Neurol* 1994;345:562–578.

187. Magill PJ, Bolam JP, Bevan MD. Neuronal activity in the globus pallidus during cortical inactivation in normal and 6-OHDA treated rats *in vivo*. *Soc Neurosci* 2000;26:961.

188. Magill PJ, Bolam JP, Bevan MD. Relationship of activity in the subthalamic nucleus-globus pallidus network to cortical electroencephalogram. *J Neurosci* 2000;20:820–833.

189. Marin O, Smeets WJ, Gonzalez A. Evolution of the basal ganglia in tetrapods: a new perspective based on recent studies in amphibians. *Trends Neurosci* 1998;21:487–494.

190. Marsden CD, Obeso JA. The functions of the basal ganglia and the paradox of stereotaxic surgery in Parkinson's disease. *Brain* 1994;117:877–897.

191. Martin DL, Rimvall K. Regulation of gamma-aminobutyric acid synthesis in the rat brain. *J Neurochem* 1993;60:395–407.

192. Matsumine H, Saito M, Shimoda-Matsubayashi S, et al. Localization of a gene for an autosomal recessive form of juvenile parkinsonism to chromosome 6q25.2-27. *Am J Hum Genet* 1997;60:588–596.

193. Maurice N, Deniau JM, Glowinski J, et al. Relationships between the prefrontal cortex and the basal ganglia in the rat: physiology of the corticosubthalamic circuits. *J Neurosci* 1998;18:9539–9546.

194. Mayberg HS, Solomon DH. Depression in Parkinson's disease: a biochemical and organic viewpoint. *Adv Neurol* 1995;65:49–60.

195. Mayeux R, Stern Y, Williams J, et al. Clinical and biochemical features of depression in Parkinson's disease. *Am J Psychiatry* 1986;143:756–759.

196. McFarland NR, Haber SN. Convergent inputs from thalamic motor nuclei and frontal cortical areas to the dorsal striatum in the primate. *J Neurosci* 2000;20:3798–3813.

197. McNaught KS, Carrupt PA, Altomare C, et al. Isoquinoline derivatives as endogenous neurotoxins in the aetiology of Parkinson's disease. *Biochem Pharmacol* 1998;56:921–933.

198. Middleton FA, Strick PL. Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science* 1994;266:458–461.

199. Miller WC, DeLong MR. Altered tonic activity of neurons in the globus pallidus and subthalamic nucleus in the primate MPTP model of parkinsonism. In: Carpenter MB, Jayaraman A, eds. *The basal ganglia II*. New York: Plenum Press, 1987:415–427.

200. Mink JW, Thach WT. Basal ganglia motor control. III. Pallidal ablation: normal reaction time, muscle cocontraction, and slow movement. *J Neurophysiol* 1991;65:330–351.

201. Mitchell IJ, Cross AJ, Sambrook MA, et al. Neural mechanisms mediating 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridineinduced parkinsonism in the monkey: relative contributions of the striatopallidal and striatonigral pathways as suggested by 2-deoxyglucose uptake. *Neurosci Lett* 1986;63:61–66.

202. Mitchell SJ, Richardson RT, Baker FH, et al. The primate globus pallidus: neuronal activity related to direction of movement. *Exp Brain Res* 1987;68:491–505.

203. Mizuno Y, Ikebe S, Hattori N, et al. Role of mitochondria in the etiology and pathogenesis of Parkinson's disease. *Biochim Biophys Acta* 1995;1271:265–274.

204. Mogoseanu D, Smith AD, Bolam JP. Monosynaptic innervation of facial motoneurones by neurones of the parvicellular reticular formation. *Exp Brain Res* 1994;101:427–438.

205. Moriizumi T, Nakamura Y, Kitao Y, et al. Ultrastructural analyses of afferent terminals in the subthalamic nucleus of the cat with a combined degeneration and horseradish peroxidase tracing method. *J Comp Neurol* 1987;265:159–174.

206. Munro-Davies LE, Winter J, Aziz TZ, et al. The role of the pedunculopontine region in basal-ganglia mechanisms of akinesia. *Exp Brain Res* 1999;129:511–517.

207. Nagaseki Y, Shibazaki T, Hirai T, et al. Long-term follow-up results of selective VIM-thalamotomy. *J Neurosurg* 1986;65:296–302.

208. Nakano K, Hasegawa Y, Tokushige A, et al. Topographical projections from the thalamus, subthalamic nucleus and pedunculopontine tegmental nucleus to the striatum in the Japanese monkey, Macaca fuscata. *Brain Res* 1990;537:54–68.

209. Nambu A, Llinas R. Electrophysiology of the globus pallidus neurons: an in vitro study in guinea pig brain slices. *Soc Neurosci Abstr* 1990;16:428.

210. Nambu A, Takada M, Inase M, et al. Dual somatotopical representations in the primate subthalamic nucleus: evidence for ordered but reversed body-map transformations from the primary motor cortex and the supplementary motor area. *J Neurosci* 1996;16:2671–2683.

211. Nambu A, Tokuno H, Hamada I, et al. Excitatory cortical inputs to pallidal neurons via the subthalamic nucleus in the monkey. *J Neurophysiol* 2000;84:289–300.

212. Nambu A, Tokuno H, Inase M, et al. Corticosubthalamic input zones from forelimb representations of the dorsal and ventral divisions of the premotor cortex in the macaque monkey: comparison with the input zones from the primary motor cortex and the supplementary motor area. *Neurosci Lett* 1997;239:13–16.

213. Nambu A, Yoshida S, Jinnai K. Projection on the motor cortex of thalamic neurons with pallidal input in the monkey. *Exp Brain Res* 1988;71:658–662.

214. Narabayashi H, Ohye C. Parkinsonian tremor in nucleus ventralis intermedius of the human thalamus. In: Desmedt JE, ed. *Progress in clinical neurophysiology: physiological tremor, pathological tremors and clonus*. New York: S. Karger, 1978:165–172.

215. Narhi L, Wood SJ, Steavenson S, et al. Both familial Parkinson's disease mutations accelerate alpha-synuclein aggregation [published erratum appears in *J Biol Chem* 1999;274(19):13728]. *J Biol Chem* 1999;274:9843–9846.

216. Nauta HJ, Cole M. Efferent projections of the subthalamic nucleus: an autoradiographic study in monkey and cat. *J Comp Neurol* 1978;180:1–16.

217. Obeso JA, Rodriguez-Oroz MC, Rodriguez M, et al. Pathophysiology of levodopa-induced dyskinesias in Parkinson's disease: problems with the current model. *Ann Neurol* 2000;47:S22–32; discussion S32–34.

218. Obeso JA, Rothwell JC, Ceballos-Bauman A, et al. The mechanism of action of pallidotomy in Parkinson's disease (PD): physiological and imaging studies. *Soc Neurosci Abstr* 1995;21:1982.

219. Ohye C, Narabayashi H. Physiological study of presumed ventralis intermedius neurons in the human thalamus. *J Neurosurg* 1979;50:290–297.

220. Osenbach RK, Burchiel KJ. Thalamotomy: Indications, techniques, and results. In: Germano IM, ed. *Neurosurgical treatment of movement disorders*. Park Ridge, IL: American Association of Neurological Surgeons, 1998:l07–130.

221. O'Suilleabhain PE, Matsumoto JY. Fluctuations of instantaneous tremor frequency and synchrony between different extremities in Parkinson's disease and essential tremor. In: *Abstracts of tremor: basic mechanisms and clinical aspects*. Kiel, Germany: 1997:2.

222. Palombo E, Porrino LJ, Bankiewicz KS, et al. Local cerebral glucose utilization in monkeys with hemiparkinsonism induced by intracarotid infusion of the neurotoxin MPTP. *J Neurosci* 1990;10:860–869.

223. Papa SM, Desimone R, Fiorani M, et al. Internal globus pallidus discharge is nearly suppressed during levodopa-induced dyskinesias. *Ann Neurol* 1999;46:732–738.

224. Pare D, Curro'Dossi R, Steriade M. Neuronal basis of the parkinsonian resting tremor: a hypothesis and its implications for treatment. *Neuroscience* 1990;35:217–226.

225. Parent A. Extrinsic connections of the basal ganglia. *Trends Neurosci* 1990;13:254–258.

226. Parent A, Charara A, Pinault D. Single striatofugal axons arborizing in both pallidal segments and in the substantia nigra in primates. *Brain Res* 1995;698:280–284.

227. Parent A, DeBellefeuille L. Organization of efferent projections from the internal segment of globus pallidus in primate as revealed by fluorescence retrograde labeling method. *Brain Res* 1982;245:201–213.

228. Parent A, Hazrati L. Anatomical aspects of information processing in primate basal ganglia. *Trends Neurosci* 1993;16:111–116.

229. Parent A, Hazrati L-N. Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Res Rev* 1995;20:91–127.

230. Parent A, Smith Y. Organization of efferent projections of the subthalamic nucleus in the squirrel monkey as revealed by retrograde labeling methods. *Brain Res* 1987;436:296–310.

231. Parthasarathy HB, Graybiel AM. Cortically driven immediate-early gene expression reflects influence of sensorimotor cortex on identified striatal neurons in the squirrel. *J Neurosci* 1997;17:2477–2491.

232. Plenz D, Kitai S. A basal ganglia pacemaker formed by the subthalamic nucleus and external globus pallidus. *Nature* 1999;400:677–682.

233. Polymeropoulos MH, Higgins JJ, Golbe LI, et al. Mapping of a gene for Parkinson's disease to chromosome 4q21-q23 [see comments]. *Science* 1996;274:1197–1199.

234. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* 1997;276:2045–2047.

235. Rabey JM, Orlov E, Spiegelman R. Levodopa-induced dyskinesias are the main feature improved by contralateral pallidotomy in Parkinson's disease. *Neurology* 1995;45:A377.

236. Rajput AH, Uitti RJ, Stern W, et al. Early onset Parkinson's disease in Saskatchewan—environmental considerations for etiology. *Can J Neurol Sci* 1986;13:312–316.

237. Rajput AH, Uitti RJ, Stern W, et al. Geography, drinking water chemistry, pesticides and herbicides and the etiology of Parkinson's disease. *Can J Neurol Sci* 1987;14:414–418.

238. Rimvall K, Martin DL. The level of GAD67 protein is highly sensitive to small increases in intraneuronal gamma-aminobutyric acid levels. *J Neurochem* 1994;62:1375–1381.

239. Rodriguez MC, Obeso JA, Olanow CW. Subthalamic nucleus-mediated excitotoxicity in Parkinson's disease: a target for neuroprotection. *Ann Neurol* 1998;44:S175–188.

240. Ross GW, Abbott RD, Petrovitch H, et al. Association of coffee and caffeine intake with the risk of Parkinson disease. *JAMA* 2000;283:2674–2679.

241. Rouiller EM, Liang F, Babalian A, et al. Cerebellothalamocortical and pallidothalamocortical projections to the primary and supplementary motor cortical areas: a multiple tracing study in macaque monkeys. *J Comp Neurol* 1994;345:185–213.

242. Russchen FT, Bakst I, Amaral DG, et al. The amygdalostriatal projections in the monkey. An anterograde tracing study. *Brain Res* 1985;329:241–257.

243. Rye DB, Lee HJ, Saper CB, et al. Medullary and spinal efferents of the pedunculopontine tegmental nucleus and adjacent mesopontine tegmentum in the rat. *J Comp Neurol* 1988;269:315–341.

244. Sadikot AF, Parent A, Francois C. Efferent connections of the centromedian and parafascicular thalamic nuclei in the squirrel monkey: a PHA-L study of subcortical projections. *J Comp Neurol* 1992;315:137–159.

245. Sadikot AF, Parent A, Smith Y, et al. Efferent connections of the centromedian and parafascicular thalamic nuclei in the squirrel monkey: a light and electron microscopic study of the thalamostriatal projection in relation to striatal heterogeneity. *J Comp Neurol* 1992;320:228–242.

246. Samuel M, Ceballos-Baumann AO, Turjanski N, et al. Pallidotomy in Parkinson's disease increases supplementary motor area and prefrontal activation during performance of volitional movements an H2(15)O PET study. *Brain* 1997;120:1301–1313.

247. Schapira AH. Evidence for mitochondrial dysfunction in Parkinson's disease—a critical appraisal. *Mov Disord* 1994;9:125–138.

248. Schapira AH. Nuclear and mitochondrial genetics in Parkinson's disease. *J Med Genet* 1995;32:411–414.

249. Schell GR, Strick PL. The origin of thalamic inputs to the arcuate premotor and supplementary motor areas. *J Neurosci* 1984;4:539–560.

250. Schultz W. The phasic reward signal of primate dopamine neurons. *Adv Pharmacol* 1998;42:686–690.

251. Schultz W, Romo R. Role of primate basal ganglia and frontal cortex in the internal generation of movement. I. Preparatory activity in the anterior striatum. *Exp Brain Res* 1992;91:363–384.

252. Schwartzman RJ, Alexander GM. Changes in the local cerebral metabolic rate for glucose in the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) primate model of Parkinson's disease. *Brain Res* 1985;358:137–143.

253. Seidler A, Hellenbrand W, Robra BP, et al. Possible environmental, occupational, and other etiologic factors for Parkinson's disease: a case-control study in Germany. *Neurology* 1996;46:1275–1284.

254. Selemon LD, Goldman-Rakic PS. Longitudinal topography and interdigitation of cortico-striatal projections in the rhesus monkey. *J Neurosci* 1985;5:776–794.

255. Shibasaki H, Tsuji S, Kuroiwa Y. Oculomotor abnormalities in Parkinson's disease. *Arch Neurol* 1979;36:360–364.

256. Shink E, Bevan MD, Bolam JP, et al. The subthalamic nucleus and the external pallidum: two tightly interconnected structures that control the output of the basal ganglia in the monkey. *Neuroscience* 1996;73:335–357.

257. Shink E, Sidibe M, Smith Y. Efferent connections of the internal globus pallidus in the squirrel monkey: II. Topography and synaptic organization of pallidal efferents to the pedunculopontine nucleus. *J Comp Neurol* 1997;382:348–363.

258. Shink E, Smith Y. Differential synaptic innervation of neurons in the internal and external segments of the globus pallidus by the GABA- and glutamate-containing terminals in the squirrel monkey. *J Comp Neurol* 1995;358:119–141.

259. Sidibe M, Bevan MD, Bolam JP, et al. Efferent connections of the internal globus pallidus in the squirrel monkey: I. Topography and synaptic organization of the pallidothalamic projection. *J Comp Neurol* 1997;382:323–347.

260. Sidibe M, Smith Y. Differential synaptic innervation of striatofugal neurones projecting to the internal or external segments of the globus pallidus by thalamic afferents in the squirrel monkey. *J Comp Neurol* 1996;365:445–465.

261. Smith ID, Grace AA. Role of the subthalamic nucleus in the regulation of nigral dopamine neuron activity. *Synapse* 1992;12:287–303.

262. Smith Y, Bevan MD, Shink E, et al. Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience* 1998;86:353–387.

263. Smith Y, Charara A, Parent A. Synaptic innervation of midbrain dopaminergic neurons by glutamate-enriched terminals in the squirrel monkey. *J Comp Neurol* 1996;364:231–253.

264. Smith Y, Hazrati LN, Parent A. Efferent projections of the subthalamic nucleus in the squirrel monkey as studied by PHA-L anterograde tracing method. *J Comp Neurol* 1990;294:306–323.

265. Smith Y, Parent A. Differential connections of caudate nucleus and putamen in the squirrel monkey (Saimiri sciureus). *Neuroscience* 1986;18:347–371.

266. Smith Y, Shink E, Sidibe M. Neuronal circuitry and synaptic connectivity of the basal ganglia. *Neurosurg Clin North Am* 1998;9:203–222.

267. Soares J, Yamamoto BK, Kliem MA, et al. Neuronal discharge and gaba levels in the external pallidal segment (gpe) and the subthalamic nucleus (stn) in an mptp-treated primate, Meeting of the Society for Neuroscience, New Orleans, 2000.

268. Soghomonian JJ, Chesselet MF. Effects of nigrostriatal lesions on the levels of messenger RNAs encoding two isoforms of glutamate decarboxylase in the globus pallidus and entopeduncular nucleus of the rat. *Synapse* 1992;11:124–133.

269. Soghomonian JJ, Pedneault S, Audet G, et al. Increased glutamate decarboxylase mRNA levels in the striatum and pallidum of MPTP-treated primates. *J Neurosci* 1994;14:6256–6265.

270. Spacey SD, Wood NW. The genetics of Parkinson's disease. *Curr Opinion Neurol* 1999;12:427–432.

271. Spann BM, Grofova I. Origin of ascending and spinal pathways from the nucleus tegmenti pedunculopontinus in the rat. *J Comp Neurol* 1989;283:13–27.

272. Spellman GG. 1962. Report of familial cases of parkinsonism: evidence of a dominant trait in a patient's family. *JAMA* 1962;179:372–374.

273. Spillantini MG, Schmidt ML, Lee VM, et al. Alpha-synuclein in Lewy bodies [letter]. *Nature* 1997;388:839–840.

274. Spooren WP, Lynd-Balta E, Mitchell S, et al. Ventral pallidostriatal pathway in the monkey: evidence for modulation of basal ganglia circuits. *J Comp Neurol* 1996;370:295–312.

275. Starkstein SE, Mayberg HS. Depression in Parkinson's disease. In: Starkstein SE, Robinson RG, eds. *Depression in neurologic disease*. Baltimore: Johns Hopkins University Press, 1993:97–116.

276. Starr PA, Vitek JL, Bakay RAE. Pallidotomy: clinical results. In: Germano IM, ed. *Neurosurgical treatment of movement disorders*. Park Ridge, IL: American Association of Neurological Surgeons, 1998:143–156.

277. Steininger TL, Rye DB, Wainer BH. Afferent projections to the cholinergic pedunculopontine tegmental nucleus and adjacent midbrain extrapyramidal area in the albino rat. I. Retrograde tracing studies. *J Comp Neurol* 1992;321:515–543.

278. Steriade M, McCormick DA, Sejnowski TJ. Thalamocortical oscillations in the sleeping and aroused brain. *Science* 1993;262:679–685.

279. Stoessl AJ. Etiology of Parkinson's disease. *Can J Neurol Sci* 1999;26(suppl 2):S5–12.

280. Strick PL. How do the basal ganglia and cerebellum gain access to the cortical motor areas? *Behav Brain Res* 1985;18:107–123.

281. Sugimoto T, Hattori T, Mizuno N, et al. Direct projections from the centre median-parafascicular complex to the subthalamic nucleus in the cat and rat. *J Comp Neurol* 1983;214:209–216.

282. Surmeier DJ, Reiner A, Levine MS, et al. Are neostriatal dopamine receptors co-localized? [see comments]. *Trends Neurosci* 1993;16:299–305.

283. Surmeier DJ, Song WJ, Yan Z. Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. *J Neurosci* 1996;16:6579–6591.

284. Taha JM, Favre J, Baumann TK, et al. Tremor control after pallidotomy in patients with Parkinson's disease: correlation with microrecording findings. *J Neurosurg* 1997;86:642–647.

285. Takada M, Li ZK, Hattori T. Long descending direct projection from the basal ganglia to the spinal cord: a revival of the extrapyramidal concept. *Brain Res* 1987;436:129–135.

286. Takahashi H, Ohama E, Suzuki S, et al. Familial juvenile parkinsonism: clinical and pathologic study in a family. *Neurology* 1994;44:437–441.

287. Tanner CM, Goldman SM. Epidemiology of Parkinson's disease. *Neurol Clin* 1996;14:317–335.

288. Tanner CM, Langston JW, Ottman R, et al. Parkinson's Disease in the National Academy of Sciences/National Resource Council World War II Veteran Twins Registry. *Mov Disord* 1998;13(suppl 2):3.

289. Tanner CM, Ottman R, Goldman SM, et al. Parkinson disease in twins: an etiologic study [see comments]. *JAMA* 1999;281:341–346.

290. Tasker RR, Lang AE, Lozano AM. Pallidal and thalamic surgery for Parkinson's disease. *Exp Neurol* 1997;144:35–40.

291. Tatton WG, Lee RG. Evidence for abnormal long-loop reflexes in rigid Parkinsonian patients. *Brain Res Amsterdam* 1975;100:671–676.

292. Tsai CH, Chen RS, Lu CS. Reciprocal inhibition in Parkinson's disease. *Acta Neurol Scand* 1997;95:13–18.

293. Turner RS, Anderson ME. The relation of pallidal arm neuron discharge to the amplitude and velocity of movement. *Soc Neurosci Abstr* 1991;17:1218.

294. Turner RS, Anderson ME. Pallidal discharge related to the kinematics of reaching movements in two dimensions. *J Neurophysiol* 1997;77:1051–1074.

295. Turner RS, DeLong MR. Corticostriatal activity in primary motor cortex of the macaque. *J Neurosci* 2000;20:7096–7108.

296. Turner RS, Grafton ST, Votaw JR, et al. Motor subcircuits mediating the control of movement velocity: a PET study. *J Neurophysiol* 1998;80:2162–2176.

297. Van Der Kooy D, Hattori T. Single subthalamic nucleus neurons project to both the globus pallidus and substantia nigra in rat. *J Comp Neurol* 1980;192:751–768.

298. Vidailhet M, Rivaud S, Gouider-Khouja N, et al. Saccades and antisaccades in parkinsonian syndromes. *Adv Neurol* 1999;80:377–382.

299. Vitek JL, Ashe J, DeLong MR, et al. Altered somatosensory response properties of neurons in the "motor" thalamus of MPTP treated parkinsonian monkeys. *Soc Neurosci Abstr* 1990;16:425.

300. Vitek JL, Ashe J, DeLong MR, et al. Physiologic properties and somatotopic organization of the primate motor thalamus. *J Neurophysiol* 1994;71:1498–1513.

301. Vitek JL, Bakay RA, DeLong MR. Microelectrode-guided pallidotomy for medically intractable Parkinson's disease. *Adv Neurol* 1997;74:183–198.

302. Vitek JL, Kaneoke Y, Turner R, et al. Neuronal activity in the internal (GPi) and external (GPe) segments of the globus pallidus (GP) of parkinsonian patients is similar to that in the MPTP-treated primate model of parkinsonism. *Soc Neurosci Abstr* 1993;19:1584.

303. Vitek JL, Wichmann T, DeLong MR. Current concepts of basal ganglia neurophysiology with respect to tremorgenesis. In: Findley LJ, Koller W, eds. *Handbook of tremor disorders*. New York: Marcel Dekker, 1994:37–50.

304. von Krosigk M, Smith Y, Bolam JP, et al. Synaptic organization of gabaergic inputs from the striatum and the globus pallidus onto neurons in the substantia nigra and retrorubral field which project to the medullary reticular formation. *Neuroscience* 1993;50:531–549.

305. Wascher E, Verleger R, Vieregge P, et al. Responses to cued signals in Parkinson's disease. Distinguishing between disorders of cognition and of activation. *Brain* 1997;120:1355–1375.

306. Watts RL, Mandir AS. The role of motor cortex in the pathophysiology of voluntary movement deficits associated with parkinsonism. *Neurol Clin* 1992;10:451–469.

307. White OB, Saint-Cyr JA, Tomlinson RD, et al. Ocular motor deficits in Parkinson's disease. II. Control of the saccadic and smooth pursuit systems. *Brain* 1983;106:571–587.

308. Wichmann T, Baron MS, DeLong MR. Local inactivation of the sensorimotor territories of the internal segment of the globus pallidus and the subthalamic nucleus alleviates parkinsonian motor signs in MPTP treated monkeys. In: Percheron G, McKenzie JS, Feger J, eds. *The basal ganglia IV: new ideas and data on structure and function*. New York: Plenum Press, 1994:357–363.

309. Wichmann T, Bergman H, DeLong MR. The primate subthalamic nucleus. I. Functional properties in intact animals. *J Neurophysiol* 1994;72:494–506.

310. Wichmann T, Bergman H, DeLong MR. The primate subthalamic nucleus. III. Changes in motor behavior and neuronal activity in the internal pallidum induced by subthalamic inactivation in the MPTP model of parkinsonism. *J Neurophysiol* 1994;72:521–530.

311. Wichmann T, Bergman H, DeLong MR. Comparison of the effects of experimental parkinsonism on neuronal discharge in motor and non-motor portions of the basal ganglia output nuclei in primates. *Soc Neurosci Abstr* 1996;22:415.

312. Wichmann T, Bergman H, Starr PA, et al. Comparison of MPTP-induced changes in spontaneous neuronal discharge in the internal pallidal segment and in the substantia nigra pars reticulata in primates. *Exp Brain Res* 1999;125:397–409.

313. Wichmann T, DeLong MR. Functional and pathophysiological models of the basal ganglia. *Curr Opinion Neurobiol* 1996;6:751–758.

314. Wichmann T, DeLong MR, Vitek JL. Pathophysiologic considerations in basal ganglia surgery: role of the basal ganglia in hypokinetic and hyperkinetic movement disorders. *Prog Neurol Surg* 2000;15:6990–6994.

315. Wichmann T, Kliem MA, DeLong MR. Antiparkinsonian and behavioral effects of inactivation of the substantia nigra pars reticulata in hemiparkinsonian primates. *Exp Neurol* 2001;167:410–424.

316. Wiesendanger M, Ruegg DG, Lucier GE. Why transcortical reflexes? *Can J Neurol Sci* 1975;2:295–301.

317. Wilson CJ, Chang HT, Kitai ST. Origins of post synaptic potentials evoked in spiny neostriatal projection neurons by thalamic stimulation in the rat. *Exp Brain Res* 1983;51:217–226.

318. Wise SP, Murray EA, Gerfen CR. The frontal cortex-basal ganglia system in primates. *Crit Rev Neurobiol* 1996;10:317–356.

319. Wurtz RH, Hikosaka O. Role of the basal ganglia in the initiation of saccadic eye movements. *Prog Brain Res* 1986;64:175–190.

320. Yeterian EH, Pandya DN. Prefrontostriatal connections in relation to cortical architectonic organization in rhesus monkeys. *J Comp Neurol* 1991;312:43–67.

321. Yeterian EH, Pandya DN. Corticostriatal connections of the superior temporal region in rhesus monkeys. *J Comp Neurol* 1998;399:384–402.

322. Yeterian EH, VanHoesen GW. Cortico-striate projections in the rhesus monkey: the organization of certain cortico-caudate connections. *Brain Res* 1978;139:43–63.

123

Pathophysiology of Parkinson's Disease

Michael J. Zigmond

Robert E. Burke

Michael J. Zigmond: Departments of Neurology and Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania.

Robert E. Burke: Department of Neurology, Columbia University, New York, New York.

Parkinson's disease (PD) is thought to affect more than 1 million people in the United States alone, 1 of every 100 individuals above the age of 55. In the two centuries since it was first described by James Parkinson, we have learned a great deal about the disorder. We have, for example, learned where the primary lesion is and what many of the clinical manifestations are. However, it has only been in the past few decades that insights have begun to emerge regarding the cause of the disease, and only now can one begin to see the possibilities of treatments emerging that will provide more than temporary symptomatic relief. Beginning with the Nobel Prize–winning work of Arvid Carlsson, which pointed to the loss of dopamine (DA) as the principal deficit in PD and to levodopa as a mode of pharmacotherapy, we have come to understand what fails in this disorder and, more recently, how we might correct that failure. Moreover, although Parkinson focused entirely on motor symptoms, we have come to realize that the disorder is much more complex and includes a panoply of psychiatric symptoms as well.

The clinical features, course, and treatment of PD are presented in detail in Chapter 122 and Chapter 124 ; thus, for the purposes of this review of etiology and pathogenesis, we only briefly highlight the more important aspects of these topics, including those of a psychiatric nature. We then describe the pathology before turning to several promising leads with regard to the underlying etiology of the disorder.

CLINICAL SIGNS AND SYMPTOMS

Part of "123 - Pathophysiology of Parkinson's Disease "

Motor Manifestations

PD is a chronic, progressive neurologic disease. It presents with four cardinal motor manifestations: tremor at rest, rigidity, bradykinesia (or slowing of movement), and postural instability. Not all patients initially present with all of the classic signs of the disorder; there may be only one or two. Often, the first complaint is one of motor weakness or stiffness, and the cause is commonly misdiagnosed. However, postural deficits and tremor may soon emerge, prompting a reconsideration of the basis of the problem. It is important to note, however, that the clinical diagnosis of PD is made on the basis of a medical history and neurologic examination; there is currently no laboratory test that can definitely establish a diagnosis. Even neuroimaging, which can be used to obtain an estimate of DA loss (15 ,128), is imperfect and in any event is too expensive to be used as a routine diagnostic tool. As a result, it has been estimated that a significant number of individuals diagnosed as having PD fail to show the histopathologic hallmarks of the disease upon autopsy (48 ,70 ,134).

A *tremor* at rest is one of the most characteristic features of the disease, occurring in 70% of patients (68). Whereas it is not required for diagnosis, the prolonged absence of tremor in the course of a patient's illness should lead to the careful consideration of other neurologic conditions that can present with signs of parkinsonism, including the multiple system atrophies, progressive supranuclear palsy, corticobasal ganglionic degeneration, and others (94). *Rigidity* is a motor sign more often appreciated by the examining physician than the patient; it is detected as a resistance to passive movement of the limbs. It is often uniform in directions of flexion and extension (“lead pipe rigidity”), but there may be a superimposed ratcheting (“cogwheel rigidity”). *Bradykinesia* refers to a slowness and paucity of movement; examples include loss of facial expression, which may be misinterpreted as a loss of affect, and associated movements such as arm swinging when walking. Bradykinesia is not due to limb rigidity; it can be observed in the absence of rigidity during treatment. When bradykinesia affects the oropharynx, it can lead to difficulties in swallowing, which in turn may cause aspiration pneumonia, a potentially life-threatening complication. Of the cardinal motor signs, *postural*

instability is the most potentially dangerous, because it can lead to falls with resulting fractures. It is also one of the manifestations that responds less well to levodopa therapy. Tremor may also be less responsive, even early in the course (95).

An additional motor feature of PD is the freezing phenomenon, also referred to as “motor block” (51). In its most typical form, freezing occurs as a sudden inability to step forward while walking. It may occur at the beginning (“start hesitation”), at a turn, or just before reaching the destination. It is transient, lasting seconds or minutes, and suddenly abates. Combined with postural instability, it can be devastating. Freezing does not always improve with levodopa, and, in fact, can be made worse.

Cognitive and Psychiatric Manifestations

It is increasingly clear that there are many parallel circuits within the basal ganglia, each subserving a different function and each modulated by DA (see Chapter 122). Thus, it is reasonable to predict that patients will have a wide variety of dysfunctions extending well beyond the classic motor disabilities associated with the disease. Indeed, patients with Parkinson's disease appear to be at increased risk for a variety of cognitive and psychiatric dysfunctions. Most common is dementia and depression. However, hallucinations, delusions, irritability, apathy, and anxiety also have been reported (1). Here we will comment on the most prevalent of these symptoms.

Dementia is now recognized as one of the cardinal nonmotor manifestations of PD. It is a major cause of disability, and, unlike the motor manifestations, there currently is no effective symptomatic treatment. Aarsland and co-workers (2) identified in dementia in 28% of PD patients. The prevalence depends on age; in a study of PD patients over the age of 85 by Mayeux et al. (108), 65% were demented. PD patients with dementia show a more rapidly progressive course (110), and are more likely to be institutionalized, than nondemented individuals (2).

Years ago, there was debate about whether *depression* is a primary manifestation of PD or a reaction to having a chronic neurologic illness. There is now little question that it is a primary manifestation. Mayeux and colleagues (109a) have found that 47% of PD patients show evidence of depression, and some have found an even higher incidence (147). Moreover, Aarsland and colleagues (2) report that major depression is much more common among PD patients who also have signs of dementia (22%) than those who did not (2%). The depression, however, is not related to the severity of motor signs; indeed, many patients are depressed prior to the onset of frank neurologic dysfunction. Moreover, the depression is often greater than that seen in individuals with comparably debilitating motor dysfunction due to other disorders.

It has long been suggested that patients with PD can have particular premorbid *personality traits* (126 ,129 ,152). For example, some have argued that they tend to follow socially approved paths, are more introverted, and have less addictive personalities (e.g., are less inclined to smoke or drink) (158). Support for this hypothesis comes from a number of studies, including several involving twins that are discordant for Parkinson's disease (65 ,128). Many of these studies suffer from such problems as small sample size and retrospective analysis. Nonetheless, as noted at the outset of this section, the anatomy of basal ganglia circuitry is consistent with a broad range of functions, and some of these could easily affect personality in subtle ways. Might one, posit, for example, a “rigid PD personality” that parallels the rigid PD motor capacity? The issue is an important one, not only for our understanding of PD and of the neurobiology of behavior, but also because of the value of developing diagnostic screens that will permit the detection of PD at the earliest possible stage. Such early detection will become increasingly important as neuroprotective strategies emerge.

Other Manifestations

In addition to these neurologic signs and symptoms, PD patients often have disturbing sensory symptoms and pain in affected limbs. Many PD patients also have signs of autonomic failure, including orthostatic hypertension, constipation, urinary hesitancy, and impotence in men (90 ,107 ,133).

PATHOLOGY

Part of "123 - Pathophysiology of Parkinson's Disease "

Neuron Loss

A hallmark pathologic feature of PD, and essential for its pathologic diagnosis, is loss of DA neurons of the substantia nigra pars compacta (SNpc). At the time of death, even mildly affected PD patients have lost about 60% of their DA neurons, and it is this loss, in addition to possible dysfunction of the remaining neurons, that accounts for the approximately 80% loss of DA in the corpus striatum. We have discussed the basis for the need for extensive loss of DA neurons before the emergence of gross neurologic deficits in the previous edition of this series (5) and elsewhere (164 ,165). Briefly, our conception is as follows: As the terminals of DA neurons degenerate, there is a reduction in high-affinity DA uptake. This, coupled with some inherent redundancy in DA terminals and DA receptors, appears to permit striatal function to continue without disruption or active compensation during early phases of the neurodegenerative process. After somewhat larger lesions, the remaining DA terminals appear to increase the amount of transmitter synthesized and delivered to the extracellular fluid. This seems to be due at least in part to a net increase in the amount of DA released in response to terminal depolarization, a consequence of the transient disruption of the homeostatic regulatory systems that exist within the affected

systems (164). Once released, a portion of the DA appears to diffuse out of the synapse and into the extracellular space, where its actions are prolonged due to the relative absence of high-affinity DA uptake sites. We hypothesize that these events permit the SNpc to continue to exert dopaminergic control over striatal cell function as long as some minimal number of DA terminals remain. However, the increased synthesis and release of DA may increase reactive metabolites formed from DA and thus contribute to the progression of the disease (see below).

Neurologic deficits emerge when the availability of DA falls below the level required for rapid compensation or when the system is subjected to certain pharmacologic, environmental, or physiologic challenges. These can subside if additional, slowly developing compensations, such as the synthesis and insertion of additional DA receptors, the induction of typrosine hydroxylase (TH) synthesis, sprouting, or regeneration, occur at a more rapid rate than does the underlying neurodegeneration. This has been observed to be the case in animal models, wherein recovery often occurs after the abrupt loss of even 90% of striatal DA, as occurs in most animal models. In patients, however, where the degenerative process is typically progressive, such recovery would not be expected to occur spontaneously. However, an important implication of the ability of brain to compensate for partial loss of DA neurons in these ways is that once deficits do occur, the task of medicine need not be to restore the entire nigrostriatal projection but only the far less daunting objective of returning the availability of DA to the level required to attain the preclinical state.

Although there are several groups of dopaminergic neurons in the central nervous system (CNS), it is the loss of DA cells in the SNpc that is believed to account for all of the motor manifestations of PD. This clinicopathologic correlation is supported by observations that the *N*-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) is selective for DA neurons of the SNpc in humans and primates, and yet can produce the full spectrum of motor signs seen in PD (96). Moreover, it is not even all of the SNpc DA neurons that appear to be involved in PD. The ventral-lateral tier is more severely affected than the dorsal tier (36), and this accounts for a more severe loss of DA in the putamen, where dorsally the loss can be as much as 95%, as compared to the caudate, where ventrally the loss can be as little as 60% (87). Some central dopaminergic systems, such as the ventral tegmental area and hypothalamic systems, are relatively spared, and descending spinal dopaminergic systems are spared entirely (6).

Although some DA neurons are spared in PD, it is also the case that neuron loss is not restricted to the dopaminergic neurons. Other catecholaminergic cell groups including the locus coeruleus are involved, as are some cells of the sympathoadrenal system and the serotoninergic neurons of the raphe nuclei (6). There is also loss of cholinergic neurons of the nucleus basalis of Meynert, and this may be responsible, at least in part, in some cases, for dementia (159).

Lewy Bodies

Another pathologic hallmark of PD is the Lewy body, an eosinophilic inclusion identified within neurons. On histologic stains, Lewy bodies have an eosinophilic core, and a surrounding pale halo. They are usually rounded, although their shape can be pleiomorphic (50), and they are generally 5 to 25 μm in diameter. They usually are observed within the cell soma, but also can be seen in neurites or free in the extracellular space. Lewy bodies are commonly observed in the brain regions showing the most neuron loss in PD, including SN, locus coeruleus, the dorsal motor nucleus of the vagus, and the nucleus basalis of Meynert, but they are also observed in neocortex, diencephalon, spinal cord, and even peripheral autonomic ganglia (50).

On ultrastructural analyses, Lewy bodies consist of an electron dense granular core and a peripheral halo consisting of radially oriented filaments 7 to 8 nm in width (28). The filaments resemble neurofilaments, and can be immunostained with antisera to neurofilament proteins (53), including the NF-L, -M, and -H forms (67). Immunostaining can be achieved with antibodies that recognize phosphorylated as well as nonphosphorylated epitopes (13 ,40). The cellular kinases responsible for the phosphorylation are not known. However, two candidates, Ca^{2+}/calmodulin–dependent protein kinase II (75) and cyclin-dependent kinase 5 (14 ,98 ,118), have both been immunolocalized to Lewy bodies.

Another major antigenic feature of Lewy bodies is the expression of cellular proteins involved in protein degradation, including ubiquitin (93), and the proteasome (37 ,71). Presence of these antigens has been hypothesized to represent efforts on the part of the cell to degrade the abnormal protein aggregate.

Following the identification of mutations in the α-synuclein gene in a few cases of familial PD (see below), it was discovered that α-synuclein is a component of Lewy bodies (11 ,142 ,143). α-Synuclein in Lewy bodies can be labeled with antibodies to the C- or N-terminal, suggesting that full-length molecules are present (142). α-Synuclein immunostaining is identified in isolated Lewy filaments, and the pattern of staining suggests a polar orientation of the molecules (142). Staining of filaments *in situ* has been confirmed by immunoelectron microscopy (10). Synphilin, a protein known to interact with α-synuclein (33), also has been identified within Lewy bodies in PD (156).

α-Synuclein is identifiable not only in Lewy bodies of PD, but also the Lewy bodies of Hallervorden–Spatz syndrome (8), familial Alzheimer's disease (99), sporadic Alzheimer's (55), Alzheimer's associated with Down syndrome (100), and diffuse Lewy body disease (10 ,11). In Alzheimer's disease, Lewy bodies demonstrated by immunostaining for α-synuclein are most often observed in the amygdala.

The demonstrations of Lewy bodies in Alzheimer patients add to the growing evidence of important pathologic overlaps between Alzheimer's and PD. This evidence is further supported by the demonstration of co-localization of phosphorylated > and α-synuclein in Lewy bodies of PD and diffuse Lewy body disease (9).

In addition to its localization in Lewy bodies in PD, abnormal α-synuclein immunostaining has been identified in axon terminals in the hippocampal dentate, hilar, and CA 2/3 regions in PD (44). Whereas immunostaining for ß-synuclein has not been observed in Lewy bodies, staining was observed in these axon terminals. In addition, although immunostaining for γ-synuclein is not present in Lewy bodies, it is observed within axonal spheroids in the hippocampal dentate molecular layer in PD (44).

ETIOLOGIC FACTORS

Part of "123 - Pathophysiology of Parkinson's Disease "

Aging

The possible role of aging in the pathogenesis of PD is suggested by its usual occurrence in late middle age, and by marked increases in its prevalence at older ages (109). The possible contribution of age to the expression of the disease is further supported by early studies showing a loss with age of striatal DA (18) and DA of cells in the SN (113). However, whereas the gradual loss of striatal dopaminergic markers (88 ,138) and SNpc neurons (36) with age has recently been confirmed, the pattern and timing of these losses differ from what occurs in PD, indicating that aging itself is not likely to play a direct role in the degenerative process. For example, although the number of dopaminergic terminals appears to decrease with age, this takes place with a different temporal and spatial pattern than occurs in PD (138). The loss of SN neurons in aging is linear and predominantly in the dorsal tier of the SNpc, whereas in PD it is exponential and predominantly in the lateral ventral tier (36 ,138). In addition, the SN in PD contains numerous reactive microglia, which are much less frequent in age-matched control brains, indicating an active destructive process that is not present in the normal aged brain (111 ,112). Thus, whereas there is no question that increased age is a risk factor for PD, it remains unclear what precise role aging plays in pathogenesis.

Environmental Factors

Consideration of a role for environmental factors in the cause of PD was given major impetus with the discovery in 1983 that exposure to MPTP is capable of inducing parkinsonism in humans (96). The role of environmental factors was given additional weight by initial results of twin studies, as discussed below, which initially appeared to exclude any important role for genetic factors. The possible role of environmental factors has been addressed by a number of epidemiologic studies that have been well reviewed by others (97 ,148). Many of these studies have shown associations between rural residence, well-water drinking, or herbicide/pesticide exposure and the risk of developing PD (148). However, the precise role played by any specific compounds has remained elusive.

Genetic Factors

For many years, genetic factors were considered unlikely to play an important role in the pathogenesis of PD. This concept was based largely on twin studies conducted in the early 1980s that demonstrated a very low rate of concordance for the disease among identical twins (157) [reviewed by Duvoisin (29)]. Nevertheless, many investigators recognized that PD could occasionally be identified in families (52). The most important advances in PD research in recent years have been the identification of specific disease-causing mutations, making it possible for the first time to begin to explore pathogenesis at the molecular level. For this review, we focus on the best documented and most widely investigated genetic causes—those in α-synuclein and parkin.

Synuclein

After mapping a disease-causing gene locus to the 4q21-q23 region (130) in a large Italian kindred (52), Polymeropoulos and co-workers (131) identified a base pair change from G to A at position 209, which resulted in an Ala to Thr substitution at position 53 in α-synuclein in this family and three small Greek kindreds. Whereas initially there was a question as to whether this may represent a benign polymorphism, that possibility was soon dispelled by the discovery of a second disease-causing mutation, an Ala to Pro substitution at position 30, in an unrelated German kindred (92). The likely role of α-synuclein in the pathogenesis of PD was further supported by the discovery of αsynuclein in Lewy bodies of sporadic PD cases, as outlined above. One of the important aspects of the discovery of these mutations in α-synuclein was that they immediately suggested a possible pathogenetic mechanism, that of protein aggregation, because α-synuclein had been identified in Alzheimer plaques (154), and a central portion of α-synuclein had been shown to have the capacity to self-aggregate (56). The role of α-synuclein in the protein aggregation hypothesis for PD is discussed in the next section.

Little is known about the normal physiologic function of α-synuclein. Scheller's group (104) was the first to discover the compound, identifying it in *Torpedo* as a brain-specific synaptic terminal protein. These investigators subsequently demonstrated that the rat protein homologue is likewise expressed in nerve terminals (105). α-Synuclein messenger RNA (mRNA) is predominantly expressed in forebrain structures, such as hippocampus and cortex, but also in a few specific midbrain-brainstem nuclei, including

the SNpc, locus coeruleus, and dorsal motor nucleus (105), which, interestingly, are typically involved in PD. The human homologue of α-synuclein was independently identified by Ueda et al. (154) and Jakes et al. (78), and demonstrated by Irizarry and colleagues (76) to be expressed in nerve terminals, as is the case for rat. These investigators also showed by fractionation studies that α-synuclein appears to be loosely associated with synaptic vesicles, and this localization has been confirmed in rat brain by ultrastructural analysis (74). Jensen and his colleagues (82) have shown that α-synuclein binds to vesicles via its amino-terminal region, and that it is carried with vesicles by the fast component of axonal transport. Interestingly, the A30P mutation appears to abolish vesicle-binding activity.

What specific physiologic role α-synuclein and its homologues may play as vesicle-binding proteins remains a mystery. George and co-investigators (46) independently identified an avian homologue of α-synuclein, synelfin, as a gene upregulated in the song control circuit during a critical period of song learning, and suggested that it plays a role in neural plasticity (46). We have shown that α-synuclein mRNA and protein are upregulated in the SNpc following early developmental striatal target lesion (85). This lesion results in the induction of apoptotic death in some, but not all, developing dopaminergic neurons (102). However, in this model α-synuclein is not expressed in apoptotic profiles; it is exclusively upregulated in normal-appearing neurons, suggesting that it plays a role either in maintaining their viability, or, alternatively, in plastic change after viability is established. Vila and co-workers (155) reached a similar conclusion in a model of chronic MPTP toxicity. In support of the possibility that α-synuclein may play a role in a plasticity response in these injury models is the observation that α-synuclein mRNA in SN is up-regulated during the first 4 postnatal weeks, a period of maximal differentiation and synaptogenesis among DA neurons (85). What precise role α-synuclein plays in the development of DA neurons remains to be established. Remarkably, homozygous α-synuclein null mice have thus far shown no obvious abnormalities in numbers or morphology of DA neurons; density of striatal dopaminergic terminals; the number, morphology, or patch/matrix distribution of striatal neurons; or the ultrastructural appearance of striatal synaptic terminals (4). These animals, however, do exhibit an increased release of DA in a paired stimulus depression paradigm. In addition, they show diminished behavioral activation following administration of amphetamine (4).

In view of α-synuclein's ability to self-aggregate, there has been a tendency to assume that the mutations cause a toxic gain of function related to aggregation. It is important to keep in mind, however, that its function is unknown, and that a loss of function may relate to disease pathogenesis. In that regard, we have found that α-synuclein mRNA levels are diminished in the SNpc of patients with sporadic PD (120). Markopoulou and colleagues (103) have shown in a large Greek family with the G209A mutation that there is diminished expression of the mutant allele in lymphoblastoid cell lines, and they suggest that the parkinsonian phenotype may arise from haploinsufficiency.

Parkin

Mutations in the parkin gene were first identified in Japanese families with a unique variant of parkinsonism (89). This form is inherited in an autosomal-recessive pattern, and typically begins at an early age; in the series of 17 patients studied by Ishikawa and Tsuji (72), the age ranged from 9 to 43 years, with a mean of 28. Many of the clinical features of patients with autosomal-recessive juvenile parkinsonism (ARJP) closely resemble those of idiopathic PD: tremor at rest, rigidity, bradykinesia, postural instability, gait freezing, and marked improvement with levodopa. However, there are differences between ARJP and idiopathic PD in addition to the age at onset. Patients with ARJP more often present with dystonia, show a marked clinical improvement after sleep, and often show hyperreflexia (72). In general, they do not show cognitive decline or autonomic failure, and the course is slowly progressive. The motor predominance of their clinical signs is in keeping with the pathologic findings, which indicate neuronal loss restricted to neurons of the SNpc and the locus coeruleus (146). Lewy bodies are not observed (146).

ARJP was found to map to the chromosome 6q25.2-27 region, and a marker for this region, D6S305, was found to be deleted in a single Japanese patient (89). Screening of complementary DNA (cDNA) libraries with a probe for a putative exon, which was also deleted in this patient, led to the identification of a sequence encoding an open reading frame for a 465 amino acid protein (89). The deduced amino acid sequence of this protein contains a ubiquitin homology domain at the N-terminal, and a ring-finger motif at the C-terminal. The gene encoding the protein is large [>500 kilobase (kb)], and contains 12 exons. Deletion mutations were identified in four other affected patients in three independent families, confirming the pathogenetic significance. A 4.5-kb mRNA transcript was identified in many human tissues, including brain. In brain, it is expressed in various regions, including the SN (89).

Subsequent molecular genetic analysis of 34 affected individuals from 18 unrelated Japanese families revealed four additional deletional mutations (64), bringing the total to six identified at that time (89). The deletions affected exon 3, exon 4, and exons 3 to 4, and a 1–base pair (bp) deletion in exon 5 resulted in a frameshift and an early stop. Further molecular analysis of non-Japanese families in Europe, revealed that in addition to deletion mutations, a variety of point mutations resulting in either truncation or missense could also cause the phenotype (3). In addition, this study identified patients with a late age of onset, up to 58 years

in one case, and indicated that in some instances the clinical phenotype was indistinguishable from idiopathic PD (3).

There is now growing recognition that mutations in parkin may cause what clinically resembles idiopathic PD. In an investigation of the scope of the molecular and clinical features in Europe, Lucking and co-workers (101) found that among 73 families with early onset (<45 years) of parkinsonism and affected family members, 49% had parkin mutations. Among early-onset patients without affected family members, 18% had mutations. The majority (77%) of these were younger than 20 years of age. Many of the patients with parkin mutations lacked the signs thought to be characteristic of ARJP such as dystonia and hyperreflexia, and were clinically difficult to distinguish from idiopathic PD. In all, 19 different rearrangements of exons mutations were identified, including multiplications as well as deletions, and there were 16 different point mutations (101).

The neurobiology of parkin is only beginning to be explored. By immunohistochemistry, the protein has been localized at the regional level to SN and locus coeruleus, and at the cellular level to the cytoplasm (139). Nuclear staining was not observed. Parkin has been shown to play a role in protein degradation as a ubiquitin-protein ligase (140). These findings suggest that abnormal accumulation of proteins or abnormal regulation of the half-life of normal cellular proteins may play a role in cell death.

PATHOGENETIC MECHANISMS

Part of "123 - Pathophysiology of Parkinson's Disease "

Free Radicals and Deficits in Energy Metabolism

The concept that free radical–mediated injury may underlie the neuronal degeneration that occurs in PD has been, and continues to be, the leading hypothesis for its pathogenesis. The free radical theory has been the subject of many excellent reviews (34 ,122), so it will be outlined here only briefly. This theory is also referred to as the oxidant stress hypothesis or the endogenous toxin hypothesis. In their review, Fahn and Cohen (34) point out that the free radical hypothesis is appealing because four aspects of the neurochemistry of DA neurons and their local environment within the SN make the concept plausible. First, a major degradative pathway for DA is its oxidative deamination by monoamine oxidases A and B. This process results in the enzymatic production of H_2O_2, which, while itself not a free radical, can nevertheless react nonenzymatically with ferrous or cupric ions via Fenton-type reactions to form highly reactive hydroxyl radicals. Second, DA can react nonenzymatically with oxygen to form quinones and semiquinones, with the production of superoxide, hydrogen peroxide, and hydroxyl radicals. Third, the SN, particularly the SN pars reticulata, is rich in iron, which as mentioned above, may in its ferrous state catalyze the formation of hydroxyl radicals from H_2O_2. Fourth, the SN contains neuromelanin formed from the auto-oxidation of DA. This auto-oxidation generates toxic quinones and reactive oxygen species. In addition, the presence of neuromelanin in the cell may alter the ability of metal ions to participate in the production of reactive oxygen species (145).

The possibility that DA neurons may undergo free radical–mediated injury in PD has received support from animal studies using one of two neurotoxins that can be used to selectively destroy DA neurons—6-hydroxydopamine (6-OHDA) and MPTP. 6-OHDA reacts with oxygen to produce superoxide anion radical, H_2O_2, and hydroxyl radical. It is a general cytotoxin but derives its specificity by virtue of its affinity for the high-affinity catecholamine transporters. Thus, when used in sufficiently low concentrations, its actions can be directed toward catecholamine neurons. Moreover, it can be limited to acting on DA neurons by pretreating animals with an inhibitor of high-affinity norepinephrine uptake, such as desipramine (19 ,66 ,166).

Interestingly, like 6-OHDA, DA itself is a selective neurotoxin for DA neurons (38 ,54 ,61 ,62 ,114 ,135). This seems to be in large part due to its ability to oxidize to form reactive oxygen species, including DA quinone, which has a high affinity for the cysteinyl residues on proteins (41 ,42 and 43 ,54 ,63). Thus, it seems possible that DA itself can be a source of oxidative stress, particularly under conditions of increased DA turnover and decreased antioxidant defenses (see below).

MPTP acts via its active product MPP^+, which is selectively taken up into DA neurons via the DA transporter, and inhibits complex I activity in mitochondria. Inhibition of complex I not only interferes with adenosine triphosphate (ATP) synthesis, but also results in augmented production of superoxide anion radical. The possible role of superoxide radical in MPTP toxicity has received direct support by the demonstration by Przedborski and co-workers (132) that transgenic mice with high Cu/Zn superoxide dismutase activity are resistant to MPTP.

The free radical hypothesis of PD has also received support from studies of human postmortem brain. Free radicals can cause injury to cells by damaging DNA, proteins, and lipids of the cell membrane. There is evidence from postmortem studies for free radical–induced modification of each of these classes of molecules. Dexter and co-workers (24) have shown that in PD brain there is a reduction in levels of polyunsaturated fatty acids, which provide an index of the amount of substrate available for lipid peroxidation, and an increase in levels of malondialdehyde, an intermediate in the lipid peroxidation process. The increase in malondialdehyde was regionally specific for the SN. These workers subsequently confirmed evidence for abnormal lipid peroxidation in PD by identifying a tenfold increase in cholesterol lipid hydroperoxide, an early marker in the lipid peroxidation process (25). Free radicals are also capable of directly damaging DNA. Sanchez-Ramos and colleagues (136) have shown that regional concentrations of 8-hydroxy-deoxyguanosine,

an index of oxyradical-mediated DNA damage, are increased in the caudate and SN of PD patients. Relatively less attention has been given to the possibilities of oxygen-mediated damage to proteins, or of advanced glycosylation changes to proteins in PD (141). The possibility that such protein changes may also occur in PD brain is supported by the demonstration that protein adducts of 4-hydroxy-2-nonenal, a cytotoxic product of lipid peroxidation, can be identified by immunohistochemistry in many nigral neurons of PD patients in contrast to age-matched controls (162).

Postmortem studies have also revealed neurochemical features that may predispose the PD brain to oxidative damage. Reduced glutathione is an important endogenous antioxidant, and it has been reported to be reduced in the SN in PD (127). Jenner and colleagues (80) have confirmed low levels of reduced glutathione in the SN of PD patients, and have shown that the alteration is disease-specific. Interestingly, they have also shown that reductions are observed in patients with incidental Lewy body disease, which may be a preclinical form of PD (49). This finding suggests that the reduced levels of glutathione may be a fundamental and primary abnormality in PD, rather than a secondary change.

A number of postmortem studies have also suggested that abnormalities of iron metabolism may underlie the neurodegeneration of PD. Iron metabolism is of particular interest in relation to the free radical hypothesis because, as noted above, it normally is found in high concentrations in SN, and is capable of catalyzing free radical formation. Dexter and colleagues (26) reported increased levels of iron in the SNpc of PD patients. This observation took on potentially greater significance when this group subsequently reported decreased levels of ferritin in PD brains (23), as ferritin normally sequesters iron in an unreactive state. However, it has become apparent that increased iron levels may be observed in many brain regions demonstrating neural degeneration in a variety of diseases of the basal ganglia (22), so the specificity of changes in iron levels in PD is less clear. Nevertheless, the possible relationship of altered iron metabolism to the pathogenesis of PD remains of interest, based on the finding of a higher density of lactoferrin receptors on neurons and microvessels of patients with PD (35). This finding suggests that lactoferrin receptors, which regulate intraneuron iron content, may be overly expressed in vulnerable dopaminergic neurons in PD.

Another postmortem finding in PD patients that is compatible with the free radical hypothesis is that of a deficiency in mitochondrial complex I. Such a defect could either result in the abnormal production of free radicals, or be the result of free radical injury (137). This defect takes on particular interest in light of the observation that MPP^+, the toxic oxidative product of MPTP, inhibits complex I (121). The defect in complex I in PD patients has been demonstrated by Schapira et al. (137) to result in a mean 37% decrease in activity. This decrease appears to be both regionally specific for the SN, and disease specific, among basal ganglia disorders, for PD.

Thus, the free radical hypothesis receives indirect support from a large number of separate lines of evidence, and, as stated above, it remains the foremost and most widely tested hypothesis of neural degeneration in PD. Nevertheless, it remains only a hypothesis, and it has its shortcomings (17). For example, there is no specific aspect of the free radical hypothesis as it is currently posed to account for the relative vulnerability of ventral tier dopaminergic neurons in PD. In addition, it must be remembered that nonaminergic neuronal groups, such as the nucleus basalis, which is cholinergic, also degenerate in PD, and aspects of the free radical hypothesis that are dependent on catecholamine metabolism are not relevant to the degeneration of these structures.

Programmed Cell Death

The concept that a genetically regulated cell death process may underlie the neuron-specific degenerations of later life has gathered great attention in recent years. The programmed cell death hypothesis in fact may be related to the concept of free radical–mediated cell death. Although traditional concepts of free radical injury have centered on the ability of toxic molecules to directly injure cellular constituents without any participation of the host cell's own genetic programs, it is now clear that in both *in vitro* and *in vivo* paradigms of free radical–mediated injury programmed cell death may occur. It is also apparent that in some settings programmed cell death may be carried out by the controlled production of free radicals.

In relation to PD, it is important first to consider the evidence that programmed cell death does in fact occur within dopaminergic neurons. As predicted by classic neurotrophic theory, some natural cell event does occur during development in the SNpc, with typical light microscopic morphology of apoptosis, demonstrated both by Nissl stain and suppressed silver staining (79), and we used a double-labeling technique to identify apoptotic natural cell death in phenotypically defined dopaminergic neurons (123). Natural cell death in these neurons has a bimodal time course. There is an initial, major peak that begins on embryonic day 20, and largely abates by the eighth postnatal day (PND). There is a second, minor peak of natural cell death on PND 14. The presence of a postnatal cell death event is in keeping with the demonstration by Tepper and colleagues (151) that there is a decrement in the number of TH-positive neurons in SN postnatally, particularly in the first postnatal week. Although there is evidence that the magnitude of the natural cell death event in DA neurons is regulated by interactions with the target striatum (see below), as classic neurotrophic theory would predict, it remains unknown which neurotrophic factors are involved. We have shown *in vitro* in postnatal primary cultures of DA neurons that GDNF is uniquely able to support viability by

suppressing spontaneous apoptotic death (16) as well as cell death initiated by 6-OHDA (see below). Whether glial cell line-derived neurotrophic factor (GDNF) plays such a role *in vivo* remains to be determined.

We have shown that natural cell death in SNpc can be regulated during development by striatal target interactions. Excitotoxic injury to the striatum on PND 7 results in an eightfold increase in the number of apoptotic profiles (102). These profiles are morphologically identical to those observed during natural cell death, and meet ultrastructural and 3′ end-labeling criteria for apoptosis. Within SNpc, induction of cell death is identified within phenotypically defined dopaminergic neurons.

Programmed cell death also occurs in DA neurons in animal models of parkinsonism. Intrastriatal injection of 6-OHDA results in an induction of apoptotic death in phenotypically defined DA neurons of the SN (106). This induction is most pronounced in a developmental setting, through PND 14, but it can also be demonstrated, at a lower level, in mature animals. Interestingly, at older ages the morphology of cell death becomes mixed, including apoptotic and nonapoptotic features (106). Although in this model 6-OHDA may lead to apoptotic death simply by the destruction of terminals and the resulting failure of target support, there is evidence that the toxin also directly mediates death. In addition to the fact that 6-OHDA is able to induce death long after it is possible to demonstrate target dependence (84), the morphologic features of activated caspase-3 expression, an important mediator of programmed cell death, differs in 6-OHDA–induced death, as compared to natural and target injury-induced cell death (83).

Other important animal models of parkinsonism, in rodents and primates, are induced by MPTP, or its active oxidized product, MPP^+. A number of investigators have shown in a variety of systems that MPP^+ can induce apoptosis *in vitro*. Dipasquale and colleagues (27) first showed that MPP^+ can induce apoptotic morphology and DNA fragmentation in postnatal cerebellar granule cells in culture. Subsequently, others have shown that MPP^+ appears to induce apoptosis in embryonic mesencephalon culture (116), in PC12 cells, both differentiated (117) and undifferentiated (57), and in a human neuroblastoma cell line (73). Tatton and Kish (149) have shown that when MPTP is administered to mice in low doses over a 5-day course, it induces apoptotic death. In this model, induction of apoptosis is dependent on the dosing regimen; acute administration of multiple doses in a single day results in nonapoptotic death (77). Further support for the role of programmed cell death in the MPTP mouse model derives from the observation that Bax, a mediator of apoptosis, is induced (60). In addition, overexpression of Bcl-2, a protein inhibitor of apoptosis, diminishes MPTP-induced injury (161).

It remains controversial whether apoptosis can be identified in the Parkinson brain. One initial report demonstrated intranuclear chromatin clumps by electron microscopy (7), but they were not clearly characteristic of apoptosis. TUNEL labeling (45) has been demonstrated in PD brains (115). However, the terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) technique is not entirely specific for apoptosis since it can also label free 3′-ends that are generated by necrotic death, producing false positives. Thus, it is essential to co-identify not only TUNEL labeling but also the classic chromatin clump morphology of apoptosis. Tompkins and co-workers (153) and Tatton and colleagues (150) have both achieved this more specific demonstration. However, many other investigators have been unable to identify specific TUNEL labeling in PD brains (12 ,86 ,91 ,160). Thus, a consensus has not been achieved, and further investigation will require methods beyond these morphologic techniques, such as the utilization of specific biochemical markers for programmed cell death. Hartmann and colleagues (58) have made such an effort utilizing an antibody that is specific for the activated form of caspase-3. They have shown that activated caspase-3 could be identified in Lewy body–containing neurons of the SN in PD brains (58). However, activated caspase-3 was also identified in control brains, in larger numbers of neurons. They attribute this staining to agonal changes. Thus, the staining was not specific for PD, and more study is needed with additional specific immunoreagents for other components and by-products of programmed cell death pathways.

Protein Aggregation

The possibility that protein aggregation may play a role in PD had long been suggested by the presence of Lewy bodies in disease brains. However, this concept was given powerful support upon discovery of the mutations in α-synuclein. Human α-synuclein was originally identified as a proteolytic fragment derived from Alzheimer senile plaques (154). The isolated fragment, termed the non-Aß component of amyloid (or NAC), corresponded to a 35 amino acid hydrophobic portion of α-synuclein. Soon after its discovery, NAC was predicted to form ß-sheet secondary structure, and shown to self-aggregate to form fibrillar amyloid *in vitro* (56). Thus, with the discovery of α-synuclein, PD was placed firmly among neurodegenerative disorders for which protein aggregation is believed to play an important pathogenetic role, including Alzheimer's, motor neuron disease, the triplet repeat diseases, and the prion diseases.

Subsequent biochemical studies have shown that full-length α-synuclein is capable of binding to Aß 1-38 and forming amyloid (163). This binding requires the hydrophobic NAC region. Other investigators have confirmed α-synuclein binding to Aß, and have shown that α-synuclein is capable of homodimerization (81 ,124). Even in the absence of Aß, full-length recombinant α-synuclein is capable of self-aggregation *in vitro* to form fibrillar amyloid material (59). A number of investigators have confirmed this observation, and have found that one or both mutant forms are more likely to form ß-sheet and aggregate, forming fibrils

that resemble those of Lewy bodies (20 ,32 ,47 ,119). Crowther and co-workers (21) have shown that C-terminally truncated α-synuclein more readily self-assembles into filaments that resemble those isolated from diseased brain.

Further analysis of the NAC fragment has shown that the N-terminal sequence is responsible for aggregation (30). NAC aggregates have been demonstrated to have cellular toxicity. They induce apoptotic death in cultured human neuroblastoma cells (31). Low concentrations of aggregated NAC are toxic to DA neurons in primary culture and neuronally differentiated PC12 cells (39). *In vivo* application of NAC aggregates induced death of SN DA neurons. Based on these demonstrations of the ability of α-synuclein and NAC to aggregate and the possible toxicity of aggregates, an abnormality of protein aggregation has become one of the principal hypotheses for the pathogenesis of PD.

FINAL COMMENT

Part of "123 - Pathophysiology of Parkinson's Disease "

Although Parkinson's disease was first described almost two centuries ago, it is only recently that we have begun to understand the complex nature of the functional deficits that it entails or its neurobiological causes. Yet, the pace of discovery is quickening. With the discovery of the genetic basis of some familiar forms of the disorder, the appreciation of trophic factors that influence DA neurons, and the development of new technologies such as the use of stem cells and viral vectors, there is every reason to believe that within a generation Parkinson's disease will become a chapter in the history of diseases of the past.

REFERENCES

1. Aarsland D, Larsen JP, Lim NG, et al. Range of neuropsychiatric disturbances in patients with Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1999;67:492–496.

2. Aarsland D, Tandberg E, Larsen JP, et al. Frequency of dementia in Parkinson disease. *Arch Neurol* 1996;53:538–542.

3. Abbas N, Lucking CB, Ricard S, et al. A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. French Parkinson's Disease Genetics Study Group and the European Consortium on Genetic Susceptibility in Parkinson's Disease. *Hum Mol Genet* 1999;8:567–574.

4. Abeliovich A, Schmitz Y, Farinas I, et al. Mice lacking alpha-synuclein display functional deficits in the nigrostriatal dopamine system. *Neuron* 2000;25:239–252.

5. Abercrombie ED, Zigmond MJ. Modification of central catecholaminergic systems by stress and injury: functional significance and clinical implications. In: Bloom EJ, Kupfer DJ. *Psychopharmacology: the fourth generation of progress*. New York: Raven Press, 1995:355–362.

6. Agid Y, Agid F, Ruberg M. Biochemistry of neurotransmitters in Parkinson's disease. In: Marsden CD, Fahn S, eds. *Movement disorders 2*. London: Butterworths, 1987:166–230.

7. Anglade P, Vyas S, Javoy-Agid F, et al. Apoptosis and autophagy in nigral neurons of patients with Parkinson's disease. *Histol Histopathol* 1997;12:25–31.

8. Arawaka S, Saito Y, Murayama S, et al. Lewy body in neurodegeneration with brain iron accumulation type 1 is immunoreactive for alpha-synuclein. *Neurology* 1998;51:887–889.

9. Arima K, Hirai S, Sunohara N, et al. Cellular co-localization of phosphorylated tau- and NACP/alpha-synuclein-epitopes in lewy bodies in sporadic Parkinson's disease and in dementia with Lewy bodies. *Brain Res* 1999;843:53–61.

10. Arima K, Ueda K, Sunohara N, et al. Immunoelectron-microscopic demonstration of NACP/alpha-synuclein-epitopes on the filamentous component of Lewy bodies in Parkinson's disease and in dementia with Lewy bodies. *Brain Res* 1998;808:93–100.

11. Baba M, Nakajo S, Tu PH, et al. Aggregation of alpha-synuclein in Lewy bodies of sporadic Parkinson's disease and dementia with Lewy bodies. *Am J Pathol* 1998;152:879–884.

12. Banati RB, Daniel SE, Blunt SB. Glial pathology but absence of apoptotic nigral neurons in long-standing Parkinson's disease. *Mov Disord* 1998;13:221–227.

13. Bancher, C, Lassman H, Budka H, et al. An antigenic profile of Lewy bodies: Immunocytochemical indication for protein phosphorylation and ubiquitination. *J Neuropathol Exp Neurol* 1989;48:81–93.

14. Brion J-P, Couck A-M. Cortical and brainstem-type Lewy bodies are immunoreactive for the cyclin-dependent kinase 5. *Am J Pathol* 1995;147:1465–1476.

15. Brooks DJ. PET studies and motor complications in Parkinson's disease. *Trends Neurosci* 2000;23:S101–108.

16. Burke RE, Antonelli M, Sulzer D. Glial cell line-derived neurotrophic growth factor inhibits apoptotic death of postnatal substantia nigra dopamine neurons in primary culture. *J Neurochem* 1998;71:517–525.

17. Calne DB. The free radical hypothesis in idiopathic parkinsonism: evidence against it. *Ann Neurol* 1992;32:799–803.

18. Carlsson A, Winblad B. Influence of age and time interval between death and autopsy on dopamine and 3-methoxytyramine levels in human basal ganglia. *J Neural Transm* 1976;38:271–276.

19. Cohen G, Heikkila RE. The generation of hydrogen peroxide, superoxide radical, and hydroxyl radical by 6-hydroxydopamine, dialuric acid, and related cytotoxic agents. *J Biol Chem* 1974;249:2447–2452.

20. Conway KA, Harper JD, Lansbury PT. Accelerated in vitro fibril formation by a mutant alpha-synuclein linked to early-onset Parkinson disease. *Nat Med* 1998;4:1318–1320.

21. Crowther RA, Jakes R, Spillantini MG, et al. Synthetic filaments assembled from C-terminally truncated alpha-synuclein. *FEBS Lett* 1998;436:309–312.

22. Dexter DT, Carayon A, Javoy-Agid F, et al. Alterations in the levels of iron, ferritin and other trace metals in Parkinson's disease and other neurodegenerative diseases affecting the basal ganglia. *Brain* 1991;114:1953–1975.

23. Dexter DT, Carayon A, Vidailhet M, et al. Decreased ferritin levels in brain in Parkinson's disease. *J Neurochem* 1990;55:16–20.

24. Dexter DT, Carter CJ, Wells FR, et al. Basal lipid peroxidation in substantia nigra is increased in Parkinson's disease. *J Neurochem* 1989;52:381–389.

25. Dexter DT, Holley AE, Flitter WD, et al. Increased levels of lipid hydroperoxides in the parkinsonian substantia nigra: An HPLC and ESR study. *Mov Disord* 1994;9:92–97.

26. Dexter DT, Wells FR, Lees AJ, et al. Increased nigral iron content and alterations in other metal ions occurring in brain in Parkinson's disease. *J Neurochem* 1989;52:1830–1836.

27. Dipasquale B, Marini AM, Youle RJ. Apoptosis and DNA degradation induced by 1-methyl-4-phenylpyridinium in neurons. *Biochem Biophys Res Commun* 1991;181:1442–1448.

28. Duffy PE, Tennyson VM. Phase and electron microscopic observations of Lewy bodies and melanin granules in the substantia nigra and locus caeruleus in Parkinson's Disease. *J Neuropathol Exp Neurol* 1965;24:398–414.

29. Duvoisin RC. Recent advances in the genetics of Parkinson's disease. In: Battistin L, Scarlato G, Caraceni T, et al., eds. *Advances in neurology, vol 69: Parkinson's disease*. Philadelphia, Lippincott-Raven, 1996:33–40.

30. El-Agnaf OM, Bodles AM, Guthrie DJ, et al. The N-terminal region of non-A beta component of Alzheimer's disease amyloid is responsible for its tendency to assume beta-sheet and aggregate to form fibrils. *Eur J Biochem* 1998;258:157–163.

31. El-Agnaf OM, Jakes R, Curran MD, et al. Aggregates from mutant and wild-type alpha-synuclein proteins and NAC peptide induce apoptotic cell death in human neuroblastoma cells by formation of beta-sheet and amyloid-like filaments. *FEBS Lett* 1998;440:71–75.

32. El-Agnaf OM, Jakes R, Curran MD, et al. Effects of the mutations Ala30 to Pro and Ala53 to Thr on the physical and morphological properties of alpha-synuclein protein implicated in Parkinson's disease. *FEBS Lett* 1998;440:67–70.

33. Engelender S, Kaminsky Z, Guo X, et al. Synphilin-1 associates with alpha-synuclein and promotes the formation of cytosolic inclusions. *Nat Genet* 1999;22:110–114.

34. Fahn S, Cohen G. The oxidant stress hypothesis in Parkinson's disease: evidence supporting it. *Ann Neurol* 1992;32:804–812.

35. Faucheux BA, Nillesse N, Damier P, et al. Expression of lactoferrin receptors is increased in the mesencephalon of patients with Parkinson's disease. *Proc Natl Acad Sci USA* 1995;92:9603–9607.

36. Fearnley JM, Lees AJ. Aging and Parkinson's disease: substantia nigra regional selectivity. *Brain* 1991;114:2283–2301.

37. Fergusson J, Landon M, Lowe J, et al. Pathological lesions of Alzheimer's disease and dementia with Lewy bodies brains exhibit immunoreactivity to an ATPase that is a regulatory subunit of the 26S proteasome. *Neurosci Lett* 1996;219:167–170.

38. Filloux F, Townsend J. Pre- and post-synaptic neurotoxic effects of dopamine demonstrated by intrastriatal injection. *Exp. Neurol* 1993;119:79–88.

39. Forloni G, Bertani I, Calella AM, et al. Alpha-synuclein and Parkinson's disease: selective neurodegenerative effect of alpha-synuclein fragment on dopaminergic neurons in vitro and in vivo. *Ann Neurol* 2000;47:632–640.

40. Forno LS, Sternberger LA, Sternberger NH, et al. Reaction of Lewy bodies with antibodies to phosphorylated and non-phosphorylated neurofilaments. *Neurosci Lett* 1986;64:253–258.

41. Fornstedt B, Brun A, Rosengren E, et al. The apparent autoxidation rate of catechols in dopamine-rich regions of human brains increases with the degree of depigmentation of substantia nigra. *J Neural Trans* 1989;1:279–295.

42. Fornstedt B, Pileblad E, Carlsson A. In vivo autoxidation of dopamine in guinea pig striatum increases with age. *J Neurochem* 1990;55:655–659.

43. Fornstedt B, Rosengren E, Carlsson A. Occurrence and distribution of 5-S-cysteinyl derivatives of dopamine dopa and dopac in the brains of eight mammalian species, *Neuropharmacol* 1986;25:451–454.

44. Galvin JE, Uryu K, Lee, VM, et al. Axon pathology in Parkinson's disease and Lewy body dementia hippocampus contains alpha-, beta-, and gamma-synuclein. *Proc Natl Acad Sci USA* 1999;96:13450–13455.

45. Gavrieli Y, Sherman Y, Ben-Sasson SA. Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. *J Cell Biol* 1992;119:493–501.

46. George JM, Jin H, Woods WS, et al. Characterization of a novel protein regulated during the critical period for song learning in the zebra finch. *Neuron* 1995;15:361–372.

47. Giasson BI, Uryu K, Trojanowski JQ, et al. Mutant and wild type human alpha-synucleins assemble into elongated filaments with distinct morphologies in vitro. *J Biol Chem* 1999;274:7619–7622.

48. Gibb WR. Accuracy in the clinical diagnosis of parkinsonian syndromes. *Postgrad Med J* 1988;64:345–351.

49. Gibb WRG, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J Neurol* 1988;51:745–752.

50. Gibb WR, Scott T, Lees AJ. Neuronal inclusions of Parkinson's disease. *Mov Disord* 1991;6:2–11.

51. Giladi N, McMahon D, Przedborski S, et al. Motor blocks in Parkinson's disease. *Neurology* 1992;42:333–339.

52. Golbe LI, Di Iorio G, Bonavita V, et al. Autosomal dominant Parkinson's disease. *Ann Neurol* 1990;27:276–282.

53. Goldman JE, Yen S-H, Chiu F-C, et al. Lewy bodies of Parkinson's disease contain neurofilament antigens. *Science* 1983;221:1082–1084.

54. Graham DG, Tiffany SM, Bell WR Jr, et al. Autoxidation versus covalent binding of quinones as the mechanism of toxicity of dopamine 6-hydroxydopamine and related compounds toward C1300 neuroblastoma cells in vitro. *Mol Pharmacol* 1978;14:644–653.

55. Hamilton RL. Lewy bodies in Alzheimer's disease: a neuropathological review of 145 cases using alpha-synuclein immunohistochemistry [In Process Citation]. *Brain Pathol* 2000;10:378–384.

56. Han H, Weinreb PH, Lansbury PTJ. The core Alzheimer's peptide NAC forms amyloid fibrils which seed and are seeded by beta-amyloid: is NAC a common trigger or target in neurodegenerative disease? *Chem Biol* 1995;2:163–169.

57. Hartley A, Stone JM, Heron C, et al. Complex I inhibitors induce dose dependent apoptosis in PC12 cells relevance to Parkinson's disease. *J Neurochem* 1994;63:1987–1990.

58. Hartmann A, Hunot S, Michel PP, et al. Caspase-3: a vulnerability factor and final effector in apoptotic death of dopaminergic neurons in Parkinson's disease. *Proc Natl Acad Sci USA* 2000;97:2875–2880.

59. Hashimoto M, Hsu LJ, Sisk A, et al. Human recombinant NACP/alpha-synuclein is aggregated and fibrillated in vitro: relevance for Lewy body disease. *Brain Res* 1998;799:301–306.

60. Hassouna I, Wickert H, Zimmermann M, et al. Increase in bax expression in substantia-nigra following 1- methyl-4-phenyl-1,2,3,6-tetrahydropyridine. (MPTP) treatment of mice. *Neurosci Lett* 1996;204:85–88.

61. Hastings TG, Lewis DA, Zigmond MJ. Role of oxidation in the neurotoxic effects of intrastriatal dopamine injections. *Proc Natl Acad Sci USA* 1996;93:1956–1961.

62. Hastings TG, Lewis DA, Zigmond MJ. Reactive dopamine metabolites and neurotoxicity: implications for Parkinson's disease. *Adv Exp Med Biol* 1996;387:97–106.

63. Hastings TG, Zigmond MJ. Identification of catechol-protein conjugates in neostriatal slices incubated with ^{3}H-dopamine: impact of ascorbic acid and glutathione. *J Neurochem* 1994;63:1126–1132.

64. Hattori N, Kitada T, Matsumine H, et al. Molecular genetic analysis of a novel Parkin gene in Japanese families with autosomal recessive juvenile parkinsonism: evidence for variable homozygous deletions in the Parkin gene in affected individuals. *Ann Neurol* 1998;44:935–941.

65. Heberlein I, Ludin HP, Scholz J, et al. Personality, depression, and premorbid lifestyle in twin pairs discordant for Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1998;64:262–266.

66. Heikkila RE, Cohen G. 6-Hydroxydopamine: evidence for superoxide radical as an oxidative intermediate. *Science* 1973;181:456–457.

67. Hill WD, Lee VM, Hurtig HI, et al. Epitopes located in spatially separate domains of each neurofilament subunit are present in Parkinson's disease Lewy bodies. *J Comp Neurol* 1991;309:150–160.

68. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology* 1967;17:427–442.

69. Hubble JP, Venkatesh R, Hassanein RES, et al. Personality and depression in Parkinson's disease. *J Nerv Ment Dis* 1993;181:657–662.

70. Hughes AJ, Daniel SE, Kilford L, et al. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992;55:181–184.

71. Ii K, Ito H, Tanaka K, et al. Immunocytochemical co-localization of the proteasome in ubiquitinated structures in neurodegenerative diseases and the elderly. *J Neuropathol Exp Neurol* 1997;56:125–131.

72. Ishikawa A, Tsuji S. Clinical analysis of 17 patients in 12 Japanese families with autosomal-recessive type juvenile parkinsonism. *Neurology* 1996;47:160–166.

73. Itano Y, Nomura Y. 1-Methyl-4-phenyl-pyridinium ion. (MPP(+)) causes DNA fragmentation and increases the Bcl-2 expression in human neuroblastoma, SH-SY5Y cells, through different mechanisms. *Brain Res* 1995;704:240–245.

74. Iwai A, Masliah E, Yoshimoto M, et al. The precursor protein of non-AB component of Alzheimer's disease amyloid is a presynaptic protein of the central nervous system. *Neuron* 1995;14:467–475.

75. Iwatsubo T, Nakano I, Fukunaga K, et al. Ca2+/calmodulin-dependent protein kinase II immunoreactivity in Lewy bodies. *Acta Neuropathol (Berl)* 1991;82:159–163.

76. Irizarry MC, Kim TW, McNamara M, et al. Characterization of the precursor protein of the non-A beta component of senile plaques. (NACP) in the human central nervous system. *J Neuropathol Exp Neurol* 1996;55(8):889–895.

77. Jackson-Lewis V, Jakowec M, Burke RE, et al. Time course and morphology of dopaminergic neuronal death caused by the neurotoxin 1-methyl-4-phenyl-1,2,3,6,-tetrahydropyridine. *Neurodegen* 1995;4:257–269.

78. Jakes R, Spillantini MG, Goedert M. Identification of two distinct synucleins from human brain. *FEBS Lett* 1994;345:27–32.

79. Janec E, Burke RE. Naturally occurring cell death during postnatal development of the substantia nigra of the rat. *Mol Cell Neurosci* 1993;4:30–35.

80. Jenner P, Dexter DT, Sian J, et al. Oxidative stress as a cause of nigral cell death in Parkinson's disease and incidental Lewy body disease. *Ann Neurol* 1992;32:S82–S87.

81. Jensen PH, Hojrup P, Hager H, et al. Binding of Abeta to alpha- and beta-synucleins: identification of segments in alpha-synuclein/NAC precursor that bind Abeta and NAC. *Biochem J* 1997;323:539–546.

82. Jensen PH, Nielsen MS, Jakes R, et al. Binding of alphasynuclein to brain vesicles is abolished by familial Parkinson's disease mutation. *J Biol Chem* 1998;273:26292–26294.

83. Jeon BS, Kholodilov NG, Oo TF, et al. Activation of caspase-3 in developmental models of programmed cell death in neurons of the substantia nigra. *J Neurochem* 1999;73:322–333.

84. Kelly WJ, Burke RE. Apoptotic neuron death in rat substantia nigra induced by striatal excitotoxic injury is developmentally dependent. *Neurosci Lett* 1996;220:85–88.

85. Kholodilov NG, Neystat M, Oo TF, et al. Increased expression of rat synuclein1 in the substantia nigra pars compacta identified by differential display in a model of developmental target injury. *J Neurochem* 1999;73:2586–2599.

86. Kingsbury AE, Mardsen CD, Foster OJ. DNA fragmentation in human substantia nigra: apoptosis or perimortem effect? [In Process Citation]. *Mov Disord* 1998;13:877–884.

87. Kish SJ, Shannak K, Hornykiewicz O. Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. Pathophysiologic and clinical implications. *N Engl J Med* 1988;318:876–880.

88. Kish SJ, Shannak K, Rajput A, et al. Aging produces a specific pattern of striatal dopamine loss implications for the etiology of idiopathic Parkinson's disease. *J Neurochem* 1992;58:642–648.

89. Kitada T, Asakawa S, Hattori N, et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism [see comments]. *Nature* 1998;392:605–608.

90. Koike Y, Takahashi A. Autonomic dysfunction in Parkinson's disease. *Eur Neurol* 1997;38(suppl 2):8–12.

91. Kosel S, Egensperger R, von Eitzen U, et al. On the question of apoptosis in the parkinsonian substantia nigra. *Acta Neuropathol* 1997;93:105–108.

92. Kruger R, Kuhn W, Muller T, et al. Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease [letter]. *Nat Genet* 1998;18:106–108.

93. Kuzuhara S, Mori H, Izumiyama N, et al. Lewy bodies are ubiquitinated. A light and electron microscopic immunocytochemical study. *Acta Neuropathol (Berl)* 1998;75:345–353.

94. Lang AE, Lozano AM. Parkinson's disease. First of two parts. *N Engl J Med* 1998;339:1044–1053.

95. Lang AE, Lozano AM. Parkinson's disease. Second of two parts. *N Engl J Med* 1998;339:1130–1143.

96. Langston JW, Ballard PA, Tetrud JW, et al. Chronic parkinsonism in humans due to a product of meperidine-analog synthesis. *Science* 1983;219:979–980.

97. Langston JW, Koller WC, Giron LT. Etiology of Parkinson's disease. In: Olanow CW, Lieberman AN, eds. *The scientific basis for the treatment of Parkinson's disease*. Park Ridge, NJ: Parthenon, 1992:33–58.

98. Lew J, Huang Q-Q, Qi Z, et al. A brain-specific activator of cyclin-dependent kinase 5. *Nature* 1994;371:423–426.

99. Lippa CF, Fujiwara H, Mann DM, et al. Lewy bodies contain altered alpha-synuclein in brains of many familial Alzheimer's disease patients with mutations in presenilin and amyloid precursor protein genes. *Am J Pathol* 1998;153:1365–1370.

100. Lippa CF, Schmidt ML, Lee VM, et al. Antibodies to alpha-synuclein detect Lewy bodies in many Down's syndrome brains with Alzheimer's disease. *Ann Neurol* 1999;45:353–357.

101. Lucking CB, Durr A, Bonifati V, et al. Association between early-onset Parkinson's disease and mutations in the parkin gene. French Parkinson's Disease Genetics Study Group. *N Engl J Med* 2000;342:1560–1567.

102. Macaya A, Munell F, Gubits RM, et al. Apoptosis in substantia nigra following developmental striatal excitotoxic injury. *Proc Natl Acad Sci USA* 1994;91:8117–8121.

103. Markopoulou K, Wszolek ZK, Pfeiffer RF, et al. Reduced expression of the G209A alpha-synuclein allele in familial Parkinsonism. *Ann Neurol* 1999;46:374–381.

104. Maroteaux L, Campanelli JT, Scheller RH. Synuclein: A neuron-specific protein localized to the nucleus and presynaptic nerve terminal. *J Neurosci* 1988;8:2804–2815.

105. Maroteaux L, Scheller RH. The rat brain synucleins; family of proteins transiently associated with neuronal membrane. *Mol Brain Res* 1991;11:335–343.

106. Marti MJ, James CJ, Oo TF, et al. Early developmental destruction of terminals in the striatal target induces apoptosis in dopamine neurons of the substantia nigra. *J Neurosci* 1997;17:2030–2039.

107. Mathias CJ. Cardiovascular autonomic dysfunction in parkinsonian patients. *Clin Neurosci* 1998;5(2):153–166.

108. Mayeux R, Chen J, Mirabello E, et al. An estimate of the incidence of dementia in idiopathic Parkinson's disease. *Neurology* 1990;40:1513–1517.

109. Mayeux R, Denaro J, Hemenegildo N, et al. A population-based investigation of Parkinson's disease with and without dementia. *Arch Neurol* 1992;49:492–497.

109a. Mayeux R, Stern Y, Rosen J, et al. Depression, intellectual impairment and Parkinson disease. *Neurology* 1981;31(6):645–650.

110. Mayeux R, Stern Y, Rosenstein R, et al. An estimate of the prevalence of dementia in idiopathic Parkinson's disease. *Arch Neurol* 1988;45:260–262.

111. McGeer PL, Itagaki S, Akiyama H, et al. Rate of cell death in parkinsonism indicates active neuropathological process. *Ann Neurol* 1988;24:574–576.

112. McGeer PL, Itagaki S, Boyes BE, et al. Reactive microglia are positive for HLA-DR in the substantia nigra of Parkinson's and Alzheimer's disease brains. *Neurology* 1988;38:1285–1291.

113. McGeer PL, McGeer EG, Suzuki JS. Aging and extrapyramidal function. *Arch Neurol* 1977;34:33–35.

114. Michel PP, Hefti F. Toxicity of 6-hydroxydopamine and dopamine for dopaminergic neurons in culture. *J Neurosci Res* 1990;26:428–435.

115. Mochizuki H, Goto K, Mori H, et al. Histochemical detection of apoptosis in Parkinson's disease. *J Neurol Sci* 1996;137:120–123.

116. Mochizuki H, Nakamura N, Nishi K, et al. Apoptosis Is induced by 1-methyl-4-phenylpyridinium ion. (MPP+) in ventral mesencephalic striatal co-culture in rat. *Neurosci Lett* 1994;170:191–194.

117. Mutoh T, Tokuda A, Marini AM, et al. 1-methyl-4-phenylpyridinum kills differentiated PC12 cells with a concomitant change in protein phosphorylation. *Brain Res* 1994;661:51–55.

118. Nakamura S, Kawamoto Y, Nakano S, et al. p35nck5a and cyclin-dependent kinase 5 colocalize in Lewy bodies of brains with Parkinson's disease. *Acta Neuropathol (Berl)* 1997;94:153–157.

119. Narhi L, Wood SJ, Steavenson S, et al. Both familial Parkinson's disease mutations accelerate alpha-synuclein aggregation. *J Biol Chem* 1999;274:9843–9846.

120. Neystat M, Lynch T, Przedborski S, et al. alpha-Synuclein expression in substantia nigra and cortex in Parkinson's Disease. *Mov Disord* 1999;14:417–422.

121. Nicklas WJ, Vyas I, Heikkila RE. Inhibition of NADH-linked oxidation in brain mitochondria by MPP^+, a metabolite of the neurotoxin MPTP. *Life Sci* 1985;36:2503–2508.

122. Olanow CW. An introduction to the free radical hypothesis in Parkinson's disease. *Ann Neurol* 1992;32:S2–S9.

123. Oo TF, Burke RE. The time course of developmental cell death in phenotypically defined dopaminergic neurons of the substantia nigra. *Dev Brain Res* 1997;98:191–196.

124. Paik SR, Lee JH, Kim DH, et al. Self-oligomerization of NACP, the precursor protein of the non-amyloid beta/A4 protein. (A beta) component of Alzheimer's disease amyloid, observed in the presence of a C-terminal A beta fragment (residues 25–35). *FEBS Lett* 1998;421:73–76.

125. Pakkenberg H, Brody H. The number of nerve cells in the substantia nigra in paralysis agitans. *Acta Neuropathol (Berl)* 1965;5:320–324.

126. Paulson GW, Dadmehr N. Is there a premorbid personality typical for Parkinson's disease? *Neurology* 1991;41(suppl 2):73–76.

127. Perry TL, Godin DV, Hansen S. Parkinson's disease: A disorder due to nigral glutathione deficiency? *Neurosci Lett* 1982;33:305–310.

128. Phelps ME. Inaugural article: positron emission tomography provides molecular imaging of biological processes. *Proc Natl Acad Sci USA* 2000;1z:9226–9233.

129. Poewe W, Gerstenbrand F, Ransmayr G, et al. Premorbid personality of Parkinson patients. *J Neural Transm Suppl* 1983;19:215–224.

130. Polymeropoulos MH, Higgins JJ, Golbe LI, et al. Mapping of a gene for Parkinson's disease to chromosome 4q21-q23. *Science* 1996;274:1197–1199.

131. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the α-synuclein gene identified in families with Parkinson's Disease. *Science* 1997;276:2045–2047.

132. Przedborski S, Kostic V, Jackson-Lewis V, et al. Transgenic mice with increased Cu/Zn-superoxide dismutase activity are resistant to *N*-methyl-4-phenyl-1,2,3,6-tetrahydropyridineinduced neurotoxicity. *J Neurosci* 1992;12:1658–1667.

133. Quigley EM. Gastrointestinal dysfunction in Parkinson's disease. *Semin Neurol* 1996;16:245–250.

134. Rajput AH, Rozdilsky B, Rajput A. Accuracy of clinical diagnosis in parkinsonism—a prospective study. *Can J Neurol Sci* 1991;18:275–278.

135. Rosenberg PA. Catecholamine toxicity in cerebral cortex in dissociated cell culture. *J Neurosci* 1988;8:2887–2894.

136. Sanchez-Ramos JR, Overvik E, Ames BN. A marker of oxyradical-mediated DNA damage. (8-hydroxy- 2′deoxyguanosine) is increased in nigro striatum of Parkinson's disease brain. *Neurodegen* 1994;3:197–204.

137. Schapira AHV, Mann VM, Cooper JM, et al. Mitochondrial function in Parkinson's disease. *Ann Neurol* 1992;32:S116–S124.

138. Scherman D, Desnos C, Darchen F, et al. Striatal dopamine deficiency in Parkinson's disease: Role of aging. *Ann Neurol* 1989;26:551–557.

139. Shimura H, Hattori N, Kubo S, et al. Immunohistochemical and subcellular localization of Parkin protein: absence of protein in autosomal recessive juvenile parkinsonism patients. *Ann Neurol* 1999;45:668–672.

140. Shimura H, Hattori N, Kubo S, et al. Familial parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat Genet* 2000;25:302–305.

141. Smith MA, Sayre LM, Monnier VM, et al. Radical AGEing in Alzheimer's disease. *Trends Neurosci* 1995;18:172–176.

142. Spillantini MG, Crowther RA, Jakes R, et al. Alpha-synuclein in filamentous inclusions of lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc Natl Acad Sci USA* 1998;95:6469–6473.

143. Spillantini MG, Schmidt ML, Lee VMY, et al. α-Synuclein in Lewy bodies. *Nature* 1997;388:839–840.

144. Srinivasan A, Roth KA, Sayers RO, et al. In situ immunodetection of activated caspase-3 in apoptotic neurons in the developing nervous system. *Cell Death Different* 1998;5:1004–1016.

145. Swartz HM, Sarna T, Zecca L. Modulation by neuromelanin of the availability and reactivity of metal ions. *Ann Neurol* 1992;32:S69–S75.

146. Takahashi H, Ohama E, Suzuki S, et al. Familial juvenile parkinsonism: clinical and pathologic study in a family. *Neurology* 1994;44:437–441.

147. Tandberg E, Larsen JP, Aarsland D, et al. The occurrence of depression in Parkinson's disease. A community-based study. *Arch Neurol* 1996;53:175–179.

148. Tanner CM. Epidemiology of Parkinson's disease. *Neurol Clin* 1992;10:317–329.

149. Tatton NA, Kish SJ. In situ detection of apoptotic nuclei in the substantia nigra compacta of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-treated mice using terminal deoxynucleotidyl transferase labelling and acridine orange. *Neuroscience* 1997;77:1037–1048.

150. Tatton NA, Maclean-Fraser A, Tatton WG, et al. A fluorescent double-labeling method to detect and confirm apoptotic nuclei in Parkinson's disease. *Ann Neurol* 1998;44:S142–S148.

151. Tepper JM, Damlama M, Trent F. Postnatal changes in the distribution and morphology of rat substantia nigra dopaminergic neurons. *Neuroscience* 1994;60:469–477.

152. Todes CJ, Lees AJ. The pre-morbid personality of patients with Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1985;48:97–100.

153. Tompkins MM, Basgall EJ, Zamrini E, et al. Apoptotic-like changes in Lewy-body-associated disorders and normal aging in substantia nigral neurons. *Am J Pathol* 1997;150:119–131.

154. Ueda K, Fukushima H, Masliah E, et al. Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc Natl Acad Sci USA* 1993;90:11282–11286.

155. Vila M, Vukosavic S, Jackson-Lewis V, et al. Alpha-synuclein up-regulation in substantia nigra dopaminergic neurons following administration of the parkinsonian toxin MPTP. *J Neurochem* 2000;74:721–729.

156. Wakabayashi K, Engelender S, Yoshimoto M, et al. Synphilin-1 is present in Lewy bodies in Parkinson's disease. *Ann Neurol* 2000;47:521–523.

157. Ward CD, Duvoisin RC, Ince SE, et al. Parkinson's disease in 65 pairs of twins and in a set of quadruplets. *Neurology* 1983;33:815–824.

158. Weiner WJ, Lang AE. *Movement disorders: a comprehensive survey*. Mt. Kisco, NY: Futura, 1989.

159. Whitehouse PJ, Hedreen JC, White CL, et al. Basal forebrain neurons in the dementia of Parkinson disease. *Ann Neurol* 1983;13:243–248.

160. Wullner U, Kornhuber J, Weller M, et al. Cell death and apoptosis regulating proteins in Parkinson's disease—a cautionary note. *Acta Neuropathol (Berl)* 1999;97:408–412.

161. Yang L, Matthews RT, Schulz JB, et al. 1-Methyl-4-phenyl-1,2,3,6-tetrahydropyride neurotoxicity is attenuated in mice overexpressing Bcl-2. *J Neurosci* 1998;18:8145–8152.

162. Yoritaka A, Hattori N, Uchida K, et al. Immunohistochemical detection of 4-hydroxynonenal protein adducts in Parkinson's disease. *Proc Natl Acad Sci USA* 1996;93:2696–2701.

163. Yoshimoto M, Iwai A, Kang D, et al. NACP, the precursor protein of the non-amyloid beta/A4 protein. (A beta) component of Alzheimer disease amyloid, binds A beta and stimulates A beta aggregation. *Proc Natl Acad Sci USA* 1995;92:9141–9145.

164. Zigmond MJ. Chemical transmission in the brain: homeostatic regulation and its functional implications. *Prog Brain Res* 1994;100:115–122.

165. Zigmond MJ, Abercrombie ED, Berger TW, et al. Compensatory responses to partial loss of dopaminergic neurons: studies with 6-hydroxydopamine. In: Schneider JS, Gupta M, eds. *Current concepts in Parkinson's disease research*. Toronto: Hogrefe & Huber, 1993:99–140.

166. Zigmond MJ, Keefe KA. 6-Hydroxydopamine as a tool for studying catecholamines in adult animals: lessons from the neostriatum. In: Kostrzewa R, ed. *Highly selective neurotoxins: basic and clinical applications*. Totowa, NJ: Humana Press, 1997:75–108.

124

Current and Experimental Therapeutics of Parkinson's Disease

Jean-Michel Gracies

C. Warren Olanow

Jean-Michel Gracies and C. Warren Olanow: Department of Neurology, Movement Disorders Program, Mt. Sinai School of Medicine, New York, New York.

Three decades after its introduction, levodopa remains the gold standard for the treatment of Parkinson's disease (PD). Levodopa is the most potent symptomatic antiparkinsonian agent, and it is associated with an increase in quality of life and longevity for patients with PD. Dopamine agonists are increasingly being used, not only as an adjunct to levodopa, but as early therapy aimed at reducing the risk of developing levodopa-induced motor complications. Catechol *O*-methyltransferase (COMT) inhibitors extend the elimination half-life of levodopa. They are useful as adjunctive treatment for patients with motor fluctuations to increase the time in which patients respond to the drug. There is now increasing interest in using these drugs from the start of levodopa therapy to deliver levodopa to the brain in a more continuous fashion and thereby, it is hoped, further reduce the risk of motor complications. In the advanced stages of the illness, surgical therapies are being performed with increasing frequency based on evidence that they can restore function when medications fail. Ablative, stimulation, and transplant procedures are all currently under investigation. Finally, there are a series of investigational drugs designed to provide neuroprotective effects and/or to block levodopa motor complications that are now being evaluated in the laboratory and in some instances in PD patients. Thus, therapies and investigational approaches to PD have been markedly expanded in the past several years and include treatments and treatment strategies aimed at restoring function to PD patients in the advanced stages of the illness, preventing the development of the motor complications that are the major source of disability for a large percentage of PD patients, and modifying the disease process so as to slow or halt disease progression.

Parkinson's disease is a progressive motor disorder caused by accelerated degeneration of selected populations of brain cells, primarily including the melanized neurons of the substantia nigra pars compacta (SNc) in the midbrain. It affects approximately 4% of the population over 65 years of age and there are 60,000 new cases each year in the United States (1). With the increasing numbers of elderly individual in modern society, the prevalence of PD is likely to increase in developed countries in generations to come. The classic clinical syndrome is composed of four cardinal features: bradykinesia (slowness of movement), rigidity (increased resistance to passive limb movement), resting tremor (i.e., tremor that is most prominent at rest and tends to abate during voluntary movement), and impairment of gait and posture. The impairment of movement in PD primarily affects “automatic” movements such as those involved during walking, speech articulation and phonation, handwriting, or swallowing. Postmortem studies indicate that approximately 25% of patients who present with a parkinsonian syndrome do not have pathologic changes of PD, but rather of an atypical parkinsonism such as multiple system atrophy (MSA), progressive supranuclear palsy (PSP), or corticobasal ganglionic degeneration (CBGD) (2 ,3 and 4). The clinical features that best predict parkinsonian pathology are resting tremor, asymmetry of motor findings, and a good response to levodopa (see below) (5).

The pathologic hallmark of PD is the loss of pigmented, dopaminergic neurons of the SNc, coupled with intracellular inclusion bodies known as Lewy bodies (6). Pathologic changes frequently including Lewy bodies can also be detected in the locus coeruleus, the nucleus basalis of Meynert, cerebral cortical regions, autonomic regions of the brainstem, the pedunculopontine nucleus, intermediolateral columns of the spinal cord, and peripheral autonomic nerves innervating the cardiovascular system and gastrointestinal tract (7). Without treatment, PD evolves over 5 to 10 years into an akinetic and rigid state in which patients are unable to care for themselves. Death commonly results from aspiration

pneumonia due to swallowing impairment, or complications of immobility such as pulmonary embolism. The introduction of levodopa over 30 years ago (8) represented a revolution in the treatment of PD as it radically altered its prognosis. Under levodopa treatment, good functional mobility can be maintained for a number of years, and the life expectancy of levodopa-treated patients is markedly increased (9 ,10). However, it soon became apparent that levodopa therapy is associated with a series of motor complications that themselves are a major source of disability to PD patients (11). In recent years there have been dramatic advances in the therapeutics of PD with the development of new medical and surgical treatments that restore function to patients with advanced disease and prevent the development of levodopa-related motor complications. The final challenge involves the development of neuroprotective or disease-modifying therapies that slow or stop disease progression and herald the end of this devastating disorder. Here, too, enormous progress has been made, and several putative neuroprotective drugs and restorative therapies are currently being tested. This chapter reviews the major therapies for PD and describes present advances and future directions in the therapeutics of PD.

MEDICAL THERAPIES FOR PARKINSON'S DISEASE

Part of "124 - Current and Experimental Therapeutics of Parkinson's Disease "

Levodopa

Since its introduction in the late 1960s (8), levodopa (L-3,4-dihydroxyphenylalanine) has remained the single most effective antiparkinsonian agent, providing benefit to virtually all patients with PD. Levodopa use is associated with improved mobility, reduced disability, and prolonged survival (9 ,10 ,12). The involvement of dopaminergic systems in PD was first suspected in the late 1950s, following the observation that patients treated with the then newly available dopamine-blocking agents (antipsychotics) developed clinical signs of parkinsonism. In the same period, an animal study showed that movement slowness in rats, due to the catecholamine depletor reserpine, could be reversed with levodopa (13). The discovery that dopamine is depleted in the striatum of PD patients soon followed (14). This in turn gave rise to the notion that a dopamine replacement strategy might be useful in PD, and the therapeutic role of levodopa in patients with PD was subsequently established in 1967 (8 ,15). Levodopa is itself largely inert, and its therapeutic and adverse effects result from the decarboxylation of the prodrug levodopa into the active product dopamine (16). After oral administration, levodopa absorption occurs in the small bowel by way of the active transport system for large neutral amino acids. Thus, it is possible that other large neutral amino acids such as lysine and phenylalanine that are present in protein-rich foods can compete with and interfere with levodopa absorption. Levodopa is used in the place of dopamine as dopamine itself cannot penetrate the blood–brain barrier and enter the central nervous system (CNS). CNS entry is also an active process mediated by the large neutral amino acid transport system, and again there may be competition for brain access between levodopa and dietary amino acids (16).

Levodopa is normally metabolized in the periphery by two enzymatic systems: amino-acid decarboxylase (AADC) and COMT. This transformation occurs in the intestinal and gastric mucosa as well as in the liver. The peripheral metabolism of levodopa is so effective that the plasma half-life is approximately 60 minutes, and only 1% of an administered oral dose reaches the CNS (16). Further, accumulating concentrations of plasma dopamine secondary to decarboxylase-mediated metabolism of levodopa can activate dopamine receptors in the area postrema that are not protected by a blood–brain barrier and cause nausea and vomiting. Indeed, nausea and vomiting are limiting side effects in as many as 50% of patients when levodopa is administered alone. To defend against this complication, levodopa is now routinely administered in combination with a peripherally acting inhibitor of AADC. In the United States, levodopa is combined with the AADC inhibitor carbidopa and marketed as Sinemet. In other parts of the world, the AADC inhibitor benserazide is also frequently used with levodopa and sold as Madopar. The combination of levodopa with an AADC inhibitor permits the use of lower doses of levodopa (by doubling its bioavailability) and reduces the incidence of peripheral dopaminergic side effects such as nausea, vomiting, and hypotension. In most patients, a daily dose of 75 mg of carbidopa is sufficient to inhibit AADC and prevent these side effects. Interestingly, even in the presence of an AADC inhibitor, 90% of levodopa is still metabolized by COMT (17). This has led to the recent introduction of COMT inhibitors (see section below).

In the CNS, dopamine is synthesized from levodopa in dopaminergic terminals, transported into storage vesicles, and released in a spike-dependent manner in association with depolarization of the presynaptic neuron. The released dopamine acts on postsynaptic dopamine receptors (possibly in a volumetric manner). Its action is terminated primarily by a very rapid presynaptic reuptake system that is antagonized by cocaine. It can be degraded either intracellularly or extracellularly by monoamine oxidase (MAO) and COMT enzymes to yield homovanillic acid (HVA)(9). MAO has two subtypes; A, which is primarily intracellular, and B, which is primarily extracellular (18).

Two classes of dopamine receptors (D1 family and D2 family) and five receptor subtypes (D1–D5) have been molecularly cloned to date (19). The D1 receptor family is characterized by positive coupling with adenylate cyclase formation, whereas D2 receptors have an affinity for neuroleptic agents and activation inhibits adenylate cyclase (20). Dopamine receptors are G-protein–coupled receptors, and activation of the different subtypes likely is associated with

a different signaling pattern and different gene and protein regulation (19). Indeed, it is becoming increasingly clear that activation of different receptors, the same receptor with different agents, and the same receptor with the same agent with a different pattern of stimulation can all lead to a different intracellular signaling cascade that potentially has different functional effects (19). Dopamine receptors are diffusely distributed throughout the CNS: motor striatum (D1, D2), hippocampus (D5), frontal cortex and amygdala (D4), hypothalamus (D3, D5), and mesolimbic system (D3). The precise role of each of these receptors in motor function remains unknown; however, it is likely that this wide distribution accounts for the diverse pattern of functional effects that can be obtained when exogenous levodopa is administered to PD patients. Although most attention has focused on the nigrostriatal dopaminergic system in PD, it is important to appreciate that there is also dopaminergic innervation of the cerebral cortex and numerous other basal ganglia regions including the substantia nigra pars reticularis (SNr), the subthalamic nucleus (STN), the globus pallidus pars interna (GPi), and the globus pallidus pars externa (GPe) (21). Activation of dopaminergic receptors in these regions might also contribute to the beneficial and adverse effects observed with levodopa administration to PD patients. Indeed, although levodopa dramatically improves the motor signs and symptoms of PD, it also has effects on vision, memory, mood, reward-related learning, and addiction (22 ,23 ,24 ,25 ,26 ,27 ,28 ,29 and 30).

Levodopa Benefits and Motor Complications

Levodopa is the most effective antiparkinsonian agent for the management of motor dysfunction in PD. Levodopa benefits can be dramatic, and improvement can be obtained in all of the cardinal signs and symptoms of PD. Levodopa has also been shown to provide a dose-dependent beneficial effect on mood and anxiety in PD patients that increases with the duration of therapy (27). These important nonmotor effects can contribute to the benefits associated with the levodopa response. Indeed, more than 30 years after its introduction, no other medication provides antiparkinsonian benefits that are superior to levodopa (30 ,31).

The acute administration of levodopa, even in the presence of a decarboxylase inhibitor can still be associated with nausea, vomiting, and orthostatic hypotension. These are usually seen during the titration phase and can be minimized by initiating levodopa at a low dose and titrating slowly to the desired clinical effect. Persistent nausea and vomiting can specifically be handled by adding supplemental doses of carbidopa (Lodosyn), or using the peripheral dopamine receptor antagonist domperidone (available in Canada and Europe) in doses of 10 mg 30 minutes before the levodopa dose. Postural hypotension can be managed by advising the patient to lie supine at night with the head of the bed elevated and rising slowly. If postural hypotension persists, pharmacologic agents such as fluorocortisone and midodrine may be helpful. If a parkinsonian patient experiences symptomatic orthostatic hypotension, the possibility that he or she suffers from MSA with autonomic involvement should be considered.

Motor complications that develop in association with chronic levodopa therapy are the most disabling side effect for most patients. In the early stages of PD, the duration of benefit following a single dose of levodopa is long lasting and far exceeds the plasma half-life of the drug (60 to 90 minutes) (32). This has been ascribed to the relatively preserved capacity of presynaptic dopaminergic terminals of nigrostriatal neurons to store dopamine and regulate its release. However, after a few years of levodopa therapy, there is further neuronal degeneration, and the duration of benefit following each dose of levodopa is shortened in duration. Thus, patients begin to fluctuate between periods of good motor function (“on” responses) and periods of poor motor function (“off” responses) (33). Further, the periods of good motor function that characterize “on” periods now becomes complicated by involuntary movements known as dyskinesia. These are usually choreiform in nature and occur in association with the peak plasma concentration of the drug. However, they may be dystonic or myoclonic in nature, and occur at the onset and termination of the “on” response. In this situation they are referred to as diphasic dyskinesia (34). Manipulating the dose and frequency of levodopa administration is the usual therapeutic approach to the onset of motor complications, but this can be difficult because doses high enough to induce a motor benefit may induce involuntary movements, and doses low enough to ameliorate dyskinesia may not be sufficient to provide antiparkinsonian benefit. Eventually, it may become virtually impossible to achieve a dose of levodopa that provides motor benefits without inducing dyskinesia, and patients may cycle between intolerable dyskinesia and intolerable parkinsonism.

In the early stages of motor fluctuation, increasing the half-life of levodopa by coadministration of a COMT inhibitor may be helpful (see COMT Inhibitors , below). Sustained-release formulations of levodopa (Sinemet CR, Madopar HBS) have been developed in the hope that they would better control motor fluctuations; however, the unpredictable intestinal absorption of these preparations makes them difficult to employ in routine practice, especially for patients with complex motor complications. Low-protein diets or redistribution diets with restriction of the protein intake until the later part of the day may provide some short-term benefits by facilitating levodopa absorption and thereby improving motor performance (36). Dyskinesias are difficult to treat medically, other than by lowering the dose of dopaminergic agent, and this in turn can be associated with worsening parkinsonism as described above. Amantadine has been reported to have an antidyskinetic effect (37) (see below). When motor complications are fully developed, medical therapies are for the most part ineffective

and patients may be considered for surgical intervention (see below). Thus, despite the best of existing medical therapy, more than 75% of PD patients eventually experience intolerable disability (35 ,38).

It is currently thought that motor complications in PD are related to both presynaptic and postsynaptic mechanisms. Chase and his colleagues initially postulated that levodopa-related motor fluctuations develop because of the progressive loss of nigrostriatal neurons and a loss of their capacity to store dopamine and buffer fluctuations in plasma levodopa. Indeed, his group demonstrated a progressive shortening of the duration of the motor response following a dose of levodopa in patients with advancing disease, despite the fact that levodopa peripheral pharmacokinetics remain stable in all stages of PD (39 ,40). This “storage hypothesis” presumed that with the loss of dopamine terminals, central buffering capacity is lost, and striatal dopamine levels become dependent on the peripheral availability of levodopa. As a result, the patient's motor state begins to fluctuate in parallel with the fluctuating plasma levodopa levels that accompany intermittent administration of oral levodopa therapy. With increasing disease severity, there is progressive degeneration of dopamine terminals with further loss of their buffering capacity and consequent exposure of striatal dopamine terminals to alternating and pathologically high and low or “pulsatile” levels of dopamine. However, the storage hypothesis cannot account for the fact that apomorphine has similar pharmacokinetic and pharmacodynamic responses with advancing disease severity as does levodopa, even though apomorphine is not stored in dopaminergic terminals (41). This implies that postsynaptic mechanisms must play some role in the pathophysiology of levodopa-related motor complications.

Current evidence indicates that levodopa-induced motor complications are related to a sequence of events that include abnormal pulsatile stimulation of the dopamine receptor by dopaminergic agents with a short plasma half-life, dysregulation of downstream genes and proteins, and altered neuronal firing patterns (see ref. 42 for complete review of this topic; also see below). In support of this concept, it has been shown that motor complications in parkinsonian monkeys are induced by short-acting dopaminergic agents such as levodopa, which induce pulsatile simulation of receptors, but not by long-acting dopamine agonists, which more closely simulate the normal tonic activation of dopamine receptors (43). Indeed, intermittent administration of a short-acting dopamine agonist induces dyskinesia, whereas continuous administration of the same short-acting agonist does not (44). Further, altered expression of genes such as preproenkephalin (PPE) in striatal neurons have been recorded in association with the development of dyskinesia in *N*-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) treated monkeys (45). Finally, levodopa-induced alterations in neuronal firing patterns have been described in dyskinetic monkeys, which include changes in firing bursts and pauses, the degree of neuronal synchrony, and neuronal firing frequency (46).

These observations have led to the development of new therapeutic options and treatment strategies designed to reverse or prevent the development of levodopa-related motor complications. One such approach is based on lesioning the output neurons in the basal ganglia so as to disrupt their abnormal neuronal firing pattern and the communication of misinformation from basal ganglia to motor cortical regions. This concept likely underlies the antidyskinesia effects that have been observed with surgical therapies for PD (47 ,48) (see Surgical Therapies for Parkinson's Disease , below). A second approach has been directed at modulating dysregulated signaling pathways in striatal neurons leading to abnormal phosphorylation of *N*-methyl-D-aspartate (NMDA) receptors (49). This concept has led to studies of NMDA receptor antagonists and other agents that interfere with intracellular signals as a means of treating levodopa-induced motor complications in PD (37 ,49 ,50). Finally, and perhaps most importantly from the standpoint of the clinician, are the use of long-acting dopaminergic agents based on attempts to provide more physiologic continuous dopaminergic stimulation to striatal dopamine receptors and avoid pulsatile stimulation of striatal dopamine receptors (51). Indeed, prospective double-blind clinical trials have now demonstrated that PD patients randomized to initiate therapy with a dopamine agonist have significantly reduced risk of developing motor complications compared to those randomized to start with levodopa (30 ,31).

Levodopa can also be associated with dose-related sedation, and neuropsychiatric problems such as hallucinations and confusion. These are most likely to occur in the elderly and in patients with preexisting cognitive impairment. The initial hallucinatory episodes usually consist of benign visual hallucinations that are often a greater source of concern to the family than the patient, who retains insight into the nature of the problem. However, patients who experience levodopa-induced hallucinations are more likely to go on and develop dementia (52). An Alzheimer-type dementia can occur in as many as one-third of PD patients, particularly if they have onset after the age of 70 years. Managing hallucinations and confusion in a levodopa-treated patient (1) involves (a) ruling out other temporary causes of mental dysfunction, such as infection, electrolyte imbalance, other brain lesions; (b) elimination of nonparkinsonian medications that are not essential and can impair cognition, (c) elimination of antiparkinsonian drugs that are prone to causing delirium such as anticholinergics, amantadine, selegiline, and dopamine agonists; thereafter the levodopa dose should be reduced to the lowest dose that provides satisfactory control of mobility; and (d) finally, low-dose therapy with atypical neuroleptics can be considered. Clozapine is an atypical neuroleptic that has minimal parkinsonian effects and has been found to be useful in the treatment of psychotic symptoms in PD patients (53). Low doses are

frequently all that is required to provide benefit for PD patients. Accordingly, treatment is initiated with a dose of 12.5 to 25 mg at night and slowly and modestly increased to the desired effect. Hallucinations can usually be controlled with doses less than 25 mg daily. Clozapine is associated with a small risk of hematologic side effects and periodic monitoring is required. Respiradone (Respiradol), olanzapine (Xyprexa), and quietapine (Seroquel) are alternative atypical neuroleptics, but they have been less thoroughly studied than clozapine for PD psychosis, and anecdotal reports suggest that they are no more, and possibly less, effective.

Sudden withdrawal of levodopa can be associated with sudden deterioration in parkinsonian features and may precipitate a life-threatening neuroleptic malignant syndrome (54 ,55). Abrupt reduction of dopaminergic therapy is rarely indicated in the modern era and should be performed in a setting where appropriate monitoring can be performed. Diagnosis is based on altered consciousness, fever, increase in rigidity and other extrapyramidal signs, autonomic instability, elevated creatine kinase level, and leukocytosis. Treatment involves supportive measures (hydration and muscle relaxants) and reintroduction of dopaminergic therapy. PD can also be associated with features that do not respond to levodopa and can themselves be a major source of disability to the patient. These include dementia, autonomic dysfunction, sensory complaints, and freezing episodes in which patients experience arrests in mobility lasting seconds to minutes in duration.

Finally, there has been some concern that despite its many benefits, levodopa might accelerate neuronal degeneration through the oxidizing species generated through its oxidative metabolism. In particular, levodopa is oxidized by MAO to form peroxides, which can combine with iron to generate the cytotoxic hydroxyl radical (56). Levodopa has been shown to induce degeneration of cultured dopaminergic neurons (57). It is less clear that levodopa induces toxicity in animal models, where it has been shown to induce SNc damage in some studies (58) but not in others (59) where there is even the suggestion that it might be protective. Levodopa has not been shown to induce damage to dopamine neurons in normal animals or humans, but the situation may be different in PD, where the SNc is in a state of oxidant stress and defense mechanisms are compromised. A recent consensus conference concluded that although the possibility that levodopa might be toxic in PD has not been excluded, there was no reason to withhold the medication for this reason based on present evidence (60).

In the United States, levodopa is most frequently administered as Sinemet, which is available in dosages of 10/100, 25/100 and 25/250 (the first number represents the dose of carbidopa in mg and the second number the dose of levodopa in mg). Madopar is available in doses of 12.5/50 and 25/100. Long-acting formulations of both of these drugs are available: Sinemet CR in doses of 25/100 and 50/200 and Madopar HBS in a dose of 25/100. Liquid formulations of levodopa can be made by adding water and ascorbate to a Sinemet tablet, but these must be made fresh and offer little additional advantage for most PD patients. Rapidly absorbed methyl and ethyl ester formulations of levodopa are currently being assessed experimentally.

In summary, levodopa continues to be an important component of the therapeutic armamentarium for PD, but it is associated with troublesome complications and some parkinsonian features do not respond. Theoretically, levodopa could accelerate neuronal degeneration through oxidizing species generated by its oxidative metabolism, but there is little evidence to suggest that this is a concern in PD, and most physicians do not restrict the use of levodopa for this reason (60). On the other hand, current research suggests that if PD therapy is initiated with a dopamine agonist and levodopa is reserved until satisfactory benefits can no longer be controlled with the agonist alone, patients can enjoy comparable motor benefits and reduced motor complications in comparison to when levodopa is administered on its own (see Dopamine Agonists , below). There is also considerable interest in administering levodopa in conjunction with a COMT inhibitor to enhance its duration of effect and thereby improve motor response and reduce the risk of the drug inducing pulsatile stimulation of the dopamine receptor (see COMT Inhibitors , below).

Dopamine Agonists

Dopamine agonists are a group of drugs that act directly on dopaminergic receptors. Historically, they have been used as adjuncts to levodopa in the treatment of PD since the 1970s (61) and offer several theoretical advantages over levodopa (62): (a) They do not depend on enzymatic conversion for activity, i.e., they do not depend on the integrity of the nigrostriatal neurons, such that they should be active even in advanced stages of PD, at which time presynaptic dopamine neurons and terminals are largely degenerated. (b) They can be designed to stimulate specific subtypes of dopamine receptors, which may lead to selective functional responses. (c) Most marketed dopamine agonists have longer half-lives and longer durations of action than levodopa. This may permit more continuous (less pulsatile) stimulation of dopamine receptors than occurs with levodopa therapy. Therefore, there has been interest in the potential of this class of drug to reduce the risk of developing levodopa-related motor complications (62). (d) They do not undergo oxidative metabolism and do not generate free radicals that might promote degeneration of remaining nigrostriatal neurons. There are data now indicating that dopamine agonists can scavenge free radicals and protect dopamine neurons in *in vitro* and *in vivo* models of PD (63 ,64). Therefore, there has been interest in the potential of dopamine agonists to provide neuroprotective effects in PD (65).

Five dopamine agonists, bromocriptine (Parlodel), pergolide

(Permax), cabergoline (Cabsar, Dostinex), pramipexole (Mirapex) and ropinirole (Requip), are currently available for use in the United States and in many other countries throughout the world; cabergoline is approved for the treatment of PD in Japan and several European countries, but it has been exclusively marketed for suppression of lactation in the United States. Lisuride, piribedil, and apomorphine are other dopamine agonists that are available in some countries but not the United States. All dopamine agonists that are marketed for the treatment of PD stimulate the D2 receptor, which is thought to underlie their antiparkinsonian effects. Dopamine and apomorphine stimulate both D1 and D2 receptors. Pergolide is also a weak agonist and bromocriptine a weak antagonist of the D1 receptor. The role of D1 receptor activation or inhibition in PD is not known, although there is some suggestion that stimulation of both D1 and D2 receptors provides enhanced motor responses. Bromocriptine, pergolide, ropinirole, and pramipexole have plasma half-lives of 6 to 15 hours, whereas cabergoline has a much longer elimination half-life of 63 to 69 hours. This contrasts with the plasma half-life of levodopa, which is 60 to 90 minutes.

Dopamine Agonists in Patients with Advanced PD

Since their introduction in the mid-1970s, dopamine agonists have primarily been used as adjuncts to levodopa in PD patients with relatively advanced disease who have begun to experience motor complications (66). As an adjunct to levodopa, numerous prospective double-blind studies have demonstrated that dopamine agonists can significantly improve PD signs and symptoms, reduce dyskinesia and motor fluctuations, and reduce the need for levodopa therapy in comparison to placebo (67 ,68 ,69 ,70 ,71 ,72 and 73). Benefits have been observed with each of the currently approved dopamine agonists and they are of approximately equal magnitude. Apomorphine stimulates both D1 and D2 receptors. It has a very short latency to onset, but also a short duration of benefit. It has been used to provide a “rescue effect” for patients who turn “off” and do not respond to their next dose of levodopa (74). Some physicians have reported benefits in advanced patients with complex motor complications with the use of continuous apomorphine (75). However, apomorphine must be administered parenterally, is associated with cutaneous ulcerations at sites of entry, and is very difficult to manage for both the physician and the patient. It therefore has relatively little role in routine practice.

Despite the benefits obtained with dopamine agonists in patients with advanced disease, they generally do not provide satisfactory control of motor function or motor complications, and sooner or later alternate therapies must be sought.

Dopamine Agonists in Patients with Early PD

As discussed in the section on levodopa-related motor complications (see above), there is growing evidence suggesting that pulsatile stimulation of dopamine receptors due to the use of short-acting dopaminergic agents contributes to the emergence of motor complications. Studies in MPTP-treated primates demonstrate that bromocriptine and ropinirole are associated with reduced frequency and severity of dyskinesia compared to levodopa, even though all groups provide comparable behavioral effects (76 ,77). These data suggest that starting treatment for PD patients with a long-acting dopamine agonist rather than levodopa might reduce the risk of developing motor complications. However, until recently dopamine agonists have not been well studied in early PD. There are now prospective double-blind controlled studies demonstrating that both pramipexole and ropinirole provide improvement in measures of motor functions and activities of daily living (ADL) in otherwise untreated PD patients that are superior to placebo (78 ,79), and almost as good as levodopa (30 ,31). Further, PD patients can be maintained on dopamine agonist monotherapy without supplemental levodopa for a mean of 3 years (80). More importantly, it has now been established in prospective double-blind long-term studies that PD patients randomized to initiate therapy with a dopamine agonist (ropinirole or pramipexole), supplemented with levodopa if necessary, have significantly fewer motor complications than patients randomized to begin therapy with levodopa alone (30 ,31). Reduced rates of both dyskinesia and motor fluctuations were observed in the agonist-treated patients. Measurements of motor function and ADL on the Unified Parkinson Disease Rating Scale (UPDRS) showed slight, but significant, benefits in favor of levodopa-treated patients in both studies. This is difficult to explain, as patients in both groups could have added open label levodopa to their blinded treatment regimen if either the physician or the patient thought it was necessary. This raises the question as to whether the UPDRS fully captures all factors that contribute to PD disability.

Based on these new studies and the concept of continuous dopaminergic stimulation, many authorities now recommend initiating symptomatic therapy for PD with a dopamine agonist, and reserving levodopa until such time as the agonist can no longer provide satisfactory clinical control (51 ,81 ,82 and 83). Others feel that the issue is still somewhat controversial and that physicians must choose between enhanced efficacy now versus delayed motor complications later. Our personal view is that the difference in motor and ADL scores between the agonist and levodopa groups is negligible, whereas the difference in the rate of developing motor complications is substantial and a much greater source of disability for the patient and frequently necessitates surgical intervention as the only means of providing satisfactory control. Accordingly, we favor initiating therapy

with a dopamine agonist in appropriate patients to diminish the risk that disabling motor complications will ensue. We still favor the use of levodopa as the initial agent in patients with cognitive impairment or who are elderly.

Adverse Effects of Dopamine Agonists

The acute side effects of dopamine agonists are similar to those observed with levodopa and include nausea, vomiting, and postural hypotension (84). These side effects tend to occur when treatment is initiated and abate over days or weeks as tolerance develops. Introducing the agonist at a low dose, and slowly titrating to the desired effect reduces the probability that they will occur. Dopamine agonists can acutely cause or intensify dyskinesias, but in the long term they have the potential to lessen dyskinesias and motor fluctuations because of their long duration of action (see above). Psychiatric complications (hallucinations, confusion) may occur and tend to be more pronounced than bioequivalent doses of levodopa (30 ,31). The ergot-derived dopamine agonists, bromocriptine, pergolide, and cabergoline, may have ergot-related side effects including pleuropulmonary and retroperitoneal fibrosis, erythromyalgia, and digital vasospasm, although these are rare (84). The newer non-ergot dopamine agonists are less likely to induce these problems, although there is anecdotal suggestion that they may still occur. Dose-related sedation may occur with dopamine agonists (69 ,78), as with other dopaminergic agents including levodopa. More recently, sudden episodes of unintended sleep while at the wheel of a motor vehicle have been described in PD patients and attributed to dopamine agonists (85). The episodes were termed “sleep attacks” because they occurred suddenly, although others have argued that there is no evidence to support the concept of a sleep attack even in narcolepsy. They have suggested that it is more likely that these patients have unintended sleep episodes as a manifestation of excess daytime sedation due to nocturnal sleep disturbances that occur in 80% to 90% of PD patients and to the sedative effect of dopaminergic medications (86). It is now apparent that these types of episodes can be associated with all dopaminergic agents including levodopa (87). Physicians should be aware of the potential of dopaminergic agents to induce sleepiness, and that patients themselves may not be aware that they are sleepy. To detect excess sleepiness and to thereby introduce appropriate management strategies, it is necessary to employ sleep questionnaires such as the Epworth sleepiness scale, which inquires into the propensity to fall asleep and does not rely upon subjective estimates of sleepiness (88).

Catechol O-Methyltransferase (COMT) Inhibitors

Orally ingested levodopa is massively transformed in the periphery by two enzymatic systems—AADC and COMT—such that only 1% of a levodopa gains access to the brain. To partially counter this effect, levodopa is routinely prescribed in combination with an inhibitor of AADC that does not cross the blood–brain barrier and blocks the peripheral decarboxylation of levodopa into dopamine. This combination reduces peripheral dopaminergic side effects associated with the administration of levodopa alone, and increases the amount of levodopa that is available to access the brain. However, even in the presence of a decarboxylase inhibitor, the bulk of levodopa is still metabolized by COMT and only 10% of a given dose is transported into the brain (17 ,89). Two new drugs that inhibit COMT, tolcapone (Tasmar) and entacapone (Comtan), have recently been introduced to the market as an adjunct to levodopa therapy. Both drugs inhibit COMT in the periphery, although tolcapone has mild central effects as well. Entacapone and tolcapone increase the elimination half-life of levodopa by approximately 40% without modifying the peak plasma concentration of levodopa (C_{max}) or the time to reach peak plasma concentration (*Tmax*) and effects are seen with both immediate and controlled release formulations (90 ,91 ,92 and 93). COMT inhibitors thus modulate peak and trough plasma levodopa concentrations, leading to a smoother plasma curve with reduced fluctuations in levodopa level (94). These pharmacokinetic effects have been shown to translate into enhanced levodopa entry into the brain on positron emission tomography (PET) (95) and clinical benefits particularly for patients experiencing mild to moderate motor fluctuations. Double-blind placebo-controlled clinical trials in fluctuating PD patients demonstrate that COMT inhibitors increase the duration of beneficial effect following a single levodopa dose (96). They also provide an increase daily “on” time of 15% to 25%, a decrease in “off” time of 25% to 40%, improvement in UPDRS motor scores, and a reduction in levodopa dose requirement of 15% to 30% (97 ,98 ,99 and 100). Benefits with COMT inhibitors have also been observed in nonfluctuating PD patients with a stable response to levodopa. Two placebo-controlled trials showed improved motor scores and reduced levodopa dose requirements in the group receiving the COMT inhibitor (101 ,102).

There has also been interest in using COMT inhibitors from the time levodopa is first initiated in order to reduce the risk of developing motor complications (103). As described in the section on motor complications, laboratory evidence supports the notion that treatment for PD patients should be employed in such a way as to try and avoid pulsatile stimulation of dopamine receptors (51). Indeed, there is now evidence indicating that initiating therapy with a long-acting dopamine agonist reduces the risk of dyskinesia and motor fluctuations (30 ,31). However, these patients eventually require levodopa, and when levodopa is administered the frequency of motor complications increases. It therefore has been postulated that administering levodopa from the time it is first introduced with a COMT inhibitor

to extend its half-life and deliver levodopa to the brain in a more continuous fashion might further reduce the risk of motor complications. Based on a similar hypothesis, studies comparing controlled-release levodopa to regular levodopa failed to demonstrate any difference between the two formulations (104 ,105). However, controlled-release formulations have variable absorption and do not provide stable plasma levels of levodopa. Further, the drug was prescribed twice daily in these studies, and that may not have been frequent enough to prevent fluctuations in plasma levodopa concentrations. Clinical trials to test this hypothesis using entacapone as an adjunct to levodopa are currently being planned.

Side effects associated with COMT inhibitors are primarily dopaminergic and reflect enhanced delivery of levodopa to the brain. Dyskinesia is the most common, but nausea, vomiting, and psychiatric complications may occasionally occur. Both the benefits and dopaminergic adverse effects develop within hours to days after initiating treatment. In general, they are easy to manage by simply reducing the dose of levodopa (by approximately 15% to 30%), not the dose of the COMT inhibitor. Dyskinesia is more likely to be a problem in patients who already experience dyskinesia, and the need for a levodopa dose reduction can be anticipated in these patients. An explosive diarrhea has been seen in 5% to 10% of tolcapone-treated and necessitates discontinuing the drug. This has been much less of a problem with entacapone and rarely requires stopping the drug. Brownish-orange urine discoloration may occur with either drug due to accumulation of a metabolite. This is a benign condition, but patients should be advised that it may occur.

Of greater seriousness is the problem of liver toxicity that has been reported in association with tolcapone (106). No evidence of liver dysfunction was detected in preclinical toxicity studies, but in clinical trials elevated liver transaminase levels were observed in 1% to 3% of patients. For this reason, liver monitoring was required. Following approval of the drug, there have been reports of four cases of severe liver dysfunction leading to the death of three of the individuals (106 ,107). These observations led to the drug being withdrawn from the market in Europe and Canada and to the issuance of a “black box” warning in the United States (108). This requires biweekly monitoring of liver enzymes for the first 12 weeks, monthly monitoring thereafter, and discontinuation of the drug if liver enzymes are elevated above normal on a single occasion. No preclinical toxicity, clinical trial, or postmarket reports of liver dysfunction have been described to date with entacapone, and no laboratory monitoring is required with its use (109).

Entacapone is typically administered in a dose of 200 mg with every scheduled dose of levodopa, whereas tolcapone is administered at a dose of 100 or 200 mg three times daily. No comparative studies between entacapone and tolcapone have been performed, but pharmacokinetic and clinical trial data indicate that tolcapone is the more potent agent. However, because of the greater risk of hepatotoxicity and diarrhea, entacapone has become the more widely employed COMT inhibitor. It should be emphasized that COMT inhibitors provide antiparkinsonian benefit only when used as an adjunct to levodopa. By themselves they have no effect.

In conclusion, COMT inhibitors represent an important advance in the medical treatment of PD and may be useful in all stages of the illness (110). Used in combination with levodopa, they extend the half-life of levodopa, smooth the plasma levodopa concentration curve, and enhance clinical dopaminergic benefits. They have been established to provide benefit in PD patients with motor fluctuations, although particular care must be taken in managing the more advanced patients with severe dyskinesia, and this is usually best left to the Parkinson specialist. There are preliminary data suggesting that they enhance motor function in the milder patient with a stable response to levodopa, and this is being further evaluated. Finally, there is good evidence to suggest that administering levodopa with a COMT inhibitor from the time it is first introduced may prevent pulsatile stimulation of dopamine receptors and minimize the risk of developing motor complications. The drugs are easy to use and require no titration. Dopaminergic side effects tend to occur within days and can be managed by tapering the levodopa dose. Because of the restrictions in the use of tolcapone due to liver toxicity, entacapone is now the COMT inhibitor of choice. It is likely that a single tablet will soon be developed that contains the combination of levodopa, an AADC inhibitor, and a COMT inhibitor.

Other Antiparkinson Agents

Anticholinergics

Anticholinergic drugs were first used as a treatment for PD in the 1860s, using extracts from the alkaloids *Atropa belladonna* and *Hyscyamus niger*, which contain hyosciamine and scopolamine (111 ,112). Synthetic anticholinergic drugs were developed in the 1940s, and they became the mainstay of PD treatment until the emergence of levodopa (113 ,114). These drugs have largely been replaced by the newer antiparkinsonian drugs, but are still used occasionally in the modern era particularly for the treatment of tremor (115). The main anticholinergic agents currently in use are trihexyphenidyl (Artane), benztropine (Cogentin), biperiden (Akineton), orphenadrine (Disipal), and procyclidine (Kemadrin). An interaction between dopaminergic and cholinergic neurons in the basal ganglia has long been recognized, and classic experiments demonstrated the capacity of cholinergic agents to worsen and anticholinergic agents to improve parkinsonian features (116). Cholinergic agents have been shown to block dopamine reuptake into presynaptic dopaminergic terminals (117) and dopamine receptor activation has been shown to regulate acetylcholine release (118).

More recent work has demonstrated that dopamine-regulated neuropeptide (preproenkephalin) expression in striatal neurons is regulated by cholinergic interneurons (119). Despite these observations, the relationship between the cholinergic and dopaminergic systems is poorly understood, as is the basis for the clinical benefits that are seen with anticholinergic agents.

Clinical studies demonstrate that anticholinergic agents provide a 10% to 25% improvement in rest tremor, whereas akinesia and postural impairment are not affected (120). In practice, anticholinergic agents can be used in early PD patients to treat tremor when it is the predominant complaint and to delay the introduction of levodopa, provided that cognitive function is preserved and that the patient does not have narrow angle glaucoma or orthostatic hypotension (see General Adverse Effects of DBS , below). Trihexyphenidyl is the most widely used anticholinergic agent in PD, although head-to-head comparisons have not been performed. The usual trihexyphenidyl doses range from 0.5 to 1 mg b.i.d. initially, with gradual increase to 2 mg t.i.d. Benztropine is also commonly used, with doses ranging from 0.5 to 2 mg b.i.d.

Side effects are a major limiting factor with respect to the use of anticholinergic drugs in PD. The most important of these are central, and consist of memory impairment, confusion, hallucinations, sedation, and dysphoria (115). These tend to be most pronounced in older individuals with some preexisting cognitive impairment, but can affect young patients with seemingly intact mentation as well. Peripheral side effects include dry mouth, dysuria, constipation, dizziness due to orthostatic hypotension, tachycardia, nausea, blurred vision, and decreased sweating. Anticholinergic agents should be avoided in patients with narrow angle glaucoma, and caution is required in using them in patients with prostatic hypertrophy because of the risk of inducing acute urinary retention. Anticholinergic drugs can enhance levodopa-induced choreiform dyskinesias, and orobuccal dyskinesias have been reported with anticholinergic therapy alone (121). If the decision is made to discontinue anticholinergics, this should always be done gradually to avoid withdrawal effects and acute exacerbation of parkinsonism (122).

Peripherally active anticholinergic drugs are also used in PD. Anticholinergic agents that are relatively selective for bladder cholinergic receptors such as tolterodine tartrate (Detrol), and oxybutynin (Ditropan) can be used to treat bladder instability (123). Anticholinergic agents that are relatively selective for salivary gland receptors such as glycopyrrolate (Robinul) can be used to treat sialorrhea.

Because of their adversity profile, and particularly their tendency to induce cognitive impairment, anticholinergic agents are not commonly used in the treatment of PD. They are perhaps most frequently used in younger PD patients with tremor-dominant PD. However, there is evidence suggesting that levodopa and other dopaminergic agents provide antitremor effects that are just as good as or superior to anticholinergic agents (124). Certainly when these agents are employed, side effects should be sought and the drug discontinued when they occur.

Amantadine

The discovery of the antiparkinson properties of the antiviral agent amantadine (Symmetrel) was fortuitous (125). The primary mechanism of action of amantadine in PD is not established with certainty. The drug has been described to increase dopamine release, block dopamine reuptake, and stimulate dopamine receptors. It has also been shown to have anticholinergic effects and weak NMDA receptor antagonist properties (126 ,127 and 128). Improvement in akinesia, rigidity, and tremor, as well as reduction in choice reaction time, have been described in uncontrolled studies, particularly in mildly affected PD patients (125 ,129 ,130 and 131). In comparison to anticholinergic drugs, amantadine was found to have a greater effect on akinesia and rigidity but lesser benefit for tremor (132).

With the recognition that amantadine provides NMDA receptor antagonism (128), there has been interest in the notion that it might have antidyskinetic and even neuroprotective effects. The potential of the drug to interfere with dyskinesia is based on the notion that dyskinesias are related to excessive phosphorylation of NMDA receptors on striatal neurons due to loss of dopamine-mediated modulatory effects (49). Studies in monkeys show that NMDA receptor antagonists can improve dyskinesia (50). Preliminary clinical trials suggest that the same is true in some PD patients (37 ,134), and this has now been confirmed in a double-blind controlled study (135). The potential of amantadine to provide neuroprotective effects is based on evidence suggesting that excitotoxicity contributes to neuronal degeneration in PD (136 ,137). Indeed, one retrospective study did suggest that there was an increase in the survival of PD patients that had been treated with amantadine (138).

The elimination half-life of amantadine is 10 to 30 hours, and the medication is typically administered in dosages of 100 mg two to three times per day. Unfortunately, amantadine is frequently associated with dose-related cognitive problems including confusion, hallucinations, insomnia, and nightmares that limit its usefulness. Amantadine has also been associated with livedo reticularis, ankle edema, and peripheral neuropathy. If amantadine must be withdrawn, it should be done gradually as some patients may experience dramatic worsening of PD on withdrawal.

In conclusion, amantadine can be used in the initial stages of PD to provide some symptomatic benefit and to delay the need for levodopa. It can also be used as an adjunct to levodopa to try to control levodopa-induced dyskinesia. Cognitive side effects limit the usefulness of this drug, and mental status must be closely monitored particularly in patients with advanced disease or preexisting cognitive impairment.

As it is difficult to withdraw in many instances, many physicians do not use this drug as a first-line therapy.

Selegiline

Selegiline (Deprenyl, Eldepryl) is a relatively selective inhibitor of monoamine oxidase-B (MAO-B). It was approved in PD as an adjunct to levodopa that provides a modest increase in “on” time in fluctuating patients with advanced PD (139). However, it is primarily used in the treatment of early PD patients as a putative neuroprotective agent. This was based on two important observations that suggested that an MAO-B inhibitor might alter the natural course of PD. First, the neurotoxin MPTP causes parkinsonism (140) by way of an MAO-B–catalyzed oxidation reaction forming the toxin MPP^+ (141), and second, dopamine is oxidized by MAO-B to generate peroxides and other potentially cytotoxic oxidizing species (56). In the laboratory, selegiline has been shown to protect nigral dopaminergic neurons in cell cultures and in MPTP-treated animals (142 ,143). Prospective double-blind clinical trials in previously untreated PD patients have demonstrated that selegiline delays the emergence of clinical dysfunction as determined by the need for levodopa and the progression of parkinsonian signs and symptoms (144 ,145). However, post hoc analyses have demonstrated that selegiline has symptomatic effects that might account for these benefits. These confound interpretation of these studies (146). In addition, the disease continues to progress, and initial benefits do not appear to persist (147 ,148).

Although there remains equipoise with respect to the possible beneficial effects of selegiline, it is now clear that the drug has clear neuroprotective effects for dopaminergic neurons in both *in vitro* and *in vivo* laboratory models (see ref. 149 for review). Further, it is now clear that neuroprotection with selegiline does not depend on MAO-B inhibition (150 ,151), and is mediated by the drug's metabolite desmethyl selegiline (DMS) (152). Work by Tatton's group (153 ,154 and 155) has now shown that DMS and other propargylamines provide neuroprotective effects by binding to the protein glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and preventing its translocation to the nucleus. GAPDH accumulation in the nucleus inhibits BCL-2 expression and promotes apoptosis (153 ,154). These findings, indicating that selegiline is an antiapoptotic drug, are particularly relevant to PD, where there is evidence that cell death occurs by way of an apoptotic process (155).

Selegiline is administered in a dose of 5 mg b.i.d. and is generally well tolerated. In levodopa-treated patients it has the potential to increase dopaminergic side effects and to possibly induce cardiovascular problems. Its amphetamine metabolite can also cause insomnia, and for this reason the second dose is usually not administered after 12 noon. The adversity profile of selegiline has been somewhat clouded by the findings of a 5-year open label study reporting increased mortality in patients receiving the combination of levodopa-carbidopa/selegiline as opposed to those treated with levodopa-carbidopa alone (156 ,157). However, the statistical methods used in this study were questioned (158), and increased mortality has not been confirmed in a metaanalysis evaluating mortality in all other prospective trials of selegiline (159).

In summary, there is theoretical evidence suggesting that selegiline might provide neuroprotective benefits in PD. Clinical trials are consistent with this notion, but might be explained by the drug's symptomatic effects. The drug is generally well tolerated, and claims of increased mortality have not been substantiated. It remains a matter of judgment and personal philosophy as to whether or not to use selegiline as a putative neuroprotective drug.

SURGICAL THERAPIES FOR PARKINSON'S DISEASE

Part of "124 - Current and Experimental Therapeutics of Parkinson's Disease "

In the past few years, the renaissance of functional neurosurgery has transformed our vision of PD therapy. Functional neurosurgery for movement disorders dates back to the beginning of the 20th century, with the introduction of pyramidal tract lesions or dorsal root sections (160 ,161 and 162). These were unfortunately characterized by their unacceptable morbidity. Lesions of the basal ganglia as a treatment for PD were introduced by Meyers in the early 1940s (163 ,164). These procedures provided some benefits for tremor and rigidity, but adverse events were common and there was an unacceptably high mortality rate ranging from 8% to 41% (162 ,163 ,164 ,165 ,166 ,167 and 168). Surgery therapies for PD became more widely accepted with the introduction of stereotactic techniques (169) and the determination that lesions of the thalamus could provide benefits with fewer adverse events (170). With the introduction of levodopa, surgery for PD was almost abandoned. However, the shortcomings of classic levodopa therapy (as discussed above), the tremendous advances in brain imaging and intraoperative monitoring techniques for target localization, and insights into the pathophysiology of the basal ganglia (171 ,172) have catalyzed a dramatic resurgence of interest in surgical procedures for PD. However, most current information is based on open trials that do not control for placebo effects and physician bias, and may thus have overstated the benefits that can be achieved (173). There is little doubt that surgical techniques offer the potential to provide benefit to PD patients with advanced disease who cannot be controlled with medical therapies, but well-designed placebo-controlled double-blind studies are required in order to determine their true value (174).

Ablative Procedures

Thalamotomy

Current knowledge of basal ganglia physiology in the normal and PD state suggest several targets for ablative procedures,

one of which is the ventral intermediate nucleus (VIM) of the thalamus. Metabolic and physiologic studies consistently indicate that the ventral anterior and ventral lateral thalamic nuclei, the STN and the GPi, are overactive in PD (175 ,176), probably reflecting increased inhibitory output from the GPi. Cooper (170) and Hassler and Riechert (177) noted in the 1950s that thalamic lesions could relieve contralateral tremor. Their experience led to thalamotomy becoming the preferred surgical procedure for the treatment of tremor-predominant forms of PD. Their choice of target was facilitated and supported by electrophysiologic studies demonstrating abnormal tremor synchronous electrical activity in this region (178 ,179 and 180). Ohye and Narabayashi (181 ,182) subsequently used these techniques to conclude that lesions in the VIM nucleus of the thalamus were most effective in reducing contralateral tremor. Studies reported a consistent reduction in contralateral tremor, but it is less certain that there are benefits with regard to rigidity, and there is virtually no benefit for more disabling features such as bradykinesia or postural impairment (183 ,184 ,185 and 186). These studies also suggested that thalamotomy has the potential to reduce or prevent the development of levodopa-induced dyskinesia (183 ,184 ,187 ,188 and 189). In recent studies, persistent morbidity associated with unilateral thalamotomy occurs in less than 10% of patients. Complications includes dysarthria, dysequilibrium, contralateral hemiparesis or hemiataxia, cognitive impairment, and personality changes (185 ,190). Bilateral thalamotomy is associated with a further increase in morbidity including dysarthria, dysphagia, and cognitive impairment, and is usually avoided (185 ,190).

In conclusion, for a select group of PD patients with disabling tremor that cannot be controlled with medications and marked unilateral predominance, thalamotomy can still be considered; however, this technique has now been largely replaced by a different surgical procedure (deep brain stimulation) and alternate targets (STN or GPi) (see below).

Pallidotomy

Despite the early encouraging reports (163 ,164), lesioning the pallidum or its efferent fibers fell out of favor and was replaced by thalamotomy. However, Leksell persisted in developing this surgery and was able to determine that lesions placed in the posteroventral portion of the globus pallidus pars interna (GPi) were the most beneficial in relieving PD signs and symptoms (191). A generation later, his pupil Laitinen performed Leksell's technique using more modern stereotactic techniques, and described benefits with respect to bradykinesia, rigidity, and levodopa-induced dyskinesias (192 ,193). Complications were observed in 14% of patients and included partial homonymous hemianopsia, transient dysphasia, and facial weakness. These results followed shortly after neurophysiologic studies demonstrating that the pallidum was hyperactive in PD (171 ,194). These results prompted a renaissance in pallidotomy as a surgical option for PD. Using the posterolateral pallidum as a target, several surgical groups have now reported benefits in PD patients (195 ,196 and 197). The most dramatic finding is a consistent long-lasting abolition of contralateral dyskinesia; antiparkinsonian benefits are more modest (198 ,199). Complications occur in 3% to 10% of patients and are primarily visual in nature, although cognitive impairment, sensory deficits, and motor weakness may all occur. Bilateral pallidotomy is associated with increased risk of disabling dysphagia, dysarthria, and cognitive impairment (200 ,201), and has largely been abandoned with the availability of stimulation procedures. Current pathophysiologic models of PD explain the improvement in parkinsonism, but do not explain the striking antidyskinetic effect of pallidotomy (202). It has been proposed that the antidyskinetic effect of pallidotomy may be due to elimination of an abnormal firing pattern in pallidal output neurons that are providing misinformation to cortical motor regions that result in the emergence of dyskinesia (42).

In summary, unilateral pallidotomy provides consistent and dramatic improvement in contralateral levodopa-induced dyskinesia. However, improvement in parkinsonian features is modest and the procedure is associated with lesion-related side effects, especially when it is performed bilaterally. Here, too, it is being replaced by stimulation procedures in many centers (173).

Subthalamotomy

Physiologic and metabolic studies demonstrate that the subthalamic nucleus (STN), similar to the GPi, is overactive in parkinsonian syndromes (171 ,175 ,202 ,203). This has led to the notion that lesions of the STN might provide benefits in PD. Indeed, subthalamotomy has been shown to improve parkinsonian features in MPTP-treated monkeys (204 ,205). However, lesions of the STN are associated with hemiballismus, and accordingly physicians have been reluctant to perform this procedure in PD patients. Deep brain stimulation procedures avoid the need to make lesions in target structures (see below), and stimulation of STN is associated with marked improvement in parkinsonian features. Preliminary studies of subthalamotomy have been performed and indicate that it also can provide excellent benefits in PD with minimal adversity (206). Nevertheless, until further experience has been gained with respect to the long-term safety and efficacy of this procedure, it must be considered experimental.

Deep Brain Stimulation Procedures

High-frequency deep brain stimulation (DBS) was introduced by Benabid and his group (207) alternate to ablative procedures. Benabid et al. noted that high-frequency stimulation of selected brain targets simulates the effects of a

lesion without the necessity of making a destructive brain lesion. In this procedure, an electrode is implanted into the desired brain target and connected to a stimulator placed subcutaneously over the chest wall. DBS has several advantages over ablative procedures: (a) It avoids the need to make a destructive brain lesion. Side effects due to stimulation can be reversed by changing the stimulator settings. (b) Bilateral procedures can be performed with relative safety. (c) Stimulator settings can be adjusted as with the doses of a medication to maximize benefit and minimize adversity. The precise mechanism of action of DBS is unknown, but it may involve jamming abnormal firing patterns of nerve cell populations within the stimulated area. Other possible mechanisms include depolarization blockade, release of inhibitory neurotransmitters, and indirect effects due to backfiring with stimulation of distant cell populations through orthodromic or antidromic firing.

Deep Brain Stimulation of the VIM of the Thalamus (DBS-VIM)

The initial trials of DBS were performed in the VIM nucleus of the thalamus. The procedure provided prominent antitremor effects in the vast majority (80% to 90%) of patients with tremor predominant PD and essential tremor (208). Tremor arrest occurrs within seconds following the onset of stimulation, and the effect is lost within seconds of its cessation. These results were confirmed in a double-blind crossover study (209) that led to the approval of unilateral DBS-VIM as a treatment for essential or parkinsonian tremor by the Food and Drug Administration (FDA) in the United States. Interestingly, stimulation slightly posterior and medial to the VIM—close to the centromedian and parafascicular complex of the thalamus—also induced reduction in levodopa-induced dyskinesias (210). Unfortunately, DBS-VIM does not meaningfully improve the more disabling features of PD such as bradykinesia and gait impairment. This shortcoming has led to consideration of other targets for DBS, such as the GPi and the STN (see below). DBS-VIM remains a very valuable procedure for PD patients for whom tremor is the main handicap.

Deep Brain Stimulation of the Subthalamic Nucleus (DBS-STN)

A large body of experimental evidence has pointed toward targeting the STN as a treatment for PD: (a) neurons in the STN are hyperactive in PD (203 ,211); (b) lesions of the STN provide benefit to MPTP-treated primates (204 ,205); (c) improvement in contralateral parkinsonism following a spontaneous hemorrhage into the STN of a PD patient (212); and (d) improvement in MPTP-treated monkeys following stimulation of the STN (213). Based on these findings, DBS-STN was introduced as a treatment for PD patients (214 ,215 and 216). Significant benefits of stimulation have been reported for all of the cardinal features of parkinsonism; these have been confirmed in a double-blind crossover study (217). Improvements in motor function range from 40% to 80%. Highly significant benefits have also been observed in home diary assessments of percent “on” time without dyskinesia, leading to a dramatic reduction in patient disability. This is all the more remarkable when one considers that these benefits have been obtained in a population of patients that could not be further improved with medical therapy. Interestingly, dyskinesias have not been a problem, which may be related to disruption of the abnormal firing pattern in STN neurons. Finally, it has recently been proposed that DBS-STN might provide neuroprotective effects by inhibiting STN-mediated excitotoxic damage in its target structures (137). Indeed, lesions of the STN have been shown to protect SNc neurons in 6-hydroxydopamine lesioned rodents (218). It is currently thought that stimulation of the STN is the most effective surgical procedure, but prospective double-blind placebo-controlled studies directly comparing stimulation of the STN to other target structures such as GPi (see below) remain to be performed (173).

Deep Brain Stimulation of the Globus Pallidus Pars Interna (DBS-Gpi)

The experimental rationale for performing stimulation of the GPi is similar to that for STN. As is the case with the STN, the GPi is also overactive in PD (203 ,211), and lesions of the GPi provide benefits in MPTP monkeys (219). Several studies have now reported that DBS-GPi can improve all of the cardinal features of parkinsonism and reduce the severity of levodopa motor complications (220 ,221 and 222). Benefits do not appear to be as potent as with DBS-STN, but a prospective controlled trial has yet to be performed to objectively compare these two targets.

General Adverse Effects of DBS

Adverse effects of DBS can be related to the surgical procedure, the device, and the stimulation itself. Surgical complications involve hemorrhage and infarction and occur in less than 3% of cases. The electrode itself does not seem to be toxic to local tissues, as in the only postmortem pathologic study available, gliosis around the electrode tip was less than 1 mm in diameter (223). Problems associated with the implanted material (infection, dislodgment, mechanical dysfunction) occur in 1% to 3% of cases and may lead to the need to replace the electrode. Stimulation-related side effects include paresthesiae, motor twitch, dysarthria, and eye movement disorders. They are usually transient and controllable by stimulator adjustment. Finally, the battery has limited longevity, ranging from 6 months to 5 years or more, depending on the electrical consumption of the stimulator

settings chosen. The battery in the chest wall can be easily replaced under local anesthesia in most cases.

Despite the potential side effects of the DBS procedure, the risk of permanent side effects is less than with ablative procedures, particularly when bilateral with procedures (224).

Management of DBS

Optimization of stimulator settings is necessary to achieve maximal benefit with DBS procedures. This is not an easy task because of the large number of stimulation variables. These include electrode configuration, amplitude, pulse width, and frequency. Determination of the optimal stimulation settings may be complicated and time consuming (hours) and may require multiple visits. Validation of a rapid and simple method for determining stimulator adjustment will enhance the utilization of these techniques.

In conclusion, DBS of selected brain targets offers PD patients the potential of experiencing clinical benefit when this cannot otherwise be attained with medical therapy. Further, this can be accomplished without the need to make a destructive brain lesion with its accompanying side effects. Studies to determine the long-term safety and efficacy of DBS and the optimal target site for individual patients remain to be performed. Nevertheless, studies performed to date indicate that this procedure has much to offer patients with advanced PD. Based on existing information, DBS-STN appears to provide the best clinical effects and is presently considered to be the stimulation target of choice. It is possible that other brain targets such as the globus pallidus pars externa and selected cortical motor regions will prove superior in the future.

Transplantation Procedures

Yet another approach to the treatment of patients with advanced PD is transplantation of dopaminergic neurons aimed at replacing host neurons that degenerate during the course of the disorder. Transplantation is a rational strategy for treating PD because (a) PD is due to specific degeneration of dopaminergic nigrostriatal neurons and its symptoms are dramatically relieved by dopaminergic treatment; and (b) the striatum, which is denervated in PD, is a well-defined target for transplantation (225). In animal models, fetal nigral neurons have been shown to survive, reinnervate the striatum, produce dopamine, and improve motor dysfunction in rodent and primate models of PD (226 ,227 ,228 and 229).

The first clinical trials in PD patients involved implantation of adrenal medullary cells into the caudate nucleus, but despite the initial encouraging reports (230), the inconsistent outcomes and the associated adverse events led to this procedure being abandoned (231 ,232). Human fetal nigral grafts provide more potent results in animal models (225), and led to the initiation of clinical trials in PD patients (233 ,234 ,235 ,236 ,237 and 238). Results were somewhat inconsistent among the different groups, but some studies noted consistent and clinically meaningful benefit. In one study using a predetermined transplant protocol, six PD patients who could not be improved with medical management experienced significant improvement over baseline in motor scores when “off” (mean of 31%) and in percent “on” time without dyskinesia (mean of approximately 250%) (238). The variability in clinical response in the different centers may have related to the use of different transplant variables (e.g., donor age, method of tissue storage, target site for transplant, volume of distribution within target site, amount of implanted tissue, use of cyclosporine). In trials documenting clinical benefit, striatal fluorodopa uptake on PET demonstrated a significant and progressive increase in striatal fluorodopa uptake (237 ,238 ,239 and 240). Benefits on PET correlated with improvement in motor scores (238 ,241). Postmortem studies have been performed on some patients who have received transplants and expired for reasons not related to the transplant procedure (242 ,243). These studies demonstrated robust survival of implanted neurons and reinnervation of the striatum in an organotypic fashion (242). In this study, there were strong correlates between the number of surviving cells and UPDRS motor scores and striatal fluorodopa uptake on PET.

Following these open studies, two prospective randomized double-blind placebo-controlled trials have been initiated. The first was a 1-year study involving 40 patients. Two donors per side were implanted into the caudate and putamen bilaterally, without immunosuppression (244). Quality of life was the primary endpoint and was not improved. However, significant improvement in UPDRS motor and ADL scores were observed in patients under 60 years. The second study is a 2-year study that compares bilateral transplantation into the postcommissural putamen with one versus four donors per side (174). Immunosuppression with cyclosporine was employed in this study. The study is still ongoing and will terminate in 2001.

Several hundreds PD patients have now undergone transplant procedures. In general, the procedure has been well tolerated, especially when performed in major university centers. There is one report of a death due to obstructive hydrocephalus caused by graft migration into the 4th ventricle. Postmortem study revealed that the migrated tissue was composed of nonneural tissue containing bone, cartilage, hair, and epithelium (243). This study illustrates the importance of developing experience in transplant biology and appropriate dissection techniques before embarking on this surgical adventure. There has also been a report in abstract form of new-onset disabling dyskinesia that persists even when levodopa is withdrawn for prolonged periods of time (245). The frequency, clinical significance, and basis for this problem remain unknown, but clearly warrant further investigation.

The role of fetal nigral transplantation in PD has not

yet been fully determined, but the only double-blind study completed so far has not shown satisfactory benefits, and there are concerning side effects that remain to be explained. Concomitant use of antioxidants, lazaroids, antiapoptotic agents, and trophic factors, or modifications in the type of donors, the amount of cells transplanted, and the site of transplantation may all enhance transplant benefits. Also, alternate sources of dopaminergic tissues will have to be found to avoid the societal and logistical problems associated with the use of fetal human tissue. Transplantation of fetal porcine nigral cells has been shown to provide some clinical benefit and postmortem cell survival (246), and a prospective double-blind clinical trial is ongoing. Other experimental approaches to repopulating the basal ganglia with dopaminergic cells include the use of stem cells and gene therapies. The concept of restoring dopaminergic innervation to the basal ganglia is appealing, and to some extent it is now clear that this can be accomplished. For the present, however, transplant therapies must still be considered experimental and not a practical option for PD patients outside of research trials.

FUTURE RESEARCH DIRECTIONS

Part of "124 - Current and Experimental Therapeutics of Parkinson's Disease "

Symptomatic Therapies: Nondopaminergic Agents

Despite the advances in the therapeutics of PD, patients continue to experience parkinsonian disability and disabling motor complications. New treatment strategies aimed at providing more continuous dopaminergic stimulation to prevent motor complications and surgical approaches to ameliorate them represent major advances. Nonetheless, many patients continue to experience disability despite these new treatment approaches. This has led to experimentation with other approaches to the symptomatic treatment of PD and its complications. Although most interest has focused on the motor aspects of PD, dementia is the greatest unmet medical need and the major reason for nursing home placement for patients with this condition (247). There are currently no treatments that are established to attenuate the decline in mental function that accompanies PD. Some physicians use central cholinergic medications such as donepezil or rivastigmine on an empiric basis, but there are no studies confirming their value in PD.

There has been increasing interest in developing new antidyskinetic therapies for PD based on activating the numerous nondopaminergic cell-surface receptor targets on basal ganglia neurons that modulate dopaminergic activity or other systems that are affected in PD (248). The development of an agent that blocks dyskinesia would permit levodopa to be used in larger doses and thereby eliminate motor fluctuations. Some possible antidyskinetic agents include drugs that are glutamate antagonists, adenosine A2A antagonists, opioid antagonists, serotoninergic 5-HT_{2C} agonists, canabanoid CB1 agonists, α_2-antagonists, dopamine uptake inhibitors, selective muscarinic antagonists, and nicotinic agonists (249). Glutamate antagonists have already been shown to have antidyskinetic effects in some PD patients (133 ,134 and 135), but they are complicated by mental side effects that limit their utility in PD. However, other agents such as riluzole that inhibit sodium channels and impair glutamate release have also been reported to improve dyskinesia and are better tolerated (250). The adenosine A2A receptor is localized to striatal cholinergic interneurons, and antagonists to the adenosine A2A receptor have been shown to increase motor activity in rodent and primate models of PD, without provoking a dyskinetic response, even when administered to levodopa-primed animals (251 ,252). Clinical trials of this agent are currently under way. Nicotine receptors are present on terminals of nigrostriatal neurons, and their stimulation has been shown to increase dopamine release in the rat nucleus accumbens (253). This may account for why cigarette smoking is addictive, and why there is a seeming reduction in the frequency of PD in smokers (254). In MPTP-treated primates, nicotine has no effect on the basal motor disability or on levodopa-induced dyskinesia, but muscarinic agonists and antagonists did influence levodopa-induced dyskinesia (255 ,256).

Restorative Therapies

The threshold for developing levodopa-induced dyskinesias appears to depend on the degree of denervation of the SNc (42 ,257). This has led to the hypothesis that increasing the number of dopaminergic terminals might better regulate dopamine storage and release and control dyskinesia. Bjorklund et al. (258) have shown that dyskinesia can be prevented in a rodent model following transplantation of dopamine neurons with restoration of greater than 20% of striatal dopamine terminals as detected by staining for dopamine transporter protein. There is considerable interest in the potential of neurotrophic factors, such as brain-derived neurotrophic factor (BDNF) or glial-derived neurotrophic factor (GDNF), to provide restorative effects and increased numbers of dopamine terminals in PD. GDNF has been shown to promote functional and anatomic recovery in MPTP-treated monkeys (259 ,260). GDNF treatment in these animals was associated with improvement in motor behavior, a reduction in levodopa-induced dyskinesia, and increased dopamine production. This approach could provide combined symptomatic and neurorestorative benefits. Clinical trials of intraventricular GDNF administration in PD patients have been stopped, presumably because of lack of efficacy. This may relate to failure of GDNF to cross the blood–brain barrier. Further studies of direct intraparenchymal injections are warranted. Administration of GDNF by gene therapy using a lentivirus vector has been shown

to prevent motor deficits and nigral degeneration in MPTP monkeys (261). This provides another opportunity for introducing GDNF to PD patients, although such trials are not likely to be started until a number of regulatory concerns have been addressed. Stem cells have tropic properties and can migrate to sites of neuronal damage. They also have the potential to be converted to dopaminergic neurons. The question is, Can the two functions be merged? This too is a promising area of research, but clinical trials may be far in the future.

Neuroprotective Therapies

Neuroprotective therapies are designed to slow or stop disease progression by rescuing or protecting vulnerable neurons. To date, no therapy has been established to be neuroprotective in PD. When a neuroprotective treatment becomes available, it will be important to define at-risk subjects or patients with very early PD so that treatment can be initiated at the earliest time possible. An ideal neuroprotective therapy would eliminate the cause of the disease. Unfortunately, it is likely that both genetic and environmental factors contribute to the etiology of PD, and they may be different in different patients (262). The recent twin study indicates that genetic factors do not play a role in the etiology of PD in the majority of patients (263). A small number of familial cases are now known to be due to mutations in the genes that code for the proteins α-synuclein and parkin (264 ,265). Although they represent a small number of individuals, these findings may yield clues for understanding the pathogenesis of PD and permit the development of therapies that are of value for the majority of cases. α-Synuclein is a protein that accumulates even in sporadic PD (266). Parkin is now known to be a ubiquitin-protein ligase that is involved in protein degradation and reduced in activity in the mutant form (267). These observations suggest that protein clearance may be a fundamental problem in the origin of nigral degeneration in PD and a source of new therapeutic opportunities. Pathogenetic factors that have been implicated in PD include oxidative stress, excitotoxicity, mitochondrial dysfunction, and inflammation (268). It is unknown to what degree each of these contributes to the initiation of cell death, but each represents an opportunity for targeting a neuroprotective therapy. There is also a growing amount of evidence supporting the notion that cell death in PD occurs through an apoptotic process (155 ,269). Apoptosis is a gradual form of cell death that is associated with intracellular signaling mechanisms (270). The knowledge of these signals and the ability to manipulate them provide another opportunity for developing neuroprotective strategies.

Thus, there are numerous possible avenues for neuroprotective therapies (82): antioxidants (free radical scavengers, glutathione, ion chelators); glutamate inhibitors (excitatory aminoacids antagonists, glutamate release inhibitors, e.g., riluzole); calcium channel blockers; mitochondrial “energizers” (creatine, coenzyme Q10, nicotinamide, gingko biloba, carnitine); antiinflammatory agents (steroids); estrogens; trophic factors (GDNF, see above); transplant strategies (human, porcine, see above); antiapoptotic agents (desmethylselegiline, TCH 346, caspase inhibitors, cyclosporine); and agents that prevent intracellular protein accumulation. To date, none has been proven to be neuroprotective in PD. Indeed, the challenge is to find sufficient funding so as to be able to evaluate so many promising new therapies (271).

ACKNOWLEDGMENTS

Part of "124 - Current and Experimental Therapeutics of Parkinson's Disease "

This work was supported in part by grants from the Lowenstein Foundation and the National Institutes of Health (5 MO1 RR00071).

REFERENCES

1. Olanow CW, Koller WC. An algorithm (decision tree) for the management of Parkinson's disease: treatment guidelines. American Academy of Neurology. *Neurology* 1998;50(3 suppl 3):S1–S57.

2. Hughes AJ, Daniel SE, Kilford L, et al. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992;55:181–184.

3. Olanow CW. Magnetic resonance imaging in parkinsonism. *Neurol Clin North Am* 1992;10:405–420.

4. Stacy M, Jankovic J. Differential diagnosis of Parkinson's disease and the Parkinsonism plus syndromes. *Neurol Clin* 1992;10:341–359.

5. Hughes AJ, Ben-Shlomo Y, Daniel SE, et al. What features improve the accuracy of clinical diagnosis in Parkinson's disease: a clinicopathologic study. *Neurology* 1992;42:1142–1146.

6. Gibb WR. Neuropathology of Parkinson's disease and related syndromes. *Neurol Clin* 1992;10(2):361–376.

7. Jellinger K. The pathology of parkinsonism. In: Marsden CD, Fahn S, eds. *Movement disorders 2*. London: Butterworths, 1987:124–165.

8. Cotzias GC, Van Woert MH, Schiffer LM. Aromatic amino acids and modification of parkinsonism. *N Engl J Med* 1967;276(7):374–379.

9. Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. *Neurology* 1967;17:427–442.

10. Diamond SG, Markham CH, Hoehn MM, et al. Multi-center study of Parkinson mortality with early versus later dopa treatment. *Ann Neurol* 1987;22(1):8–12.

11. Marsden CD, Parkes JD. “On-off” effects in patients with Parkinson's disease on chronic levodopa therapy. *Lancet* 1976;1:292–295.

12. Poewe WH, Wenning GK. The natural history of Parkinson's disease. *Neurology* 1996;47(6 suppl 3):S146–152.

13. Carlsson A, Lindquist M, Magnusson T. 3,4-Dihydroxyphenylalanine and 6-hydroxytryptophan as reserpine antagonist. *Nature* 1957;180:1200–1201.

14. Ehringer H, Hornykiewicz O. Distribution of noradrenaline and dopamine (3-hydroxytyramine) in the human brain and their behavior in diseases of the extrapyramidal system (German). *Klin Wochenschr* 1960;38:1236–1239.

15. Yahr MD, Duvoisin RC, Schear MJ, et al. Treatment of parkinsonism with levodopa. *Arch Neurol* 1969;21:343–354.

16. Standaert DG, Young AB. Treatment of central nervous system disorders. In: Hardman JG, Limbird LE, eds. *Goodman and Gilman's the pharmacological basis of therapeutics*, ninth ed. New York: Macmillan, 1995:503–519.

17. Nutt JG, Fellman JH. Pharmacokinetics of levodopa. *Clin Neuropharmacol* 1984;7(1):35–49.

18. Cooper JR, Bloom FE, Roth HR. *The biochemical basis of neuropharmacology*. New York: Oxford University Press, 1991.

19. Sealfon SC, Olanow CW. Dopamine receptors: from structure to behavior. *Trends Neurosci* 2000;23:S34–S40.

20. Kebabian JW, Calne DB. Multiple receptors for dopamine. *Nature* 1979; 277:93–96.

21. Smith Y, Kieval JZ. Anatomy of the dopamine system in the basal ganglia. *Trends Neurosci* 2000;23:S28–S33.

22. Hutton JT, Morris JL, Elias JW. Levodopa improves spatial contrast sensitivity in Parkinson's disease. *Arch Neurol* 1993;50(7):721–724.

23. Newman RP, Weingartner H, Smallberg SA, et al. Effortful and automatic memory: effects of dopamine. *Neurology* 1984;34(6):805–807.

24. Luciana M, Collins PF, Depue RA. Opposing roles for dopamine and serotonin in the modulation of human spatial working memory functions. *Cereb Cortex* 1998;8(3):218–226.

25. Thomas MJ, Malenka RC, Bonci A. Modulation of long-term depression by dopamine in the mesolimbic system. *J Neurosci* 2000;20(15):5581–5586.

26. Maricle RA, Nutt JG, Valentine RJ, et al. Dose-response relationship of levodopa with mood and anxiety in fluctuating Parkinson's disease: a double-blind, placebo-controlled study. *Neurology* 1995;45(9):1757–1760.

27. Maricle RA, Valentine RJ, Carter J, et al. Mood response to levodopa infusion in early Parkinson's disease. *Neurology* 1998;50(6):1890–1892.

28. Corrigan MH, Denahan AQ, Wright CE, et al. Comparison of pramipexole, fluoxetine, and placebo in patients with major depression. *Depress Anxiety* 2000;11(2):58–65.

29. Schultz W, et al. A neural substrate of prediction and reward. *Science* 1997;275:1593.

30. Rascol O, Brooks DJ, Korczyn AD, et al. A five-year study of the incidence of dyskinesia in patients with early Parkinson's disease who were treated with ropinirole or levodopa. *N Engl J Med* 2000;342:1484–1491.

31. Parkinson Study Group. A randomized controlled trial comparing the agonist pramipexole with levodopa as initial dopaminergic treatment for Parkinson's disease. *JAMA* 2000;284(15):1571–1573.

32. Nutt JG, Holford NHG. The response to levodopa in Parkinson's disease: imposing pharmacologic law and order. *Ann Neurol* 1996;39:561–573.

33. Obeso JA, Olanow CW, Nutt JG. Levodopa motor complications in Parkinson's disease. *Trends Neurosci* 2000;23:S2–S7.

34. Muenter MD, Sharpless NS, Tyce GM, et al. Patterns of dystonia ("I.D.I." and "D.I.D") in response to L-dopa therapy for Parkinson's disease. *Mayo Clin Proc* 1977;52:163–174.

35. Fahn S. Adverse effects of levodopa. In: Olanow CW, Lieberman AN, eds. *The scientific basis for the treatment of Parkinson's disease*. United Kingdom: Parthenon, 1992:89–112.

36. Tsui JK, Ross S, Poulin K, et al. The effect of dietary protein on the efficacy of L-dopa: a double-blind study. *Neurology* 1989;39(4):549–552.

37. Verhagen Metman L, Del Dotto P, van den Munchkof P, et al. Amantadine as treatment for dyskinesias and motor fluctuations in Parkinson's disease. *Neurology* 1998;50:1323–1326

38. Lang AP, Lozano AE. Parkinson's disease. *N Engl J Med* 1998;339:1044–1053.

39. Fabbrini G, Mouradian MM, Juncos JL, et al. Motor fluctuations in Parkinson's disease: central pathophysiological mechanisms, part I. *Ann Neurol* 1988;24(3):366–371.

40. Mouradian MM, Juncos JL, Fabbrini G, et al. Motor fluctuations in Parkinson's disease: central pathophysiological mechanisms, part II. *Ann Neurol* 1988;24(3):372–378.

41. Bravi D, Mouradian MM, Roberts JW, et al. Wearing-off fluctuations in Parkinson's disease: contribution of post-synaptic mechanisms. *Ann Neurol* 1994;36:27–31.

42. Olanow CW, Obeso JA. Preventing levodopa-induced dyskinesias. *Ann Neurol* 2000;47(4 suppl 1):S167–S176.

43. Jenner P. Factors influencing the onset and persistence of dyskinesia in MPTP treated primates. *Ann Neurol* 2000;47(4 suppl 1):S90–S99.

44. Blanchet PJ, Calon F, Martel JC, et al. Continuous administration decreases and pulsatile administration increases behavioral sensitivity to a novel dopamine D2 agonist (U-91356A) in MPTP-exposed monkeys. *J Pharmacol Exp Ther* 1995;272:854–859.

45. Calon F, Grondin R, Morissette M, et al. Molecular basis of levodopa-induced dyskinesias. *Ann Neurol* 1998;47:S70–S78.

46. Filion M, Tremblay L, Bedard PJ. Effects of dopamine agonists on the spontaneous activity of globus pallidus neurons in monkeys with MPTP-induced parkinsonism. *Brain Res* 1991;547:152–161.

47. Lang AE, Lozano A, Montgomery E, et al. Posteroventral medial pallidotomy in advanced Parkinson's disease. *N Engl J Med* 1997;337:1036–1042.

48. Baron MS, Vitek JL, Bakay RA, et al. Treatment of advanced Parkinson's disease by posterior GPi pallidotomy: 1-year results of a pilot study. *Ann Neurol* 1996;40:355–366.

49. Chase TN, Oh JD. Striatal mechanisms and pathogenesis of parkinsonian signs and motor complications. *Ann Neurol* 2000;47:(4 suppl 1):S122–S129.

50. Papa SM, Chase TN. Levodopa-induced dyskinesias improved by a glutamate antagonist in Parkinsonian monkeys. *Ann Neurol* 1996;39:574–578.

51. Olanow CW, Schapira AHV, Rascol O. Continuous dopamine receptor stimulation in early Parkinson's disease. *Trends Neurosci* 2000;23:S117–S126.

52. Stern Y, Marder K, Tang MX, et al. Antecedent clinical features associated with dementia in Parkinson's disease. *Neurology* 1993;43:1690–1692.

53. Friedman JH, Factor SA. Atypical antipsychotics in the treatment of drug-induced psychosis in Parkinson's disease. *Mov Disord* 2000;15:201–211.

54. Friedman JH, Feinberg SS, Feldman RG. A neuroleptic malignant-like syndrome due to levodopa therapy withdrawal. *JAMA* 1985;254:2792–2795.

55. Keyser DL, Rodnitzky RL. Neuroleptic malignant syndrome in Parkinson's disease after withdrawal or alteration of dopaminergic therapy. *Arch Intern Med* 1991;151:794–796.

56. Olanow CW. Oxidation reactions in Parkinson's disease. *Neurology* 1990;40:32–37.

57. Walkinshaw G, Waters CM. Induction of apoptosis in catecholaminergic PC 12 cells by L-dopa. Implications for the treatment of Parkinson's disease. *J Clin Invest* 1995;95:2458–2464.

58. Blunt SB, Jenner P, Marsden CD. Suppressive effect of L-dopa on dopamine cells remaining in the ventral tegmental area of rats previously exposed to the neurotoxin 6-hydroxydopamine. *Mov Disord* 1993;8:129–133.

59. Murer MG, Dziewczapolski G, Menalled LB, et al. Chronic levodopa is not toxic for remaining dopamine neurons, but instead promotes their recovery, in rats with moderate nigrostriatal lesions. *Ann Neurol* 1998;43:561–575.

60. Agid Y, Ahlskog E, Albanese A, et al. Levodopa in the treatment of Parkinson's disease: a consensus meeting. *Mov Disord* 1999;14:911–913.

61. Calne DB. Treatment of Parkinson's disease. *N Engl J Med* 1993;329:1021–1027.

62. Olanow CW A rationale for dopamine agonists as primary therapy for Parkinson's disease. *Can J Neurol Sci* 1992;19:108–112.

63. Iida M, Miyazaki I, Tanaka K, et al. Dopamine D2 receptor-mediated antioxidant and neuroprotective effects of ropinirole, a dopamine agonist. *Brain Res* 1999;838:5159.

64. Ogawa N, Tanaka K, Asanuma M, et al. Bromocriptine protects mice against 6-hydroxydopamine and scavenges hydroxyl free radicals in vitro. *Brain Res* 1994;657:207–213.

65. Olanow CW, Jenner P, Brooks D. Dopamine agonists and neuroprotection in Parkinson's Disease. *Ann Neurol* 1998;44:167–174.

66. Calne DB, Burton K, Beckman J, et al. Dopamine agonists in Parkinson's disease. *Can J Neurol Sci* 1984;11:221–224.

67. Kartzinel R, Teychenne P, Gillespie MM, et al. Bromocriptine and levodopa (with or without carbidopa) in parkinsonism. *Lancet* 1976;2(7980):272–275.

68. Lieberman A, Kupersmith M, Estey E, et al. Treatment of parkinson's disease with bromocriptine. *N Engl J Med* 1976;295:1400–1404.

69. Olanow CW, Fahn S, Muenter M, et al. A multicenter double-blind placebo-controlled trial of pergolide as an adjunct to Sinemet in Parkinson's disease. *Mov Disord* 1994;9:40–47.

70. Hutton JT, Morris JL, Brewer MA. Controlled study of the antiparkinsonian activity and tolerability of cabergoline. *Neurology* 1993;43:613–616.

71. McDonald RJ, Horowski R. Lisuride in the treatment of parkinsonism. *Eur Neurol* 1983;22:240–255.

72. Lieberman A, Olanow CW, Sethi K, et al. A multicenter trial of ropinirole as adjunct treatment for Parkinson's disease. *Neurology* 1998;51:1057–1062.

73. Lieberman A, Ranhosky A, Korts D. Clinical evaluation of pramipexole in advanced Parkinson's disease: results of a double-blind, placebo-controlled, parallel-group study. *Neurology* 1997;49:162–168.

74. Stibe CM, Lees AJ, Kempster PA, et al. Subcutaneous apomorphine in parkinsonian on-off oscillations. *Lancet* 1988;1:403–406.

75. Pietz K, Hagell P, Odin P. Subcutaneous apomorphine in late stage Parkinson's disease: a long term follow up. *J Neurol Neurosurg Psychiatry* 1998;65:709–716.

76. Bedard PJ, Di Paolo T, Falardeau P, et al. Chronic treatment with L-DOPA, but not bromocriptine induces dyskinesia in MPTP-parkinsonian monkeys. Correlation with [3H]spiperone binding. *Brain Res* 1986;379:294–299.

77. Pearce RK, Banerji T, Jenner P, et al. De novo administration of ropinirole and bromocriptine induces less dyskinesia than L-dopa in the MPTP-treated marmoset. *Mov Disord* 1998;13:234–241.

78. Parkinson Study Group. Safety and efficacy of pramipexole in early Parkinson disease: a randomized dose-ranging study. *JAMA* 1997;278:125–130.

79. Adler CH, Sethi KD, Hauser RA, et al. Ropinirole for the treatment of early Parkinson's disease. *Neurology* 1997;49:393–399.

80. Bressman SB, Shulman LM, Tanner CM et al. Long-term safety and efficacy of pramipexole in early Parkinson's disease. *Mov Disord* 2000;15(suppl 3):112–113.

81. Montastruc JL, Olivier R, Senard JM. Treatment of Parkinson's disease should begin with a dopamine agonist. *Mov Disord* 1999;14:725–730.

82. Olanow CW, Watts R, Koller WC. An algorithm (decision tree) for the management of Parkinson's disease (2001): treatment guidelines. *Neurology* 2001;56(suppl 5):1–88.

83. Chase TN, Baronti F, Fabbrini G, et al. Rationale for continuous dopaminomimetic therapy of Parkinson's disease. *Neurology* 1989;39(11 suppl 2):7–10.

84. Rajput AH. Adverse effects of ergot-derivative of dopamine agonists. In: Olanow CW, Obeso JA, eds. *Dopamine agonists in early Parkinson's disease*. Kent, UK: Wells Medical, 1997:209–216.

85. Frucht S, Rogers JD, Greene PE, et al. Falling asleep at the wheel: motor vehicle mishaps in persons taking pramipexole and ropinirole. *Neurology* 1999;52(9):1908–1910.

86. Olanow CW, Schapira AH, Roth T. Waking up to sleep episodes in Parkinson's disease. *Mov Disord* 2000;15(2):212–215.

87. Ferreira JJ, Galitzky M, Montastruc JL, et al. Sleep attacks and Parkinson's disease treatment. *Lancet* 2000;355:1333–1334.

88. Johns MW. New method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991;14:540–545.

89. Dingemanse J, Jorga KM, Schmitt M, et al. Integrated pharmacokinetics and pharmacodynamics of the novel catechol-O-methyltransferase inhibitor tolcapone during first administration to humans. *Clin Pharmacol Ther* 1995;57(5):508–517.

90. Keranen T, Gordin A, Karlsson M, et al. Inhibition of soluble catechol-O-methyltransferase and single-dose pharmacokinetics after oral and intravenous administration of entacapone. *Eur J Clin Pharmacol* 1994;46:151–157.

91. Jorga K, Fotteler B, Sedek G, et al. The effect of tolcapone on levodopa pharmacokinetics is independent of levodopa/carbidopa formulation. *J Neurol* 1998;245:223–230.

92. Myllyla VV, Sotaniemi KA, Illi A, et al. Effect of entacapone, a COMT inhibitor, on the pharmacokinetics of levodopa and on cardiovascular responses in patients with Parkinson's disease. *Eur J Clin Pharmacol* 1993;45:419–423.

93. Nutt JG, Woodward WR, Beckner RM, et al. Effect of peripheral catechol-O-methyltransferase inhibition on the pharmacokinetics and pharmacodynamics of levodopa in parkinsonian patients. *Neurology* 1994;44:913–919.

94. Nutt JG. Effects of catechol-O-methyltransferase (COMT) inhibition on the pharmacokinetics of L-DOPA. *Adv Neurol* 1996;69:493–496.

95. Sawle GV, Burn DJ, Morrish PK, et al. The effect of entacapone (OR-611) on brain (^{18}F)-6-fluorodopa metabolism: implications for levodopa therapy in Parkinson's disease. *Neurology* 1994;44:1292–1297.

96. Ruottinen HM, Rinne UK. Entacapone prolongs levodopa response in a one month double blind study in parkinsonian patients with levodopa related fluctuations. *J Neurol Neurosurg Psychiatry* 1996;60(1):36–40.

97. Rajput AH, Martin W, Saint-Hilaire MH, et al. Tolcapone improves motor function in parkinsonian patients with the “wearing-off” phenomenon: a double-blind, placebo-controlled, multicenter trial. *Neurology* 1997;49:1066–1071.

98. Parkinson Study Group. Entacapone improves motor fluctuations in levodopa-treated Parkinson's disease patients. *Ann Neurol* 1997;42:747–755.

99. Rinne UK, Larsen JP, Siden A, et al. Entacapone enhances the response to levodopa in parkinsonian patients with motor fluctuations. Nomecomt Study Group. *Neurology* 1998;51:1309–1314.

100. Kurth MC, Adler CH, Hilaire MS, et al. Tolcapone improves motor function and reduces levodopa requirement in patients with Parkinson's disease experiencing motor fluctuations: a multicenter, double-blind, randomized, placebo-controlled trial. Tolcapone Fluctuator Study Group I. *Neurology* 1997;48:81–87.

101. Dupont E, Burgunder JM, Findley LJ, et al. Tolcapone added to levodopa in stable parkinsonian patients: a double-blind placebo-controlled study. Tolcapone in Parkinson's Disease Study Group II (TIPS II). *Mov Disord* 1997;12:928–934.

102. Waters CH, Kurth M, Bailey P, et al. Tolcapone in stable Parkinson's disease: efficacy and safety of long-term treatment. The Tolcapone Stable Study Group. *Neurology* 1997;49:665–671.

103. Olanow CW, Obeso JA. Pulsatile stimulation of dopamine receptors and levodopa-induced motor complications in Parkinson's disease: implications for the early use of COMT inhibitors. *Neurology* 2000;55(11 suppl 4):S72–S77.

104. Koller WC, Hutton JT, Tolosa E et al. Immediate-release and controlled-release carbidopa/levodopa in PD: a 5-year randomized multi-center study. *Neurology* 1999;53:1012–1019.

105. Dupont E, Andersen A, Boas J, et al. Sustained-release Madopar HBS compared with standard Madopar in the long-term treatment of de novo parkinsonian patients. *Acta Neurol Scand* 1996;93:14–20.

106. Olanow CW, and the Tasmar advisory panel. Tolcapone and hepatotoxic effects. *Arch Neurol* 2000;57:263–267.

107. Assal F, Spahr L, Hadengue A, et al. Tolcapone and fulminant hepatitis. *Lancet* 1998;352:958.

108. Anonymous. *New warnings for parkinson's drug Tasmar*. FDA talk paper. Rockville, MD: Food and Drug Administration, 1998.

109. Olanow CW. Hepatic safety of the COMT inhibitor entacapone. *Neurology* 2000;54:279.

110. Schapira A, Obeso JA, Olanow CW. The place of COMT inhibitors in the armamentarium of drugs for the treatment of Parkinson's disease. *Neurology* 2000;55(11 suppl 4):S65–S68.

111. Charcot J-M. *Leçons sur les maladies du système nerveux*. Paris: Delahaye, 1872–1873.

112. Ordenstein L. *Sur la paralysie agitante et la sclérose en plaques généralisée*. Doctoral Thesis, Martinet, Paris, 1967.

113. Schwab RS, Leigh D. Parpanit in the treatment of Parkinson's disease. *JAMA* 1939;139:629–634.

114. Corbin KB. Trihexyphenidyl (Artane): the evaluation of a new agent in the treatment of parkinsonism. *JAMA* 1949;141:377–382.

115. Lang AE. Treatment of Parkinson's disease with agents other than levodopa and dopamine agonists: controversies and new approaches. *Can J Neurol Sci* 1984;11(1 suppl):210–220.

116. Duvoisin RC. Cholinergic-anticholinergic antagonism in parkinsonism. *Arch Neurol* 1967;17:124–136.

117. Coyle JT, Snyder SH. Antiparkinsonian drugs: inhibition of dopamine uptake in the corpus striatum as a possible mechanism of action. *Science* 1969;166(907):899–901.

118. Stoof JC, Kebabian JW. Independent in vitro regulation by the D-2 dopamine receptor of dopamine-stimulated efflux of cyclic AMP and K+-stimulated release of acetylcholine from rat neostriatum. *Brain Res* 1982;250(2):263–270.

119. Pollack AE, Wooten GF. D2 dopaminergic regulation of striatal preproenkephalin mRNA levels is mediated at least in part through cholinergic interneurons. *Brain Res Mol Brain Res* 1992;13:35–41.

120. Quinn NP. Anti-parkinsonian drugs today. *Drugs* 1984;28:236–262.

121. Hauser RA, Olanow CW. Orobuccal dyskinesia associated with trihexyphenidyl therapy in a patient with Parkinson's disease. *Mov Disord* 1993;8:512–514.

122. Horrocks PM, Vicary DJ, Rees JE, et al. Anticholinergic withdrawal and benzhexol treatment in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1973;36(6):936–941.

123. Stocchi F, Corazziari E, Carbone A, et al. Urinary and anorectal function evaluation in Parkinson's disease. *Adv Neurol* 1999;80:347–351.

124. Koller WC. Pharmacologic treatment of parkinsonian tremor. *Arch Neurol* 1986;43:126–127.

125. Schwab RS, England AC Jr, Poskanzer DC, et al. Amantadine in the treatment of Parkinson's disease. *JAMA* 1969;208:1168–1170.

126. Bailey EV, Stone TW. The mechanism of action of amantadine in Parkinsonism: a review. *Arch Int Pharmacodyn Ther* 1975;216:246–262.

127. Stone TW, Bailey EV. Responses of central neurones to amantadine: comparison with dopamine and amphetamine. *Brain Res* 1975;85:126–129.

128. Kornhuber J, Weller M, Schoppmeyer K, et al. Amantadine and memantine are NMDA receptor antagonists with neuroprotective properties. *J Neural Transm* 1994;43(suppl):91–104.

129. Schwab RS, Poskanzer DC, England AC Jr, et al. Amantadine in Parkinson's disease. Review of more than two years' experience. *JAMA* 1972;222:792–795.

130. Pinter MM, Birk M, Helscher RJ, et al. Short-term effect of amantadine sulphate on motor performance and reaction time in patients with idiopathic Parkinson's disease. *J Neural Transm* 1999;106:711–724.

131. Parkes JD, Baxter RC, Curzon G, et al. Treatment of Parkinson's disease with amantadine and levodopa. A one-year study. *Lancet* 1971;1:1083–1086.

132. Parkes JD, Baxter RC, Marsden CD, et al. Comparative trial of benzhexol, amantadine, and levodopa in the treatment of Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1974;37:422–426.

133. Deleted in press.

134. Metman LV, Del Dotto P, LePoole K, et al. Amantadine for levodopa-induced dyskinesias: a 1-year follow-up study. *Arch Neurol* 1999;56:1383–1386.

135. Snow BJ, Macdonald L, Mcauley D, et al. The effect of amantadine on levodopa-induced dyskinesias in Parkinson's disease: a double-blind, placebo-controlled study. *Clin Neuropharmacol* 2000;23(2):82–85.

136. Greenamyre JT, O'Brien CF. N-methyl-D-aspartate antagonists in the treatment of Parkinson's disease. *Arch Neurol* 1991;48:977–981.

137. Rodriguez MC, Obeso JA, Olanow CW. Subthalamic nucleus-mediated excitotoxicity in Parkinson's disease: a target for neuroprotection. *Ann Neurol* 1998;44(3 suppl 1):S175–S88.

138. Uitti RJ, Rajput AH, Ahlskog JE, et al. Amantadine treatment is an independent predictor of improved survival in Parkinson's disease. *Neurology* 1996;46:1551–1556.

139. Golbe LI, Lieberman AN, Muenter MD, et al. Deprenyl in the treatment of symptom fluctuations in advanced Parkinson's disease. *Clin Neuropharmacol* 1988;11:45–55.

140. Langston JW, Ballard P, Tetrud JW, et al. Chronic parkinsonism in humans due to a product of meperidine-analog synthesis. *Science* 1983;219:979–980.

141. Chiba K, Trevor A, Castagnoli N Jr. Metabolism of the neurotoxic tertiary amine, MPTP, by brain monoamine oxidase. *Biochem Biophys Res Commun* 1984;120:457–478.

142. Heikkila RE, Manzino L, Cabbat FS, et al. Protection against the dopaminergic neurotoxicity of 1-methyl-4-phenyl-1,2,5,6-tetrahydropyridine by monoamine oxidase inhibitors. *Nature* 1984;311:467–469.

143. Cohen G, Pasik P, Cohen B et al. Pargyline and deprenyl prevent the neurotoxicity of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) in monkeys. *Eur J Pharmacol* 1985;106:209–210.

144. The Parkinson Study Group. Effects of tocopherol and deprenyl on the progression of disability in early Parkinson's disease. *N Engl J Med* 1993;328:176–183.

145. Olanow CW, Hauser RA, Gauger L, et al. The effect of deprenyl and levodopa on the progression of Parkinson's disease. *Ann Neurol* 1995;38:771–777.

146. Olanow CW, Calne D. Does selegiline monotherapy in Parkinson's disease act by symptomatic or protective mechanisms? *Neurology* 1992;42:13–26.

147. Parkinson Study Group. Impact of deprenyl and tocopherol treatment on Parkinson's disease in DATATOP patients requiring levodopa. *Ann Neurol* 1996;39:37–45.

148. Parkinson Study Group. Impact of deprenyl and tocopherol treatment on Parkinson's disease in DATATOP subjects not requiring levodopa. *Ann Neurol* 1996;39:29–36.

149. Olanow CW. Selegiline: current perspectives on issues related to neuroprotection and mortality. *Neurology* 1996;47:S210–S216.

150. Tatton WG, Greenwood CE. Rescue of dying neurons: a new action for deprenyl in MPTP parkinsonism. *J Neurosci Res* 1991;30:666–677.

151. Mytilineou C, Radcliffe P, Leonardi EK, et al. L-deprenyl protects mesencephalic dopamine neurons from glutamate receptor-mediated toxicity. *J Neurochem* 1997;68:33–39.

152. Mytilineou C, Kototos Leonardi E, Radcliffe P, et al. Deprenyl and desmethylselegiline protect mesencephalic neurons from toxicity induced by glutathione depletion. *J Pharmacol Exp Ther* 1998;284:700–706.

153. Carlile GW, Chalmers-Redman RME, Tatton N, et al. Reduced apoptosis after nerve growth factor and serum withdrawal; conversion of tetrameric glyceraldehyde-3-phosphate dehydrogenase to a dimmer. *Mol Pharmacol* 2000;57:2–12.

154. Chalmers-Redman RME, Ju W, Carlile GW, et al. Deprenyl-related propargylamines induce anti-apoptotic new protein synthesis in NGF-differentiated PC12 cells. *J Neurosci* (*in press*).

155. Tatton NA, Maclean-Fraser A, Tatton WG, et al. A fluorescent double-labeling method to detect and confirm apoptotic nuclei in Parkinson's disease. *Ann Neurol* 1998;44(3):142–148.

156. Lees AJ. Comparison of therapeutic effects and mortality data of levodopa and levodopa combined with selegiline in patients with early, mild Parkinson's disease. Parkinson's Disease Research Group of the United Kingdom. *BMJ* 1995;311:1602–1607.

157. Ben-Shlomo Y, Churchyard A, Head J, et al. Investigation by Parkinson's Disease Research Group of United Kingdom into excess mortality seen with combined levodopa and selegiline treatment in patients with early, mild Parkinson's disease: further results of randomised trial and confidential inquiry. *BMJ* 1998;316:1191–1196.

158. Olanow CW, Fahn S, Langston JW, et al. Selegiline and mortality in Parkinson's disease. *Ann Neurol* 1996;40(6):841–845.

159. Olanow CW, Myllyla VV, Sotaniemi KA, et al. Effect of selegiline on mortality in patients with Parkinson's disease: a meta-analysis. *Neurology* 1998;51:825–830.

160. Horsley V. The functions of the so-called motor areas of the brain. *BMJ* 1909;124:5–28.

161. Bucy PC, Case JT. Tremor: physiologic mechanism and abolition by surgical means. *Arch Neurol Psychiatry* 1939;41:721–746.

162. Leriche R. Ueber chirurgische Eingriffe bei Parkinsonischer Krankheit. *Neurol Zbl* 1912;13:1093–1096.

163. Meyers R. Surgical procedure for postencephalitic tremor, with notes on the physiology of premotor fibres. *Arch Neurol Psychiatry* 1940;44:455–457.

164. Meyers R. The modification of alternating tremors, rigidity and festination by surgery of the basal ganglia. *Proc Assoc Nerv Ment Dis* 1942;21:602–665.

165. Hamby WB. Discussion to the communication of Browder J “Parkinsonism, is it a surgical problem?” *NY State J Med* 1947;47:2592.

166. Browder J. Section of fibres of anterior limb of internal capsula in Parkinsonism. *Am J Surg* 1948;75:264–268.

167. Fenelon F. Essais de traitement neurochirurgical du syndrome parkinsonian par intervention directe sur les voies extrapyramidales immément sous-striopallidales (anse lenticulaire). *Rev Neurol* 1950;83:437–439.

168. Guiot G, Brion S. Traitement des mouvements anormaux par la coagulation pallidale. Technique et résults. *Rev Neurol* 1953;89:578–580.

169. Spiegel EA, Wycis HT, Marks M, et al. Stereotaxic apparatus for operations on the human brain. *Science* 1946;106:349–359.

170. Cooper IS. Ligation of the anterior choroidal artery for involuntary movements and Parkinsonism. *Psychiatr Q* 1953;27:317–319.

171. DeLong MR. Primate models of movement disorders of basal ganglia origin. *Trends Neurosci* 1990;13:281–285.

172. Alexander GE, Crutcher MD, DeLong MR. Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, “prefrontal” and “limbic” functions. *Prog Brain Res* 1990;85:119–146.

173. Olanow CW, Brin MF. Surgery for Parkinson's disease: a physician's perspective. *Adv Neurol* 2001;86:421–433.

174. Freeman TB, Vawter DE, Leaverton PE, et al. The use of a surgical placebo-control in evaluating cellular-based therapy for Parkinson's disease. *N Engl J Med* 1999;341:988–992.

175. Crossman AR, Mitchell IJ, Sambrook MA. Regional brain uptake of 2-deoxyglucose in N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-induced parkinsonism in the macaque monkey. *Neuropharmacology* 1985;24:587–591.

176. Eidelberg D, Moeller JR, Dhawan V, et al. The metabolic topography of parkinsonism. *J Cereb Blood Flow Metab* 1994;14:783–801.

177. Hassler R, Riechert T. Indikationen und Lokalisation Methode der gezielten Hirmoperationen. *Nervenarzt* 1954;25:441–447.

178. Albe-Fessard D, Arfel G, Guiot G, et al. Electrophysiological studies of some deep cerebral structures in man. *J Neurol Sci* 1966;3:37–51.

179. Taren J, Guiot G, Derome P, et al. Hazards of stereotaxic thalamectomy. Added safety factor in corroborating x-ray target localization with neurophysiological methods. *J Neurosurg* 1968;29:173–182.

180. Kelly PJ, Derome P, Guiot G. Thalamic spatial variability and the surgical results of lesions placed with neurophysiologic control. *Surg Neurol* 1978;9:307–315.

181. Ohye C, Narabayashi H. Physiological study of presumed ventralis intermedius neurons in the human thalamus. *J Neurosurg* 1979;50:290–297.

182. Narabayashi H. Stereotaxic Vim thalamotomy for treatment of tremor. *Eur Neurol* 1989;29(suppl 1):29–32.

183. Matsumoto K, Shichijo F, Fukami T. Long-term follow-up review of cases of Parkinson's disease after unilateral or bilateral thalamotomy. *J Neurosurg* 1984;65:1033–1044.

184. Narabayashi H, Maeda T, Yokochi F. Long-term follow-up study of nucleus ventralis intermedius and ventrolateralis thalamotomy using a microelectrode technique in parkinsonism. *Appl Neurophysiol* 1987;50:330–337.

185. Tasker RR. Thalamotomy. *Neurosurg Clin North Am* 1990;1:841–864.

186. Fox MW, Ahlskog JE, Kelly PJ. Stereotactic ventrolateralis thalamotomy for medically refractory tremor in post-levodopa era Parkinson's disease patients. *J Neurosurg* 1991;75:723–730.

187. Tasker RR, Siqueira J, Hawrylyshyn P, et al. What happened to VIM thalamotomy for Parkinson's disease? *Appl Neurophysiol* 1983;46:68–83.

188. Narabayashi H, Yokochi F, Nakajima Y. Levodopa-induced dyskinesia and thalamotomy. *J Neurol Neurosurg Psychiatry* 1984;47:831–839.

189. Page RD. The use of thalamotomy in the treatment of levodopa-induced dyskinesia. *Acta Neurochir (Wien)* 1992;114:77–117.

190. Burchiel KJ. Thalamotomy for movement disorders. *Neurosurg Clin North Am* 1995;6:55–71.

191. Svennilson E, Torvik A, Lowe R, et al. Treatment of Parkinsonism by Stereotactic thermolesions in the pallidal region. *Acta Psychiatr Neurol Scand* 1960;35:358–377.

192. Laitinen LV, Bergenheim AT, Hariz MI. Leksell's posteroventral pallidotomy in the treatment of Parkinson's disease. *J Neurosurg* 1992;76:53–61.

193. Laitinen LV, Bergenheim AT, Hariz MI. Ventroposterolateral pallidotomy can abolish all parkinsonian symptoms. *Stereotact Funct Neurosurg* 1992;58:14–21.

194. Penney JB Jr, Young AB. Striatal inhomogeneities and basal ganglia function. *Mov Disord* 1986;1(1):3–15.

195. Bakay RA, DeLong MR, Vitek JL. Posteroventral pallidotomy for Parkinson's disease. *J Neurosurg* 1992;77:487–488.

196. Dogali M, Sterio D, Fazzini E, et al. Effects of posteroventral pallidotomy on Parkinson's disease. *Adv Neurol* 1996;69:585–590.

197. Lozano AM, Lang AE, Galvez-Jimenez N, et al. Effect of GPi pallidotomy on motor function in Parkinson's disease. *Lancet* 1995;346:1383–1387.

198. Baron MS, Vitek JL, Bakay RA, et al. Treatment of advanced Parkinson's disease by unilateral posterior GPi pallidotomy: 4-year results of a pilot study. *Mov Disord* 2000;15:230–237.

199. Fine J, Duff J, Chen R, et al. Long-term follow-up of unilateral pallidotomy in advanced Parkinson's disease. *N Engl J Med* 2000;342:1708–1714.

200. Favre J, Burchiel KJ, Taha JM, et al. Outcome of unilateral and bilateral pallidotomy for Parkinson's disease: patient assessment. *Neurosurgery* 2000;46:344–353.

201. Ghika J, Ghika-Schmid F, Fankhauser H, et al. Bilateral contemporaneous posteroventral pallidotomy for the treatment of Parkinson's disease: neuropsychological and neurological side effects. Report of four cases and review of the literature. *J Neurosurg* 1999;91:313–321.

202. Obeso JA, Rodriguez-Oroz MC, Rodriguez M, et al. Pathophysiology of levodopa-induced dyskinesias in Parkinson's disease: problems with current models of the basal ganglia. *Ann Neurol* 2000;47:22–34.

203. Mitchell IJ, Clarke CE, Boyce S, et al. Neural mechanisms underlying parkinsonian symptoms based upon regional uptake of 2-deoxyglucose in monkeys exposed to 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine. *Neuroscience* 1989;32(1):213–226.

204. Bergman H, Wichmann T, DeLong MR. Reversal of experimental parkinsonism by lesions of the subthalamic nucleus. *Science* 1990;249:1436–1438.

205. Guridi J, Herrero MT, Luquin R, et al. Subthalamotomy improves MPTP-induced parkinsonism in monkeys. *Stereotact Funct Neurosurg* 1994;62:98–102.

206. Alvarez A, Macias R, Guridi J, et al. Dorsal subthalamotomy for Parkinson's disease. *Mov Disord* 2001;16(1):72–78.

207. Benabid AL, Pollak P, Louveau A, et al. Combined (thalamotomy and stimulation) stereotactic surgery of the VIM thalamic nucleus for bilateral Parkinson disease. *Appl Neurophysiol* 1987;50:344–346.

208. Benabid AL, Pollak P, Gervason C, et al. Long-term suppression of tremor by chronic stimulation of the ventral intermediate thalamic nucleus. *Lancet* 1991;337:403–406.

209. Koller W, Pahwa R, Busenbark K, et al. High-frequency unilateral thalamic stimulation in the treatment of essential and parkinsonian tremor. *Ann Neurol* 1997;42:292–299.

210. Caparros-Lefebvre D, Blond S, Feltin MP, et al. Improvement of levodopa induced dyskinesias by thalamic deep brain stimulation is related to slight variation in electrode placement: possible involvement of the centre median and parafascicularis complex. *J Neurol Neurosurg Psychiatry* 1999;67:308–314.

211. Bergman H, Wichmann T, Karmon B, et al. The primate subthalamic nucleus. II. Neuronal activity in the MPTP model of parkinsonism. *J Neurophysiol* 1994;72:507–520.

212. Sellal F, Hirsch E, Lisovoski F, et al. Contralateral disappearance of parkinsonian signs after subthalamic hematoma. *Neurology* 1992;42:255–256.

213. Benazzouz A, Gross C, Feger J, et al. Reversal of rigidity and improvement in motor performance by subthalamic high-frequency stimulation in MPTP-treated monkeys. *Eur J Neurosci* 1993;5:382–389.

214. Benabid AL, Pollak P, Gross C, et al. Acute and long-term effects of subthalamic nucleus stimulation in Parkinson's disease. *Stereotact Funct Neurosurg* 1994;62(1–4):76–84.

215. Limousin P, Pollak P, Benazzouz A, et al. Effect on parkinsonian signs and symptoms of bilateral subthalamic nucleus stimulation. *Lancet* 1995;345:91–95.

216. Limousin P, Krack P, Pollak P, et al. Electrical stimulation of the subthalamic nucleus in advanced Parkinson's disease. *N Engl J Med* 1998;339:1105–1111.

217. Kumar R, Lozano AM, Kim YJ, et al. Double-blind evaluation of subthalamic nucleus deep brain stimulation in advanced Parkinson's disease. *Neurology* 1998;51:850–855.

218. Piallat B, Benazzouz A, Benabid AL. Subthalamic nucleus lesion in rats prevents dopaminergic nigral neuron degeneration after striatal 6-OHDA injection: behavioural and immunohistochemical studies. *Eur J Neurosci* 1996;8:1408–1414.

219. Brotchie JM, Mitchell IJ, Sambrook MA, et al. Alleviation of parkinsonism by antagonist of excitatory amino acid transmission in the medial segment of the globus pallidus in rat and primate. *Mov Disord* 1991;6:133–138.

220. Siegfried J, Lippitz B. Bilateral chronic electrostimulation of ventroposterolateral pallidum: a new therapeutic approach for alleviating all parkinsonian symptoms. *Neurosurgery* 1994;35:1126–1129.

221. Bejjani B, Damier P, Arnulf I, et al. Pallidal stimulation for Parkinson's disease. Two targets? *Neurology* 1997;49:1564–1569.

222. Pahwa R, Wilkinson S, Smith D, et al. High-frequency stimulation of the globus pallidus for the treatment of Parkinson's disease. *Neurology* 1997;49:249–253.

223. Caparros-Lefebvre D, Ruchoux MM, Blond S, et al. Long-term thalamic stimulation in Parkinson's disease: postmortem anatomoclinical study. *Neurology* 1994;44:1856–1860.

224. Hariz MI. Complications of movement disorder surgery and how to avoid them. In: Lozano AM, ed. *Movement disorder surgery*. Basel Switzerland: Karger 2000:246–265.

225. Olanow CW, Kordower JH, Freeman TB. Fetal nigral transplantation as a therapy for Parkinson's disease. *Trends Neurosci* 1996;19:102–109.

226. Bjorklund A, Stenevi U. Reconstruction of the nigrostriatal dopamine pathway by intracerebral nigral transplants. *Brain Res* 1979;177:555–560.

227. Perlow MJ, Freed WJ, Hoffer BJ, et al. Brain grafts reduce motor abnormalities produced by destruction of nigrostriatal dopamine system. *Science* 1979;204:643–647.

228. Bankiewicz KS, Plunkett RJ, Jacobowitz DM, et al The effect of fetal mesencephalon implants on primate MPTP-induced parkinsonism..Histochemical and behavioral studies. *J Neurosurg* 1990;72:231–244.

229. Redmond DE, Sladek JR Jr, Roth RH, et al. Fetal neuronal grafts in monkeys given methylphenyltetrahydropyridine. *Lancet* 1986;1:1125–1127.

230. Madrazo I, Drucker-Colin R, Diaz V, et al. Open microsurgical autograft of adrenal medulla to the right caudate nucleus in two patients with intractable Parkinson's disease. *N Engl J Med* 1987;316:831–834.

231. Lindvall O, Backlund EO, Farde L, et al. Transplantation in Parkinson's disease: two cases of adrenal medullary grafts to the putamen. *Ann Neurol* 1987;22:457–468.

232. Goetz CG, Olanow CW, Koller WC, et al. Multicenter study of autologous adrenal medullary transplantation to the corpus striatum in patients with advanced Parkinson's disease. *N Engl J Med* 1989;320:337–341.

233. Lindvall O, Brundin P, Widner H, et al. Grafts of fetal dopamine neurons survive and improve motor function in Parkinson's disease. *Science* 1990;247:574–577.

234. Spencer DD, Robins RJ, Naftolin F, et al. Unilateral transplantation of human fetal mesencephalic tissue into the caudate nucleus in patients with Parkinson's disease. *N Engl J Med* 1992;327:1541–1548.

235. Freed CR, Breeze RE, Rosenberg NI, et al. Survival of implanted fetal dopamine cells and neurologic improvement 12 to 46 months after transplantation for Parkinson's disease. *N Engl J Med* 1992;327:1549–1555.

236. Peschanski M, Defer G, N'Guyen JP, et al Bilateral motor improvement and alteration of L-dopa effect in two patients with Parkinson's disease following intrastriatal transplantation of foetal ventral mesencephalon. *Brain* 1994;117:487–499.

237. Freeman TB, Olanow CW, Hauser RA, et al. Bilateral fetal nigral transplantation into the postcommissural putamen in Parkinson's disease. *Ann Neurol* 1995;38:379–388.

238. Hauser RA, Freeman TB, Snow BJ, et al. Long-term evaluation of bilateral fetal nigral transplantation in Parkinson's disease. *Arch Neurol* 1999;56:179–187.

239. Lindvall O, Widner H, Rehncrona S, et al. Transplantation of fetal dopamine neurons in Parkinson's disease: one year clinical and neurophysiological observations in two patients with putaminal implants. *Ann Neurol* 1992;31:155–165.

240. Sawle GV, Bloomfield PM, Bjorklund A, et al. Transplantation of fetal dopamine neurons in Parkinson's disease: PET [18F]-6-L- Fluorodopa studies in two patients with putaminal implants. *Ann Neurol* 1992;31:166–173.

241. Rémy P, Samson Y, Hantrave P, et al. Clinical correlates of [18] fluorodopa uptake in five grafted parkinsonian patients. *Ann Neurol* 1995;38:580–588.

242. Kordower JH, Freeman TB, Snow BJ, et al. Post mortem evidence of dopamine graft survival and striatal reinnervation I a Parkinson's disease patient displaying improved motor function following fetal nigral transplantation. *N Engl J Med* 1995;332:1118–1124.

243. Folkerth RD, Durso R. Survival and proliferation of non-neuronal tissues, with obstruction of cerebral ventricles, in a parkinsonian patient streated with fetal allografts. *Neurology* 1996;46:1219–1225.

244. Freed CR, Breeze RE, Greene P, et al. Double-blind placebo-controlled human fetal dopamine cell transplants in advanced Parkinson's disease. *Soc Neurosci Abstr* 1999;25:212 (abst).

245. Greene PE, Fahn S, Tsai WY, et al. Severe spontaneous dyskinesias: a disabling complication of embryonic dopaminergic tissue implants in a subset of transplanted patients with advanced Parkinson's disease. *Mov Disord* 1999;14:904.

246. Deacon T, Schumacher J, Dinsmore J, et al. Histological evidence of fetal pig neural cell survival after transplantation into a patient with Parkinson's disease. *Nat Med* 1997;3:350–353.

247. Goetz CG, Stebbins GT. Risk factors for nursing home placement in advanced Parkinson's disease. *Neurology* 1993;43:2227–2229.

248. Brotchie JM. Adjuncts to dopamine replacements: a pragmatic approach to reducing the problem of dyskinesia in Parkinson's disease. *Mov Disord* 1998;13:871–876.

249. Jenner P. Pathophysiology and biochemistry of dyskinesia: clues for the development of non-dopaminergic treatments. *J Neurol* 2000;247(suppl 2):1143–1150.

250. Merims D, Ziv I, Djaldetti R, et al. Riluzole for levodopa-induced dyskinesia in advanced Parkinson's disease. *Lancet* 1999;353:1764–1765.

251. Kanda T, Shiozaki S, Schimada J, et al. KF17387: a novel selective adenosine A2a receptor antagonist with anticataleptic activity. *Eur J Pharmacol* 1994;256:263–268.

252. Kanda T, Jackson MJ, Smith LA, et al. Adenosine A2a antagonist: a novel antiparkinsonian agent that does not provoke dyskinesia in parkinsonian monkeys. *Ann Neurol* 1998;43:507–513.

253. Rowall PP, Carr LA, Garner AD. Stimulation of [H3] dopamine release by nicotine in rat nucleus accumbens. *J Neurochem* 1987;49:1449–1454.

254. Morens DM, Grandinetti A, Reed D, et al. Cigarette smoking and protection from Parkinson's disease: false association or etiologic clue? *Neurology* 1995;45:1041–1051.

255. Schneider JS, Pope-Coleman A, Van Velson M, et al. Effects of S1B-1508Y, a novel neuronal nicotinic acetyl choline receptor agonist, on motor behavior in parkinsonian monkeys. *Mov Disord* 1998;13:637–642.

256. Banerji T, Pearce RKB, Jackson MJ. Cholinergic manipulation of L-Dopa-induced dyskinesias in the MPTP-treated common marmoset (Callithrix Jaccus). *Br J Pharmacol* 1996;117(suppl):24.

257. Boyce S, Rupniak NMJ, Steventon MJ, et al. Nigrostriatal damage is required for induction of dyskinesias by L-Dopa in squirrel monkeys. *Clin Neuropharmacol* 1990;13:448–458

258. Lee CS, Cenci MA, Schulzer M, et al. Embryonic ventral mesencephalic grafts improve levodopa-induced dyskinesia in a rat model of PD. *Brain* 2000;123:1365–1379.

259. Gash DM, Zhang Z, Ovadia A. Functional recovery in GDNF treated parkinsonian monkeys. *Nature* 1996;380:252–255

260. Zhang Z, Miyoshi Y, Lapohak PA. Dose response to intraventricular glial cell line-derived neurotrophic factor administration in parkinsonian monkeys. *J Pharmacol Exp Ther* 1997;282:1396–1401.

261. Kordower JH, Emborg M, Bloch J, et al. Lente viral vetor-mediated expression of GDNF prevents motor deficits and nigro-striatal degeneration in non-human primate models of parkinson's disease. *Science* 2000;290(5492):767–773.

262. Olanow CW, Tatton WG. Etiology and pathogenesis of Parkinson's disease. *Annu Rev Neurosci* 1999;22:123–144.

263. Tanner CM, Ottman R, Goldman SM, et al. Parkinson disease in twins: an etiologic study. *JAMA* 1999;281:341–346.

264. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* 1997;276:2045–2047.

265. Kitada T, Asakawa S, Hattori N, et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* 1998;392:605–608.

266. Spillantini MG, Crowther RA, Jakes R, et al. alpha-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with lewy bodies. *Proc Natl Acad Sci USA* 1998;95:6469–6473.

267. Shimura H, Hattori N, Kubo S, et al. Familial parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nature Genet* 2000;25:302–305.

268. Olanow CW, Jenner P, Beal F. Neurodegenration and neuroprotection in Parkinson's disease. *Ann Neurol* 1998;44(suppl).

269. Hirsch EC, Hunot S, Faucheux B, et al. Dopaminergic neurons degenerate by apoptosis in Parkinson's disease. *Mov Disord* 1999;14:383–385.

270. Tatton WG, Olanow CW. Apoptosis in neurodegenerative diseases: the role of mitochondria. *Biochem Biophys Acta* 1999;1410:195–213.

271. Marsden CD, Olanow CW. Neuroprotection in Parkinson's disease: the causes of Parkinson's disease are being unravelled and rational neuroprotective therapy is close to reality. *Ann Neurol* 1998;44:189–196.

125

Huntington Disease

Christopher A. Ross

Russell L. Margolis

Christopher A. Ross: Departments of Psychiatry and Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland.

Russell L. Margolis: Department of Psychiatry, Johns Hopkins University School of Medicine, Baltimore, Maryland.

Huntington disease (HD) is a progressive neurodegenerative disorder with an established genetic origin and symptoms that are referable to specific regions of brain disease. Cellular and molecular techniques are rapidly elucidating the pathogenesis of the disorder and are leading to approaches designed to develop rational treatments. Thus, HD serves as a model for the future study of those psychiatric disorders in which abnormal brain function is thought to arise from predominantly genetic factors.

CLINICAL FEATURES

Part of "125 - Huntington Disease "

HD can be described as a triad of motor, cognitive, and emotional disturbances (1 ,2). Symptoms usually begin between the ages of 35 and 50 years, although the onset may occur at any time from childhood to old age. Death occurs an average of 15 to 20 years after symptoms first appear, with some patients dying earlier from falls or suicide and others surviving for 30 to 40 years (Fig. 125.1).

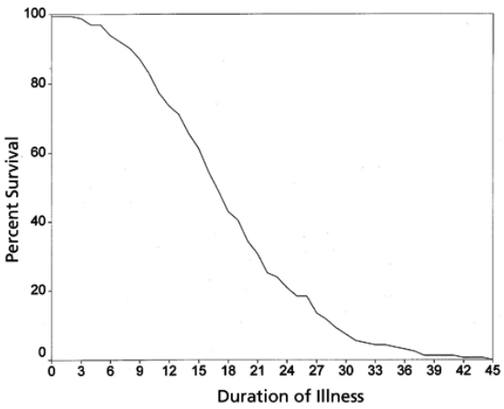


FIGURE 125.1. HD survival curve. The percentage of patients surviving as a function of years since disease onset. The curve is derived from 163 patients enrolled in the Baltimore Huntington's Disease Center with a record of both age at disease onset and age at death. (From Ross CA, Margolis RL, Rosenblatt A, et al. Reviews in molecular medicine: Huntington disease and the related disorder, dentatorubral-pallidoluysian atrophy (DRPLA). *Medicine (Baltimore)* 1997;76:305–338, with permission.)

Movement Disorders

The *movement disorder* of HD consists of two components: *involuntary* movements and abnormal *voluntary* movements. *Chorea*, or choreoathetosis, is the movement abnormality most frequently associated with HD. It consists of continuous and irregular jerky or writhing motions. Disturbances of voluntary movement, however, are more highly correlated with functional disability and disease severity, as measured by the degree of brain disease. The disordered voluntary movements observed in HD include the following: abnormal eye movements, such as slow, hypometric saccades and catchy pursuit; uncoordinated, arrhythmic, and slow fine motor movements; dysphagia and dysarthria; dysdiadochokinesis; rigidity; and gait disturbances.

The nature of the motor symptoms changes over time. The onset is usually insidious. Early complaints include clumsiness, difficulty with balance, and jerky movements or tremor. In addition to limb and truncal movements, patients may have motor tics or chorea involving respiratory, laryngeal, pharyngeal, oral, or nasal musculature. Chorea often plateaus and even wanes in the later stages of the disease, but disturbances in voluntary movement continue to progress (3). In late-stage HD, patients typically become akinetic and largely nonverbal, with severe rigidity and joint contractures. At this point, they may have few involuntary movements except for occasional movements of the entire body, resembling myoclonic jerks, when disturbed. Difficulties with swallowing commonly lead to death in HD, either directly from suffocation or aspiration or indirectly from starvation.

When HD begins in childhood or adolescence (juvenile-onset HD), the presentation is often somewhat different, with prominent bradykinesia, rigidity and dystonia, and minimal chorea. Involuntary movements may take the form of tremors, and patients may develop seizures and myoclonus.

Cognitive Disorders

Cognitive difficulties usually begin about the same time and proceed at the same rate as the abnormal movements (4), although some patients may have considerable motor impairment with very little dementia, or the reverse. Early in the course of HD, aphasia and agnosia are usually much less obvious than in the cortical dementias such as Alzheimer disease, whereas deficits in cognitive speed and flexibility are more common. In contrast to Alzheimer disease, patients with HD seem to have trouble with retrieval rather than storage of memories. They are more apt than patients with Alzheimer disease to recognize words from a previously memorized list or to respond to other cues to help them recall information. This distinction has led to the classification of HD as a *subcortical dementia* (5). Cognitive losses accumulate progressively. Deficits in memory, visuospatial abilities, and judgment develop, and patients with late-stage

HD demonstrate profound global impairment similar to patients with late-stage Alzheimer disease, although their paucity of speech makes assessment difficult.

Psychiatric Disorders

Patients with HD frequently develop *psychiatric symptoms*, most commonly depression, irritability, and apathy (3). The behavioral expression of these symptoms varies considerably, and it may include aggressive outbursts, impulsiveness, social withdrawal, and suicide. This aspect of HD can be devastating to both the patient and his or her family. The suicide rate alone, estimated at up to 12.7%, indicates the magnitude of the problem. Yet of all the complications of HD, the psychiatric manifestations are the most amenable to treatment.

Affective (mood) disorder is extremely common. Epidemiologic and phenomenologic evidence indicates that affective disorder in HD is a function of the brain disease itself, rather than a reaction to changes in life circumstance (2). HD-related major depression resembles the idiopathic form of major depression. Prominent symptoms include feelings of worthlessness or guilt, self-blame, changes in sleep and appetite, anxiety, anhedonia, loss of energy, hopelessness, and diurnal variation of mood with more severe symptoms in the morning. Delusions and hallucinations, when present, tend to be mood congruent: delusions of poverty, illness, or guilt; auditory hallucinations of derogatory or threatening voices. The diagnosis of major depression may be more difficult in patients with advanced disease, but the condition is often signaled by a departure from baseline levels of activity or functional capacity.

Severe irritability is another common symptom, present in one-third of patients in the Maryland HD survey (2). Irritability and aggression may occur in patients without a prior history of a short temper, but these symptoms are more common in patients who have had these traits all their lives. Apathy may become evident at any time in the course of the disease. Once present, it tends to persist or worsen. Irritability can coexist with apathy. Either apathy or irritability may exist independently or as part of an affective syndrome.

Patients with HD occasionally develop classic obsessive-compulsive disorder, with typical symptoms such as fear of contamination or excessive hand washing. More commonly, however, patients may display an obsessive preoccupation with particular ideas or plans (e.g., obtaining cigarettes, getting a refill of coffee) and may become irritable when these requests are not honored. Rarely, patients develop a schizophrenia-like syndrome, with prominent delusions, hallucinations, or thought disorder in the absence of an abnormal mood.

Clinical Course

In summary, adult-onset HD falls roughly into three stages. Early in the disease, manifestations include subtle changes in coordination, perhaps some minor involuntary movements, difficulty thinking through problems, and, often, a depressed or irritable mood. In the middle stage, chorea usually becomes prominent, and difficulty with voluntary motor activities becomes more evident with worsening dysarthria and dysphagia. As cognitive deficits increase, the patient becomes unable to hold a job or carry out most household responsibilities. Patients with late-stage disease may have severe chorea, but they are more often rigid and bradykinetic. They are largely nonverbal and bedridden, with a more global dementia, although they retain a significant degree of comprehension.

SYMPTOMATIC TREATMENT

Part of "125 - Huntington Disease "

There are no currently accepted specific treatments to slow the rate of clinical progression of HD (6). However, symptomatic management of both movement and emotional disturbances is possible (2). A detailed handbook of HD treatment options was prepared by Rosenblatt and colleagues (6).

Treatment of Movement Abnormalities

Chorea may be a disabling symptom, leading to bruises, fractures, or falls and impairing the ability of patients to

feed themselves. Other patients find the chorea of major cosmetic concern. Treatment with high-potency neuroleptics, such as haloperidol and fluphenazine, may be indicated in such cases, but with important caveats. These medicines may exacerbate the disturbance of voluntary movement, which, as noted earlier, correlates best with functional disability. Furthermore, neuroleptics increase morbidity by making patients more rigid, sedated, and apathetic.

If pharmacologic treatment for chorea is initiated, starting doses of neuroleptics should be low, for example, 0.5 to 1 mg of haloperidol or fluphenazine per day. Doses higher than 10 mg per day of haloperidol yield little or no benefit over lower doses. If patients experience unacceptable rigidity, akathisia or dystonic reactions to high-potency neuroleptics, lower-potency agents such as thioridazine may be better tolerated. However, use of lower-potency neuroleptics increases the risk of sedation, anticholinergic side effects, and postural hypotension.

Dopamine-depleting agents such as reserpine and tetrabenazine represent another option in the treatment of chorea. Reserpine is a known cause of drug-induced depression, and the affective state of patients receiving this agent should be monitored carefully. The benzodiazepine clonazepam may also be useful in the treatment of chorea, and it may be of benefit in the later stages of the disease, when neuroleptic medication often has little effect.

Treatment of Cognitive Abnormalities

There is no known effective pharmacologic treatment for the dementia of HD. Cholinergic agents have not been systematically assessed in HD, but the rationale for using these medications is less compelling than in Alzheimer disease, because cholinergic neurons are relatively spared in HD. Patients can be instructed to jot down notes and reminders and to sequence tasks so they can concentrate on one at a time. Complex cognitive tasks should be minimized, and, as the disease progresses, questions should be framed in a choice format with the provision of frequent cues to assist recall.

Treatment of Psychiatric Disorders

Major depression in HD responds to the same treatments used in idiopathic depression. In general, depression in HD is underdiagnosed and undertreated, perhaps because of the propensity of clinicians to see it as an understandable reaction to having the disease. Although no controlled studies exist, our experience is that both tricyclic antidepressants and selective serotonin reuptake inhibitors are effective. As with any neuropsychiatric disorder, patients should be started on low doses that are slowly increased while the patient is closely monitored for adverse effects, particularly delirium. It is important to remain with a medication for a full therapeutic trial at adequate doses and blood levels. Depressed patients should always be questioned about thoughts of suicide. When suicide is a concern, the patient should receive as few pills as possible, especially if they are to be kept in the patient's care.

The addition of antipsychotic (neuroleptic) medication is indicated as an *adjunct* to antidepressant treatment in depressed patients with hallucinations or delusions. Clozapine and other atypical neuroleptics may have the advantage over traditional neuroleptic medications, such as haloperidol, of causing fewer extrapyramidal side effects and therefore not worsening aspects of the voluntary movement disturbance. Electroconvulsive therapy is indicated for depressed patients who are refractory to treatment with medication, for patients with delusions, for those who are not eating or drinking because of their depression, or for those who are at high risk of suicide. For patients with bipolar disorder, carbamazepine, divalproex sodium, or lithium may the initial treatment of choice; again, it is prudent to start with low doses that are gradually increased until symptoms respond, side effects make further dose increases counterproductive, or therapeutic blood levels have been reached.

In treating irritability, it is important to attempt to identify and to minimize precipitants such as hunger, pain, inability to communicate, frustration with failing capabilities, boredom, difficult interpersonal relationships, and minor unexpected changes in routine. Pharmacologic treatment can be very effective. We have had success using selective serotonin reuptake inhibitors (7) and divalproex. Sexual disorders in HD, particularly aggressive hypersexuality, can be treated with antiandrogenic medications. Obsessive-compulsive disorder in HD can be treated with standard antiobsessional agents, such as selective serotonin reuptake inhibitors and clomipramine.

DIFFERENTIAL DIAGNOSIS

Part of "125 - Huntington Disease "

The clinical features of HD are often characteristic, and the diagnosis is not difficult in a patient with a known family history, typical choreiform movements, and cognitive dysfunction. The diagnosis is less clear in patients with uncharacteristic presentations or a lack of family history (1 ,8). For instance, patients may present with very little chorea or with movements that are predominantly athetoid, dystonic, or even ticlike. All the affected members of a pedigree may manifest atypical features of the disorder, such as prominent brainstem involvement, a finding contributing to diagnostic confusion. Occasional patients (particularly with late onset) may have only subtle movement abnormality and relatively little cognitive disorder (1 ,8). Fortunately, with the availability of the HD gene test, it is now possible to establish the diagnosis of HD definitively even in patients with no family history or an atypical presentation. Most patients thought to have HD on clinical examination but who do

not have the triplet repeat expansion appear to have atypical features, more characteristic of spinocerebellar ataxias or other multisystem atrophies.

HD is now recognized as part of a family of related neurodegenerative disorders, all caused by expansions of CAG repeats encoding glutamine (9). The diseases share certain clinical features, especially ataxia and dementia, and can be confused with each other. Among these diseases, Machado–Joseph disease (MIM no. 109150) and dentatorubral pallidoluysian atrophy (MIM no. 125370) are most like HD. Various other diseases may also present with HD-like symptoms, including Wilson disease, Creutzfeldt–Jakob disease, forms of ceroid neuronal lipofuscinoses, chorea with red blood cell acanthocytosis, hereditary nonprogressive chorea, paroxysmal choreoathetosis, mitochondrial disorders, corticobasal degeneration, basal ganglia calcification, forms of hereditary dystonia, Sydenham chorea, vitamin E deficiency, and cerebral vascular disease (10).

GENETIC ETIOLOGY

Part of "125 - Huntington Disease "

Discovery of the HD Mutation

HD was the first disease mapped (to chromosome 4p) using the techniques of linkage analysis with anonymous DNA probes (11). The techniques of linkage analysis were used to demonstrate that HD exhibited almost complete genetic dominance (13) and locus homogeneity (14). Many of the techniques of positional cloning, including marker development, recombination and haplotype analysis, linkage disequilibrium, physical mapping, and exon amplification, were first developed or tested during the search for the HD gene (12).

Using exon amplification and cDNA cloning, the actual gene (IT15 or huntingtin) was identified in 1993 (15). The mutation proved to be an expansion of a CAG repeat, making HD a member of a group of similar triplet repeat disorders (9 , 16 ,17 and 18). The IT15 gene is composed of 67 exons in both mice and humans (19), and in human it is located between markers D4S127 and D4S180 on chromosome 4p16.3. It spans a genomic region of more than 200 kb and is transcribed into two versions of mRNA, varying only in the length of their 3′ untranslated region, The open reading frame encodes a protein of about 350 kd with no significant homology to known proteins (15 ,19).

Genetic Diagnosis

Diagnosis of HD has been greatly simplified by the direct triplet repeat gene test. Previously, genetic diagnosis required cumbersome linkage analysis, impossible if family members were not available or were not heterozygous for the linkage markers. By contrast, the direct gene test, typically a single polymerase chain reaction, enables the length of the repeat in each allele to be measured. The test is highly sensitive and specific, although several variations in the sequence surrounding the repeat can lead to erroneous results (20).

On a population basis, there is a clear distinction between expanded and normal length repeats in huntingtin. Repeats with fewer than 29 triplets are within the normal range. The rare repeats with 29 to 35 triplets are considered of intermediate length, prone to expansion but not in themselves of sufficient length to produce a phenotype. Repeats with 36 or more triplets are considered expansions (21).

Before the discovery of the causative expansion mutation, HD was considered 100% penetrant. However, it is now clear that *penetrance* (currently defined as the presence of signs or symptoms of HD by the age of 65 years) is less than 100% in persons carrying an allele with 36 to 40 triplets. For instance, four of seven persons who were more than 70 years old and who had a 36 triplet allele had no signs or symptoms of HD. One person with a 39 triplet allele died at the age of 95 years with no definite clinical or pathologic evidence of HD (21). Because alleles with 36 to 40 triplets are quite uncommon, the frequency of nonpenetrance is difficult to estimate reliably. However, for repeats of 36 or 37 triplets, it may be on the order of 50%.

The discovery of the HD mutation revolutionized genetic counseling of presymptomatic persons at risk of HD. Between 1983 and 1993, genetic testing was based on linkage analysis, a complex procedure requiring the cooperation of several family members. HD emerged as a model for presymptomatic genetic testing, and most testing has been carried out under careful protocols involving extensive counseling and patient education (22). Most persons, including those testing positive, state that they are relieved to know the results, and this knowledge enables them to face the future with less uncertainty. Not all patients, however, have had unequivocally positive experiences with testing (23). The central difficulty stems from the availability of a presymptomatic test for a disorder with limited therapeutic interventions (at present), a dilemma termed the *Tiresias complex* from Tiresias' words to Oedipus: “It is but sorrow to be wise when wisdom profits not” (23).

Repeat Length Instability and HD Clinical Genetics

Analysis of the triplet repeat has clarified the issue of new mutations in HD. The previous belief, that new HD mutations do not occur, was disproved with the discovery of the HD repeat expansion. The lengths of normal repeats change in fewer than 1% of intergenerational transmissions; when the length does change, it is typically by only one triplet. The frequency of changes in length increases with transmission of longer repeats and becomes appreciable for alleles with repeats of intermediate length (29 to 35 triplets). During paternal transmission, these alleles are more likely to expand then to contract, and the change, unlike in normal repeats, is often of more than one or two triplets. On occasion,

the repeat can expand to a length sufficient to cause clinical HD. Therefore, new HD mutations do occur, arising from alleles in the intermediate range (24 ,25). Factors increasing instability include change of a CAA triplet adjacent to the CAG repeat into a CAG and advanced paternal age.

The nature of the HD mutation now provides a molecular basis for understanding *anticipation,* the phenomenon of increasing disease severity or decreasing age of onset in successive generations (26). Anticipation in HD has long been recognized (Fig. 125.2). In our clinical sample of affected parent-child pairs, there was no significant change in the age of onset in maternal transmission, but a mean advance of 8 years in paternal transmission. In addition, it is now clear that most patients with juvenile-onset HD arise from paternal transmission (19 ,27 ,28). Two features of the molecular genetics of HD explain the phenomenon of anticipation (29 ,30). First, the age of HD onset is inversely correlated with repeat length, a quite striking phenomenon (Fig. 125.3). Second, the length of the expanded triplet repeat is unstable in vertical transmission (Fig. 125.4). Paternal alleles more frequently expand then contract during transmission, whereas maternal alleles have an equal probability of expanding and contracting. Instability increases as repeat length increases (27 ,31). The net result, driven by paternal transmissions, is a skew toward earlier ages of onset in successive generations of a family.

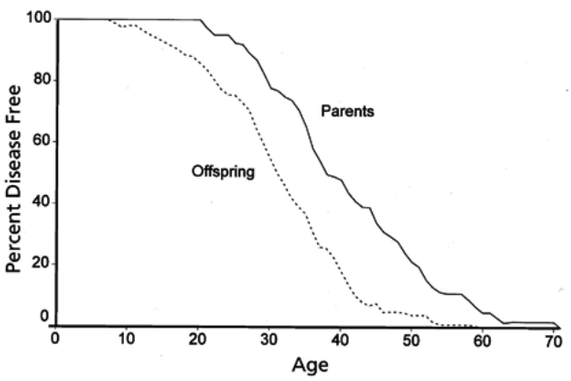


FIGURE 125.2. Anticipation in Huntington disease. The age at which affected parents and their affected children first manifest disease symptoms is depicted as a survival curve. The younger generation is affected at a substantially earlier age. Data from 61 parents and 82 children. (From Margolis RL, McInnis MG, Rosenblatt A, et al. Trinucleotide repeat expansion and neuropsychiatric disease. *Arch Gen Psychiatry* 1999;56:1019–1031, with permission.)

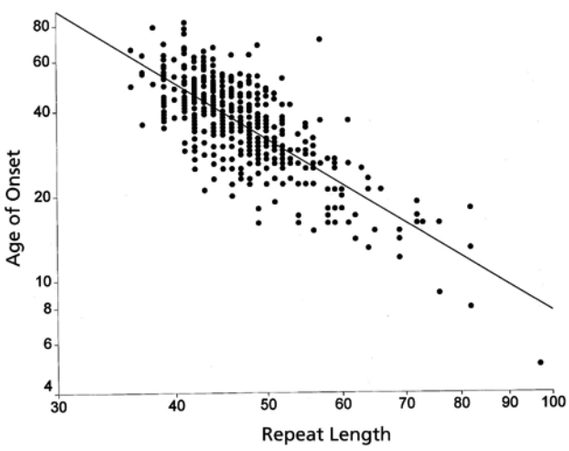


FIGURE 125.3. Correlation of repeat length with age at onset of HD. As repeat length increases, age at onset of disease decreases (n = 480; $r^2 = 0.57$). (From Margolis RL, McInnis MG, Rosenblatt A, et al. Trinucleotide repeat expansion and neuropsychiatric disease. *Arch Gen Psychiatry* 1999;56:1019–1031, with permission.)

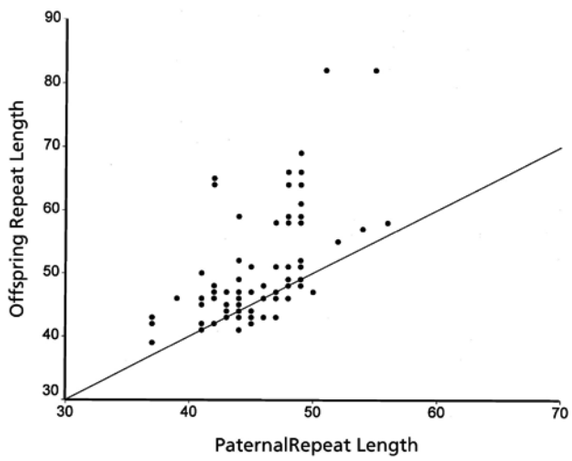


FIGURE 125.4. Increase in repeat length with paternal transmission of the HD disease allele. *Points above the diagonal line* represent cases in which the repeat length increased during transmission from father to child (n = 84 pairs; mean increase of repeat length ±SD = 4.2±0.8 triplets). (From Margolis RL, McInnis MG, Rosenblatt A, et al. Trinucleotide repeat expansion and neuropsychiatric disease. *Arch Gen Psychiatry* 1999;56:1019–1031, with permission.)

The increasing length of the repeat in paternal transmission appears to arise during spermatogenesis (32). In particular, persons with longer expanded repeats (as assayed from DNA obtained from leukocytes) have dramatically variable repeat lengths in individual sperm. The variability in oocytes appears to be less, perhaps because mature spermatocytes have undergone, on average, more cell divisions than mature oocytes. The increased likelihood of transmission of an expanded repeat as paternal age increases is consistent with this observation.

The relationship between the length of the repeat and the *rate* of clinical progression has not been resolved, although this relationship is of considerable importance in determining methods for slowing the course of HD. The rate of disease progression may be more rapid in cases with longer repeats (33), but this is not a universal finding (34). On postmortem examination, pathologic changes in cases with longer triplet repeats were more advanced than in cases with the same duration of illness but shorter repeats, a finding providing support for a correlation between repeat length and rate of progression (35). The related issue whether earlier age of onset correlates with more rapid disease progression has also not been resolved. There is no apparent correlation between repeat length and the presence of psychiatric symptoms.

PATHOLOGY

Part of "125 - Huntington Disease "

The only known pathologic changes of HD are specific to the brain, and they are characterized by striking regional selectivity of atrophy and neuronal loss (18). The most prominent atrophy is found in the caudate nucleus and putamen, which together comprise the corpus striatum within the basal ganglia. Striatal atrophy leads to hydrocephalus *ex vacuo* and marked dilatation of the lateral ventricles. In addition, there is overall atrophy of the brain. Total brain weight is reduced by 25% to 30% in advanced cases, a finding reflecting atrophy of the cerebral cortex and underlying white matter in addition to the basal ganglia.

Within the striatum, there is selective neuronal vulnerability both in the anatomic pattern of regions affected and in the particular neurons lost. Loss of neurons in the caudate and putamen shows a gradient, with early and most severe loss in the dorsal and medial regions and progressive loss of neurons in ventral and lateral regions as the disease progresses (36). There is severe loss of medium spiny projection neurons, especially those synthesizing enkephalin and γ-aminobutyric acid, but relative preservation of large and medium aspiny interneurons (37 ,38). Neuronal loss is accompanied by reactive astrocytosis (gliosis). Other areas of the basal ganglia, especially the globus pallidus and subthalamic nucleus, also become atrophic, although less than the striatum. A 0 to 4 rating scale of gross and microscopic neurodegeneration, based primarily on changes in the caudate and putamen, has been used semiquantitatively to grade the severity of HD (39). Patients with more severe neurodegeneration have greater clinical impairment before they die, and, as more recently observed, they tend to have longer expanded repeats (35 ,40).

Quantitative studies of cell number and morphology have demonstrated shrinkage and loss of neurons in the cerebral cortex (41), consistent with previous evidence from gross pathology and neuroimaging studies. Large cortical neurons appear to be most severely affected, and there is laminar specificity, with greatest loss in layer VI and significant loss in layers III and V. The neurons lost in the greatest numbers appear to project to the thalamus, whereas most neurons that project to the caudate and putamen lie in more superficial regions of layer V. In addition, the extent of cortical degeneration does not closely correlate to the severity of striatal degeneration. This set of observations indicates that the loss of neurons in the cortex does not arise simply from retrograde changes beginning in the striatum.

Whereas the atrophy and neuronal cell loss of HD have been extensively studied, less attention has focused on the morphology of the surviving neurons. Contrary to expectations, application of the Golgi metal impregnation method to study neuronal morphology in the caudate and cortex from HD cases revealed evidence, in the surviving neurons, of “regenerative” or “plastic” changes (42 ,43). Relative to neurons in these regions from normal brains, surviving neurons in HD cases had more dendrites, more long recurved dendrites, greater density and larger size of dendritic spines, and greater somatic area. A complete understanding of the pathogenesis of HD will need to encompass an explanation of these regenerative changes as well as neuronal death and brain atrophy.

Many of the clinical symptoms of HD correlate to pathologic features in the basal ganglia and cerebral cortex. The motor changes are believed to stem from interruption of a set of neuronal circuits interconnecting the cerebral cortex, the basal ganglia, and the thalamus. The involuntary and voluntary movement abnormalities in HD may arise from early degeneration of specific populations of medium spiny neurons. Changes in the basal ganglia also likely underlie the early and relatively mild “subcortical” cognitive changes seen in HD. The more severe global dementia seen late in the illness may relate to the widespread loss of neurons in the cerebral cortex, which does not appear to occur until relatively late in the disease.

Although it is relatively straightforward to correlate clinical symptoms with atrophy or neuronal cell loss, it is more difficult to determine the role of functional changes in neurons in the production of symptoms. In most cases, symptomatic HD is reflected in neuronal loss at autopsy. However, occasional cases arise of patients with clinically diagnosed HD but no discernible cell loss. These have been termed “grade zero” in the Vonsattel severity scale (39 ,44): there is no appreciable gliosis or neuronal loss. Therefore, it is possible that symptoms in these patients arise more from functional changes than from actual neuronal loss.

Degeneration is also observed, although less consistently and less prominently, in other regions of the brain, including the brainstem, the cerebellum, the lateral tuberal nucleus of the hypothalamus, the amygdala, and portions of the thalamus. The relationship between these changes and clinical features is not clear.

Studies of human postmortem HD brain tissue using antibodies directed at the N-terminus of huntingtin have

revealed small intranuclear inclusion bodies present in neurons but not in glia (45 ,46). Although they are present in other brain regions, inclusions are most abundant in the cortex and the caudate, and in those cell types (medium spiny neurons of the striatum) and cell layers (III, V, and VI of the cortex) most severely affected in HD. The density of inclusions is significantly correlated with the length of the CAG repeat in IT15 (47). Inclusions were not present in the brain of one presymptomatic person with the HD mutations.

The inclusions cannot be detected by antibodies directed at internal epitopes of huntingtin, a finding suggesting that huntingtin within the inclusions is abnormal, most likely truncated but possibly misfolded such that internal epitopes are sequestered. The inclusions can, however, be detected with antibodies to ubiquitin, a tag for proteins undergoing proteolytic degradation. This could mean that huntingtin within inclusions has been targeted for degradation but cannot be removed by proteolysis. Ultrastructural analysis of the inclusions indicates that they are composed of a mixture of granules, straight and tortuous filaments, and masses of parallel and randomly oriented fibrils, not enclosed by an intracellular membrane. Similar inclusions have been detected in transgenic mouse models of HD (48).

In addition, huntingtin has been found in aggregates in dystrophic neurites in HD-affected brains (45). These were present predominantly in cortical layers V and VI and appeared to be contained within neurofilament labeled axonal processes. Such dystrophic neurites may reflect dysfunction of retrograde axonal transport.

PATHOGENESIS

Part of "125 - Huntington Disease "

Certain lines of evidence indicate that the major pathogenic mechanisms of HD involve a toxic gain of function by the mutant protein; in other words, the abnormal length of the polyglutamine repeat gives huntingtin a toxic property not found in the wild-type protein. First, HD, like all the other polyglutamine repeat disorders, has a dominant mode of inheritance, which is typically the result of gain-of-function mutations. Second, the age of onset for homozygotes for the HD mutation generally is not markedly less than the age of onset for cases with only one copy of comparable repeat length (13), although this is not necessarily the case in the other glutamine repeat diseases. Third, no cases of HD or related polyglutamine disorders have been identified with deletions or point mutations in any of the causative genes. In contrast, the fragile X phenotype can be caused by a triplet repeat expansion leading to impaired transcription, a deletion, or a point mutation; all three types of mutations result in loss of normal protein function (49). Finally, mice with targeted deletions of the HD gene resulting in expression that is a small fraction of normal demonstrate developmental abnormalities rather than a progressive neurologic disorder (50).

Neurotoxicity Models

Before the discovery of the genetic etiology of HD, animal models of HD had been generated using neurotoxins. Injections of *N*-methyl-D-aspartate–receptor agonists, such as quinolinic acid, into the striatum induce HD-like disease, with loss of medium spiny projection neurons and sparing of cholinergic and reduced nicotinamide-adenine dinucleotide phosphate diaphorase neurons (51). Peripheral injections into rodents or primates of several mitochondrial toxins, including 3-nitropropionic acid, also reproduce aspects of striatal disease found in HD (52). Other metabolic poisons cause preferential toxicity in different regions of the brain, often those regions affected in other glutamine repeat diseases.

These neurotoxin experiments suggest several pathways that could be involved in HD cell death (53). For instance, both excitotoxicity and metabolic poisoning may be mediated, in part, by damage from free radicals (54). In addition, neurotoxic stimuli can give rise to *apoptosis*, a controlled form of cell death under control of cellular machinery that plays an essential role in normal development. The process is triggered by a group of aspartate proteases termed caspases, and glyceraldehyde phosphate dehydrogenase (GAPDH) and other metabolic enzymes may also serve as initiating factors (55 ,56).

Studies of subjects with HD have yielded results consistent with some of these neurotoxicologic mechanisms. For instance, nuclear magnetic resonance spectroscopy suggests the presence of metabolic compromise within neurons (57). Marked biochemical defects of mitochondrial complex II and complex III activity, and moderate defects of complex IV activity, have been detected in mitochondria isolated from the brain tissue of individuals with HD. Evidence of free radical activation is also present in HD postmortem tissue (58 ,59).

Huntingtin Biochemistry

The *huntingtin* protein is widely expressed in both the brain and peripheral tissues (60 ,61). The CAG repeat, even when expanded, is translated into polyglutamine (62). Immunocytochemical and subcellular fractionation studies indicate that huntingtin is present in neuronal perikarya, dendrites, and terminals, with a generally cytoplasmic localization. Huntingtin has not been consistently detected in the nucleus. The protein appears to associate with cytoskeletal elements and intracellular vesicles, with enrichment in endosomal compartments and Golgi complex membranes (63), and it is detected at all stages of embryonic and postnatal brain development. Within the striatum, huntingtin may be enriched in the medium spiny neurons, the neuronal

population most severely affected in HD; expression level may therefore contribute to the selective vulnerability of the medium spiny neurons.

To provide clues about the normal function of huntingtin and about HD pathogenesis, an intensive effort has been devoted to finding proteins that interact with huntingtin. Interactors of particular interest include HIP1, HAP1, GAPDH, and SH3GL3 (64). Some of these proteins are directly or indirectly associated with microtubule motor proteins and intracellular vesicles, findings suggesting a role for huntingtin in cytoskeletal function or vesicular transport. The interaction of huntingtin and other proteins containing glutamine repeats with the metabolic enzyme GAPDH is of potential significance, given the possible role of GAPDH in apoptosis (65). Huntingtin also interacts with the nuclear co-repressor protein, and the strength of the interaction correlates to the length of the huntingtin glutamine repeat (66). This interaction suggests that huntingtin may have some role in transcriptional regulation, although relatively little huntingtin is detected in the nucleus of normal cells under most conditions.

Polyglutamine Biochemistry

Many proteins contain stretches of polyglutamine, and such tracts are more common than repeats of other amino acids (67). However, the normal function of glutamine repeats remains unknown. Proteins containing glutamine repeats often appear to have a role in the regulation of development and neurogenesis, and certain proteins with glutamine repeats are transcription factors (68). Glutamine-rich regions may function as factor interaction domains in transcription factors, but it is unclear whether glutamine repeats serve this or more specialized functions. The lengths of glutamine repeats tend to vary considerably in homologous genes from different species; mouse huntingtin has only seven consecutive glutamines, and the puffer fish homologue has only four.

One hypothesis for the role of glutamine repeats in human disease is based on the “polar zipper” model proposed by Perutz (69). He suggested that two antiparallel ß strands of polyglutamine can be linked together by hydrogen bonds between their main chain and side chain amides, to form ß sheets and potentially to lead to protein aggregation and precipitation. Circular dichroism, electron microscopic and x-ray defraction studies of synthetic peptides, and an engineered protein provide *in vitro* evidence supporting the formation of ß strands and possibly ß sheets by glutamine repeats. Alternatively, it has been suggested that the covalent modification of glutamines by an isopeptide linkage to lysine by the enzyme transglutaminase could also lead to an insoluble precipitate of proteins containing long stretches of glutamine (70).

In support of the polar zipper hypothesis, an *in vitro* filter assay was used to demonstrate that a short truncation (exon 1 only) of the huntingtin protein with an expanded glutamine repeat can aggregate to form amyloid-like fibrils (71 ,72). These fibrils showed green birefringence when they were stained with Congo red and viewed by polarized light microscopy, consistent with the presence of ß sheets. Aggregation did not occur when the polyglutamine repeat was of normal length. In addition, aggregation only occurred when a huntingtin fragment with an expansion of typical length was first cleaved from the carrier protein to which it was fused during synthesis for the assay. The implication, consistent with cell and mouse models described later, is that generation of a proteolytic fragment of HD may be an important step in HD pathogenesis. In fact, consensus cleavage sites for caspase-3 (also termed apopain or CPP32) exist at approximately position 513 and 530 of huntingtin, and huntingtin can be cleaved by purified caspase-3, caspase-6, and caspase-8.

Cell Models

Research into the pathogenesis of HD has been greatly facilitated by the development of cell models. Two approaches have been used. In the first, huntingtin is introduced into cells through transient transfection; in the other, cell lines are engineered that stably express huntingtin (77 ,78 ,79 and 80). In general, short truncations of huntingtin containing the expanded polyglutamine appear to be much more toxic than full-length huntingtin and more liable to aggregation (81 ,82 and 83). However, aggregate formation and cellular toxicity can be disassociated, a finding suggesting that cell toxicity is not related in a simple way to aggregation (79 ,84 ,85). Elimination of caspase cleavages sites may reduce the toxicity of mutant huntingtin (9 ,73 ,74 ,75 and 76). Cell death does not correspond to all characteristics of apoptosis, but it can be decreased or blocked in several models with caspase inhibitors (79 ,84).

A role of nuclear localization for huntingtin toxicity has been suggested, but it is still not proved. Transfection of primary neurons with constructs incorporating a nuclear export signal diminished toxicity, whereas the addition of nuclear localization signals appeared to enhance toxicity (77 ,79) However, other studies suggested that both the nucleus and the cytoplasm can be the site of pathogenesis (86).

Transgenic Mouse Models

Transgenic animal models have provided some of the most striking evidence for the *gain-of-function hypothesis* of HD pathogenesis. The first animal model of HD was constructed using exon 1 of huntingtin with a very long expanded repeat (87). These animals developed progressive neurologic deficits strikingly similar to those of HD, including incoordination, abnormal involuntary movements, seizures, and weight loss (88 ,89). However, unlike patients with HD, neuronal cell loss is not prominent. These mice

also developed intranuclear inclusions containing the truncated huntingtin transgene product, but not the endogenous huntingtin protein (48). The intranuclear inclusions are present at the time, and perhaps before, the animals have neurologic signs or brain or body weight loss. The intranuclear inclusions are clearly distinct from the nucleolus, and no membrane separates them from the rest of the nucleus.

Several other HD mouse models have been constructed. A truncated N-terminal fragment of huntingtin driven by the prion protein promoter resulted in mice with features very similar to the initial model (90): the mice develop progressive hypoactivity, incoordination, and weight loss, and on neuropathologic examination they have both intranuclear inclusions and neuritic aggregates. A transgene consisting of a full-length huntingtin cDNA driven by the CMV promoter resulted in a line of mice with a rather different phenotype, characterized by early weight gain and hyperactivity followed later by hypoactivity. These mice have both intranuclear inclusions and some loss of neurons (91).

Perhaps the most promising mouse model of HD involves the use of YAC constructs, so the transgene consists of the entire human HD gene, including the human HD promotor and all introns, with an expanded repeat. These mice develop neurologic signs, electrophysiologic abnormalities, and a shortened life span (92). A single founder with a long repeat had striking evidence of selective striatal neurodegeneration and nuclear localization of N-terminal epitopes of huntingtin in striatal neurons. If additional lines can be generated, this model may be the closest to the human disease of any model yet generated.

Another model of potential utility was generated by inserting an expansion of polyglutamine into the mouse huntingtin gene, thus avoiding the confounding factor of the presence of the human transgene. So far, these mice have not developed neurologic signs, and no neuronal loss has been detected (93). There is evidence of translocation of huntingtin into the nucleus in striatal neurons. Thus, these mice may model early aspects of HD pathogenesis and could provide a useful model for studying the early features of the disease.

The construction of an inducible mouse model of HD has yielded insight into HD pathogenesis (94). A transgene containing exon 1 of huntingtin with an expanded glutamine repeat under the control of the tet-off system was inserted so that the timing of transgene expression could be externally controlled by the presence or absence of an antibiotic in the animals' food. With the transgene on, mice developed neurologic signs and neuropathologic changes including nuclear inclusions. Remarkably, when the expression of huntingtin was turned off, these abnormalities partially reversed. This surprising result suggests that the brain may have more restorative and plastic ability than previously appreciated, and that if the pathologic changes of HD could be halted, substantial repair would perhaps be possible.

Invertebrate Models

Invertebrate models offer the potential of using powerful genetic techniques to search for genetic factors that enhance or suppress an experimentally induced phenotype. Several *Drosophila* models of polyglutamine-induced neurodegeneration have been generated (95 ,96 and 97), with many of the same features of neuronal degeneration observed in mammalian cell models and mouse models. Genetics screens have been used to demonstrate that molecular chaperones such as HDJ1 and HSP70 can suppress the phenotype (97 ,98). *Caenorhabditis elegans* models may also prove to be of similar value (99 ,100).

A MODEL OF POLYGLUTAMINE PATHOGENESIS

Part of "125 - Huntington Disease "

A model for HD pathogenesis is depicted in Figure 125.5 . Several of the steps are speculative, and, as has been evident in this summary, the data supporting the model are at times conflicting. Here we highlight areas of uncertainty:

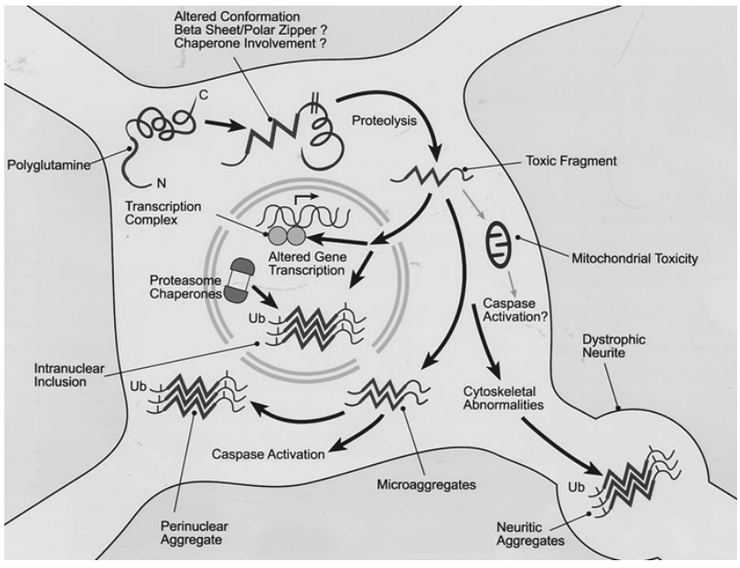


FIGURE 125.5. A model of HD pathogenesis. We propose that HD pathogenesis begins with altered conformation of the protein containing the expanded polyglutamine repeat. Proteolysis generates a fragment that leads to toxicity through several pathways. Nuclear importation may lead to altered gene transcription with a detrimental effect on cell survival. Inclusions also form in the nucleus but may not be a major cause of cell death. Huntingtin fragments may interfere with mitochondrial energy metabolism, either directly, or more likely indirectly, perhaps by altered gene transcription. Microaggregation of the fragment may lead to caspase activation and the consequent initiation of cell death pathways. Fragments may be transported into neurites, interfering with cytoskeletal function. Toxicity may also be mediated by the full-length huntingtin protein with the expanded repeat. As discussed in the text, many of the steps remain speculative.

Evidence for pathogenetically significant proteolytic cleavage of huntingtin remains indirect. The finding that nuclear inclusions can be labeled for N-terminal epitopes, but not for internal or C-terminal epitopes, is consistent with proteolytic cleavage, but it could also result from masking of epitopes. In addition, details of the cleavage, including whether it is processive or endoproteolytic and where in the cell it may occur, are uncertain. The figure shows cleavage taking place in the cytoplasm, but this is speculative.

Another uncertainty is the role of aggregation. It seems likely that the mutant protein adopts an abnormal confirmation and that this is a necessary step in pathogenesis. However, the large inclusion bodies visible by light microscopy may represent a downstream event, an epiphenomenon, or even a protective reaction, and therefore they may not be directly tied to pathogenesis. Our hypothesis is that mutant huntingtin adopts an abnormal confirmation, leading to abnormal interaction with other proteins, including the sequestration of proteins that contain polyglutamine repeats. An example of such protein is CREB binding protein, an important transcriptional regulator; sequestration of this and similar proteins could have marked effects on neuronal function and survival.

The relative importance of nuclear and cytoplasmic events also remains unclear. In our model, the location of huntingtin in the nucleus leads to altered gene transcription. However, huntingtin could act also within the cytoplasm to interfere with proteins that could otherwise be imported into the nucleus. Furthermore, huntingtin may interact with cytoplasmic molecules, including microtubules or microtubule motors, caspase adaptors, and other proteins, to yield toxicity.

THERAPY

Part of "125 - Huntington Disease "

Recent biochemical, cell, and animal studies are beginning to suggest approaches for development of rational therapeutics (Fig. 125.6). As described earlier, current therapeutics for HD are limited to symptomatic treatments, so any intervention that can stop or slow disease progression would be a major advance.

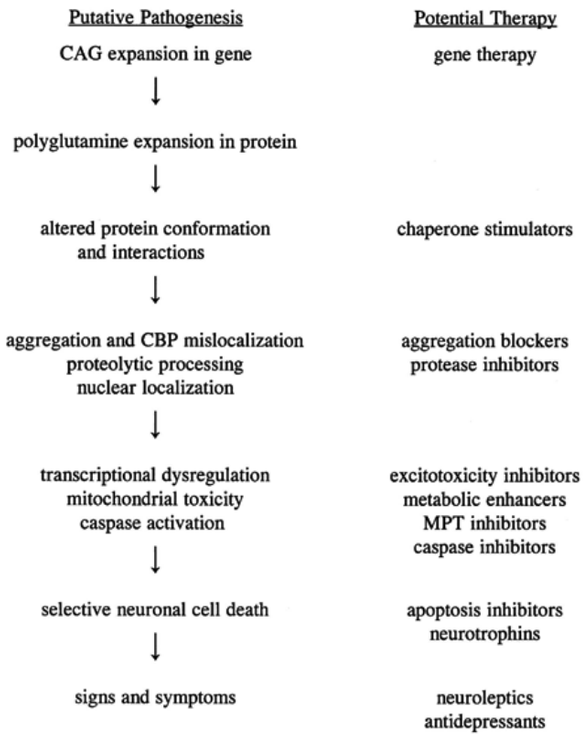


FIGURE 125.6. Approaches to Huntington disease (HD) therapy. The pathogenetic model outlined in Fig. 125.5 suggests certain strategies for slowing, stopping, or even preventing the manifestations of HD.

The first generation of agents designed to slow the progression or delay the onset of HD has emerged from neurotoxicologic models of HD. As of this writing, a major multicenter trial is under way, termed the CARE-HD Study (Coenzyme Q and Remacemide Evaluation in Huntington's Disease). Coenzyme Q is a mitochondrial cofactor, and remacemide is a glutamate-receptor antagonist, and both drugs have efficacy in neurotoxicologic mouse models. The CARE-HD Study is the first multicenter HD drug trial and involves 340 patients treated for 30 months under the sponsorship of the Huntington Study Group. Neurotoxicologic models have also led to smaller trials of other agents, including vitamin E, idebenone, and lamotrigine. None of these agents has demonstrated any clear efficacy (101 ,102 ,103 ,104 and 105).

Another approach to HD therapeutics involves transplantation either of fetal striatal cells or of cells secreting growth factors. These methods have met with at best limited success so far, but they may have promise for the future. Stem cells, if such cells can be differentiated into striatal neurons in a controlled fashion, may have great potential as therapeutic agents.

A screen for therapeutic agents that may alter polyglutamine aggregation is currently in progress using the filter assay developed by Wanker and colleagues. Early results indicate that Congo red and related dyes that intercalate into ß sheets (106) and members of the heat shock protein–chaperone family (16) can reduce aggregation. The effect of these agents in animal or cell models of HD is unknown, and even if effective, it is unclear whether these agents would themselves be good candidates for therapeutic compounds. Nonetheless, these results are very encouraging. New compounds may emerge based on those already shown to be effective *in vitro*, or entirely new classes of effective agents may be discovered with further screening.

Mouse models are also in use to screen for therapeutic compounds. Based on work in cells, the role of caspase inhibitors on disease progression has been investigated. Caspase inhibition led to a modest but significant beneficial effect in the exon 1 HD transgenic mouse model (107). A genetic cross of these transgenic mice with a line of mice overexpressing a caspase 1 dominant negative construct (and hence deficient in caspase-1 activity) also suggested that caspase inhibition can slow disease progress. Creatine, chosen based on its effect in neurotoxicologic models of neuronal cell death, also has a significant effect on disease progression in the exon 1 HD transgenic mouse (108). Minocycline, which has shown some efficacy in ischemic models, also has a beneficial effect on this mouse line (109). The magnitude of the effect on disease progression observed in these studies is significant but modest. A similar effect can be induced by altering the environment in which the mice are raised (110). It is possible that a combination of several of these agents may have enhanced effectiveness. Moreover, most of the agents tested so far have been targeted at relatively downstream points in the presumed pathogenetic pathway; agents that are targeted at earlier steps in the pathogenetic pathway may be more effective. As understanding of the pathogenesis of HD advances, the development of therapeutics should follow soon thereafter.

ACKNOWLEDGMENTS

Part of "125 - Huntington Disease "

This work is supported by National Institutes of Health grants NS 16375, 34172, 38144, and 38054, and grants from the Huntington's Disease Society of America (HDSA) and Hereditary Disease Foundation (HDF). We thank the staff of the Baltimore Huntington's Disease Center, founded by Drs. Paul McHugh, Susan Folstein, and Marshal Folstein. We thank Ms. Debbie Pollard and Ms. Marie Sonderman for expert assistance in manuscript preparation and Dr. Adam Rosenblatt for assistance with data analysis and figure construction. Finally, we thank our patients with HD and their families for their inspiration, patience, and cooperation. Parts of this review are adapted from previously published material (10 ,111 ,112), with permission from the publishers.

DISCLAIMER

Part of "125 - Huntington Disease "

Dr. Ross has received research support from Astra-Zeneca and serves as a consultant for Amgen.

REFERENCES

1. Harper PS. *Huntington disease*. London: WB Saunders, 1996.

2. Folstein SE. *Huntington's disease: a disorder of families*. Baltimore: Johns Hopkins University Press, 1989.

3. Penney JB Jr, Young AB, Shoulson I, et al. Huntington's disease in Venezuela: seven years of follow-up on symptomatic and asymptomatic individuals. *Mov Disord* 1990;5:93–99.

4. Brandt J, Butters N. The neuropsychology of Huntington's disease. *Trends Neurosci* 1986;9:118–120.

5. Brandt J, Folstein SE, Folstein MF. Differential cognitive impairment in Alzheimer disease and Huntington's disease. *Ann Neurol* 1988;23:555–561.

6. Rosenblatt A, Ranen N, Nance M, et al. *A physician's guide to the management of Huntington's disease*, second ed. New York: Huntington's Disease Society of America, 1999.

7. Ranen NG, Lipsey JR, Treisman G, et al. Setraline in the treatment of severe aggressiveness in Huntington's disease. *J Neuropsychiatry Clin Neurosci* 1996;8:338–340.

8. Hayden MR. *Huntington's chorea*. New York: Springer-Verlag, 1981.

9. Zoghbi HY, Orr HT. Glutamine repeats and neurodegeneration. *Annu Rev Neurosci* 2000;23:217–247.

10. Ross CA, Margolis RL, Rosenblatt A, et al. Reviews in molecular medicine: Huntington disease and the related disorder, dentatorubral-pallidoluysian atrophy (DRPLA). *Medicine (Baltimore)* 1997;76:305–338.

11. Gusella JF, Wexler NS, Conneally PM, et al. A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* 1983;306:234–238.

12. Nance MA. Huntington disease: another chapter rewritten. *Am J Hum Genet* 1996;59:1–6.

13. Wexler NS, Young AB, Tanzi RE, et al. Homozygotes for Huntington's disease. *Nature* 1987;326:194–197.

14. Folstein SE, Conneally PM, Wexler NS, et al. Localization of the Huntington's disease gene to a small segment of chromosome 4 flanked by D4S10 and the telomere. *Cell* 1987;50:565–571.

15. Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 1993;72:971–983.

16. Muchowski PJ, Schaffar G, Sittler A, et al. Hsp70 and Hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. *Proc Natl Acad Sci USA* 2000;97:7841–7846.

17. Gusella JF, Persichetti F, MacDonald ME. The genetic defect causing Huntington's disease: repeated in other contexts? *Mol Med* 1997;3:238—246.

18. Ross CA. When more is less: pathogenesis of glutamine repeat neurodegenerative diseases. *Neuron* 1995;15:493–496.

19. Ambrose CM, Duyao MP, Barnes G, et al. Structure and expression of the Huntington's disease gene: evidence against simple inactivation due to an expanded CAG repeat. *Somat Cell Mol Genet* 1994;20:2744-28.

20. Margolis RL, Stine OC, Callahan C, et al. Two novel single-base-pair substitutions adjacent to the CAG repeat in the huntington disease gene (IT15): implications for diagnostic testing [Letter]. *Am J Hum Genet* 1999;64:323–326.

21. Rubinsztein DC, Leggo J, Coles R, et al. Phenotypic characterization of individuals with 30–40 CAG repeats and apparently normal elderly individuals with 36–39 repeats. *Am J Hum Genet* 1996;59:16–22.

22. Brandt J, Quaid KA, Folstein SE, et al. Presymptomatic diagnosis of delayed-onset disease with linked DNA markers: the experience in Huntington's disease. *JAMA* 1989;261:3108–3114.

23. Wexler NS. The Tiresias complex: Huntington's disease as a paradigm of testing for late-onset disorders. *FASEB J* 1992;6:2820–2825.

24. Goldberg YP, Kremer B, Andrew SE, et al. Molecular analysis of new mutations for Huntington disease: intermediate alleles and sex of origin effects. *Nat Genet* 1993;5:174–179.

25. Myers RH, MacDonald ME, Koroshetz WJ, et al. *De novo* expansion of a $(CAG)_n$ repeat in sporadic Huntington disease. *Nat Genet* 1993;5:169–273.

26. McInnis MG. Anticipation an old idea in new genes. *Am J Hum Genet J* 1996;59:973–979.

27. Duyao M, Ambrose C, Myers R, et al. Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nat Genet* 1993;4:387–392.

28. Stine OC, Pleasant N, Franz ML, et al. Correlation between the onset of Huntington's disease and length of the trinucleotide repeat in IT-15. *Hum Mol Genet* 1993;2:1547–1549.

29. Ranen NG, Stine OC, Abbott MH, et al. Anticipation and instability of (CAG)n repeats in IT-15 in parent-offspring pairs with Huntington's disease. *Am J Hum Genet* 1995;57:593-602.

30. Andrew SE, Goldberg YP, Kremer B, et al. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat Genet* 1993;4:398–403.

31. Durr A, Dode C, Hahn V, et al. Diagnosis of “sporadic” Huntington's disease. *J Neurol Sci* 1995;129:51–55.

32. MacDonald ME, Barnes G, Srinidhi J, et al. Gametic but not somatic instability of CAG repeat length in Huntington's disease. *J Med Genet* 1993;30:982–986.

33. Aylward EH, Li Q, Stine OC, et al. Longitudinal change in basal ganglia volume in patients with Huntington's disease. *Neurology* 1997;48:394–399.

34. Kieburtz K, MacDonald M, Shih C, et al. Trinucleotide repeat length and progression of illness in Huntington's disease. *J Med Genet* 1994;31:872–874.

35. Furtado S, Suchowersky O, Rewcastle B, et al. Relationship between trinucleotide repeats and neuropathological changes in Huntington disease. *Ann Neurol* 1996;39:132–136.

36. Ferrante RJ, Kowall NW, Beal MF, et al. Selective sparing of a class of striatal neurons in Huntington's disease. *Science* 1985;230:561–563.

37. Albin RL, Reiner A, Anderson KD, et al. Preferential loss of striato external pallidal projection neurons in presymptomatic Huntington's disease. *Ann Neurol* 1992;31:425–430.

38. Kowall NW, Ferrante RJ, Martin JB. Patterns of cell loss in Huntington's disease. *Trends Neurosci* 1987;10:24–29.

39. Vonsattel J-P, Myers RH, Stevens TJ, et al. Neuropathological classification of Huntington's disease. *J Neuropathol Exp Neurol* 1985;44:559–577.

40. Myers RH, Vonsattel J-P, Stevens TJ, et al. Clinical and neuropathologic assessment of severity in Huntington disease. *Neurology* 1988;38:341–347.

41. Hedreen JC, Peyser CE, Folstein SE, et al. Neuronal loss in layers V and VI of cerebral cortex in Huntington's disease. *Neurosci Lett* 1991;133:257–261.

42. Graveland GA, Williams RS, DiFiglia M. A Golgi study of the human neostriatum: neurons and afferent fibers. *J Comp Neurol* 1985;234:317–333.

43. Sotrel A, Williams RS, Kaufmann WE, et al. Evidence for neuronal degeneration and dendritic plasticity in cortical pyramidal neurons of Huntington's disease: a quantitative Golgi study. *Neurology* 1993;43:2088–2986.

44. Folstein SE, Phillips JA III, Meyers DA, et al. Huntington's disease: two families with differing clinical features show linkage to the G8 probe. *Science* 1985;229:776–779.

45. DiFiglia M, Sapp E, Chase KO, et al. Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* 1997;277:1990–1993.

46. Becher MW, Kotzuk JA, Sharp AH, et al. Intranuclear neuronal inclusions in Huntington's disease and dentatorubral and pallidoluysian atrophy: correlation between the density of inclusions and IT15 CAG triplet repeat length. *Neurobiol Dis* 1998;4:387–397.

47. Ross CA, Margolis RL, Becher MW, et al. Pathogenesis of polyglutamine neurodegenerative diseases: towards a unifying mechanism. In: Wells RD, Warren ST, eds. *Genetic instabilities and hereditary neurological diseases*. San Diego: Academic, 1998.

48. Davies SW, Turmaine M, Cozens BA, et al. Formation of neuronal intranuclear inclusions (NII) underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* 1997;90:537–548.

49. Kaufmann WE, Reiss AL. Molecular and cellular genetics of fragile X syndrome. *Am J Med Genet* 1999;88:11–24.

50. White JK, Auerbach W, Duyao MP, et al. Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. *Nat Genet* 1997;17:404–410.

51. Beal MF, Kowall NW, Ellison DZW, et al. Replication of the neurochemical characteristics of Huntington's disease by quinolinic acid. *Nature* 1986;321:168–171.

52. Beal MF, Brouillet E, Jenkins B, et al. Neurochemical and histologic characterization of striatal excitotoxic lesions produced by the mitochondrial toxin 3-nitropropionic acid. *J Neurosci* 1993;13:4181–4192.

53. Grunewald T, Beal MF. Bioenergetics in Huntington's disease. *Ann NY Acad Sci* 1999;893:203–213.

54. Facchinetti F, Dawson VL, Dawson TM. Free radicals as mediators of neuronal injury. *Cell Mol Neurobiol* 1998;18:667–682.

55. Utz PJ, Anderson P. Life and death decisions: regulation of apoptosis by proteolysis of signaling molecules. *Cell Death Differ* 2000;7:589–602.

56. Wolozin B, Behl C. Mechanisms of neurodegenerative disorders. II. Control of cell death. *Arch Neurol* 2000;57:801–804.

57. Koroshetz WJ, Jenkins BG, Rosen BR, et al. Energy metabolism defects in Huntington's disease and effects of coenzyme Q_{10}. *Ann Neurol* 1997;41:160–165.

58. Browne SE, Bowling AC, MacGarvey U, et al. Oxidative damage and metabolic dysfunction in Huntington's disease: selective vulnerability of the basal ganglia. *Ann Neurol* 1997;41:646–653.

59. Gu M, Gash MT, Mann VM, et al. Mitochondrial defect in Huntington's disease caudate nucleus. *Ann Neurol* 1996;39:385–389.

60. Sharp AH, Loev SJ, Schilling G. et al. Widespread expression of the Huntington's disease gene (IT15) protein product. *Neuron* 1995;14:1065–1074.

61. DiFiglia M, Sapp E, Chase K, et al. Huntingtin is a cytoplasmic protein associated with vesicles in human and rat brain neurons. *Neuron* 1995;14:1075–1081.

62. Trottier Y, Lutz Y, Stevanin G, et al. Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature* 1995;378:403–406.

63. Velier J, Schwarz C, Young C, et al. Wild-type and mutant huntingtin localize to the golgi complex and to vesicles in the peripheral cytoplasm in fibroblasts of control and HD patients. *Soc Neurosci Abstr* 1996;22:226.

64. Gusella JF, MacDonald ME. Huntingtin: a single bait hooks many species. *Curr Opin Neurobiol* 1998;8:425–430.

65. Burke JR, Enghild JJ, Martin ME, et al. Huntingtin and DRPLA proteins selectively interact with the enzyme GAPDH. *Nat Med* 1996;347:350.

66. Boutell JM, Thomas P, Neal JW, et al. Aberrant interactions of transcriptional repressor proteins with the Huntington's disease gene product, huntingtin. *Hum Mol Genet* 1999;8:1647–1655.

67. Green H. Human genetic diseases due to codon reiteration: relationship to an evolutionary mechanism. *Cell* 1993;74:955–956.

68. Margolis RL, Abraham MA, Gatchell SB, et al. cDNAs with long CAG trinucleotide repeats from human brain. *Hum Genet* 1997;100:114–122.

69. Perutz M. Polar zippers: Their role in human disease. In: *Protein science*, third ed. London: Cambridge University Press; 1994:1629–1637.

70. Kahlem P, Terre C, Green H, et al. Peptides containing glutamine repeats as substrates for transglutaminase-catalyzed cross-linking: relevance to diseases of the nervous system [see Comments]. *Proc Natl Acad Sci USA* 1996;93:14580–14585.

71. Georgalis Y, Starikov EB, Hollenbach B, et al. Huntingtin aggregation monitored by dynamic light scattering. *Proc Natl Acad Sci USA* 1998;95:6118-6121.

72. Scherzinger E, Sittler A, Schweiger K, et al. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. *Proc Natl Acad Sci USA* 1999;96:4604–4609.

73. Wellington CL, Ellerby LM, Hackam AS, et al. Caspase cleavage of gene products associated with triplet expansion disorders generates truncated fragments containing the polyglutamine tract. *J Biol Chem* 1998;273:9158–9167.

74. Wellington CL, Hayden MR. Caspases and neurodegeneration: on the cutting edge of new therapeutic approaches. *Clin Genet* 2000;57:1–10.

75. Wellington CL, Singaraja R, Ellerby L, et al. Inhibiting caspase cleavage of huntingtin reduces toxicity and aggregate formation in neuronal and nonneuronal cells. *J Biol Chem* 2000;275:19831–19838.

76. Sanchez I, Xu CJ, Juo P, et al. Caspase-8 is required for cell death induced by expanded polyglutamine repeats [see Comments]. *Neuron* 1999;22:623–633.

77. Peters MF, Nucifora FC Jr, Kushi J, et al.. Nuclear targeting of mutant huntingtin increases toxicity. *Mol Cell Neurosci* 1999;14:121–128.

78. Lunkes A, Mandel JL. A cellular model that recapitulates major pathogenic steps of Huntington's disease. *Hum Mol Genet* 1998;7:1355–1361.

79. Saudou F, Finkbeiner S, Devys D, et al. Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. *Cell* 1998;95:55–66.

80. Moulder KL, Onodera O, Burke JR, et al. Generation of neuronal intranuclear inclusions by polyglutamine-GFP: analysis of inclusion clearance and toxicity as a function of polyglutamine length. *J Neurosci* 1999;19:705–715.

81. Hackam AS, Singaraja R, Wellington CL, et al. The influence of huntingtin protein size on nuclear localization and cellular toxicity. *J Cell Biol* 1998;141:1097–1105.

82. Martindale D, Hackam A, Wieczorek A, et al. Length of huntingtin and its polyglutamine tract influences localization and frequency of intracellular aggregates. *Nat Genet* 1998;18:150–154.

83. Cooper JK, Schilling G, Peters MF, et al. Truncated N-terminal fragments of huntingtin with expanded glutamine repeats form nuclear and cytoplasmic aggregates in cell culture. *Hum Mol Genet* 1998;7:783–790.

84. Kim M, Lee HS, LaForet G, et al. Mutant huntingtin expression in clonal striatal cells: dissociation of inclusion formation and neuronal survival by caspase inhibition. *J Neurosci* 1999;19:964–973.

85. Li H, Li SH, Cheng AL, et al. Ultrastructural localization and progressive formation of neuropil aggregates in Huntington's disease transgenic mice. *Hum Mol Genet* 1999;8:1227–1236.

86. Ellerby LM, Andrusiak RL, Wellington CL, et al. Cleavage of atrophin-1 at caspase site aspartic acid 109 modulates cytotoxicity. *J Biol Chem* 1999;274:8730–8736.

87. Mangiarini L, Sathasivam S, Seller M, et al. Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* 1996;87:493–506.

88. Carter RJ, Lione LA, Humby T, et al. Characterization of progressive motor deficits in mice transgenic for the human Huntington's disease mutation. *J Neurosci* 1999;19:3248–3257.

89. Murphy KP, Carter RJ, Lione LA, et al. Abnormal synaptic plasticity and impaired spatial cognition in mice transgenic for exon 1 of the human Huntington's disease mutation. *J Neurosci* 2000;20:5115–5123.

90. Schilling G, Becher MW, Sharp AH, et al. Intranuclear inclusions and neuritic aggregates in transgenic mice expressing a mutant N-terminal fragment of huntingtin. *Hum Mol Genet* 1999;8:397–407.

91. Reddy PH, Williams M, Charles V, et al. Behavioral abnormalities and selective neuronal loss in HD transgenic mice expressing mutated full-length HD cDNA. *Nat Genet* 1998;20:198–202.

92. Hodgson JG, Agopyan N, Gutekunst CA, et al. A YAC mouse model for Huntington's disease with full-length mutant huntingtin, cytoplasmic toxicity, and selective striatal neurodegeneration. *Neuron* 1999;23:181–192.

93. Wheeler VC, White JK, Gutekunst CA, et al. Long glutamine tracts cause nuclear localization of a novel form of huntingtin in medium spiny striatal neurons in HdhQ92 and HdhQ111 knock- in mice. *Hum Mol Genet* 2000;9:503–513.

94. Yamamoto A, Lucas JJ, Hen R. Reversal of neuropathology and motor dysfunction in a conditional model of Huntington's disease [see Comments]. *Cell* 2000;101:57–66.

95. Warrick JM, Paulson HL, Gray-Board GL, et al. Expanded polyglutamine protein forms nuclear inclusions and causes neural degeneration in *Drosophila. Cell* 1998;93:939–949.

96. Jackson GR, Salecker I, Dong X, et al. Polyglutamine-expanded human huntingtin transgenes induce degeneration of *Drosophila* photoreceptor neurons. *Neuron* 1998;21:633–642.

97. Kazemi-Esfarjani P, Benzer S. Genetic suppression of polyglutamine toxicity in *Drosophila. Science* 2000;287:1837–1840.

98. Warrick JM, Chan HY, Gray-Board GL, et al. Suppression of polyglutamine-mediated neurodegeneration in *Drosophila* by the molecular chaperone HSP70. *Nat Genet* 1999;23:425–428.

99. Faber PW, Alter JR, MacDonald ME, et al. Polyglutamine-mediated dysfunction and apoptotic death of a *Caenorhabditis elegans* sensory neuron. *Proc Natl Acad Sci U S A* 1999;96:179–184.

100. Satyal SH, Schmidt E, Kitagawa K, et al. Polyglutamine aggregates alter protein folding homeostasis in *Caenorhabditis elegans. Proc Natl Acad Sci USA* 2000;97:5750–5755.

101. Shoulson I, Odoroff C, Oakes D, et al. A controlled clinical trial of baclofen as protective therapy in early Huntington's disease. *Ann Neurol* 1989;25:252–259.

102. Kieburtz K, Feigin A, McDermott M, et al. A controlled trial of remacemide hydrochloride in Huntington's disease. *Mov Disord* 1996;11:273–277.

103. Feigin A, Keiburtz K, Como P, et al. Assessment of coenzyme Q10 tolerability in Huntington's disease. *Mov Disord* 1996;11:321–323.

104. Ranen NG, Peyser CE, Coyle J, et al. A controlled trial of idebenone in Huntington's disease. *Mov Disord* 1996;11:549–554.

105. Peyser CE, Folstein MF, Chase GA, et al. Trial of *d*-alpha-tocopherol in Huntington's disease. *Am J Psychiatry* 1995;152:1771–1775.

106. Heiser V, Scherzinger E, Boeddrich A, et al. Inhibition of huntingtin fibrillogenesis by specific antibodies and small molecules: implications for Huntington's disease therapy. *Proc Natl Acad Sci USA* 2000;97:6739–6744.

107. Ona VO, Li M, Vonsattel J-P, et al. Inhibition of caspase-1 slows disease progression in a mouse model of Huntington's disease [see Comments]. *Nature* 1999;399:263–267.

108. Ferrante RJ, Andreassen OA, Jenkins BG, et al. Neuroprotective effects of creatine in a transgenic mouse model of Huntington's disease. *J Neurosci* 2000;20:4389–4397.

109. Chen M, Ona VO, Li M, et al. Minocycline inhibits caspase-1 and caspase-3 expression and delays mortality in a transgenic mouse model of Huntington disease. *Nat Med* 2000;6:797–801.

110. van Dellen A, Blakemore C, Deacon R, et al. Delaying the onset of Huntington's in mice. *Nature* 2000;404:721–722.

111. Margolis RL, McInnis MG, Rosenblatt A, et al. Trinucleotide repeat expansion and neuropsychiatric disease. *Arch Gen Psychiatry* 1999;56:1019–1031.

112. Ross CA, Hayden MR. Huntington's disease. In: Rubinsztein DC, Hayden MR, eds. *Analysis of triplet repeat disorders*. Oxford: Bios, 1998:169–208.

126

Clinical Course and Cellular Pathology of Tardive Dyskinesia

Carol A. Tamminga

Margaret G. Woerner

Carol A. Tamminga: Maryland Psychiatric Research Center, University of Maryland School of Medicine, Baltimore, Maryland.

Margaret G. Woerner: Department of Psychiatry Research, Hillside Hospital, Long Island Jewish Health System, Glen Oaks, New York.

Tardive dyskinesia (TD) is an iatrogenic human hyperkinetic movement disorder associated with chronic antipsychotic drug treatment. The cause of the syndrome, namely, chronic dopamine-receptor blockade, is specifically known and is straightforward to model. Thus, disease pathophysiology has been approached with surrogate preparations. Traditionally, factors mediating TD have been sought in striatal dopaminergic transmission. However, several incompatibilities have developed between the characteristics of the dopaminergic model and the presentation of the disorder (10). γ-aminobutyric acid (GABA)–ergic transmission within the basal ganglia has also been targeted as the putative pathophysiologic agent in TD. Here, although biochemical characteristics between the model and the disease have appeared compatible, therapeutic implications have not been fully met, possibly because of inadequate pharmacologic tools (10).

Research has suggested that both these transmitter alterations, dopaminergic and GABAergic, may reflect an action of antipsychotic drugs on neuronal activity within the basal ganglia thalamocortical motor circuit. Clinically, antipsychotic drug action overall tends to normalize mental status in psychosis and produces symptom remission in diverse psychotic illnesses, including schizophrenia, bipolar disease, and dementia (8); parkinsonism characteristically occurs with traditional antipsychotics as a major side effect (42). On a cellular basis, gradual alterations produced by these drugs over time, within selected subcortical brain regions and at selected synapses of this circuit, likely result in TD. The mechanism of all these clinical actions is thought to begin with dopamine-receptor blockade, but thereafter it is critically mediated by altered neuronal activity (including GABAergic) in gray matter regions in the segregated, parallel frontal-subcortical modulatory motor circuits (35).

CLINICAL FEATURES AND COURSE OF TARDIVE DYSKINESIA

Part of "126 - Clinical Course and Cellular Pathology of Tardive Dyskinesia "

The presentation of TD is typical of a dyskinesia, with orofacial, axial, and extremity hyperkinetic movements. The movements worsen with stress and concomitant physical activity and diminish or disappear during sleep. TD onset is defining, in that all dyskinesias with their first onset within 6 months of ongoing antipsychotic treatment are diagnosed as TD (50). Consequently, it is expected that the diagnosis will be confounded by other categories of dyskinesia, including dyskinesias of schizophrenia and of the elderly. TD is distinguished from parkinsonism, the other major category of antipsychotic-induced movements, by being hyperkinetic, with delayed onset and delayed resolution after medication discontinuation, and by its contrasting pharmacology.

Movements in TD are typically suppressed by antipsychotic drugs, and they are suppressed or even show remission with potent GABA agonists; dopamine agonists exacerbate the movements, as can anticholinergic drugs (11). Although TD has a characteristic pharmacology, none of the drugs that suppress movements in probe studies are potent enough to be used as a treatment (36). Treatment approaches for TD are discussed later.

Although TD risk is inevitable with the traditional antipsychotics, the risk appears to have dropped significantly with the second-generation antipsychotic drugs, including clozapine, olanzapine, and risperidone, with which reduced risk has been demonstrated (38 ,63 ,66), and with quetiapine, with which reduced risk is probable. Therefore, even though TD may be diminishing as a significant clinical risk with modern antipsychotic treatment, its study has enabled

identification of cellular and circuit determinants of dyskinesias in mammalian brain.

The clinical course of TD varies by psychiatric diagnosis, age, sex, and concomitant medical or neurologic illness. Data from a prospective study of 971 young adult psychiatric patients followed for up to 20 years indicated a increasing cumulative incidence rate of TD over the 20-year interval (Table 126.1). The cumulative incidence of persistent TD was lower but increased proportionally (40). These data are consistent with a declining rate of TD over time, illustrated by the declining hazard rate (Table 126.1). A similar pattern develops in the hazard rates over 20 years for persistent TD. A decline in hazard rate over time was also reported by Morganstern and Glazer (46), although their data show a sharper decline after the first 5 years.

Neuroleptic Exposure (y)	Tardive Dyskinesia Cumulative Incidence, (95% CI)	Persistent Tardive Dyskinesia, (85% CI) Cumulative Incidence	Tardive Dyskinesia Hazard Rate[a]
1	.05 (.04–.07)	.03 (.02–.04)	—
5	.27 (.23–.31)	.20 (.17–.23)	6.1%
10	.43 (.38–.47)	.34 (.30–.39)	4.7%
15	.52 (.46–.57)	.42 (.36–.47)	3.3%
20	.56 (.50–.63)	.49 (.41–.57)	2.1%

[a]Hazard Rate is the rate of tardive dyskinesia occurrence per year among those remaining at risk. The number represents the average yearly hazard rate over the 5-year block of time.

TABLE 126.1. CUMULATIVE INCIDENCE OF TARDIVE DYSKINESIA

Rates of TD are significantly affected by the age and other characteristics of the several samples studied. An increased vulnerability to TD associated with increasing age is the most consistently reported finding in TD research. A prospective study of 261 geriatric patients with a mean age of 77 years revealed cumulative rates of 25%, 34%, and 53% after, 1, 2, and 3 years of antipsychotic drug treatment (74). Similar findings were reported by Jeste et al. (37) and by Yassa et al. (75). In the younger adult sample (40), the hazard rate for TD was lowest for patients in their twenties and thirties at study entry, increased for those in their forties, and sharply increased for those aged 50 to 60 years.

Other variables associated with a significantly increased TD risk in the prospective study (Table 126.1) included the presence of extrapyramidal symptoms during antipsychotic treatment, diagnosis of unipolar depression, the number of peaks in dosage of antipsychotic drug (more than 1,000 mg in chlorpromazine equivalents), and intermittent antipsychotic treatment. Treatment with lithium is associated with a lower risk of TD development.

In this prospective study of 971 psychiatric patients (Table 126.1), the initial episode of TD was persistent for at least 3 months for half of the cases. Of these persistent cases, 68% remitted within the next 2 years. For the half whose initial episode was transient, there was a high risk (32%) of developing a persistent episode within 1 year of additional antipsychotic exposure. Regardless of the duration of the initial TD episode, if it remitted, there was a high likelihood (61%) of developing a second episode within 1 year.

Preliminary data suggest that prognosis for TD remission is better for patients who are treated for shorter times within the follow-up period after TD diagnosis and for those treated with lower doses of antipsychotic during follow-up (41). These facts about TD encourage future study in animals and humans on TD mechanisms and treatment and facilitate the research by providing a firm baseline and a clear description of course.

FUNCTIONAL HUMAN NEUROANATOMY OF ANTIPSYCHOTIC DRUG ACTION

Part of "126 - Clinical Course and Cellular Pathology of Tardive Dyskinesia "

The human central nervous system (CNS) has been the focus of studies to determine the mechanism of antipsychotic drug action with both acute and chronic antipsychotic drug administration. These antipsychotic data are relevant to TD mechanisms insofar as the localization of the neural pathways subserving antipsychotic drug action can suggest the regions that likely contain the neurochemical disorder underlying TD. The firm association between striatal dopamine-receptor blockade and antipsychotic drug action was established early and has been consistently confirmed in subsequent study (4 ,7 ,53). It has been suggested that antipsychotic drugs deliver their full antipsychotic action by blocking the D2 dopamine receptors in limbic cortex (25 ,54). Alternatively, as argued here, antidopaminergic drugs could deliver their antipsychotic action by altering activity in neuronal populations within the long-loop feedback neurons in the basal ganglia thalamocortical pathways, at least in part, thereby modulating neocortical activity indirectly. The mechanisms responsible for TD may well occur at sites within these modulatory pathways.

Early investigators observed an elevation of neuronal activity in the human caudate and putamen with antipsychotic drug treatment using functional *in vivo* imaging (33 ,69). To refine the localization of the signal, our laboratory

conducted a within-subject crossover study comparing haloperidol (0.3 mg/kg/day for 4 weeks) to placebo (0 mg/kg/day for 4 weeks), with a positron emission tomography scan using [^{18}F]fluorodeoxyglucose carried out at the end of each treatment period (35). Regional cerebral metabolic rates of glucose, calculated using the usual analytic techniques, were used for pixel-by-pixel regional comparisons. Haloperidol significantly activated neuronal activity in the basal ganglia (caudate and putamen) and thalamus, whereas the frontal cortex (especially the middle and inferior regions) and the anterior cingulate cortex demonstrated a reduction in regional cerebral metabolic rates of glucose with haloperidol (Fig. 126.1). Activational differences between the on-drug and off-drug conditions were surprisingly restricted to these areas, despite the systemic manner of drug delivery and steady-state kinetic conditions at testing. The striatal changes had been previously reported (6 ,68). Subsequent evaluation of haloperidol action using regional cerebral blood flow analysis pharmacodynamically verified these regional actions.

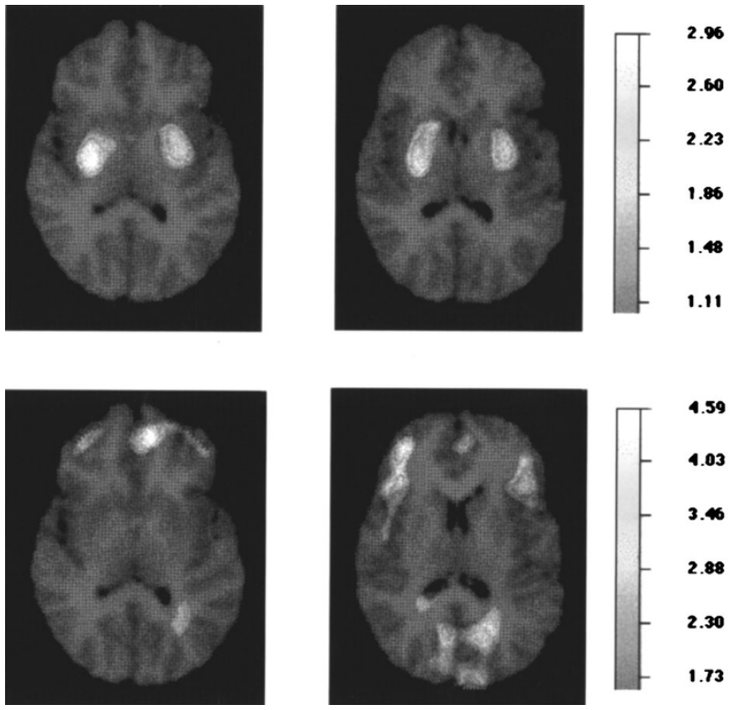


FIGURE 126.1. The regions of color indicate quantitatively the areas of regional cerebral blood flow (rCBF) activation (ON-30 day OFF) or inhibition (30 day OFF-ON) with subchronic haloperidol treatment (0.3 mg/kg/day). In the **top two figures**, at -04 and 00 mm from the anterior commisure/posterior commissure line (ACPC) plane, the basal ganglia show significant rCBF activation with haloperidol. In the **bottom two figures**, at -04 and +04 mm from the ACPC plane, both the anterior cingulate cortex and the middle frontal cortex show significant diminished rCBF with haloperidol. See color version of figure.

The regions identified in the foregoing experiment, whose activity is perturbed by haloperidol, are the same ones already known to be connected within the basal ganglia thalamocortical pathway (15) and involved in modulation of motor and cognitive function (2). In an animal preparation, Abercrombie and DeBoer demonstrated the principle that a pharmacologic perturbation delivered to a restricted region in the basal ganglia (e.g., striatum) can exert distant neurochemical and functional effects (1). The model of antipsychotic drug action we propose suggests that a primary effect of dopamine-receptor blockade occurs in the caudate and putamen with antipsychotic drug treatment, and the transmission of this primary action through the basal ganglia thalamocortical pathway to limbic and neocortex mediates antipsychotic and motor aspect of the drug actions in humans. The full mechanism of antidopaminergic actions therefore could include altered GABAergic transmission in the globus pallidus (GP), which alters activity in the basal ganglia output nuclei; these changes then modify GABAergic transmission from substantia nigra pars reticulata (SNR) to the thalamus, inhibit thalamic nuclei, and reduce the overall excitatory glutamatergic signal to the cortex. These observations suggest the hypothesis that the same basal ganglia thalamocortical circuits that mediate

basal ganglia modulatory influence on prefrontal cortex may also mediate the therapeutic effect of antidopaminergic antipsychotic compounds in schizophrenia. We have reasoned that a drug-induced regional change, occurring over time within this basal ganglia–thalamocortical pathway, associated with a regional increase in the thalamocortical signal could be associated with TD.

ROLE OF THE BASAL GANGLIA THALAMOCORTICAL PATHWAY IN MEDIATING AND TRANSMITTING THE ANTIDOPAMINERGIC ACTION IN STRIATUM TO CORTEX

Part of "126 - Clinical Course and Cellular Pathology of Tardive Dyskinesia "

Basal ganglia and thalamic structures modulate functions of the frontal cortex through parallel segregated circuits, a process that has been most fully studied for motor function. This topic is directly addressed in Chapter 122 . For the CNS motor system, specific areas of primary motor cortex, primary sensory cortex, and supplementary motor cortex project topographically to the putamen. These projections are thought to remain segregated but parallel throughout the full course of the circuit, but they are subject to basal ganglia and thalamic influences within each region. Investigators have proposed a family of these frontal circuits whose pathways originate in specific frontal cortex areas, course through the basal ganglia and thalamus, and return to the same areas of cortex, to modulate regional frontal cortical function (3). For the motor system, these subcortical structures appear to contribute to the planning and execution of body movements; for other frontal systems, these same subcortical structures may contribute to maintenance and switching of behaviors and aspects of cognition (2 ,26).

The thalamus exerts an excitatory effect on frontal cortex pyramidal cell activity, partially delivered within each frontal region by the paired parallel segregated circuit. The basal ganglia output nuclei with their high rates of spontaneous discharge keep their own target nuclei of the thalamus under tonic GABAergic inhibition. These inhibitory output nuclei stimuli are, in turn, regulated by two parallel but opposing pathways from the caudate and putamen, one excitatory and the other inhibitory. The primary cortical signal to basal ganglia is mediated by an excitatory glutamatergic pathway. It would be rational to suspect some role of these parallel segregated motor circuits in antipsychotic drug–induced dyskinesias, as well, especially because the primary action of antipsychotic drugs is D2 dopamine-family–receptor blockade in striatum.

ANIMAL MODELS OF TARDIVE DYSKINESIA

Part of "126 - Clinical Course and Cellular Pathology of Tardive Dyskinesia "

The cause of TD in antipsychotic drug–treated patients is, by definition, long-term drug treatment. Thus, putative models of the condition have been developed in nonhuman primates and in small animals using long-term administration of antipsychotics and applied to the study of TD mechanisms. The structural and functional brain characteristics of nonhuman primates are reasonably similar to those of human primates; hence, primate preparations may make more valid TD models (9). Yet, it is the reality of nonhuman primate models that many treatment years are required for dyskinesia development (2 to 6 years), and monkey care is involved and expensive. Alternatively, putative rat models have also been proposed; rat oral dyskinesias develop faster with chronic treatment (6 to 12 months), yet they retain many phenomenologic and pharmacologic characteristics of human TD (61 ,70). These models often provide a more realistic experimental platform, even if not more valid, than the nonhuman primate for developing hypotheses about human TD.

Animal models of all human diseases have been sought for pursuing pathophysiologic hypotheses and for identifying new therapeutics. Investigators have suggested the characteristics of good animal models for an expressed human illness as similarities in (a) origin, (b) phenomenology, (c) biochemical characteristics, and (d) pharmacology (45 ,73). Similarities of these features provide greater validation of the model. Across all the TD models, both primate and rodent, origin and phenomenology approximate characteristics of human TD. Biochemical determinants are unknown for the TD models as well as for the disease itself. However, extensive similarities exist in pharmacologic response between the human illness (TD) and the model preparations. Because rodent models are more practical to pursue, their pharmacology has been more broadly described than the primate model. It would be obvious to suggest that the use of both rat and monkey models would be ideal, as the efficiency and specificity of the questions dictate.

Nonhuman Primate Models

Antipsychotic treatment in the nonhuman primate has been studied to define the mechanism of acute drug-induced parkinsonism and of chronic drug-induced TD (22 ,31). Gunne et al. (28 ,31) showed not only the partial penetration of the syndrome in the nonhuman primate, but also differences in glutamic acid decarboxylase synthesis, the rate-limiting synthetic enzyme for GABA, and GABA levels in GP and SNR between drug-treated monkeys with and without the dyskinesia. Based on these data, Gunne et al. proposed that the mechanism of TD involved a reduction in GABAergic transmission in these regions of the basal ganglia, GP, and SNR. This idea correlated with the known clinical pharmacology of TD, namely, that GABA agonist treatment can improve drug-induced dyskinesias (65).

Rodent Models

Results from many laboratories suggested that rats treated chronically with traditional antipsychotics (e.g., fluphenazine, haloperidol) exhibit seemingly involuntary, irregular, and purposeless oral chewing movements (CMs) over time, often called *vacuous chewing movements* (13 ,16 ,18 ,23 ,27 ,30 ,56 ,61). The phenomenology of CMs resembles TD, in that movements have a gradual onset (61), partial penetrance (34), and a delayed offset, and they are sensitive to stress (49). However, the movements in rats remain limited to the oral region and rarely extend to other body parts. The pharmacology of CMs resembles that of TD: CMs are suppressed by antipsychotics, but not by anticholinergics (52); they are reduced by GABAmimetics (20), and they are attenuated with benzodiazepines. Rat CMs are associated with molecular and cellular changes in CNS histology, including an increase in perforated synapses in striatum and an alteration in relative synaptic number in striatum (47). Administration of the antipsychotic on an irregular schedule can advance the onset and severity of the rat CMs (24). The similarities across phenomenology and pharmacology are close enough between human TD and rat CMs for investigators to pursue the biochemical basis of CMs as a clue to pathophysiology in TD. Moreover, the pharmacology of the two are similar enough for the use of this model as a screen for new antipsychotic drugs to rule out TD potential.

Early pharmacologic studies in the rodent preparation reported that although all traditional antipsychotics are associated with CMs (70 ,71), clozapine is not (19 ,27). Subsequently, the other “new” antipsychotics have been tested and have generated results consistent with clinical data, demonstrating low TD potential for the second-generation antipsychotics (29 ,39). Neither olanzapine nor sertindole produce the CM syndrome at drug doses that produce human therapeutic plasma levels in the animals (21); risperidone at low doses is not associated with CMs, whereas high doses produce haloperidol-like CMs (Gao, unpublished observations). Data using quetiapine or ziprasidone in this animal model have not been reported.

NEUROCHEMICAL CHANGES WITHIN THE BASAL GANGLIA THALAMOCORTICAL PATHWAYS IN A RODENT MODEL OF TARDIVE DYSKINESIA

Part of "126 - Clinical Course and Cellular Pathology of Tardive Dyskinesia "

We designed and carried out a series of studies in a putative rodent model of TD based on the broadly accepted, functional architecture of the basal ganglia and thalamus already described. These studies were based not only on the existence of these theoretic models, but also on early experimental data in nonhuman primates with chronic antipsychotic treatment implicating GABAergic transmission in TD (30 ,31). The drug- and time-induced changes in GABAergic transmission in nonhuman primate basal ganglia directly affected the output nuclei, and from there, the thalamic and frontal regions associated with the segregated motor circuit.

With chronic (6 months) haloperidol treatment, some but not all laboratory rats acquired hyperkinetic oral CMs over time (61 ,63). In comparison, the newer antipsychotics, including clozapine, olanzapine, sertindole, and low-dose risperidone, failed to induce the rat “syndrome” of CMs (21 ,39). Can these preparations contribute to knowledge of TD pathophysiology? Can they contribute unique information to the mechanism of antipsychotic drug action? Comparison of several different animal treatment groups has been useful in addressing these questions: (a) haloperidol-treated rats, *with* versus *without* rat CMs and (b) haloperidol-treated rats versus newer antipsychotic drug–treated rats.

Chronic haloperidol induces similar D2-receptor up-regulation in striatum in the CM compared with the non-CM rat (Fig. 126.2), a finding discouraging consideration of this feature as a correlate of the CMs. Antipsychotic drugs block the inhibitory D2 receptor and disinhibit the medium spiny neuronal projections to the GP. In these studies, striatal disinhibition is reflected in the glutamic acid decarboxylase mRNA increases in GP, especially in the CM rats (Table 126.2). At the same time, activity in the direct striatonigral pathway appears also to be altered possibly by the haloperidol-induced increase in dopamine release in striatum and its action there on the unblocked D1 receptor (12). In the SNR, a primary basal ganglia output nuclei in the rat, abnormalities

also occur. Here the CM rats show reduced nigral D1-receptor numbers, whereas the non-CM treated rats show no change in D1-receptor density (Fig. 126.3). Moreover, this CM-associated receptor change in SNR is blocked (along with the CMs) by concomitant chronic treatment with the GABA agonist progabide (Fig. 126.3), a finding strengthening the association between altered SNR D1 receptors and CMs (45). The reduction in D1-receptor number in SNR could be associated with an antipsychotic-induced increase in the dendritic release of dopamine (Fig. 126.4). D1 receptors in SNR mediate the release of GABA. Hence, an increase in dopamine release within SNR could mediate the release of GABA at striatonigral terminals and subsequently could inhibit activity in the GABA-mediated efferent pathway to thalamus. A reduction in GABA-mediated transmission from SNR to the target nucleus in the thalamus could produce a compensatory up-regulation of thalamic $GABA_A$ receptors, hence marking altered SNR activity. In the haloperidol-treated animals, a significant elevation of $GABA_A$ receptor occurred, and a significant correlation developed between the elevated receptor number and CMs in the mediodorsal thalamus (Fig. 126.5). This positive correlation implicates a nigral D1 defect along with an overinhibition of the nigrothalamic efferent GABAergic pathway as an important mediator of CMs in rat (56). The idea that a reduction in the activity of the basal ganglia output nuclei disinhibits the thalamus and is associated with drug-induced rat hyperkinetic oral CMs is consistent with the already established functional models of these interactions (2).

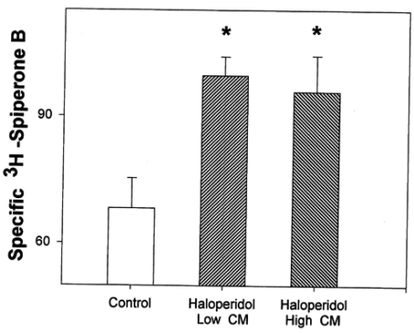


FIGURE 126.2. Specific binding of ^{3}H-spiperone to D2-family dopamine receptors in the nucleus accumbens of control and chronically haloperidol-treated rats. D2 binding data were similar in the caudate and putamen. Significant dopamine-receptor up-regulation was apparent in the haloperidol-treated rats with chewing movements (CMs) and in those without CMs; there was no apparent difference between CM and non-CM rats in the magnitude of increase. *$p < .05$.

	Haloperidol + VCMs	Haloperidol – VCMs	Olanzapine, no VCMs	Sertindole, no VCMs
Striatum	↑D_2R, ↑GAD	↑D_2R, ↑GAD	S1 ↑D_2R No Δ GAD	No Δ D_2R ↑GAD
Globus pallidus	↓GAD, ↓$GABA_AR$	No Δ GAD, No Δ $GABA_AR$	No Δ GAD No Δ $GABA_AR$	No Δ GAD, ↓$GABA_AR$
Substantia nigra pars reticulata	↑$GABA_AR$ ↓D_1R	Tr ↑$GABA_AR$Δ No Δ D_1R	Nl, $GABA_AR$ No Δ D_1R	Nl, $GABA_AR_1$ No Δ D_1R
MD thalamus	↑$GABA_AR$ $GABA_AR$/VCM correlation	No Δ $GABA_AR$ No correlation	No Δ $GABA_AR$ No correlation	No Δ $GABA_AR$ No correlation
Right thalamus	↑GAD mRNA	↑GAD mRNA	↑GAD mRNA	↑GAD mRNA

arrow, significant change; D_1R, D_1 family dopamine receptor; D_2R, D_2 family dopamine receptor; $GABA_AR$, $GABA_A$ receptor; GAD, glutamic acid decarboxylase mRNA; Tr, trend; VCMS, vacuous chewing movements.

TABLE 126.2.

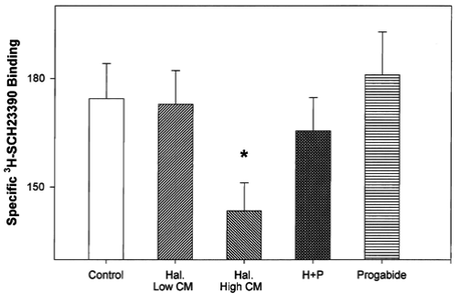


FIGURE 126.3. Specific binding of 3H-SCH23390 to D1-family dopamine receptors in the substantia nigra pars reticulata (SNR). D1 binding was down-regulated by chronic haloperidol (HAL) treatment only in the rats that displayed the vacuous chewing movements (CMs). This reduction was blocked by the GABA agonist progabide (P), as were the CMs themselves. Progabide alone had no effect on the D1 binding. *$p < .05$.

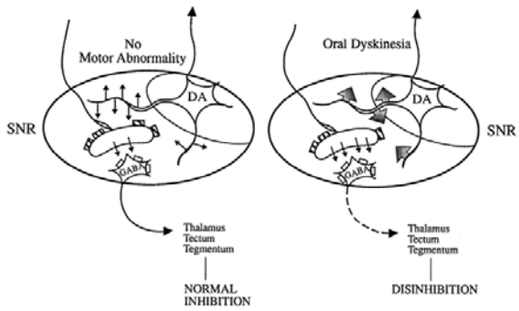


FIGURE 126.4. Hypothetical model for the mechanism of chronic haloperidol-induced chewing movements (CMs) in rat. The data collected can be parsimoniously explained by postulating an increase in dendritic dopamine release associated with the development of dyskinetic movements in the CM rats. An increase in dopamine could stimulate the release of GABA presynaptically from the striatonigral GABAergic neurons and thereby increase the overall GABAergic inhibition on the GABA-containing nigral efferent fibers, including those to the thalamus. This alteration would serve to decrease the level of GABAergic inhibition on critical nuclei of the thalamus. This interpretation of the data, even though parsimonious, is speculative.

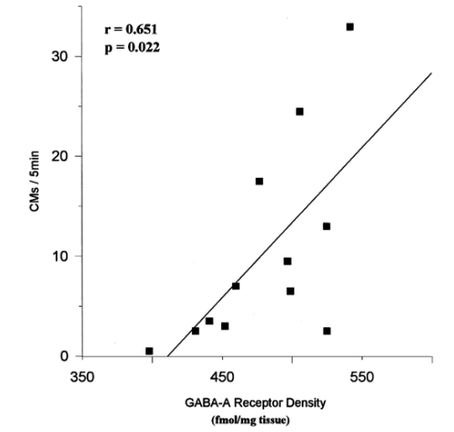


FIGURE 126.5. Correlation of $GABA_A$ receptor density with chewing movements (CMs) in mediodorsal thalamus in rats treated chronically with haloperidol. The significant positive correlation between these two factors suggests their relationship. This correlation supports the hypothesis that CMs are associated with diminished GABAergic transmission from the SNR to significant thalamic nuclei, evidenced by the up-regulation in $GABA_A$-receptor density associated with the putatively reduced GABAergic nigral signal with CMs. However, no other thalamic nuclei show this type of correlation, even the ventral nuclei traditionally associated with motor control in thalamus.

Two second-generation antipsychotics tested in the same animal chronic treatment paradigm differed from haloperidol in their actions. Both olanzapine and sertindole, each at two doses, were compared with haloperidol after 6 months of treatment (56). Neither olanzapine nor sertindole substantially up-regulated striatal D2 binding in the rat, even though we know from human studies that D2 blockade of some strength and duration occurs with each of these drugs (21). Because of the relatively high receptor affinities of these drugs at the D2 receptor, the data suggest that any regional reduction in blockade may occur only at some of the D2 receptors, and the resultant antidopaminergic action is weaker or of a reduced duration than with haloperidol. Nonetheless, olanzapine shows mild, haloperidol-like actions in striatum, and sertindole shows mild, haloperidol-like actions in GP and STN. Still, in SNR, neither new compound is associated with D1-receptor down-regulation or $GABA_A$ up-regulation (Table 126.2), nor are $GABA_A$ receptors altered in mediodorsal thalamus by either new drug in contrast to the haloperidol effect. It is possible

that the mechanism whereby these two effective antipsychotics (each antidopaminergic) fail to induce CMs is different in striatum, but both result in a common sparing of a critical change in SNR. It may be that the common serotonergic influence exerted within the basal ganglia nuclei by both these compounds spares SNR changes.

WORKING HYPOTHESIS OF THE CM-TD MECHANISM

Part of "126 - Clinical Course and Cellular Pathology of Tardive Dyskinesia "

The data summarized here are consistent with many reports in the literature and confirm the central role of the basal ganglia output nuclei and thalamus in mediating hyperkinetic movements in chronic antipsychotic-treated animals. It would be our current notion that traditional antipsychotic drugs alter the dynamic balance of neurotransmitter activity within the indirect and direct striatonigral pathway. This change, perhaps associated with sustained increase in nigral dendritic dopamine release, results in rodent hyperkinetic oral movements through the feedback of this information to motor regions of neocortex through thalamus. This antipsychotic-induced alteration acutely would merely inhibit activity within the indirect pathway and would be associated with parkinsonism. As treatment progresses and CMs begin, the indirect pathway inhibition could be progressively counterbalanced by direct pathway overactivity in the vulnerable animals. We postulate that the changes in SNR in D1-receptor (decrease) and $GABA_A$-receptor (increase) numbers reflect CM-related pathophysiology. These available data so far derive from animal model studies and provide a putative framework for TD pathophysiology. Work now must proceed with the human illness itself, by testing these and other ideas in *in vivo* brain imaging studies and postmortem tissue analysis.

These findings illustrate not only the basal ganglia mechanisms of one type of dyskinesia, but also one outcome of CNS plasticity in response to chronic dopamine-receptor blockade. The difference between those animals who are vulnerable to develop CMs and those who are not remains obscure, as is the vulnerability to TD with antipsychotic treatment in humans.

TREATMENT APPROACHES

Part of "126 - Clinical Course and Cellular Pathology of Tardive Dyskinesia "

There is still no definitive treatment for TD. Nonetheless, therapeutic strategies are often used for patients with TD symptoms. Prevention, reversal, and suppression or (clinical management) all need to be considered (57). Prevention and reversal used to be only theoretic possibilities, but this is no longer the case. The newer generation of antipsychotic drugs, with their low TD incidence, has introduced these various new options.

Prevention

Patients who require antipsychotic treatment for extended times today have the opportunity of treatment with one of the newer antipsychotic drugs and thus are at reduced risk of developing TD, probably at considerably reduced risk. Clozapine (48), olanzapine (50), risperidone (58), and quetiapine (67) all have been reported to have reduced association with TD. Data are not yet available for new antipsychotic treatment of neuroleptic-naive persons, in whom the expectation may be for an even lower association with the syndrome. However, these data will be generated in time.

Whether treatment with very low-dose traditional antipsychotic drugs will result in the same low incidence of TD is a possibility, but it is nowhere nearly as probable as with the use of the newer drugs. Very low-dose traditional drug treatment, such as haloperidol, 2 to 3 mg/day, is being tested in several centers for efficacy in psychosis and for side effects. However, low-dose haloperidol at 4 mg/day has the same incidence of acute parkinsonism and akathisia as a haloperidol dose of 16 mg/day (76), a finding suggesting that low-dose haloperidol still has considerable potency in modifying motor function. Nonetheless, the economic advantage of traditional antipsychotic treatments, when necessity demands it, deserves thorough testing.

The cellular basis for the advantage of the newer antipsychotics with respect to TD risk is being examined and has been inferred from what is known of their pharmacology. Clozapine has a complex pharmacology, and consequently its TD advantage can be theoretically associated with several transmitter mechanisms (14). The possibilities include a D1 antagonist action, a serotonin$_{2A}$ antagonist action, an antimuscarinic action, particularly at the M1 receptor, and even the antihistaminic action that can spare drug-induced TD. Parsimoniously, because this TD advantage appears to be a property of several or all of the newer antipsychotics, one could propose a common mechanism for the TD sparing. Speculations would then center most strongly on the role of the serotonin$_{2A}$-receptor blockade in striatum to attenuate the tardive motor effects of the dopamine-receptor blockade. This could be mediated by a blockade of serotonin action on striatal neurons, and a resultant modulation of dopamine-receptor blockade. Alternatively, the newer drugs may modulate dopamine blockade regionally to attenuate the cellular changes produced by the drugs. The results of the chronic treatment experiments in rat would favor this latter explanation. Additional possibilities exist, including a thalamic or cortical action of the drug at the serotonin 2A receptor.

Reversal

Clinical data with traditional antipsychotic treatment suggests that dyskinesia reversal occurs in persons with TD during ongoing treatment (Table 126.1), however, still conferring future risk. Data suggest that this reversal may happen faster with newer drug treatment than with the continued use of traditional drugs. TD reversal occurs frequently, although not inevitably, with cessation of antipsychotic treatment (43). This reversal appears to be more likely in the young rather than in the older patient, presumably because of greater tissue or system plasticity. The reversal occurs over the course of months to years, not in the range of weeks, so the phenomenon is challenging to document. Based on the simplistic formulation that relieving the brain of the inducing agent will allow the drug-induced tissue changes to reverse, then drugs such as clozapine or other newer antipsychotics, which have clinical efficacy with reduced dopamine-receptor occupancy, may reverse existent TD. Clozapine has been tested in a double-blind protocol comparing dyskinesia scores between two treatment populations during ongoing drug treatment (haloperidol versus clozapine) over the course of 12 months (63). Dyskinesia in the clozapine-treated group tended to be reduced after clozapine compared with haloperidol treatment ($p < .057$). Of some significance is that the rebound dyskinesia with drug withdrawal after a year of haloperidol treatment was significant, whereas after a year of clozapine treatment, the previously sensitive group failed to show any dyskinesia rebound after drug withdrawal, a finding suggesting the lack of system sensitivity (perhaps dopamine-receptor sensitivity) with clozapine. Olanzapine is currently being tested in persons with TD for its ability to relieve dyskinesia symptoms over 12 months compared with haloperidol.

Suppression

The feature that most consistently characterizes the results of suppression trials in TD is the variability within patients

and among clinical centers conducting trials. No drugs have been reliably demonstrated to suppress TD across all patient groups and ages, although some drug classes show more promise and consistency in this regard than other approaches (32).

Benzodiazepines most reliably reduce the dyskinesias of TD, even at doses that do not produce sedation (64). Clonazepine has been widely used and seems to be one of the more effective benzodiazepines (51 ,64 ,72). Even when effective, clonazepam shows tolerance over time and requires continual dose increases to sustain efficacy. Thus, treatment must be occasionally withdrawn to "resensitize" the system to its effects for treatment to sustain action. The mechanism of therapeutic action has always been thought related to the GABA-enhancing drug action. Because the biology of TD has suggested regional GABA reductions, a GABA agonist is rational as a therapeutic choice.

Based on a continuation of this reasoning, other GABAmimetics have been tested in TD (60). It has been challenging to find a therapeutic window for a GABA agonist in TD, in which the GABA agonist action is potent enough to be antidyskinetic but the side effects are not limiting. Valproic acid has not been shown to be an effective therapeutic agent in TD, presumably because of its low potency (44). More potent GABAmimetic, such as γ-vinyl-GABA and tetrahydroisoxazolopyridinol, have shown antidyskinetic efficacy, but in some cases with limiting side effects (62 ,65). The studies that have shown efficacy of GABAmimetics in TD have used younger patients (i.e., less than 50 years old). Studies in older volunteers have not demonstrated efficacy with GABAmimetics (12). Although the question of age effect on TD reversal has not been directly tested, it is generally believed that symptoms are more likely to be suppressed with a GABAergic drug in the young patient than in the older one.

Other drug treatment strategies include vitamin E (17 ,48 ,59), melatonin (55), noradrenergic-receptor antagonists such as propranolol or clonazepan, amantadine (5), and calcium channel blockers, especially nifedipine (58). Whereas each of these approaches provides clinical interest, none has developed into a therapeutic approach.

Treatments for TD will always be needed, even though the incidence of new cases may fall significantly with time. Treatments will always be more effective in younger than in older patients. Opportunities for drug development are likely to be found in the GABAergic or the glutamatergic neurotransmitter systems.

REFERENCES

1. Abercrombie ED, DeBoer P. Substantia nigra D1 receptors and stimulation of striatal cholinergic interneurons by dopamine: a proposed circuit mechanism. *J Neurosci* 1997;17:8498–8505.

2. Alexander GE, Crutcher MD. Functional architecture of basal ganglia circuits: neural substrates of parallel processing [Review]. *Trends Neurosci* 1990;13:266–271.

3. Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex [Review]. *Annu Rev Neurosci* 1986;9:357–381.

4. Anden NE, Butcher SG, Corrodi H, et al. Receptor activity and turnover of dopamine and noradrenaline after neuroleptics. *Eur J Pharmacol* 1970;11:303–314.

5. Angus S, Sugars J, Boltezar R, et al. A controlled trial of amantadine hydrochloride and neuroleptics in the treatment of tardive dyskinesia. *J Clin Psychopharmacol* 1997;17:88–91.

6. Buchsbaum MS, Potkin SG, Siegel BV Jr, et al. Striatal metabolic rate and clinical response to neuroleptics in schizophrenia. *Arch Gen Psychiatry* 1992;49:966–974.

7. Bymaster FP, Hemrick-Luecke SK, Perry KW, et al. Neurochemical evidence for antagonism by olanzapine of dopamine, serotoin, adrenergic and muscarinic receptors *in vivo* in rats. *Psychopharmacology (Berl)* 1996;124:87–94.

8. Carpenter WT Jr, Buchanan RW. Schizophrenia [Review]. N. Engl. J. Med. 1994;330:681–690.

9. Casey DE. Extrapyramidal syndromes in nonhuman primates: typical and atypical neuroleptics. *Psychopharmacol Bull* 1991;27:47–50.

10. Casey DE. The nonhuman primate model: focus on dopamine D_2 and serotonin mechanisms. In: Fog R, Gerlach J, Hemmingsen R, eds. *Schizophrenia: an integrated view*. Coppenhagen: Munksgaard, 1995.

11. Casey DE, Denney D. Pharmacological characterization of tardive dyskinesia. *Psychopharmacology (Berl)* 1977;54:1–8.

12. Casey DE, Gerlach J, Magelund G, et al. Gamma-acetylenic GABA in tardive dyskinesia. *Arch Gen Psychiatry* 1980;37:1376–1379.

13. Clow A, Theodorou A, Jenner P, et al. Changes in cerebral dopamine function induced by a year's administration of trifluoperazine or thioridazine and their subsequent withdrawal. *Adv Biochem Psychopharmacol* 1980;24:335–340.

14. Coward DM. General pharmacology of clozapine. [Review]. *Br J Psychiatry* 1992;17[Suppl]:5–11.

15. DeLong MR. Primate models of movement disorders of basal ganglia origin [Review]. *Trends Neurosci* 1990;13:281–285.

16. Egan MF, Hurd Y, Ferguson J, et al. Pharmacologic and neurochemical differences between acute and tardive vacuous chewing movements induced by haloperidol. *Psychopharmacology (Berl)* 1996;127:337–345.

17. Elkashef AM, Wyatt RJ. Tardive dyskinesia: possible involvement of free radicals and treatment with vitamin E. *Schizophr Bull* 1999;25:731–740.

18. Ellison G, See R, Levin E, et al. Tremorous mouth movements in rats administered chronic neuroleptics. *Psychopharmacology (Berl)* 1987;92:122–126.

19. Gao XM, Hashimoto T, Cooper TB, et al. The dose-response characteristics of rat oral dyskinesias with chronic haloperidol or clozapine administration. *J Neural Transm* 1997;104:97–104.

20. Gao XM, Kakigi T, Friedman MB, et al. Tiagabine inhibits haloperidol-induced oral dyskinesias in rats. *J Neural Transm Gen Sect* 1994;95:63–69.

21. Gao XM, Sakai K, Tamminga CA. Chronic olanzapine or sertindole treatment results in reduced oral chewing movements in rats compared to haloperidol. *Neuropsychopharmacology* 1998;19:428–433.

22. Gerlach J, Reisby N, Randrup A. Dopaminergic hypersensitivity and cholinergic hypofunction in the pathophysiology of tardive dyskinesia. *Psychopharmacologia* 1974;34:21–35.

23. Glenthoj B. Persistent vacuous chewing in rats following neuroleptic treatment: relationship to dopaminergic and cholinergic function. *Psychopharmacology (Berl)* 1993;113:157–166.

24. Glenthoj B, Hemmingsen R. Intermittent neuroleptic treatment induces long-lasting abnormal mouthing in the rat. *Eur J Pharmacol* 1989;164:393–396.

25. Grace AA, Bunney BS, Moore H, et al. Dopamine-cell depolarization block as a model for the therapeutic actions of antipsychotic drugs. [Review]. *Trends Neurosci* 1997;20:31—37.

26. Graybiel AM: Neurotransmitters and neuromodulators in the basal ganglia [Review]. *Trends Neurosci* 1990;13:244–254.

27. Gunne LM, Andersson U, Bondesson U, et al. Spontaneous chewing movements in rats during acute and chronic antipsychotic drug administration. *Pharmacol Biochem Behav* 1986;25:897–901.

28. Gunne LM, Barany S. Haloperidol-induced tardive dyskinesia in monkeys. *Psychopharmacology (Berl)* 1976;50:237–240.

29. Gunne LM, Growdon J, Glaeser B. Oral dyskinesia in rats following brain lesions and neuroleptic drug administration. *Psychopharmacology (Berl)* 1982;77:134–139.

30. Gunne LM, Haggstrom JE. Reductions of nigral glutamic acid decarboxylase in rats withneuroleptic-induced oral dyskinesia. *Psychopharmacology (Berl)* 1983;81:191–194.

31. Gunne LM, Haggstrom JE, Sjoquist B. Association with persistent neuroleptic- induced dyskinesia of regional changes in brain GABA synthesis. *Nature* 1984;309:347–349.

32. Gupta S, Mosnik D, Black DW, et al. Tardive dyskinesia: review of treatments past, present, and future. *Ann Clin Psychiatry* 1999;11:257–266.

33. Gur RE, Mozley PD, Resnick SM, et al. Resting cerebral glucose metabolism in first-episode and previously treated patients with schizophrenia relates to clinical features. *Arch Gen Psychiatry* 1995;52:657–667.

34. Hashimoto T, Ross DE, Gao XM, et al. Mixture in the distribution of haloperidol-induced oral dyskinesias in the rat supports an animal model of tardive dyskinesia. *Psychopharmacology (Berl)* 1998;137:107–112.

35. Holcomb HH, Cascella NG, Thaker GK, et al. Functional sites of neuroleptic drug action in the human brain: PET/FDG studies with and without haloperidol. *Am J Psychiatry* 1996;153:41–49.

36. Jeste DV. In search of treatment for tardive dyskinesia: review of the literature. *Schizophr Bull* 1979;5(2):251–292.

37. Jeste DV, Caligiuri MP, Paulsen JS, et al. Risk of tardive dyskinesia in older patients: a prospective longitudinal study of 266 outpatients. *Arch Gen Psychiatry* 1995;52:756–765.

38. Jeste DV, Okamoto A, Napolitano J, et al. Low incidence of persistent tardive dyskinesia in elderly patients with dementia treated with risperidone. *Am J Psychiatry* 2000;157:1150–1155.

39. Kakigi T, Gao XM, Tamminga CA. Drug-induced oral dyskinesias in rats after traditional and new neuroleptics. *J Neural Transm Gen Sect* 1995;101:41–49.

40. Kane JM, Woerner M, Borenstein M, et al. The incidence of tardive dyskinesia: the Hillside Prospective Study. *Arch Gen Psychiatry*, in press.

41. Kane JM, Woerner M, Borenstein M, et al. Integrating incidence and prevalence of tardive dyskinesia. *Psychopharmacol Bull* 1986;22:254–258.

42. Klein DF, Davis JM. *Diagnosis and drug treatment of psychiatric disorders*. Baltimore: Williams & Wilkins, 1969.

43. Lahti AC, Thaker GK, Nguyen JA, et al. Reversibility of tardive dyskinesia following neuroleptic withdrawal. *Neurosci Abstr* 1998;14:498(abstr).

44. Linnoila M, Viukari M, Kietala O. Effect of sodium valproate on tardive dyskinesia. *Br J Psychiatry* 1976;129:114–119.

45. McKinney WT, Bunney WE. Animal models of depression. *Arch Gen Psychiatry* 1969;21:240–248.

46. Morgenstern H, Glazer WM. Identifying risk factors for tardive dyskinesia among long-term outpatients maintained with neuroleptic medications: results of the Yale Tardive Dyskinesia Study. *Arch Gen Psychiatry* 1993;50:723–733.

47. Roberts RC, Gaither LA, Gao XM, et al. Ultrastructural correlates of haloperidol-induced oral dyskinesias in rat striatum. *Synapse* 1995;20:234–243.

48. Sachdev P, Saharov T, Cathcart S. The preventative role of antioxidants (selegiline and vitamin E) in a rat model of tardive dyskinesia. *Biol Psychiatry* 1999;46:1672–1681.

49. Sakai K, Gao XM, Tamminga CA. Scopolamine fails to diminish chronic haloperidol-induced purposeless chewing in rats. *Psychopharmacology (Berl)* 2001;153:191–195.

50. Schooler NR, Kane JM. Research diagnoses for tardive dyskinesia [Letter]. *Arch Gen Psychiatry* 1982;39:486–487.

51. Sedman G. Clonazepam in treatment of tardive oral dyskinesia [Letter]. *BMJ* 1976;2:583.

52. See RE, Chapman MA. Cholinergic modulation of oral activity in drug-naive and chronic haloperidol-treated rats. *Pharmacol Biochem Behav* 1991;39:49–54.

53. Seeman P, Chau-Wong M, Tedesco J, et al. Brain receptors for antipsychotic drugs and dopamine: direct binding assays. *Proc Natl Acad Sci USA* 1975;72:4376–4380.

54. Seeman P, Van Tol HH. Dopamine receptor pharmacology [Review]. *Trends Pharmacol Sci* 1994;15:264–270.

55. Shamir E, Barak Y, Plopsky I, et al. Is melatonin treatment effective for tardive dyskinesia? *J Clin Psychiatry* 2000;61:556–558.

56. Shirakawa O, Tamminga CA. Basal ganglia GABAA and dopamine D1 binding site correlates of haloperidol-induced oral dyskinesias in rat. *Exp Neurol* 1994;127:62–69.

57. Simpson GM. The treatment of tardive dyskinesia and tardive dystonia. *J Clin Psychiatry* 2000;61[Suppl 4]:39–44.

58. Soares KV, McGrath JJ. Diltiazem, nifedipine, nimodipine or verapamil for neuroleptic-induced tardive dyskinesia. *Cochrane Database Syst Rev* 2000;CD000206.

59. Soares KV, McGrath JJ. Vitamin E for neuroleptic-induced tardive dyskinesia. *Cochrane Database Syst Rev* 2000;CD000209.

60. Soares KV, McGrath JJ, Deeks JJ. Gamma-aminobutyric acid agonists for neuroleptic-induced tardive dyskinesia. *Cochrane Database Syst Rev* 2000;CD000203.

61. Tamminga CA, Dale JM, Goodman L, et al. Neurolepticinduced vacuous chewing movements as an animal model of tardive dyskinesia: a study in three rat strains. *Psychopharmacology (Berl)* 1990;102:474-478.

62. Tamminga CA, Thaker GK, Ferraro TN, et al. GABA agonist treatment improves tardive dyskinesia [Letter]. *Lancet* 1983;2:97–98.

63. Tamminga CA, Thaker GK, Moran M, et al. Clozapine in tardive dyskinesia: observations from human and animal model studies. *J Clin Psychiatry* 1994;55[Suppl B]:102–106.

64. Thaker GK, Nguyen JA, Strauss ME, et al. Clonazepam treatment of tardive dyskinesia: a practical GABAmimetic strategy. *Am J Psychiatry* 1990;147:445–451.

65. Thaker GK, Tamminga CA, Alphs LD, et al. Brain gamma-aminobutyric acid abnormality in tardive dyskinesia: reduction in cerebrospinal fluid GABA levels and therapeutic response to GABA agonist treatment. *Arch Gen Psychiatry* 1987;44:522–529.

66. Tollefson GD, Beasley CM Jr, Tamura RN, et al. Blind, controlled, long-term study of the comparative incidence of treatment-emergent tardive dyskinesia with olanzapine or haloperidol. *Am J Psychiatry* 1997;154:1248–1254.

67. Vesely C, Kufferle B, Brucke T, et al. Remission of severe tardive dyskinesia in a schizophrenic patient treated with the atypical antipsychotic substance quetiapine. *Int Clin Psychopharmacol* 2000;15:57–60.

68. Volkow ND, Brodie JD, Wolf AP, et al. Brain metabolism in patients with schizophrenia before and after acute neuroleptic administration. *J Neurol Neurosurg Psychiatry* 1986;49:1199–1202.

69. Volkow ND, Brodie JD, Wolf AP, et al. Brain organization in schizophrenia. *J Cereb Blood Flow Metab* 1986;6:441–446.

70. Waddington JL. Spontaneous orofacial movements induced in rodents by very long-term neuroleptic drug administration: phenomenology, pathophysiology and putative relationship to tardive dyskinesia. *Psychopharmacology (Berl)* 1990;101:431–447.

71. Waddington JL, Cross AJ, Gamble SJ, et al. Spontaneous orofacial dyskinesia and dopaminergic function in rats after 6 months of neuroleptic treatment. *Science* 1983;220:530–532.

72. Weber SS, Dufresne RL, Becker RE, et al. Diazepam in tardive dyskinesia. *Drug Intell Clin Pharmacol* 1983;17:523–527.

73. Weiss JM, Kilts CD. Animal models of depression and schizophrenia. In: Schatzberg AF, Nemeroff CV, eds. *Textbook of psychopharmacology*. Washington, DC: American Psychiatric Press, 1995.

74. Woerner MG, Alvir JM, Saltz BL, et al. Prospective study of tardive dyskinesia in the elderly: rates and risk factors. *Am J Psychiatry* 1998;155:1521–1528.

75. Yassa R, Nastase C, Dupont D, et al. Tardive dyskinesia in elderly psychiatric patients: a 5-year study. *Am J Psychiatry* 1992;149:1206–1211.

76. Zimbroff DL, Kane JM, Tamminga CA, et al. A controlled, dose-response study of sertindole and haloperidol in schizophrenia. *Am J Psychiatry* 1997;154:782–791.

127

Temporal Lobe Epilepsy: Renewed Emphasis on Extrahippocampal Areas

Robert Schwarcz

Helen E. Scharfman

Edward H. Bertram

Robert Schwarcz: Maryland Psychiatric Research Center, University of Maryland School of Medicine, Baltimore, Maryland.

Helen E. Scharfman: Neurology Research Center, Helen Hayes Hospital, West Haverstraw, New York.

Edward H. Bertram: Department of Neurology, University of Virginia Medical Center, Charlottesville, Virginia.

Epilepsy is a chronic condition characterized by spontaneously recurring seizures. Although often viewed and discussed as a single clinical entity, epilepsy is a symptom of several disorders that affect the brain. The variety of causes is quite extensive and includes tumors, congenital malformations, genetic alterations in receptors or channels, and acquired structural abnormalities such as those following trauma or infection. Some of the epilepsies with inborn causes, such as rolandic epilepsy, are self-limited and benign, whereas others are progressive (1). The seizures in some forms of epilepsy may arise from the entire brain at one time, whereas in other forms they start in a particular region or focus. Any region of the brain can serve as a seizure focus, but seizure onset is commonly observed in the temporal lobe. Although there are multiple causes for epilepsy originating in the temporal lobe, the most common form is the *mesial temporal lobe epilepsy syndrome* (MTLE), sometimes also termed *limbic epilepsy* because of the apparent involvement of one or more limbic brain regions (2).

MTLE has been recognized as a distinct entity for many years. The common clinical pattern during the seizure episode includes staring and lack of responsiveness, frequently accompanied by automatisms of hand activity and mastication. There is often (but not always) a history of prolonged febrile convulsions in early childhood and common pathologic features of atrophy and neuronal loss in the hippocampus, a structure located on the medial border of the temporal lobe. Seizure onset is frequently detected in the hippocampus when EEG depth recordings are made directly from this area in patients undergoing evaluation for surgical treatment of uncontrolled seizures. Removal of the hippocampus (together with adjacent structures) often successfully controls seizures in “intractable” patients whose seizures do not respond to medication (2). For these reasons, a general consensus has developed that the hippocampus is the key to understanding and treating limbic epilepsy, and much of the research directed at MTLE has focused on this area of the brain. However, there is increasing evidence that other structures of the limbic system, such as the amygdala, parts of the neocortex, and the entorhinal cortex, which is a phylogenetically older part of the cortex that controls the information flow into and from the hippocampus (3), also play important roles in the initiation and propagation of seizures in MTLE.

Support for the involvement of nonhippocampal limbic sites in MTLE comes from a variety of sources. As reviewed in this chapter, extrahippocampal areas frequently show pathologic structural changes on histologic examination. Intracranial recordings from patients undergoing evaluation for therapeutic epilepsy surgery often present with a pattern of diffuse limbic onset without a regional predilection (4 ,5). Imaging studies have also indicated that there is atrophy or metabolic change in medial temporal structures other than the hippocampus, as well as in subcortical structures with limbic connections (6 ,7 and 8). Surgical results have suggested that it is necessary to remove more than the hippocampus to achieve successful outcome (9 ,10). Finally, there is ample evidence from a variety of animal models of limbic epilepsy that extrahippocampal sites participate in epileptogenesis, demonstrate anatomic and neuropathologic changes, and show alterations in cellular physiology. In many instances, changes in these regions are greater than those seen in the hippocampus.

The advent of new imaging techniques, the development of several useful new animal models of limbic epilepsy that closely parallel the human condition, and improved means

for studying neuronal physiology *in vivo* and *in vitro* have provided new opportunities to explore the role and fate of extrahippocampal brain regions in MTLE. In this chapter, we review current knowledge of the neuropathology, physiology, anatomy, and neurochemistry of MTLE and compare the hippocampus with extrahippocampal limbic regions. In closing, we briefly explore how new research directions resulting from recent data may lead to novel and improved treatments of MTLE.

NEUROPATHOLOGY

Part of "127 - Temporal Lobe Epilepsy: Renewed Emphasis on Extrahippocampal Areas "

Hippocampus

The notion of a central role of the hippocampus in MTLE can be traced to a highly influential review article by Sommer, which was published toward the end of the nineteenth century (11). As a young physician at the “insane asylum” in Allenberg, Germany, Sommer reviewed the anamneses and corresponding neuropathologic features of 90 patients with epilepsy, only five of whom had been patients at his institution. After pointing out that “there can be no question that epileptic symptoms are frequently associated with a disease of the Ammon's horn,” as others had suggested before, he made the ground-breaking observation that patients presented with a preferential degeneration of a band of pyramidal cells termed the CA1 subfield (the *Sommer sector*) of the hippocampus proper (12). With astounding foresight, he concluded that the cell loss was a consequence of prolonged seizure activity and that both head trauma and developmental malformations ought to be considered etiologically important factors in epilepsy. Sommer also noted that neurodegenerative changes in patients with epilepsy are frequently detected in areas surrounding the hippocampus, namely, in the subiculum and in the temporal and occipital cortices. He remarked that “it is not unlikely that a single underlying defect spreading through several anatomically linked brain areas accounts for the various clinical manifestations of the epileptic syndrome.”

During the first half of the twentieth century, Sommer's pioneering insights and predictions were confirmed and refined by many investigators studying brain pathology post mortem. This period was increasingly dominated by the question whether “idiopathic” seizure activity causes neurodegeneration or, conversely, epilepsy occurs only as a sequela of brain damage. The chicken-egg debate, which has not been entirely resolved to this day, was accompanied by uncertainties about the nature of putatively epileptogenic insults, such as respiratory difficulties and asphyxia, infectious diseases leading to encephalitis, and, in particular, vascular abnormalities (13). The advent, in the 1930s, of surgical interventions for the treatment of MTLE revolutionized clinical management of the disease and at the same time provided invaluable information for research purposes (14).

Histologic analysis of excised brain tissue and correlations between the nature of removed temporal lobe structures and surgical outcome unequivocally established the centrality of the hippocampus and associated “limbic” brain regions in MTLE *(psychomotor epilepsy)*. By the 1950s, support had developed among leading epilepsy researchers for the concepts that MTLE (a) was a “network” disorder, caused by abnormal interactions between a limited number of highly interconnected areas within the temporal lobe, and (b) in many cases originated in, and could be treated by removal of, extrahippocampal structures (15). Thus, irrespective of the cellular and molecular events underlying epileptogenesis and the eventual development of spontaneously recurring seizures, which soon were to become the focal points of epilepsy research, brain structures such as the entorhinal and perirhinal cortices and the amygdala were viewed as the critical components of the seizure network.

Studies of the hippocampus of patients with MTLE that used more recent neuroanatomic techniques revealed several signature abnormalities, which, alone or in concert, are suspected to play a critical role in the pathophysiology of the disease. These changes are often seen in conjunction with pronounced neurodegeneration in area CA1 and in the so-called end folium, which includes polymorphic neurons in the hilus of the dentate gyrus and proximal “CA4” pyramidal cells (16) (Fig. 127.1A).

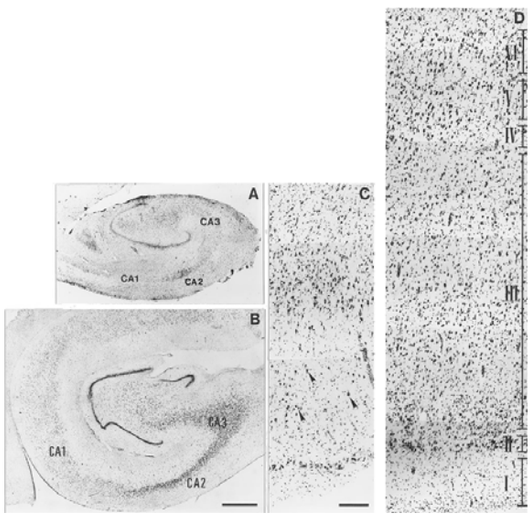


FIGURE 127.1. A and **B:** Nissl-stained coronal sections through the rostral portion of the hippocampus taken from a patient with MTLE **(A)** and a control subject **(B)**. Note the pronounced neuronal loss and gliosis in areas CA1 and CA3 in the patient with epilepsy. Moreover, the epileptic hippocampus is substantially reduced in size compared with the control. **C** and **D:** High magnification of portions of the olfactory field of the entorhinal cortex from a patient with MTLE **(C)** and a control subject **(D)**. Note the substantially reduced width of layers I to III of the entorhinal cortex of the patient with epilepsy. *Arrowheads* in **C** indicate surviving neurons. I to VI, layers of the entorhinal cortex. Scale bars: 1.5 mm in **B** and also in **A;** 200 μm in **C** and also in **D**. (From Du F, Whetsell WO Jr, Abou-Khalil B, et al. Preferential neuronal loss in layer III of the entorhinal cortex in patients with temporal lobe epilepsy. *Epilepsy Res* 1993;16:223–233, with permission.)

Probably the most reliable changes occur in glial cells, which increase in size and number (astrocytes) or change their shape to resemble phagocytotic macrophages (microglia) more closely (17 ,18 and 19). Because of their special biochemical and biophysical properties, these abnormal glial cells are believed to play an active role in the disease process, either by containing or actively enhancing seizure spread (see later).

Of particular interest is the sprouting of mossy fiber axons of surviving excitatory granule cells, which can be visualized by Timm staining or by immunohistochemistry using antibodies against the neuropeptide dynorphin (20). It is often assumed that the new, sprouted fiber network, which has been shown to form recurrent excitatory connections with preexisting neurons (21), may predispose the region to hyperexcitability (22). However, axons also sprout from surviving inhibitory neurons (23), and these novel axons have been proposed to impede chronic hippocampal hyperexcitability (24).

Extrahippocampal Areas

Triggered by new information about the anatomic interconnections in the normal brain and by the results of noninvasive imaging studies of MTLE patients (see later), the 1990s also witnessed a renewed interest in the neuropathology of extrahippocampal temporal lobe lesions in MTLE. As mentioned earlier, lesions in the parahippocampal region or the amygdala, in particular, had been considered among the defining hallmarks of MTLE in earlier days, but they were

less appreciated during a more than 20-year hiatus, ranging essentially from the landmark publication of Margerison and Corsellis (16) to the equally influential monograph of Bruton (25).

In the entorhinal cortex, neuropathologic changes are most readily observed in the superficial layers of the anterior portion of this six-layered parahippocampal structure. Patients frequently present with a characteristic pattern of neuronal loss and associated gliosis, with layer III being preferentially affected and layer II showing pronounced disorganization and some cell loss (26) (Fig. 127.1C). Neuron loss in layer III of the rostral entorhinal cortex can reach dramatic proportions, with only a few, probably γ-aminobutyric acid (GABA)–ergic, cells surviving. This lesion is almost singularly responsible for the substantial tissue shrinkage often described in the epileptic entorhinal cortex (15 ,26). Because pyramidal cells of layer III normally give rise to the monosynaptic "temporoammonic" pathway to area CA1 of the hippocampus (27), their degeneration in MTLE may lead to deafferentation-induced changes in hippocampal excitability, and such changes have indeed been observed in animals (28). Neuropathologic changes in layer

II of the entorhinal cortex, the origin of the major input to the granule cells of the dentate gyrus (the “perforant path”), may also contribute to hippocampal hyperexcitability in MTLE.

Neuronal loss and gliosis in the amygdala are frequently seen in MTLE and often occur in conjunction with lesions in other parts of the limbic system (29 ,30). Although the pattern of cell loss has so far not been analyzed in great detail, degenerative events appear primarily to affect the ventromedial aspects of the lateral amygdaloid nucleus and the parvicellular region of the basal nucleus (31). Based on published studies, this relatively restricted damage not only impedes processing of sensory information in intraamygdaloid circuits, but may also account for the impairment of memory processing in MTLE by interrupting information flow to the hippocampal formation (31 ,32).

It is likely that neuropathologic changes also occur in other areas that are connected to the reverberating seizure network underlying MTLE (33). These structures, which include the thalamus (34), have been shown to be atrophied in patients, but the precise nature and distribution of the degenerative changes, as well as their relation to the pathophysiology of MTLE, have not been elucidated to date.

IN VIVO IMAGING

Part of "127 - Temporal Lobe Epilepsy: Renewed Emphasis on Extrahippocampal Areas "

Structural and functional neuroimaging techniques provide noninvasive means to identify brain abnormalities and have thus become indispensable tools to guide therapeutic interventions in neurologic and psychiatric diseases. These methods have also been, and continue to be, of critical importance for the generation and testing of hypotheses related to pathogenesis and disease progression. In the case of MTLE, techniques such as computed tomography, measurements of regional glucose use and receptor densities by positron emission tomography, single photon emission computed tomography measuring regional cerebral blood flow, and, more recently, volumetric and functional magnetic resonance imaging are now widely used—often in concert—to complement classic EEG and neuropsychological patient evaluation. Imaging test results are increasingly used for diagnostic purposes and, specifically, to provide guidance for neurosurgical procedures.

Improvements in the spatial resolution of most imaging techniques have made it possible to study regional brain abnormalities in MTLE with increasing accuracy. In early studies, hypometabolism and decreases in cerebral blood flow in the temporal lobe were demonstrated even when no structural damage was detectable by computed tomography. These methods were unable to localize the changes and to provide quantitative data adequately (8 ,35), but results obtained by magnetic resonance imaging are remarkably informative on both counts. Thus, using various modifications of the technique, it became feasible to visualize hippocampal atrophy in MTLE (36 ,37) (Fig. 127.2A). In the 1990s, magnetic resonance imaging studies also revealed shrinkage in other areas of the seizure circuit, namely, the amygdala (38 ,39), the entorhinal cortex (40), and the thalamus (6), findings demonstrating that the extrahippocampal changes in tissue volume known to exist in many MTLE patients can be visualized noninvasively (Fig. 127.2B-D). Further methodologic developments are likely to permit the imaging of increasingly smaller lesions and thereby to shed light on the roles of other extrahippocampal brain areas in the pathophysiology of MTLE.

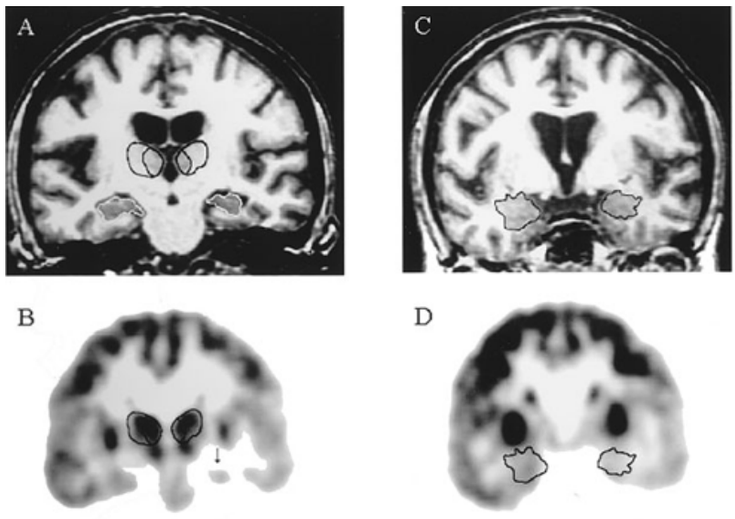


FIGURE 127.2. Imaging studies illustrating changes in extrahippocampal brain regions in MTLE. MRI scans (**A** and **C**) and fluorodeoxyglucose PET scans (**B** and **D**) in a patient with left MTLE. The demarcations on the two sets of scans represent the co-registration of the two techniques so comparable sites are illustrated in the two studies. **A** and **B** show changes in hippocampus (*circled in white* in MRI, indicated with an *arrow* in the PET scan) and thalamus *(outlined in black)* with the medial dorsal nucleus indicated separately in **A**. MRI demonstrates atrophy in the left hippocampus and the left mediodorsal nucleus, and PET shows hypometabolism specific to the temporal lobe and the medial dorsal thalamus. **C** and **D** show atrophy in the left amygdala. The left brain hemisphere corresponds to the right side of the images. (Micrographs courtesy of Drs. C. Juhász and H. Chugani, Wayne State University, Detroit, MI.)

STUDIES IN EXPERIMENTAL ANIMALS

Part of "127 - Temporal Lobe Epilepsy: Renewed Emphasis on Extrahippocampal Areas "

Kindling

Kindling, a phenomenon first described in 1969 (41), has become a major research tool to study seizures involving the limbic system. In this model, a single site in the brain is stimulated electrically with sufficient intensity to induce an electrical after-discharge, but hardly any behavioral changes. With repeated focal stimulations for days or weeks, there is a gradual lengthening of the after-discharge, and behavioral seizures develop. During the seizure, the animal's behavior progresses to the point of a full convulsion. After a number of stimulations, the seizures reach a plateau of

consistent duration and behavioral severity, at which point the animal is considered fully kindled. The number of stimulations required to achieve this plateau depends on several factors, such as the frequency and duration of focal stimulation and the temporal interval between stimuli (42). In addition, however, the site of stimulation is critical for the development of kindling. Studies of several limbic sites, as well as comparisons of neocortical and subcortical regions, demonstrated that some extrahippocampal sites, for example, the amygdala, achieve the fully kindled state much faster than the hippocampus (43). In addition, focal pharmacologic manipulations revealed that both the development and the expression of kindling are significantly influenced by certain subcortical structures, including the midline thalamus and the substantia nigra (44 ,45).

The variable rates of kindling from different stimulation sites suggest that some regions are more *epileptogenic*, that is, more able to generate and support seizure activity than other sites. An alternative explanation is that the areas with more rapid rates of kindling are "closer" to the pathways of secondary generalization. This implies that seizure activity spreads by gradual recruitment of adjacent cortex (similar to the classic *jacksonian march*), exemplified by the proposed role of the perirhinal cortex as the major route from the limbic system to the neocortex (46). However, this idea is confounded by the observation that seizures can appear simultaneously at sites that have few direct interactions (e.g., hippocampus and amygdala) (33). Taken together, these findings therefore suggest the existence of a subcortical "control center," for example, the midline thalamic region, with direct connections to several of the limbic sites involved (Fig. 127.4B).

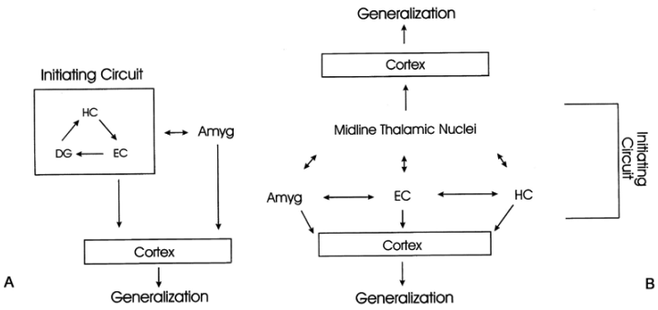


FIGURE 127.4. Schematic representation of the functional anatomy underlying MTLE. **A:** In the conventional model, seizures originate in a reverberating loop consisting of entorhinal cortex, dentate gyrus, and area CA1 of the hippocampus. At some point, seizure activity spreads to the amygdala and other adjacent areas before involving the neocortex and resulting in secondary generalization. **B:** Newer studies suggest an alternative model, in which several limbic sites initially interact with one another independently. Because of their reciprocal connections to all these regions, the midline thalamic nuclei constitute part of the initiating circuit, acting as a subcortical synchronizing site. Generalization can occur through gradual recruitment of adjacent neocortex or through the thalamus, which secondarily recruits the neocortex. Amyg, amygdala; DG, dentate gyrus; EC, entorhinal cortex; HC, hippocampus.

The effect of pharmacologic intervention on the rate of kindling has been of great interest to epilepsy research, because drugs that prolong or abolish kindling acquisition could ameliorate or prevent the development of chronic epilepsy, that is, act as antiepileptogenic agents. Indeed, agents such as phenobarbital, carbamazepine, and several *N*-methyl-D-aspartate (NMDA)–receptor antagonists attenuate the kindling rate, whereas other commonly used antiepileptic compounds, such as phenytoin, do not have such an effect (47). This clinically relevant use of the kindling model will eventually also provide a better definition of the molecular events that underlie the progressive lengthening of seizure activity on repetitive stimulation.

Animal Models of MTLE

The kindling model has been, and continues to be, valuable for our understanding of the role of particular brain regions in the generation and propagation of seizures. However, with a few exceptions, kindled animals do not develop spontaneous seizures. Since the 1980s, the development of animal models with recurrent, spontaneous convulsions has therefore become a major focus of epilepsy research. These models are commonly based on an inciting event of limbic status epilepticus that is precipitated by various methods, including the systemic administration of chemoconvulsants (e.g., pilocarpine or kainate) (48 ,49), focal microinjection of a chemoconvulsant (e.g., kainate) in a limbic brain region (50), or focal, prolonged (up to 90 minutes) electrical stimulation of a limbic structure (hippocampus, perforant path, amygdala) (51 ,52 and 53). In all these models, the animals recover after status epilepticus and, after a latent period of weeks to months, develop spontaneous seizures that continue intermittently throughout the animals' lives.

These rat models parallel the human condition in several noteworthy ways. First, seizures appear spontaneously and, like the seizures of human MTLE, have a clear predilection to occur during daytime hours. This pattern suggests that the seizures are likely to be under subcortical influence. Second, as in the case of human MTLE, there is a latent period during which a previously nonepileptic brain evolves into one that generates recurring seizures. Third, the histopathologic changes in chronically epileptic animals, such as gliosis and neuronal loss in the hippocampus, amygdala, entorhinal cortex, and medial dorsal thalamus, closely resemble those seen in the human condition. Finally, the EEG patterns and the locations of seizure onset are similar in human MTLE and in these animal models, a finding suggesting similarities in the underlying pathophysiology (Fig. 127.3).

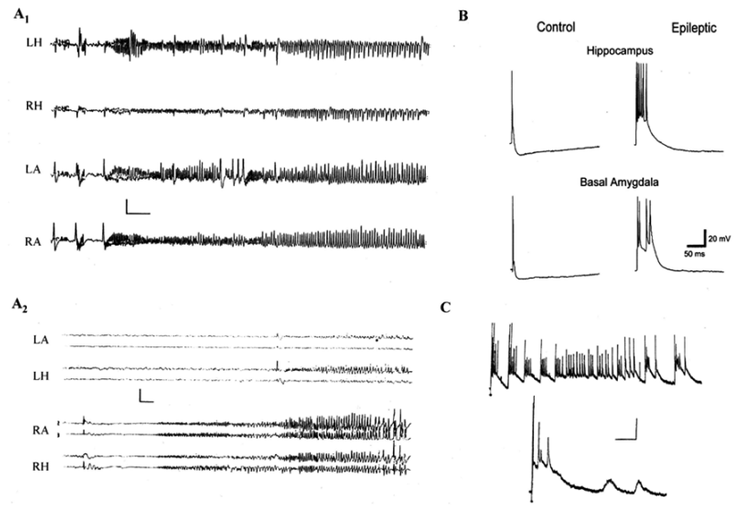


FIGURE 127.3. Seizures in human MTLE and in rat models of MTLE: illustrations of altered neuronal physiology. Chronic epilepsy in animals was induced by prolonged electrical stimulation or chemoconvulsants (see text). **A:** Bilateral recordings from hippocampus (LH, RH) and amygdala (LA, RA) in rat (A1) and human (A2). Both sites are involved at the onset, a finding suggesting a synchronizing pulse from an external site. The human recording shows a later but regionally simultaneous onset on the left hemisphere. This time difference is not seen in rats, probably as a result of stronger interhemispheric connections. **B:** Intracellular recordings from hippocampal and amygdalar neurons in normal and epileptic rats. At both sites, after a brief (0.1-ms) electrical pulse, the neurons from the epileptic animals show prolonged depolarization with multiple superimposed action potentials as compared with normal controls. **C:** Intracellular recordings from the entorhinal cortex of epileptic rats, also illustrating the prolonged depolarizations and multiple superimposed action potentials. Overall, this figure emphasizes the changes throughout the limbic system in nonhippocampal regions.

For these reasons, rat models of MTLE have become increasingly important for the study of limbic epilepsy. The animals offer several important advantages, which largely offset the limitations of the rodent brain, that is, its small size and less developed cortical architecture. Rats are not only relatively inexpensive and easy to manipulate, but allow well-controlled, multidisciplinary studies that can be readily compared across laboratories. The availability of "MTLE-like" animals also makes it possible to study the progression of events that eventually result in a chronic epileptic condition, that is, to assess the silent, latent period between status epilepticus and the first spontaneous limbic seizure. Finally, these animal models provide opportunities to test innovative clinical interventions for the treatment and possible cure of MTLE (54 ,55).

These models of limbic epilepsy have been exploited to test novel concepts related to the pathophysiology of MTLE. For example, hyperexcitable neurons, both projection and interneurons, are often found in many limbic sites associated with MTLE. In some cases, the synaptically mediated responses in these neurons show prolonged excitatory depolarizations with multiple superimposed action potentials (23 ,56 ,57) (Fig. 127.3). Underlying these profound changes are multiple presynaptic and postsynaptic alterations in neurotransmitter receptors and ion channels. Judged by physiologic criteria, there is evidence that changes in both inhibitory and excitatory receptors, namely GABA and glutamate receptors, contribute to the abnormal physiology (58 ,59). These changes are necessarily complex because

they reflect a composite picture of receptors and receptor subunits and their reactive up-regulation or down-regulation (60 ,61 and 62). In addition, the models also involve many other biochemical changes in one or more parts of the limbic seizure network. This ever-increasing list includes ions, enzymes, hormones and their receptors, and a host of other chemical entities. In most instances, however, the possible functional significance of these changes has so far not been elucidated.

Experimental MTLE also provides an opportunity to evaluate the existence and functional significance of synaptic rearrangements, which occur in response to epileptogenic insults or neuronal damage and may be involved in the development of spontaneously recurring seizures (20 ,21 ,22 ,23 and 24 ,63). The prototype for this process, the sprouting of mossy fibers originating from hippocampal granule cells, has garnered substantial attention, in part because of the seeming simplicity of the new circuit and the easily visualized anatomic change (64). Because the animal models of MTLE show extensive neuronal loss in multiple brain areas, it is quite likely that seizure-related synaptic and circuit rearrangements also occur in extrahippocampal regions (65).

CURRENT CONCEPTS OF EPILEPTOGENESIS: MOLECULES AND MECHANISMS

Part of "127 - Temporal Lobe Epilepsy: Renewed Emphasis on Extrahippocampal Areas "

Neurotransmitters

A long-standing theme in epilepsy research has been the role of the major inhibitory neurotransmitter, GABA, in

both epileptogenesis and chronic epilepsy. Defects of the GABA system have been implicated since the 1970s, when a loss of GABA neurons was suspected to underlie seizures (66), but physiologic studies in animal models of epilepsy did not necessarily reveal a reduction in GABAergic inhibition (67). Other studies proposed a specific defect in the excitability of otherwise normal (“dormant”) GABAergic neurons resulting from their deafferentation (68). This hypothesis is still rather controversial, attesting to the complexity of GABAergic neurotransmission (63). However, it seems unlikely that a single deficit in the GABAergic systems of the brain underlies limbic epilepsy.

More recent studies have expanded our understanding of GABAergic transmission. For example, we now have a greatly improved view of GABAergic inhibitory neurons *(interneurons)*. These cells are morphologically heterogeneous, have diverse connections, and contain different colocalized transmitters (69). Thus, many GABAergic neurons also contain the peptides neuropeptide Y and somatostatin or calcium-binding proteins such as parvalbumin or calretinin. Although many anatomic studies have focused exclusively on the hippocampus and the neocortex, some common themes also seem to apply to other sites, such as the entorhinal and the perirhinal cortex. The physiologic characteristics of these different cells and their respective roles in epileptogenesis in both hippocampus and extrahippocampal areas will need to be elaborated in the future. These studies must also be complemented by a detailed molecular and functional analysis of GABA receptors and GABA transporters (70 ,71).

Because MTLE is frequently conceptualized as the result of an imbalance between excitation and inhibition in the brain, the abundant excitatory neurotransmitter glutamate is also likely to play an important role in the pathophysiology of MTLE. Indeed, both ionotropic [NMDA, 2-amino-(3-hydroxy-5-methylisoxazol-4-yl) propanoate (AMPA) and kainate] and metabotropic (i.e., G-protein–coupled) glutamate receptors have been implicated in epileptogenesis and chronic epilepsy (44 ,48 ,50 ,58 ,61 ,62). For example, a decrease in magnesium, an endogenous modulator of the NMDA receptor, greatly reduces the receptor's voltage dependency and allows channel opening at central neurons even at normal, hyperpolarized membrane potentials. The result is spontaneous epileptiform activity (72). NMDA-receptor subunit composition changes in response to seizure activity and may be responsible for the increase in excitability that often follows an initial seizure (61 ,62). Such changes occur in MTLE as well as in relevant animal models and are seen both in the hippocampus, so far the most thoroughly investigated brain region, and in extrahippocampal areas such as the entorhinal cortex (56).

Blockade of AMPA or kainate receptors is one of the most effective means to reduce epileptiform activity. These receptors, which mediate most fast excitatory neurotransmission, are also composed of an array of subunits, which assemble to form distinct receptor subtypes (73). Receptor compositions are abnormal in epileptic brain tissue of both humans and animals (60 ,61), although their role in epileptogenesis is currently unclear. This not only suggests that AMPA and kainate receptors are involved in spontaneously recurrent seizure activity but also offers promising new ideas for antiepileptic drug development (74).

Pharmacologic manipulation of metabotropic glutamate receptors and interference with the function of glutamate transporters, a group of several distinct proteins that control the extracellular concentration of glutamate in the brain, too, have profound consequences on neuronal excitability. These sites are therefore under careful scrutiny both for their potential role in epileptogenesis and as novel therapeutic targets (75 ,76).

A considerable body of evidence suggests that other classic neurotransmitters, such as monoamines and acetylcholine, also play critical roles in MTLE, and that the proconvulsant or anticonvulsant effects of these agents manifest themselves in several regions of the epileptic network. Drugs that selectively influence, for example, noradrenergic or cholinergic neurotransmission are therefore useful experimental tools and of potential benefit in the treatment of MTLE (77 ,78).

Neuromodulators: Neuropeptides, Growth Factors, Cytokines

Neuropeptides have long constituted an enigma to neuroscientists. They are present in many different brain areas, where they are localized in various neuronal types, often in cells expressing GABA as a neurotransmitter (69), yet their physiologic function remains obscure. Some of these agents, such as neuropeptide Y and dynorphin, show increased or reduced brain content after prolonged seizure activity (20 ,79). Physiologic studies have demonstrated that neuropeptide Y has an inhibitory action, suppressing neurotransmission in single hippocampal cells and seizures *in vivo*. It is therefore conceivable that neuropeptide Y, portrayed here as a prototype of a relatively large group of neuroactive peptides, plays an important role as an endogenous modulator of seizure generation and in epilepsy. Its physiologic actions, as well as its chemical and anatomic changes in MTLE, are not confined to the hippocampus but also affect other parts of the seizure network (79).

Several peptides and small proteins with no previously recognized relevance to epilepsy may also turn out to play a role in epileptogenesis. Examples include growth factors and cytokines, which have been shown to be neuroactive in various test systems. Thus, various members of the neurotrophin family, including nerve growth factor, brain-derived growth factor (BDNF), and others, for example, neurotrophin 3 and neurotrophin 4/5, influence neurotransmission, although their effects vary widely in both qualitative and quantitative terms (80). All reports of these neuroactive effects are so far based on exogenously applied neurotrophins; that is, the results could be compromised by the finding

that the concentrations used for experimentation exceeded physiologic levels. Still, it is certainly of interest that the brain concentration of some neurotrophins, such as BDNF, increases dramatically after seizures, whereas others, such as neurotrophin 3, decrease. The high-affinity trk receptors and low-affinity p75 receptors for neurotrophins also appear to change after seizures, setting the stage for potentially important, interactive roles of the various neurotrophins in epileptogenesis and epilepsy (80). Because normal expression and seizure-induced changes of these putative neuromodulators occur throughout the limbic system, critical seizure-related effects may not only take place in the hippocampus, where most studies have been performed to date. For example, potent effects of BDNF in the entorhinal cortex have been described (81).

Cytokines were originally characterized as mediators of inflammatory responses and were discussed as messengers of the immune system. However, some of these compounds, for example, interleukin 1 (IL-1), IL-6, and tumor necrosis factor-α, are expressed in the brain, influence neuronal activity, and therefore have potential links to seizure mechanisms. In particular, the mRNA for IL-1ß and IL-6, as well as IL receptors, is increased by seizures (82). Moreover, IL-1α is elevated in tissue from patients with epilepsy (83). Finally, the finding that cytokines are expressed and released by microglia underscores the role of nonneuronal cells in epilepsy (see later).

Neuron-Glia Interactions

The importance of glia to neuronal function has been appreciated since the early period of neuroscience research. Yet to this day, the interrelationship of glia and neurons continues to unfold. Besides other roles in brain physiology, glial cells may play a significant role in the modulation of seizure phenomena. Thus, both astrocytes and microglial cells, the two major glial cell types in the brain, are rapidly activated by seizure activity in the limbic system (84). It is currently unclear, however, whether this cellular reaction has proconvulsive or anticonvulsive effects. Astrocytes have the ability to buffer extracellular potassium and can avidly accumulate the excitatory amino acid glutamate (76). Moreover, astrocytes increase the production and release of the endogenous neuroinhibitory and anticonvulsant compound kynurenate, possibly as an early defensive response to seizures (85). These and many conceptually related data indicate a protective function of astrocytes in epileptogenesis and, perhaps, in chronic epilepsy. Conversely, both astrocytes and microglia also synthesize endogenous proconvulsive agents such as quinolinate (86) and cytokines (87), which may exacerbate seizure activity during any stage of the epileptic process.

Of possible relevance to their role in epilepsy, glial cells in limbic brain areas are heterogeneous with regard to both structure and function. For example, both the electrophysiologic properties and the histochemical staining pattern of astrocytes in area CA1 of the hippocampus differ from those in area CA3 (88). Similar differences are likely to exist in other brain areas as well, adding another layer of complexity to the study of neuron-glia interactions as they pertain to mechanisms of epileptogenesis and chronic epilepsy.

Seizure-Induced Changes in Gene Expression

Prolonged seizure activity, especially episodes of status epilepticus, often has dramatic effects on gene expression, inducing a bewildering array of new genes in the brain. The expression of a variety of proteins and peptides is increased, whereas that of others is reduced. These changes involve neurotransmitters, neuromodulators and their receptors, growth factors, cytokines, and additional classes of compounds. A popular hypothesis to explain these changes is that seizures induce, or influence the expression of, genes that are normally expressed during development. Sequelae of seizures may therefore mimic or, to use a more teleologic term, *recapitulate* development. Examples include the seizure-induced up-regulation of growth factors and proteins that are involved in synaptogenesis (89). Other investigators have proposed, again arguing teleologically, that seizure-induced changes merely constitute compensatory mechanisms of the adult brain, that is, attempts of the system to counter a potentially dangerous increase in excitability or impending cellular damage. This implies the up-regulation of systems that inhibit neuronal activity. Indeed, seizures lead to changed expression of glutamate decarboxylase, the enzyme responsible for GABA synthesis, and GABA receptors (71 ,90).

Conversely, seizure-induced changes in gene expression may be part of the development of the epileptic state. Thus, a first seizure may induce gene expression of substances that will contribute to further hyperexcitability. One example is the neurotrophin BDNF, which is normally expressed in dentate gyrus granule cells and, to a lesser extent, other areas of the hippocampus and other brain regions (91). After a single seizure, BDNF message, protein, and the high-affinity trkB receptor all increase in granule cells (92 ,93). Because BDNF enhances neuronal activity in the hippocampus, the increase in expression could have functional consequences; that is, it could lead to a reduction in seizure threshold. BDNF also has effects on neuronal structure and could thus contribute to structural changes occurring after seizures that, in turn, increase susceptibility to seizures (81 ,94). Because BDNF, and other neurotrophins and neuromodulators, are expressed in extrahippocampal regions, this hyperexcitability may also occur in these areas.

Synaptic Reorganization

As mentioned above, seizures induce many genes in the brain. These genes express a variety of different proteins,

which often closely resemble or duplicate those that are preferentially expressed during brain development. It is therefore hardly surprising that growth and synaptic reorganization occur as a consequence of seizure activity. In experimental animals, this phenomenon has been studied in great detail in the hippocampus, although it can also be observed in extrahippocampal areas such as the entorhinal cortex (65) and the neocortex (95). In the hippocampus, mossy fiber axons of dentate granule cells grow new collaterals that innervate an abnormal lamina *(mossy fiber sprouting)* (96). It has been suggested that these connections are functional (97 ,98), but opinions are still divided on whether and to what extent they contribute to the precipitation of spontaneously recurring seizures (20 ,21 ,22 ,23 and 24).

Neurogenesis

Confirmation of neurogenesis in the adult mammalian nervous system has galvanized interest in the potential role of newly born cells in the mature brain. This issue may be of particular relevance to the study of epilepsy because neurogenesis is stimulated in adult animals after seizures. This has been documented in some detail in the hippocampus, where an increase in newly born dentate granule cells was detected in response to single seizures (99), kindling (100), or status epilepticus (101). In rats treated with pilocarpine, newly born granule cells develop intrinsic electrophysiologic properties that are identical to normal granule cells. However, the integration of these cells into the host circuitry appears to be different from adult granule cells, because they become synchronized with epileptiform activity in the CA3 cell layer (102). These studies indicate that cells born in response to seizures can become fully functioning neurons and may develop abnormal electrical activity.

More recent studies, although still somewhat controversial, have demonstrated that neurogenesis in the adult brain can also occur in the neocortex (103). Moreover, many newly born cells, possibly neurons, can also be detected in the entorhinal cortex of pilocarpine-treated rats (104). Rather than an isolated, region-specific phenomenon, the birth of new neurons in response to seizures may therefore take place in several parts of the limbic system and possibly elsewhere. Alone or together, these cells may eventually cause enhanced excitability and may decrease seizure threshold. It is tempting to speculate that this may underlie the heightened excitability in patients with cortical dysplasias, which can arise as a result of aberrant migration of newly born cells.

Transgenic Animals

The availability of transgenic mice has provided exciting new research opportunities for the study of MTLE. Thus, targeted genetic manipulation has permitted the detailed examination of the effects of changes in one specific gene product. This selectivity is particularly important in the study of epileptic phenomena, which are associated with many concurrent molecular and cellular changes. Transgenics may be engineered to overexpress or delete a given gene product, and recent techniques have made it possible to modify gene expression conditionally, that is, in a brain region– or age-specific fashion (105).

To name only one of several relevant examples, BDNF transgenics show increased excitability and, in some cases, exhibit spontaneous seizures. BDNF overexpression in these animals is most obvious in mossy fiber axons of dentate gyrus granule cells, and excitability is indeed increased in the postsynaptic targets of mossy fibers. However, BDNF is also overexpressed in other limbic regions such as the entorhinal cortex, and these areas, too, are hyperexcitable (94). These and qualitatively similar results from other genetically manipulated animals illustrate the value of this novel experimental approach for delineating the respective roles of the hippocampus and extrahippocampal areas in epileptogenesis and chronic epilepsy.

MTLE AS A NETWORK DISEASE: IMPLICATIONS FOR FURTHER RESEARCH AND THERAPY

Part of "127 - Temporal Lobe Epilepsy: Renewed Emphasis on Extrahippocampal Areas "

As briefly reviewed here, the available data indicate that both the hippocampus and interconnected extrahippocampal limbic regions play a role in the pathophysiology of MTLE (Fig. 127.4). Methodologic advances, ranging from novel and improved *in vivo* imaging techniques and neurosurgical approaches to increasingly sophisticated assessments of cellular physiology and chemistry, have resulted in a refinement of nineteenth-century concepts without fundamentally altering the major premises of Sommer and his contemporaries. In humans, any of a number of primary insults, such as severe febrile convulsions during childhood, head trauma, infections, tumors, or developmental malformations, have been proposed to be epileptogenic, leading to MTLE after a characteristic latency period. Studies in several new animal models have shown that these pathogenic injuries preferentially affect one or more limbic brain areas, including, but not necessarily limited to, hippocampus, entorhinal cortex, amygdala, thalamus, and neocortical regions. This neuronal injury or degeneration may constitute a major factor in the establishment of epileptogenic circuits within the limbic system and may cause further structural and functional changes. Eventually, after a “silent” interval characterized by functionally significant yet currently still unspecified changes, a hyperexcitable epileptogenic circuit develops, resulting in MTLE.

This concept of epileptogenesis and chronic limbic epilepsy has several ramifications for the treatment of MTLE. First, it is possible that the development of an epileptic circuit will be inhibited if the multiple changes in limbic

circuitry that follow the initial insult are minimized. Second, it raises the possibility that hippocampal or extrahippocampal interventions during the “silent” period may prevent the evolution of spontaneously recurring seizures. Third, it indicates that any of a number of limbic brain areas, or even certain neuronal populations within a given region, could conceivably be targeted for surgical or pharmacologic intervention.

Although the pathophysiologically important role of extrahippocampal brain regions in MTLE is supported by the successful outcome of various surgical interventions in patients with medically intractable epilepsy (5 ,9 ,10 ,14 ,15 and 16), several critical questions need to be resolved before the therapeutic approaches listed earlier can be adequately evaluated in clinical settings. For example, we need to clarify whether neuronal death, gliosis, and synaptic reorganizations promote or attenuate seizure evolution and whether, on the cellular level, pharmacologic or genetic manipulation of receptors and ion channels can influence the development of chronic hyperexcitability. Because circuit disruption is widespread throughout the limbic system, we need to elaborate those subcortical structures that are particularly effective in controlling seizure spread and generalization. Other studies will decipher cell- and region-specific and time-dependent changes in the composition of neurotransmitter receptors and of other proteins that determine the action of endogenous neuroactive agents. These analyses must be complemented by examining shifts in the concentration of endogenous neuromodulators that can alter neuronal excitability. To optimize the development of novel strategies for the treatment of MTLE, these molecular studies should be performed in parallel in extrahippocampal areas and in the hippocampal region whenever possible. This approach should eventually provide significant advances in the therapy of this debilitating disorder.

ACKNOWLEDGMENTS

Part of "127 - Temporal Lobe Epilepsy: Renewed Emphasis on Extrahippocampal Areas "

Work described in this article was in part supported by United States Public Health Service grants NS 16102, NS 25605, and NS 37562. We are grateful to Drs. C. Juhasz and H. Chugani (Wayne State University, Detroit, MI) for providing the micrographs shown in Fig. 127.2 .

REFERENCES

1. Wyllie E. *The treatment of epilepsy: principles and practice*. Philadelphia: Lea & Febiger, 1993.

2. Engel J, ed. *Surgical treatment of the epilepsies*. New York: Raven, 1993.

3. Lopes Da Silva FH, Witter MP, Boeijinga PH, et al. Anatomic organization and physiology of the limbic cortex.*Physiol Rev* 1990;70:453–511.

4. So NK. Depth electrode studies in mesial temporal epilepsy. In: Lüders H, ed. *Epilepsy surgery*. New York: Raven, 1991:371–384.

5. Spencer DD, Spencer SS. Hippocampal resections and the use of human tissue in defining temporal lobe epilepsy syndromes. *Hippocampus* 1994;4:243–249.

6. Juhász C, Nagy F, Watson C, et al. Glucose and [^{14}C]flumazenil positron emission tomography abnormalities of thalamic nuclei in temporal lobe epilepsy. *Neurology* 1999;53:2037–2045.

7. Sperling MR, Gur RC, Alavi A, et al. Subcortical metabolic alterations in partial epilepsy. *Epilepsia* 1990;31:145-155.

8. Theodore WH, Brooks R, Margolin R, et al. Positron emission tomography in generalized seizures. *Neurology* 1985;35:684–690.

9. Siegel AM, Wieser HG, Wichmann W, et al. Relationships between MR-imaged total amount of tissue removed, resection scores of specific mediobasal limbic subcompartments and clinical outcome following selective amygdalohippocampectomy. *Epilepsy Res* 1990;6:56–65.

10. Parrent AG, Blume WT. Stereotactic amygdalohippocampotomy for the treatment of medial temporal lobe epilepsy. *Epilepsia* 1999;40:1408–1416.

11. Sommer W. Erkrankung des Ammonshorns als aetiologisches Moment der Epilepsie. *Arch Psychiatr Nervenkrankh* 1880;10:631–675.

12. Lorente de Nó R. Studies on the structure of the cerebral cortex. *J Psychol Neurol* 1933;45:381–438.

13. Zimmerman HM. The histopathology of convulsive disorders of children. *J Pediatr* 1938;13:839–890.

14. Penfield W, Baldwin M. Temporal lobe seizures and the technic of subtotal temporal lobectomy. *Ann Surg* 1952;136:625–634.

15. Gastaut H. Colloque sur les problemes d'anatomie normale et pathologique posés par les décharges épileptiques. *Acta Med Belg* 1956:5–66.

16. Margerison JH, Corsellis JAN. Epilepsy and the temporal lobes: a clinical, electroencephalographic and neuropathological study of the brain in epilepsy, with particular reference to the temporal lobes. *Brain* 1966;89:499–530.

17. O'Connor ER, Sontheimer H, Spencer DD, et al. Astrocytes from human hippocampal epileptogenic foci exhibit action potential-like responses. *Epilepsia* 1998;39:347–354.

18. Krishman B, Armstrong DL, Grossman RG, et al. Glial cell nuclear hypertrophy in complex partial seizures. *J Neuropathol Exp Neurol* 1994;53:502–507.

19. Beach TG, Woodhurst WB, MacDonald DB, et al. Reactive microglia in hippocampal sclerosis associated with human temporal lobe epilepsy. *Neurosci Lett* 1995;191:27–30.

20. Houser CR, Miyashiro JE, Swartz BE, et al. Altered patterns of dynorphin immunoreactivity suggest mossy fiber reorganization in human hippocampal epilepsy. *J Neurosci* 1990;10:267–282.

21. Okazaki MM, Evenson DA, Nadler JV. Hippocampal mossy fiber sprouting and synapse formation after status epilepticus in rats: visualization after retrograde transport of biocytin. *J Comparative Neurology* 1995;352:515–534.

22. Wuarin JP, Dudek FE. Electrographic seizures and new recurrent excitatory circuits in the dentate gyrus of hippocampal slices from kainate-treated epileptic rats.*J Neurosci* 1996;16:4438–4448.

23. Davenport CJ, Brown WJ, Babb TL. Sprouting of GABAergic and mossy fiber axons in dentate gyrus following intrahippocampal kainate in the rat. *Exp Neurol* 1990;109:180–190.

24. Sloviter RS. Possible functional consequences of synaptic reorganization in the dentate gyrus of kainate-treated rats. *Neurosci Lett* 1992;137:91–96.

25. Bruton CJ. The neuropathology of temporal lobe epilepsy. *Maudsley monograph no. 31*. Oxford: Oxford University Press, 1988.

26. Du F, Whetsell WO Jr, Abou-Khalil B, et al. Preferential neuronal loss in layer III of the entorhinal cortex in patients with temporal lobe epilepsy. *Epilepsy Res* 1993;16:223–233.

27. Witter MP, Griffioen AW, Jorritsma-Byham B, et al. Entorhinal projections to the hippocampal CA1 region in the rat: an underestimated pathway. *Neurosci Lett* 1988;85:193–198.

28. Scharfman HE, Goodman JH, Du F, et al. Chronic changes in synaptic responses of entorhinal and hippocampal neurons after entorhinal cortical cell loss produced by intracortical amino-oxyacetic acid (AOAA) injection: an *in vitro* and *in vivo* study in the rat. *J Neurophysiol* 1998;80:3031–3046.

29. Cavanagh JB, Meyer A. Aetiological aspects of Ammon's horn sclerosis associated with temporal lobe epilepsy.*BMJ* 1956;2:1403–1407.

30. Hudson LP, Monoz DG, Miller L, et al. Amgdaloid sclerosis in temporal lobe epilepsy. *Ann Neurol* 1993;33:622–631.

31. Pitkänen A, Tuunanen J, Kalviainen R, et al. Amygdala damage in experimental and human temporal lobe epilepsy. *Epilepsy Res* 1998;32:233–253.

32. Cahill L, Haier RJ, Fallon J, et al. Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proc Natl Acad Sci USA* 1996;93:8016–8021.

33. Bertram EH. Functional anatomy of spontaneous seizures in a rat model of limbic epilepsy. *Epilepsia* 1997;38:95–105.

34. De Carli C, Hatta J, Fazilat S, et al. Extratemporal atrophy in patients with complex partial seizures of left temporal origin. *Ann Neurol* 1998;43:41–45.

35. Engel J Jr, Kuhl DE, Phelps ME, et al. Interictal cerebral glucose metabolism in partial epilepsy and its relation to EEG changes. *Ann Neurol* 1982;12:510–517.

36. Berkovic SF, Andermann F, Olivier A, et al. Hippocampal sclerosis in temporal lobe epilepsy demonstrated by magnetic resonance imaging. *Ann Neurol* 1991;29:175–182.

37. Lencz T, McCarthy G, Bronen RA, et al. Quantitative magnetic resonance imaging in temporal lobe epilepsy: relationship to neuropathology and neuropsychological function. *Ann Neurol* 1992;31:629–637.

38. Cendes F, Andermann F, Gloor P, et al. Atrophy of mesial structures in patients with temporal lobe epilepsy: cause or consequence of repeated seizures? *Ann Neurol* 1993;34:795–801.

39. Van Paesschen W, Connelly A, Johnson CL, et al. The amygdala and intractable temporal lobe epilepsy: a quantitative magnetic resonance imaging study. *Neurology* 1996;47:1021–1031.

40. Bernasconi N, Bernasconi A, Andermann F, et al. Entorhinal cortex in temporal lobe epilepsy. *Neurology* 1999;52:1870–1876.

41. Goddard GV, McIntyre DC, Leech CK. A permanent change in brain function resulting from daily electrical stimulation. *Exp Neurol* 1969;25:295–330.

42. Lothman EW, Williamson JK. Rapid kindling with recurrent hipocampal seizures: effect of stimulus frequency and train duration. *Epilepsy Res* 1993;14:209–220.

43. Racine R. Kindling: the first decade. *Neurosurgery* 1978;3:234–252.

44. Hirayasu Y, Wada JA. *N*-methyl-D-aspartate injection into the massa intermedia facilitates development of limbic kindling in rats. *Epilepsia* 1992;33:965–970.

45. Gale K. Subcortical structures and pathways involved in convulsive seizure generation. *J Clin Neurophysiol* 1992;9:264–277.

46. McIntyre DC, Kelly ME, Dufresne C. FAST and SLOW amygdala kindling rat strains: comparison of amygdala, hippocampal, piriform and perirhinal cortex kindling. *Epilepsy Res* 1999;35:197–209.

47. Löscher W. Animal models of intractable epilepsy. *Prog Neurobiol* 1997;53:239–258.

48. Ben-Ari Y. Limbic seizure and brain damage produced by kainic acid: mechanisms and relevance to human temporal lobe epilepsy. *Neuroscience* 1985;14:375–403.

49. Cavalheiro EA, Leite JP, Bortolotto ZA, et al. Long-term effects of pilocarpine in rats: structural damage of the brain triggers kindling and spontaneous recurrent seizures. *Epilepsia* 1991;32:778–782.

50. Wilson CL, Maidment NT, Shomer MH, et al. Comparison of seizure related amino acid release in human epileptic hippocampus versus a chronic, kainate rat model of hippocampal epilepsy. *Epilepsy Res* 1996;26:245–254.

51. Lothman EW, Bertram EH, Kapur J, et al. Recurrent spontaneous hippocampal seizures in the rat as a chronic sequela to limbic status epilepticus. *Epilepsy Res* 1990;6:110–118.

52. Nissinen J, Halonen T, Koivisto E, et al. A new model of chronic temporal lobe epilepsy induced by electrical stimulation of the amygdala in rat. *Epilepsy Res* 2000;38:177–205.

53. Mazarati AM, Wasterlain CG, Sankar R, et al. Self-sustaining status epilepticus after brief electrical stimulation of the perforant path. *Brain Res* 1998;801:251–253.

54. Leite JP, Cavalheiro EA. Effects of conventional antiepileptic drugs in a model of spontaneous recurrent seizures in rats. *Epilepsy Res* 1995;20:93–104.

55. Bertram, EH, Riddick G. Model of chronic spontaneous limbic seizures can be used for testing antiepileptic drugs. *Epilepsia* 1993;34(suppl 6):12.

56. Fountain NB, Bear J, Bertram EH III, et al. Responses of deep entorhinal cortex are epileptiform in an electrogenic rat model of chronic temporal lobe epilepsy. *J Neurophysiol* 1998;80:230–240.

57. Scharfman HE. Hyperexcitability of entorhinal cortex and hippocampus after application of amino-oxyacetic acid (AOAA) to medial entorhinal cortex in combined entorhinal and hippocampal slices. *J Neurophysiol* 1997;76:2986–3001.

58. Bradford HF. Glutamate, GABA and epilepsy. *Prog Neurobiol* 1995;47:477–511.

59. Mangan P, Bertram E. Shortened duration GABA A receptor mediated synaptic potentials underlie enhanced CA1 excitability in a chronic model of temporal lobe epilepsy. *Neuroscience* 1997;80:1101–1111.

60. de Lanerolle NC, Eid T, von Campe G, et al. Glutamate receptor subunits GluR1 and GluR2/3 distribution shows reorganization in the human epileptogenic hippocampus. *Eur J Neurosci* 1998;10:1687–1703.

61. Zilles K, Qu MS, Kohling R, et al. Ionotropic glutamate and GABA receptors in human epileptic neocortical tissue: quantitative *in vitro* receptor autoradiography. *Neuroscience* 1999;94:1051–1061.

62. Musshof U, Schünke U, Köhling R, et al. Alternative splicing of the NMDAR1 glutamate receptor subunit in human temporal lobe epilepsy. *Mol Brain Res* 2000;76:377–384.

63. Bernard C, Esclapez M, Hirsch JC, et al. Interneurones are not so dormant in temporal lobe epilepsy: a critical reappraisal of the dormant basket cell hypothesis. *Epilepsy Res* 1998;32:93–103.

64. Tauck DL, Nadler JV. Evidence of functional mossy fiber sprouting in hippocampal formation of kainic acid treated rats. *J Neurosci* 1985;5:1016–1022.

65. Eid T, Schwarcz R, Ottersen OP. Ultrastructure and immunocytochemical distribution of GABA in layer III of the rat medial cortex following aminoxyacetic acid-induced seizures. *Exp Brain Res* 1999;125:463–475.

66. Ribak CE. Morphological, biochemical and immunocytochemical changes of the cortical GABAergic system in epileptic foci. In: Ward AA, Penry JK, Purpura DP, eds. *Epilepsy*. New York: Raven, 1983:109–129.

67. Prince DA, Jacobs KM, Salin PA, et al. Chronic focal neocortical epileptogenesis: does disinhibition play a role? *Can J Physiol Pharmacol* 1997;75:500–507.

68. Sloviter RS. Permanently altered hippocampal structure, excitability, and inhibition after experimental status epilepticus in the rat: the “dormant basket cell” hypothesis and its possible relevance to temporal lobe epilepsy. *Hippocampus* 1991;1:41–66.

69. Freund TF, Buzsaki G. Interneurons of the hippocampus.*Hippocampus* 1996;6:347–470.

70. Kapur J, Coulter DA. Experimental status epilepticus alters γ-aminobutyric acid type A receptor function in CA1 pyramidal neurons. *Ann Neurol* 1995;38:893–900.

71. Sperk G, Schwarzer C, Tsunashima K, et al. Expression of GABA A receptor subunits in the hippocampus of the rat after kainic acid-induced seizures. *Epilepsy Res* 1998;32:129–139.

72. Stanton PK, Jones RSG, Mody I, et al. Epileptiform activity induced by lowering extracellular [Mg^{2+}] in combined hippocampal-entorhinal cortex slices: modulation by receptors for norepinephrine and *N*-methyl-D-aspartate. *Epilepsy Res* 1987;1:53–62.

73. Seeburg PH. The TINS/TIPS lecture The molecular biology of mammalian glutamate receptor channels. *Trends Neurosci* 1993;16:359–365.

74. Swedberg MDB, Jacobsen P, Honoré T. Anticonvulsant, anxiolytic and discriminative effects of the AMPA antagonist 2,3-dihydroxy-6-nitro-7-sulfamoyl-benzo(f)quinoxaline (NBQX). *J Pharmacol Exp Ther* 1995;274:1113–1121.

75. Aronica EM, Gorter JA, Paupard M-C, et al. Status epilepticus-induced alterations in metabotropic glutamate receptor expression in young and adult rats. *J Neurosci* 1997;17:8588–8595.

76. Tessler S, Danbolt NC, Faull RLM, et al. Expression of the glutamate transporters in human temporal lobe epilepsy. *Neuroscience* 1999;88:1083–1091.

77. Nagao T, Alonso A, Avoli M. Epileptiform activity induced by pilocarpine in the rat hippocampal/entorhinal slice preparation. *Neuroscience* 1996;72:399–408.

78. Cavalheiro EA, Fernandes MJ, Turski L, et al. Spontaneous recurrent seizures in rats: amino acid and monamine determination in the hippocampus.*Epilepsia* 1994;35:1–11.

79. Vezzani A, Sperk G, Colmers WF. Neuropeptide Y: emerging evidence for a functional role in seizure modulation. *Trends Neurosc* 1999;22:25–30.

80. Lindvall O, Kokaia Z, Bengzon J, et al. Neurotrophins and brain insults. *Trends Neurosci* 1994;17:490–496.

81. Scharfman HE. Hyperexcitability in combined entorhinal/hippocampal slices of adult rat after exposure to brain-derived neurotrophic factor. *J Neurophysiol* 1997;78:1082–1095.

82. Minami M, Kuraishi Y, Satoh M. Effects of kainic acid on messenger RNA levels of IL1 IL6, TNFα, and Lif in the rat brain. *Biochem Biophys Res Commun* 1991;176:593–598.

83. Sheng JG, Boop FA, Mrak RE, et al. Increased neuronal ß-amyloid precursor protein expression in human temporal lobe epilepsy: association with interleukin 1 immunoreactivity. *J Neurochem* 1994;63:1872–1879.

84. Steward O, Torre ER, Tomasulo R, et al. Neuronal activity up-regulates astroglial gene expression. *Proc Natl Acad Sci USA* 1991;88:6819–6823.

85. Wu H-Q, Schwarcz R. Seizure activity causes elevation of endogenous extracellular kynurenic acid in the rat brain. *Brain Res Bull* 1996;39:155–162.

86. Okuno E, Köhler C, Schwarcz R. Rat 3-hydroxyanthranilic acid oxygenase: purification from the liver and immunocytochemical localization in the brain. *J Neurochem* 1987;49:771–780.

87. Vezzani A, Conti M, Luigi AD, et al. Interleukin 1 immunoreactivity and microglia are enhanced in the rat hippocampus by focal kainate application: functional evidence for enhancement of electrographic seizures.*J Neurosci* 1999;19:5054–5065.

88. D'Ambrosio R, Wenzel J, Schwartzkroin PA, et al. Functional specialization and topographic segregation of hippocampal astrocytes. *J Neurosci* 1998;18:4425–4438.

89. Liu M, Pleasure SJ, Collins AE, et al. Loss of B2/NeuroD leads to malformation of the dentate gyrus and epilepsy. *Proc Natl Acad Sci USA* 2000;97:865–870.

90. Schwarzer C, Sperk G. Hippocampal granule cells empress glutamic acid decarboxylase-67 after limbic seizures in the rat. *Neuroscience* 1995;69:705–709.

91. Yan Q, Rosenfeld RD, Matheson CR, et al. Expression of brainderived neurotrophic factor protein in the adult rat central nervous system. *Neuroscience* 1997;78:431–448.

92. Schmidt-Kastner R, Humpel C, Wetmore C, et al. Cellular hybridization for BDNF, trkB, and NGF mRNAs and BDNF-immunoreactivity in rat forebrain after pilocarpine induced status epilepticus. *Exp Brain Res* 1996;107:331–347.

93. Gall CM. Seizure-induced changes in neurotrophin expression: implications for epilepsy. *Exp Neurol* 1993;124:150–166.

94. Croll SD, Suri C, Compton DL, et al. Brain-derived neurotrophic factor (BDNF) transgenic mice exhibit passive avoidance deficits, increased seizure severity, and *in vitro* hyperexcitability in the hippocampus and entorhinal cortex. *Neuroscience* 1999;93:1491–1506.

95. Salin P, Tseng GF, Hoffman S, et al. Axonal sprouting in layer V pyramidal neurons of chronically injured cerebral cortex. *J Neurosci* 1995;12:8234–8245.

96. Sugino T, Nozaki K, Tokime T, et al. 3-Nitropropionic acid induces poly (ADP-ribosyl)ation and apoptosis related gene expression in the striatum *in vivo. Neurosci Lett* 1997;237:121–124.

97. Buckmaster PS, Dudek FE. Neuron loss, granule cell axon reorganization, and functional changes in the dentate gyrus of epileptic kainate-treated rats. *J Comp Neurol* 1997;385:385–404.

98. Esclapez M, Hirsch JC, Ben-Ari Y, et al. Newly formed excitatory pathways provide a substrate for hyperexcitability in experimental temporal lobe epilepsy. *J Comp Neurol* 1999;408:449–460.

99. Bengzon J, Kokaia Z, Elmer E, et al. Apoptosis and proliferation of dentate gyrus neurons after single and intermittent limbic seizures. *Proc Natl Acad Sci USA* 1997;94:10432–10437.

100. Gray WP, Sundstrom LE. Kainic acid increases the proliferation of granule cell progenitors in the dentate gyrus of the adult rat. *Brain Res* 1998;790:52–59.

101. Parent JM, Yu TW, Leibowitz RT, et al. Dentate granule cell neurogenesis is increased by seizures and contributes to aberrant network reorganization in the adult rat hippocampus. *J Neurosci* 1997;17:3727–3728.

102. Scharfman HE, Goodman JH. Granule-like neurons at the hilar/CA3 border after status epilepticus and their synchrony with area CA3 pyramidal cells: functional implications of seizure-induced neurogenesis. *J Neurosci* 2000;20:6144–6158.

103. Gould E, Reeves AJ, Graziano MSA, et al. Neurogenesis in the neocortex of adult primates. *Science* 1999;286:548–552.

104. Scharfman HE. Epileptogenesis in the parahippocampal region: parallels with the dentate gyrus. In: Scharfman HE, Witter MP, Schwarcz R, eds. *The parahippocampal region: basic science and clinical implications*. New York: New York Academy of Sciences, 2000:305–327.

105. Tsien JZ, Huerta PT, Tonegawa S. The essential role of hippocampal CA1 NMDA receptor-dependent synaptic plasticity in spatial memory. *Cell* 1996;87:1327–1338.

Section XIII
Chronobiology and Sleep Disorders

David J. Kupfer

David J. Kupfer: Department Of Psychiatry, University Of Pittsburgh Medical Center, Western Psychiatric Institute And Clinic, Pittsburgh, Pennsylvania.

Chronobiology and Sleep Disorders - Introduction

In this section on chronobiology and sleep disorders, the seven chapters describe considerable advances in this field since the *Fourth Generation of Progress* was published. This group of chapters seeks to integrate our basic and clinical knowledge and to convey the high level of current excitement in sleep and biological rhythm research. All these chapters present certain cross-cutting themes providing a basic and clinical integration of both primary sleep disorders and those disorders in which sleep alterations represent important aspects of serious neuropsychiatric disorders.

In Chapter 128, on basic mechanisms, Pace-Schott and Hobson add a wealth of new detail to our knowledge of the brain structures involved in the control of sleep and waking, as well as the cellular level mechanisms that orchestrate the sleep cycle by neuromodulation. This chapter focuses on the brainstem neuromodulatory systems and the more specific organization of those systems in controlling the alternation of wake, non–rapid eye movement (NREM) sleep, and REM sleep. Interactions of diverse neuromodulatory systems operate in widespread subcortical areas to amplify or suppress REM sleep generation, as well as to facilitate onset and offset of the control of behavioral state by the pontine REM-NREM oscillator. This existence of an executive aminergic-cholinergic reciprocal interaction system controlling REM-NREM alternation in the pontine brainstem has been strongly confirmed by recent findings.

In Chapter 129, Lewy discusses circadian phase sleep and mood disorders and covers the following areas: circadian anatomy and physiology; the shifting of circadian phase using bright light; an update on seasonal affective disorders; discussion of other circadian phase disorders such as shift work and jet lag; and, finally, a section on melatonin and circadian phase disorders. In this review, the author discusses phase typing sleep and mood disorders, including both advanced and delayed types, phase shifts with both bright light and melatonin administration, and whets our appetite on the considerable activity on melatonin research. Optimal dosing of melatonin will depend on minimizing its soporific side effect while maximizing its phase-shifting effects. This may entail using a low-dose sustained-release formulation to smooth out any sharp spikes in melatonin levels that appear to cause sleepiness in some people. Another useful product that we may look forward to is a delayed-release sustained-release formulation that can be taken at bedtime to produce increases in melatonin conveniently throughout the night.

Chapter 130, by Kloss, Szuba, and Dinges, covers sleep loss and sleepiness and reviews the four major types of sleep disorder: pathophysiology of difficulty initiating or maintaining sleep; pathophysiology of disorders of excessive somnolence; neurobehavioral and physiologic effects of sleep loss; and treatment for sleep loss and sleepiness. This chapter discusses the causes, consequences, and mechanisms of sleep disruption and concomitant daytime sequelae, namely, sleepiness and neurobehavioral performance decrements. Several advances in the psychopharmacologic and behavioral treatments of the causes and consequences of sleep loss have evolved. Technologies are rapidly developing and show promise for effective evaluation of these highly prevalent problems. The authors also present recent advances in assessment and prevention technology and go well beyond Multiple Sleep Latency Test (MSLT) and pupillometer evaluations. As examples, they also discuss the sleep switch and other operator-centered fatigue monitoring technologies.

Mignot and Nishino's Chapter 131, on the pathophysiologic and pharmacologic aspects of narcolepsy, provides an exciting

set of insights to the advances in this particular area. Narcolepsy is frequently both overdiagnosed and underdiagnosed. However, the condition is not rare and has a population prevalence similar to that of multiple sclerosis. With the availability of validated animal models and as the only known disorder with a complete disorganization of sleep and REM sleep, narcolepsy is also a unique disease model for basic sleep researchers. Our understanding of the pathophysiology of the disorder is rapidly emerging as a result of the discovery that narcolepsy or cataplexy is associated with a deficiency in the hypocretin (orexin) neuropeptide system. The discovery that a deficit in hypocretin neurotransmission, as revealed by cerebrospinal fluid hypocretin studies, frequently causes human narcolepsy opens the door to new diagnostic and therapeutic strategies. Measuring hypocretin levels in the cerebrospinal fluid or other biological fluids may soon be used as a diagnostic test for narcolepsy. The finding that human narcolepsy is HLA associated also suggests a possible autoimmune mediation in many cases.

In Chapter 132, Mendelson discusses certain basic mechanisms on how hypnotics act. Thus, we are beginning to have some insight into an early issue in sleep research: how administration of sedative-hypnotic compounds from such diverse pharmacologic classes can result in sleep induction. It appears that most or all of them produce their pharmacologic effects by altering the function of various moieties of the γ-aminobutyric $acid_A$ ($GABA_A$)-benzodiazepine receptor complex. One possibility that has received little attention has been that classic hypnotics such as benzodiazepines or barbiturates may alter the ascending histaminergic arousal system, which is presumably the mechanism by which antihistamines produce sedating effects. Certainly, one area of interest would be the tuberomamillary nucleus (TMN), which lies adjacent to the mamillary bodies, just above the ventral surface of the hypothalamus. Another focus for dysregulation in insomnia may involve the ventrolateral preoptic area (VLPO) and its interactions with the TMN in the posterior hypothalamus. The GABAergic VLPO has been identified as one of the few “sleep-active” areas of the brain; dysregulation in this nucleus and its efferent projections to histaminergic, cholinergic, and noradrenergic nuclei could conceivably shift the sleep-wake balance in the direction of wakefulness. In principle, benzodiazepine or other hypnotic compounds may act by enhancing GABAergic inhibition of the TMN and thereby decreasing its arousing effects.

In Chapter 133, Buysse and Dorsey provide a superb review on experimental therapeutics of insomnia. Although considerable progress has been made with regard to the epidemiology of insomnia, further work needs to be done regarding its consequences for health and role functioning. Persons with insomnia complain not only of sleep disturbance, but of daytime consequences as well. In addition, investigations into the neurobiology of insomnia are clearly needed. This will help to define the underlying pathophysiology of insomnia in the general sense, but it will also help to define the boundaries of specific insomnia disorders. Several issues also remain with regard to treatment aspects of insomnia. First, the relative benefits and risks of treatment in terms of symptomatic relief, health-related quality of life, and morbidity remain to be defined. These issues are of considerable importance, given the potential for some insomnia treatments to cause significant adverse effects, such as cognitive impairment and injurious falls. The optimal duration of treatment and the conceptualization of potential “maintenance” treatments for insomnia are also areas open for further investigation. With regard to behavioral treatments, one of the major challenges is designing well-manualized and “exportable” treatments that can be applied more readily in a variety of treatment settings, including primary care settings. Data from several studies examining the optimal combination of behavioral and medication treatment approaches suggest better durability of treatment effects with behavioral treatment alone. However, sequential treatments and concurrent treatments need to be investigated. In addition, treatment strategies for nonresponders to either behavioral or pharmacologic interventions must be developed.

The final chapter in this section, Chapter 134, reviews our current understanding of sleep disturbances associated with neuropsychiatric disease. Nofzinger and Keshavan provide a brief review of the advances relating basic research on sleep with clinical sleep findings in major neuropsychiatric diseases such as depression, schizophrenia, Alzheimer disease, and other disorders across the life span. As one of the earlier tools available to psychiatric research for discovering the biological basis of mental disorders, EEG sleep recordings have been used extensively to characterize alterations in brain function across diverse mental disorders. Newer tools available include refinements in electrophysiologic recordings using automated EEG and the concurrent use of electrical recordings of cognitive processes such as evoked responses to characterize changes in information processing during sleep in relation to mental disorders. Advances in functional neuroimaging could provide us with dynamic images of brain function as it makes transitions throughout the sleep-wake cycle. In this manner, the functional neuroanatomic basis of the electrophysiologic abnormalities could be determined, and interventions could be designed targeting not only specific neurotransmitter systems but also systems that are specific to a discrete brain region responsible for the sleep-wake disturbance.

128

Basic Mechanisms of Sleep: New Evidence on the Neuroanatomy and Neuromodulation of the NREM-REM Cycle

Edward F. Pace-Schott

J. Allan Hobson

Edward F. Pace-Schott and J. Allan Hobson: Laboratory of Neurophysiology, Department of Psychiatry, Harvard Medical School. Boston, Massachusetts.

The 1990s brought a wealth of new detail to our knowledge of the brain structures involved in the control of sleep and waking and in the cellular level mechanisms that orchestrate the sleep cycle through neuromodulation. This chapter presents these new findings in the context of the general history of research on the brainstem neuromodulatory systems and the more specific organization of those systems in the control of the alternation of wake, non–rapid eye movement (NREM), and REM sleep.

Although the main focus of the chapter is on the our own model of reciprocal aminergic-cholinergic interaction, we review new data suggesting the involvement of many more chemically specific neuronal groups than can be accommodated by that model. We also extend our purview to the way in which the brainstem interacts with the forebrain. These considerations inform not only sleep-cycle control *per se*, but also the way that circadian and ultradian rhythms resonate to regulate human behavior including the intensity and form of conscious awareness.

RECIPROCAL INTERACTION AND ITS RECENT MODIFICATIONS

Part of "128 - Basic Mechanisms of Sleep: New Evidence on the Neuroanatomy and Neuromodulation of the NREM-REM Cycle "

Behavioral State–Dependent Variations in Neuromodulation

A paradigm shift in thinking about sleep-cycle control was forced by the discovery of the chemically specific neuromodulatory subsystems of the brainstem (for reviews, see refs. 1 ,2 ,3 and 4) and of their differential activity in waking (noradrenergic, serotonergic, and cholinergic systems on), NREM sleep (noradrenergic, serotonergic, and cholinergic systems damped), and REM sleep (noradrenergic and serotonergic systems off, cholinergic system undamped) (1 ,2 ,3 and 4).

Original Reciprocal Interaction Model: An Aminergic-Cholinergic Interplay

The model of reciprocal interaction (5) provided a theoretic framework for experimental interventions at the cellular and molecular level that has vindicated the notion that waking and REM sleep are at opposite ends of an aminergically dominant to cholinergically dominant neuromodulatory continuum, with NREM sleep holding an intermediate position (Fig. 128.1). The reciprocal interaction hypothesis (5) provided a description of the aminergic-cholinergic interplay at the synaptic level and a mathematic analysis of the dynamics of the neurobiological control system. In this section, we review ongoing recent findings of the essential roles of both acetylcholine (ACh) and the monoamines serotonin (5-HT) and norepinephrine (NE) in the control of the NREM-REM cycle as well as work that has led to the alteration (Fig. 128.2) and elaboration (Fig. 128.3) of the original reciprocal interaction model.

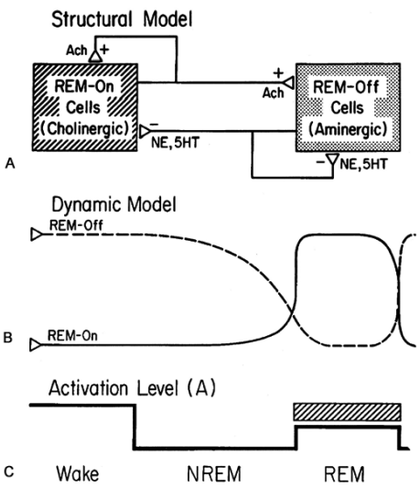


FIGURE 128.1. The original reciprocal interaction model of physiologic mechanisms determining alterations in activation level. **A:** Structural model of reciprocal interaction. REM-on cells of the pontine reticular formation are cholinoceptively excited or cholinergically excitatory (ACH+) at their synaptic endings. Pontine REM-off cells are noradrenergically (NE) or serotonergically (5-HT) inhibitory (-) at their synapses. **B:** Dynamic model. During waking, the pontine aminergic system is tonically activated and inhibits the pontine cholinergic system. During NREM sleep, aminergic inhibition gradually wanes, and cholinergic excitation reciprocally waxes. At REM sleep onset, aminergic inhibition is shut off, and cholinergic excitation reaches its high point. **C:** Activation level. As a consequence of the interplay of the neuronal systems shown in **A** and **B**, the net activation level of the brain **(A)** is at equally high levels in waking and REM sleep and at about half this peak level in NREM sleep. (From Hobson JA, Stickgold R, Pace-Schott EF. The neuropsychology of REM sleep dreaming. *Neuroreport* 1998;9:R1–R14, with permission.)

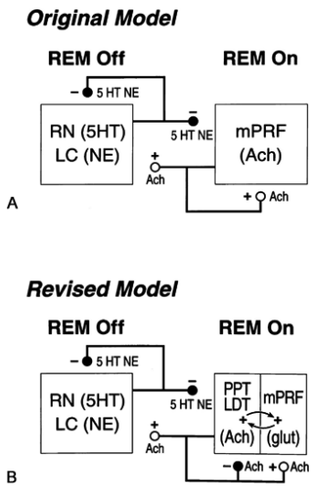


FIGURE 128.2. Synaptic modifications of the original reciprocal interaction model based on recent findings. **A:** The original model proposed by McCarley and Hobson (5). **B:** Synaptic modifications of the original reciprocal interaction model based on recent findings of self-inhibitory cholinergic autoreceptors in mesopontine cholinergic nuclei and excitatory interactions between mesopontine cholinergic and noncholinergic neurons (Fig. 128.3 and 128.4C, for more detail and references). The exponential magnification of cholinergic output predicted by the original model *(A)* can also occur in this model with mutually excitatory cholinergic-noncholinergic interactions taking the place of the previously postulated, mutually excitatory cholinergic-cholinergic interactions. In the revised model, inhibitory cholinergic autoreceptors would contribute to the inhibition of laterodorsal tegmental nucleus (LDT) and pedunculopontine tegmental nucleus (PPT) cholinergic neurons that is also caused by noradrenergic and serotonergic inputs to these nuclei. Therefore, the basic shape of reciprocal interaction's dynamic model (Fig. 128.1B) and its resultant alternation of behavioral state (Fig. 128.1C) could also result from the revised model. *Open circles*, excitatory postsynaptic potentials; *closed circles*, inhibitory postsynaptic potentials; Ach, acetylcholine; glut, glutamate; 5-HT, serotonin; LC, locus ceruleus; mPRF, medial pontine reticular formation; NE, norepinephrine; RN, dorsal raphe nucleus. (From Hobson JA, Stickgold R, Pace-Schott EF. The neuropsychology of REM sleep dreaming. *Neuroreport* 1998;9:R1–R14, with permission.)

Although there is abundant evidence for a cholinergic mechanism of REM-sleep generation centered in the pedunculopontine (PPT) and laterodorsal tegmental (LDT) nuclei of the mesopontine tegmentum (for reviews, see refs. 2 ,3 and 4 and 6 ,7 and 8), not all PPT-LDT neurons are cholinergic (9 ,10 ,11 and 12), and cortical ACh release may be as high during wakefulness as during sleep (13).

Recently, reciprocal interaction (5) and reciprocal inhibition (14) models for control of the REM-NREM sleep cycle by brainstem cholinergic and aminergic neurons have been questioned (10). Specifically, the hypothesized self-stimulatory role of ACh on those mesopontine neurons associated with the characteristic pontogeniculooccipital (PGO) waves

of REM sleep has not been confirmed in *in vitro* slice preparations of the rodent brainstem (10). For example, ACh has been shown to hyperpolarize cell membranes in slice preparations of the rodent parabrachial nucleus (15), LDT (16), and PPT (10). Similarly, those LDT-PPT neurons with burst discharge properties most like those hypothesized to occur in PGO-burst neurons (“type I” neurons) may not be cholinergic (9). Much evidence remains, however, that the reciprocal interaction model accurately describes essential elements of REM-NREM sleep-cycle control even though a few assumptions in its detailed synaptic mechanisms and connectivity are heuristic (see Fig. 128.2 and Fig. 128.3).

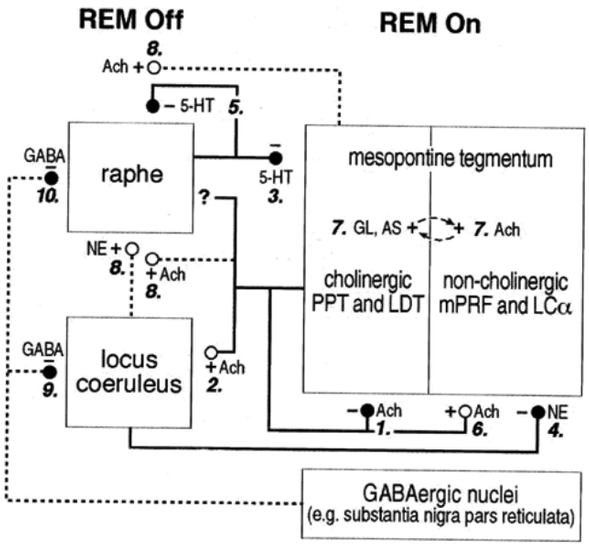


FIGURE 128.3. Additional synaptic details of the revised reciprocal interaction model shown in Fig. 128.2B derived from data reported *(solid lines)* and hypothesized relationships suggested *(dotted lines)* in recent experimental studies (numbered on figure and below). See the text for a discussion of these findings. Additional synaptic details can be superimposed on the revised reciprocal interaction model without altering the basic effects of aminergic and cholinergic influences on the REM sleep cycle. Excitatory cholinergic-noncholinergic interactions using acetylcholine (Ach) and the excitatory amino acid transmitters enhance firing of REM-on cells *(6* and *7)*, whereas inhibitory noradrenergic (4), serotonergic (3), and autoreceptor cholinergic (1) interactions suppress REM-on cells. Cholinergic effects on aminergic neurons are excitatory (2), as hypothesized in the original reciprocal interaction model, and they may also operate through presynaptic influences on noradrenergic-serotonergic as well as serotonergic-serotonergic circuits (8). GABAergic influences *(9* and *10)*, as well as other neurotransmitters such as adenosine and nitric oxide (see text), may contribute to the modulation of these interactions. *Open circles*, excitatory postsynaptic potentials; *closed circles*, inhibitory postsynaptic potentials; AS, aspartate; glut, glutamate; 5-HT, serotonin; LCa, peri-locus ceruleus α; LDT, laterodorsal tegmental nucleus; mPRF, medial pontine reticular formation; NE, norepinephrine; PPT, pedunculopontine tegmental nucleus. (For the specific references corresponding to the interactions numbered *1* to *10*, please refer to ref. 2.)

New Findings Supporting the Cholinergic Enhancement of REM Sleep

Numerous findings confirm the hypothesis that cholinergic mechanisms are essential to the generation of REM sleep and its physiologic signs (for reviews, see refs. 1 ,2 ,3 and 4 , 6 , 7 and 8 , 11 ,12 and 13 , 14 , 17 , and 18).

A selection of the many recent examples follows.

Experimental REM Sleep Induction and Suppression

Microinjection of cholinergic agonist (e.g., carbachol) or cholinesterase inhibitors into many areas of the paramedian pontine reticular formation (PRF) of the cat induces REM sleep (for reviews, see refs. 6 ,7 and 8). The endogenous ACh released into these areas originates in neurons of the mesopontine tegmentum (6). In addition to these short-term REM induction sites, carbachol injection into a more lateral pontine site in the caudal peribrachial area has been shown to induce long-term (more than 7 days) REM enhancement (19) and long-term PGO enhancement but without REM enhancement (20). *In vivo* cholinergic REM enhancement and a specific carbachol-sensitive site in the dorsal locus subceruleus of rats have been described (21).

In addition to the well-known suppression of REM by muscarinic antagonists (1), presynaptic anticholinergic agents have also been shown to block REM (22). Activation of muscarinic M2 receptors (M2AChR) in the pontine reticular formation has been shown to be the primary mechanism for REM induction with carbachol, and such activation has been shown to increase G-protein binding in brainstem nuclei associated with ACh release in the PRF and REM sleep (7, 23).

Cholinergic Neurons and REM Sleep

As reviewed by Semba (6), studies have demonstrated a physiologically meaningful heterogeneity among mesopontine neurons both *in vivo* and *in vitro*. In behaving cats and rats, separate populations of REM-and-wake-on, REM-on, and PGO wave–associated neurons can be identified in the LDT and PPT with strong evidence that a subset of these are cholinergic (6). In rodent brainstem slices, mesopontine neurons can be divided according to their membrane current characteristics into types I, II and III, with types II and III being cholinergic and projecting to the thalamus (6).

Many recent experimental findings associate REM sleep generation with mesopontine cholinergic neurons. For example, although type I bursting neurons are noncholinergic, cholinergic type II and III PPT-LDT neurons have firing properties that make them well suited for the *tonic* maintenance of REM (9). Three supportive experimental studies by Robert McCarley's group at Harvard (reviewed in ref. 2) are as follows: (a) PPT-LDT neurons specifically show immediate-early gene (e.g., *c-fos*) immunoreactivity after carbachol-induced REM sleep; (b) low-amplitude electrical stimulation of the LDT enhances subsequent REM sleep; and (c) electrical stimulation of the cholinergic LDT evokes excitatory postsynaptic potentials in PRF neurons that can be blocked by scopolamine. Finally, the excitatory amino acid, glutamate, when microinjected into the cholinergic PPT, increases REM sleep in a dose-dependent manner (24).

Acetylcholine Release and REM Sleep

Microdialysis studies show enhanced release of endogenous ACh in the medial PRF during natural REM sleep (25). Moreover, stimulation of the PPT causes increased ACh release in the PRF (26). Thalamic ACh concentration of mesopontine origin is higher in both wake and REM than

in NREM (27), and a REM-specific increase of ACh in the lateral geniculate body (LGB) has also been observed (28). Both muscarinic and nicotinic receptors participate in the depolarization of thalamic nuclei by the cholinergic brainstem (29).

Cholinergic Mediation of Specific REM Signs: PGO Waves, Muscle Atonia, Cortical Desynchronization, and Hippocampal Theta

PGO input to the LGB of the thalamus is cholinergic (12), and it can be antidromically traced to pontine PGO-burst neurons (30). In turn, stimulation of mesopontine neurons induces depolarization of cortically projecting thalamic neurons (29). Notably, retrograde tracers injected into the thalamus label 50% or more of cholinergic PPT-LDT neurons (31). Neurotoxic lesions of pontomesencephalic cholinergic neurons reduce the rate of PGO spiking (32), and PGO waves can be blocked by cholinergic antagonists (8). A long history of microinjection studies has shown that, at the level of the pons, cholinergic mechanisms at a variety of sites participate in the suppression of muscle tone accompanying REM (for review, see ref. 33). In addition to brainstem-mediated cholinergic mediation of PGO waves and atonia, cholinergic basal forebrain (BF) nuclei control other distinctive signs of REM including cortical desynchronization and hippocampal theta (see the section on the BF). It may therefore not be an exaggeration to state that the evidence of cholinergic REM sleep generation is now so overwhelming and so well accepted that this tenet of the reciprocal interaction model is an established principle.

New Findings Supporting the Serotonergic and Noradrenergic Suppression of REM Sleep

Aminergic Inhibition of the Cholinergic REM Generator

At the heart of the reciprocal interaction concept is the idea that cholinergic REM sleep generation can only occur when the noradrenergic and serotonergic mediators of waking release their inhibitory constraint of the cholinergic REM generator. The evidence for such inhibitory serotonergic and noradrenergic influences on cholinergic neurons and REM sleep is also now quite strong. (For reviews, see refs. 18 , 34 , and 35 for 5-HT and ref. 36 for NE.)

Serotonin in Natural REM Sleep

Serotonergic neurons from the dorsal raphe (DR) have been shown to synapse on LDT-PPT neurons (37). Extracellular levels of 5-HT are higher in waking than in NREM and higher in NREM than REM in the brainstem and cortex of rats (38) and the DR (39) and medial PRF (40) of cats. Moreover, reduced extracellular 5-HT concentration in REM sleep has been demonstrated in the human amygdala, hippocampus, orbitofrontal cortex, and cingulate cortex (41).

There is also strong evidence that specific physiologic signs of REM sleep are inhibited by endogenous 5-HT (34). For example, in sleeping cats, the firing of DR neurons is inversely correlated with the occurrence of PGO waves (34). Similarly, hippocampal theta activity, another specific sign of REM sleep, is suppressed by serotonergic activity of the median raphe nucleus (42).

Experimental Serotonergic Suppression of Cholinergic Systems and REM Sleep

Numerous experimental findings have shown that 5-HT and its agonists inhibit mesopontine cholinergic cells as well as REM sleep itself. For example, 5-HT has been shown both to hyperpolarize rat cholinergic LDT cells *in vitro* (10) and to reduce REM sleep percentage *in vivo* (43). Experimentally administered 5-HT has also been shown to suppress specific physiologic signs of REM. For example, 5-HT has been shown to counteract the REM-like carbachol-induced atonia of hypoglossal motor neurons (44).

Microinjection of the 5-HT agonist 8-OH-DPAT into the peribrachial region impedes PGO waves and REM sleep initiation in cats (45). Simultaneous unit recording has shown that microinjection of 8-OH-DPAT selectively suppressed the firing of REM-on but not REM-and-wake-on cells of the cholinergic LDT-PPT (46). *In-vivo* microdialysis of 5-HT agonists into the dorsal raphe nucleus (DRN) decreased DRN levels of serotonin (presumably by 5-HT autoreceptors on DRN cells) which, in turn, increased REM sleep percentage (47). Mesopontine injection of a 5-HT agonist depressed ACh release in the lateral genicolate body (28).

Such findings conclusively show brainstem involvement in the serotonergic suppression of REM sleep. However, localization of this effect solely to the brainstem has been challenged in favor of an amygdala-pontine interaction (48).

Suppression of REM by Endogenous Norepinephrine and Its Agonists

Much recent evidence also implicates NE in the inhibitory control of REM sleep. For example, locus ceruleus (LC) neurons have been shown to become quiescent during REM in the monkey (49), as well as in the cat and rat (1). Electrical stimulation of the pons in the vicinity of the (noradrenergic) LC reduced REM sleep in rats (50), and the noradrenergic antagonist idazoxan increases REM when injected into the PRF of cats (51).

Combined Effects of Serotonin and Norepinephrine on REM Sleep

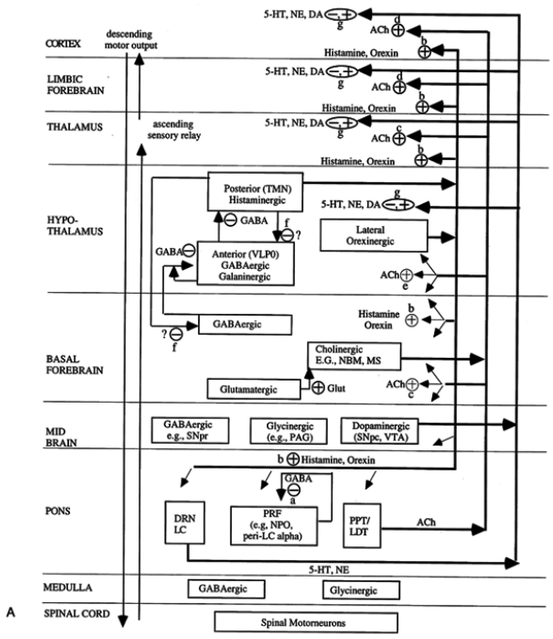
The REM suppressive effects of 5-HT and NE are likely to be additive. This is suggested by the finding that reuptake

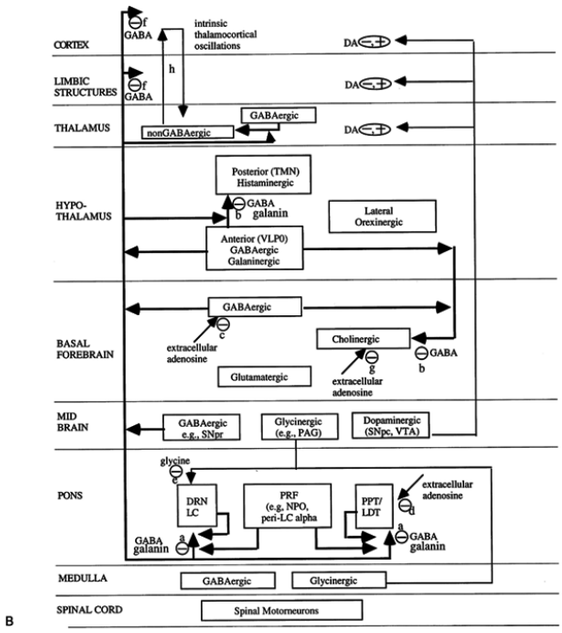
inhibitors targeting primarily either 5-HT or NE transporters all suppress REM sleep in humans (52). Unlike the other brainstem monoamines, the REM sleep effects of dopamine (DA) are more complex (see later).

Therefore, like cholinergic enhancement, aminergic suppression of REM sleep is now an established principle. The 5-HT_{1A} receptor may be of the greatest importance in the inhibition of cholinergic firing in the cat PPT (45) and LDT (53), and mesopontine postsynaptic 5-HT_{1A} receptors may be the active site for serotonergic inhibition of REM (35). Although 5-HT_2 receptors may also be involved in modulating the REM-NREM cycle, their roles are unclear because both 5-HT_2 agonists and 5-HT_2 antagonists suppress REM, whereas 5-HT_2 agonists suppress but 5-HT_2 antagonists increase slow-wave sleep (SWS) (35). Both α_1 (54) and α_2 receptors (55) may be sites of adrenergic REM suppression.

Modification of the Original Reciprocal Interaction Hypothesis to Accommodate New Findings

Modifications of simple reciprocal inhibition or interaction models, which are consonant with recent findings, have been proposed for the brainstem control of REM sleep. For example, Leonard and Llinas suggested in regard to the McCarley and Hobson (5) model that "… 'indirect feedback' excitation via cholinergic inhibition of an inhibitory input or cholinergic excitation of an excitatory input or some combination of the two could replace direct feedback excitation in their model" (10). Mutually excitatory or mutually inhibitory interactions between REM-on cholinergic and REM-on noncholinergic mesopontine neurons have also been proposed in the cat (11). Similarly, Semba suggested that naturally occurring REM sleep is instigated when cholinergic LDT-PPT neurons increase their cholinergic stimulation of PRF networks known to be associated with carbachol-induced REM (6). In turn, Semba suggested that PRF neurons may provide a glutamatergic, excitatory feedback to cholinergic neurons in the LDT-PPT thereby maintaining REM sleep (6). Representative hypothetical cholinergic-noncholinergic mechanisms are illustrated in Fig. 128.2B , Fig. 128.3 , and Fig. 128.4C-a . [**Please note** that, with regard to Fig. 128.4A–C , neuronal interactions will be identified in the subsequent main body of the text with *lower-case letters* (e.g., 128.4C-a) that also designate the corresponding excitatory or inhibitory synaptic interaction in the figure itself as well as designating this same interaction in the corresponding figure legend. The illustration of a particular synaptic interaction in a schematic of a particular behavioral state (i.e., Fig. 128.4A -wake, Fig. 128.4B -NREM sleep, and Fig. 128.4C -REM sleep), is not meant to imply that this interaction only occurs in this behavioral state. For example, many of the subcortical excitatory interactions associated with ascending forebrain arousal are shared by both waking and REM sleep (6).]


CORTEX
descending motor output
5-HT, NE, DA
ACh
Histamine, Orexin
LIMBIC FOREBRAIN
THALAMUS
ascending sensory relay
HYPO-THALAMUS
Posterior (TMN) Histaminergic
GABA
Lateral Orexinergic
Anterior (VLP0) GABAergic Galaninergic
GABAergic
Histamine Orexin
BASAL FOREBRAIN
Cholinergic E.G., NBM, MS
Glutamatergic
Glut
MID BRAIN
GABAergic e.g., SNpr
Glycinergic (e.g., PAG)
Dopaminergic (SNpc, VTA)
PONS
DRN LC
PRF (e.g, NPO, peri-LC alpha)
PPT/ LDT
MEDULLA
GABAergic
Glycinergic
SPINAL CORD
Spinal Motorneurons
A


CORTEX
intrinsic thalamocortical oscillations
DA
LIMBIC STRUCTURES
GABA
THALAMUS
GABAergic
nonGABAergic
HYPO-THALAMUS
Posterior (TMN) Histaminergic
GABA galanin
Lateral Orexinergic
Anterior (VLP0) GABAergic Galaninergic
BASAL FOREBRAIN
extracellular adenosine
Cholinergic
Glutamatergic
MID BRAIN
GABAergic e.g., SNpr
Glycinergic (e.g., PAG)
Dopaminergic (SNpc, VTA)
PONS
glycine
DRN LC
PRF (e.g, NPO, peri-LC alpha
PPT/ LDT
MEDULLA
SPINAL CORD
Spinal Motorneurons
B

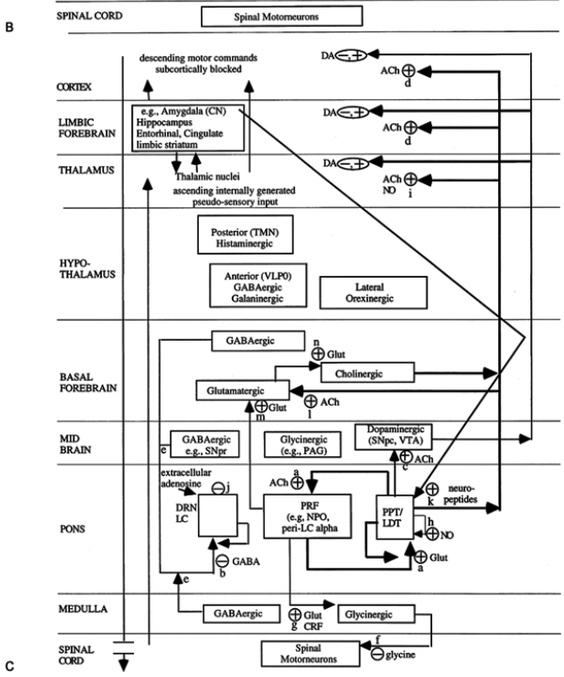

descending motor commands subcortically blocked
CORTEX
LIMBIC FOREBRAIN
e.g., Amygdala (CN) Hippocampus Entorhinal, Cingulate limbic striatum
THALAMUS
Thalamic nuclei
ascending internally generated pseudo-sensory input
DA
ACh
NO
HYPO-THALAMUS
Posterior (TMN) Histaminergic
Anterior (VLP0) GABAergic Galaninergic
Lateral Orexinergic
BASAL FOREBRAIN
GABAergic
Glut
Cholinergic
Glutamatergic
MID BRAIN
GABAergic e.g., SNpr
Glycinergic (e.g., PAG)
Dopaminergic (SNpc, VTA)
PONS
extracellular adenosine
DRN LC
PRF (e.g, NPO, peri-LC alpha
PPT/ LDT
neuro-peptides
GABA
MEDULLA
GABAergic
Glycinergic
CRF
SPINAL CORD
Spinal Motorneurons
glycine
C

FIGURE 128.4. Critical neuromodulatory systems for the initiation and maintenance of the behavioral states of wake **(A)**, NREM sleep **(B)**, and REM sleep **(C)**. Illustrated circuits are those hypothesized to play executive roles in state initiation or maintenance or to mediate cardinal physiologic signs of that state (e.g., REM atonia). The major defining neuromodulatory features of each state are described, and the most important of these are depicted with the *heaviest lines* in diagrams. The exclusion of particular circuits in individual state diagrams does not imply that circuit's inactivity during that state, and illustration of a particular synaptic interaction in a particular behavioral state is not meant to imply that this interaction only occurs in this behavioral state. For example, much of the sleep-associated GABAergic neuronal inhibition illustrated in NREM **(B)** is probably maintained in REM **(C)**. Similarly, many of the subcortical excitatory interactions associated with ascending forebrain arousal are shared by both waking and REM sleep (6). Potentially influential peptidergic neuromodulation (e.g., VIP), and behavioral state–related changes in basal ganglia activity (see text) are left out for the sake of clarity. Finally, neuromodulatory changes occurring with the alternations of substages within NREM sleep are not illustrated. Details of neuronal interactions are provided in the text and are cross-referenced using lower-case letters appearing adjacent to the excitatory or inhibitory synaptic interaction illustrated in **A–C**. These neuronal interactions are also summarized at the end of each sublegend with a representative citation.

A: Wake: diverse ascending activation. During the wake state, the full complement of ascending arousal systems classified by Saper et al. (76) actively modulates the forebrain with the chemical products of their respective brainstem and diencephalic nuclei *(heavy lines in diagram)*, including: *serotonin* (5-HT) from the dorsal raphe (DRN) nucleus of the pons innervating the entire forebrain; *norepinephrine* (NE) from the locus ceruleus (LC) of the pons innervating the entire forebrain; dopamine (DA) from the ventral tegmental area (VTA) and substantia nigra pars compacta (SNpc) of the midbrain innervating the entire forebrain; *acetylcholine* (ACh) from the mesopontine laterodorsal tegmental and pedunculopontine nuclei (LDT-PPT), which chiefly modulates the diencephalon; and ACh from the magnocellular cholinergic cells of the basal forebrain that innervate limbic forebrain (e.g., medial septal area or MS to hippocampus) and the cortex (nucleus basalis of Meynert or NBM); *histamine* from the posterior hypothalamus, especially the tuberomamillary nucleus (TMN), which both promotes arousal of the entire forebrain and facilitates other ascending arousal systems in the brainstem; and *orexin* from the lateral hypothalamus that promotes arousal of the forebrain and brainstem arousal systems in a manner similar to histamine. *Selected neuronal interactions: a*, GABAergic inhibition promoting wakefulness versus REM (77); *b*, histaminergic and orexinergic activation of forebrain and brainstem (83,86); *c*, cholinergic activation of diencephalon (see also e) by mesopontine nuclei (6); *d*, cholinergic activation of limbic forebrain and cortex by basal forebrain nuclei (69); *e*, cholinergic innervation of hypothalamic structures by mesopontine nuclei (6); *f*, reciprocal inhibition by wake-promoting posterior hypothalamus of sleep-promoting anterior hypothalamic and basal forebrain nuclei (83); *g*, serotonergic, noradrenergic, and dopaminergic arousal of diencephalon, limbic forebrain, and neocortex by aminergic brainstem nuclei by both inhibitory and excitatory synapses; and *h*, GABAergic inhibition of sleep-promoting anterior hypothalamic ventrolateral preoptic area (VLPO) cells by basal forebrain and other anterior hypothalamic cells (83).

B: NREM sleep: widespread GABAergic inhibition. The attenuation of wake-associated arousal of the brainstem, diencephalon, limbic forebrain, and neocortex by ascending and descending cholinergic, noradrenergic, serotonergic, histaminergic, and orexinergic systems allows inhibitory GABAergic systems (heavy lines) to become prominent at various levels of the neuraxis. Maintained dopaminergic activation of the forebrain (61,62) is insufficient to maintain arousal sufficient to support wakefulness or REM sleep. At the thalamocortical level, such inhibition allows emergence of intrinsic oscillatory rhythms characterized in the EEG by sleep spindles, delta waves, and very slow (less than 1 Hz) oscillations (108). *Selected neuronal interactions: a*, GABAergic and galaninergic inhibition by the anterior hypothalamus (e.g., VLPO) of brainstem aminergic and cholinergic ascending arousal systems (195); *b*, GABAergic and galaninergic inhibition by the anterior hypothalamus (e.g., VLPO) of diencephalic aminergic (especially the histaminergic TMN) and cholinergic ascending arousal systems (83); *c*, high levels of extracellular adenosine accumulated during waking trigger sleep onset by inhibiting specific GABAergic basal forebrain cells that have, in turn, been inhibiting sleep-associated VLPO neurons during waking (83); *d*, extracellular adenosine accumulated during waking inhibits mesopontine cholinergic ascending activating systems (18); *e*, glycinergic midbrain and medullary cells exert tonic inhibition over pontine aminergic ascending activating systems (56); *f*, GABAergic inhibition of limbic and neocortical forebrain (56); *g*, extracellular adenosine accumulated during waking inhibits basal forebrain cholinergic ascending activating systems (18); and *h*, aminergic demodulation and GABAergic inhibition allows emergence of intrinsic thalamocortical oscillations (108).

C: REM sleep: selective cholinergic activation: During REM sleep, cholinergic activation of the forebrain from the brainstem (mesopontine) and basal forebrain ascend ing cholinergic systems is reinstated *(heavy lines)* in the absence of the aminergic (5-HT, NE, histamine) and orexinergic activation present during waking. Such selective cholinergic activation favors the distinctive forebrain-brainstem interactions of REM sleep in which internally generated ascending pseudosensory signals (e.g., pontogeniculooccipital or PGO waves) impinge on thalamocortical relay circuits and descending cortical and subcortical motor commands are blocked (3). Specific brainstem and basal forebrain circuits promoting the distinctive physiological features of REM sleep (cholinergic forebrain activation, PGO waves, skeletal muscle atonia, and rapid eye movements) are detailed in the text and are summarized here. *Selected neuronal interactions: a*, exponentially increasing activity of mesopontine PPT-LDT cholinergic neurons results in part from positive feedback *(heavy lines)* involving cholinergic excitation of pontine reticular formation (PRF) cells by the PPT-LDT and reciprocal glutamatergic excitation of the PPT-LDT by the PRF (6); *b*, GABAergic inhibition of pontine aminergic nuclei withdraws their inhibitory influence on PPT-LDT cholinergic cells (56) (see also *e*); *c*, cholinergic stimulation from the PPT-LDT enhances activity of midbrain dopaminergic cells (31), which, in turn, enhance cortical release of ACh (68); *d*, cholinergic activation of the limbic and neocortical forebrain by the basal forebrain occurs (69); *e*, additional inhibition of pontine aminergic (REM-off) nuclei by midbrain and medullary GABAergic neurons occurs (56); *f*, inhibition of spinal motor neurons by medullary glycinergic neurons is primary source of REM atonia (33); *g*, medullary glycinergic neurons are, in turn, activated by pontine glutamatergic and corticotropin-releasing factor (CRF)–ergic neurons (33); *h*, nitric oxide (NO) co-released from cholinergic cells facilitates self-activation of LDT-PPT (80); *i*, cholinergic activation of the thalamus by the LDT-PPT is facilitated by co-released NO (81); *j*, extracellular adenosine accumulated during waking provides additional inhibition of pontine aminergic (REM-off) nuclei (18); *k*, descending signals from the limbic forebrain may contribute to the initiation of REM by neuropeptidergic efferent projections to the mesopontine tegmentum from central nucleus (CN) of the amygdala (104); *l*, the LDT-PPT cholinergically stimulates basal forebrain glutamatergic cells, which, in turn, excite cholinergic basal forebrain neurons projecting to the limbic and neocortical forebrain (6); *m*, PRF glutamatergic neurons excite basal forebrain glutamatergic cells, which, in turn, excite cholinergic basal forebrain neurons projecting to the limbic and neocortical forebrain (6); and *n*, intrinsic basal forebrain glutamatergic cells excite cholinergic basal forebrain neurons projecting to the limbic and neocortical forebrain (69). SNpr, substantia nigra pars reticulata.

Other suggested modifications have invoked the contribution of inhibitory neuromodulators such as γ-aminobutyric acid (GABA) in the control of the REM-NREM cycle (for review, see ref. 56). For example, from *in vivo* microdialysis studies of GABA in the cat, Nitz and Siegel (57) suggested that GABA suppresses noradrenergic LC neurons during REM (Fig. 128.3 and Fig. 128.4C-b). Similarly, other investigators have suggested that GABAergic inhibition is responsible for the quiescence of serotonergic DR neurons during REM (34). The role of GABA in the suppression of aminergic neurons during REM sleep is strengthened by findings that natural concentrations of 5-HT in the DR during REM are decreased, thereby arguing against serotonergic DR self-inhibition by 5-HT_{1A} somatodendritic autoreceptors (35).

Still other modifications have caused investigators to postulate a role for presynaptic autoreceptors and heteroreceptors. For example, *in vitro* studies in the rat suggested the following modification of reciprocal interaction (58) (Fig. 128.3). During waking, presynaptic nicotinic facilitation of excitatory LC noradrenergic inputs to the DR enhances serotonergic firing. During REM, when the LC is silent, this same presynaptic nicotinic input may facilitate serotonergic self-inhibition by raphe neurons themselves. Notably, all such modifications retain one or both of the major tenets of the reciprocal interaction model: cholinergic facilitation and adrenergic inhibition of REM.

Many of the studies questioning reciprocal interaction or reciprocal inhibition (9 ,10 ,15 ,16) have been carried out on *in vitro* rodent models, and the relation of these findings to findings on the *in vivo* generation of REM sleep signs in the cat is only in its early stages (11). Moreover, the hyperpolarization by ACh of cholinergic cells (9 ,10 and 11 ,16 ,28 ,59) (Fig. 128.3) may result from recently identified ACh M2 autoreceptors that contribute to the homeostatic control of cholinergic activity (7 ,59).

In contrast to the hyperpolarization of some mesopontine cholinergic neurons by cholinergic agonists, many medial PRF neurons are depolarized by carbachol (6). This finding suggests that the exponential self-stimulatory activation that can be triggered by cholinergic stimulation in diverse mesopontine and medial pontine sites (1 ,2 ,3 and 4) may involve noncholinergic excitatory intermediary neurons. Such cholinergic self-regulation, combined with cholinergic-noncholinergic mutual excitation, is illustrated in Fig. 128.2B , Fig. 128.3 , and Fig. 128.4C-a .

Conclusions Regarding the Current Status of Reciprocal Interaction

We conclude that the two central ideas of the reciprocal interaction model are strongly supported by subsequent research: (a) noradrenergic and serotonergic influences enhance waking and impede REM by anticholinergic mechanisms; and (b) cholinergic mechanisms are essential to REM sleep and come into full play only when the serotonergic and

nbnoradrenergic systems are inhibited. Because many different synaptic mechanisms could mediate these effects, we now turn our attention to some intriguing possibilities.

OTHER NEUROTRANSMITTER SYSTEMS

Beyond the originally proposed cholinergic and aminergic neuronal populations, many additional neurotransmitter systems may participate in the neuromodulation of REM and NREM sleep. Since 1975, significant progress has been made in the identification of other chemically specific neuromodulatory systems showing differential activation with particular behavioral states or with specific physiologic signs within a behavioral state. These other neuromodulatory systems may interact with aminergic and cholinergic systems in the generation of REM sleep and its signs. We now discuss these new findings and the ways that they modify and extend the reciprocal interaction model, first in terms of the chemically specific systems and then in terms of the neuroanatomic networks subserving this physiology. In the following sections, we also address current findings on the diencephalic neuromodulation of NREM sleep and the wake-sleep transition.

Dopaminergic Systems

Given the key role of the other monoamines in the control of behavioral state and the powerful alerting effects of dopaminergic drugs, the potential role of DA in the control of the sleep-wake and the REM-NREM cycle has also been examined (see ref. 60 for review). DA release does not vary dramatically in phase with the natural sleep cycle as do 5-HT, NE, and ACh (61 ,62). However, REM sleep deprivation appears to enhance DA levels and DA receptor sensitivities (63).

Experimental manipulation of dopaminergic systems also gives varying results. For example, a DA agonist reduced REM sleep at low doses but enhanced it at higher ones (64), whereas a DA reuptake inhibitor had the opposite effect (65). In addition, although many human studies report REM suppression by DA reuptake inhibitors and indirect agonists (52 ,66), a DA-enhancing agent, bupropion, has been shown to human enhance REM sleep (67). Moreover, studies on the administration of dopaminergic drugs have suggested that DA may play a role in the induction or intensification of nightmares (60). Therefore, the effects of DA on sleep appear to be variable and are in much need of further study.

As is the case with many other neuromodulators, the sleep effects of DA may be mediated by dopaminergic effects on the aminergic and cholinergic systems involved in the executive control of the REM-NREM cycle. For example, DA has been shown to enhance cortical ACh release (68), whereas cholinergic mesopontine neurons have been shown to enhance mesolimbic DA release (31) (Fig. 128.4C-c). Such mutual facilitation between cholinergic and dopaminergic systems may serve to maintain or intensify REM sleep, especially given DA neurons' continued activity during REM (61 ,62) (Fig. 128.4C-c).

GABAergic Systems

Numerous findings have implicated GABA, the most ubiquitous central nervous system inhibitory neurotransmitter,

in the control of the sleep-wake and the REM-NREM cycle (8 ,17 ,56 ,57 ,69 ,70) (for reviews, see refs. 56 and 71). GABAergic inhibition has been hypothesized to play both REM-facilitory and REM-inhibitory roles in its mediation of executive circuits controlling the REM-NREM cycle. In addition, during NREM sleep, GABA has been hypothesized to play key roles in the deactivation of wake-related arousal systems and in the generation of intrinsic thalamocortical oscillations such as the slow oscillations of NREM sleep (see the later section on the thalamus).

REM-facilitory roles of GABA include inhibition of those aminergic neurons, which, in turn, exert tonic suppression of pontine cholinergic REM-generation networks. For example, during REM, GABA may suppresses noradrenergic LC (57) and serotonergic DR neurons (34 ,72) (Fig. 128.4C-b). In addition, iontophoretic injection of GABA antagonists into the LC induced increased LC neuronal activity in all behavioral states, with an especially dramatic rise seen during REM (56). GABA may regulate specific elements of REM activation such as PGO wave activity. For example, in the initial stages of PGO wave generation, GABAergic and glycinergic cells may inhibit aminergic cells and thus may release the cholinergic PGO-triggering or transmitting cells (8 ,17 ,56 ,57 ,71 ,72) (Fig. 128.4C-b).

The REM-inhibitory roles of GABA may include direct inhibition of these same pontine cholinergic REM-generation networks. More specifically, GABAergic afferents to the PPT and LDT originating in the substantia nigra pars reticulata or in GABAergic neurons of the mesopontine tegmentum itself may exert direct inhibitory influences on PGO-related cells of these nuclei (3 ,8 ,17) (Fig. 128.4B-a). For example, the spike-bursting pattern in pontine PGO-burst cells may be the result of excitatory signals impinging on cells that are tonically inhibited by GABA (70). Such excitatory signals may include corollary discharge from ocular premotor neurons commanding REMs (70).

GABAergic inhibition is also important in sleep-promoting dysfacilitation of the brainstem and diencephalic arousal networks that maintain wakefulness and prevent NREM sleep. For example, GABAergic cells of the BF may inhibit hypothalamic and brainstem arousal systems (73) (Fig. 128.4B-b), whereas GABAergic input from a variety of brainstem nuclei may be responsible for decreased activity of the LC and DR during SWS (56) (Fig. 128.4B-a). Similarly, sleep-active GABAergic cells of the VLPO of anterior hypothalamus may maintain or initiate sleep by inhibiting histaminergic arousal networks in the posterior hypothalamus (74 ,75 and 76) (Fig. 128.4B-b) as well as in the LC and DR (56) (Fig. 128.4B-a). Luppi et al. (56) proposed that although VLPO GABAergic neurons inhibit the LC and DR during SWS, a second population of GABAergic cells in midbrain and medulla are recruited to accomplish more complete LC-DR inhibition of REM (Fig. 128.4C-e).

In contrast to the foregoing sleep-promoting effects of GABAergic transmission in the diencephalon, it has been reported that a specific GABAergic mechanism in the nucleus pontis oralis of the cat PRF promotes wakefulness versus REM sleep (77) (Fig. 128.4A-a). Such opposing sleep effects highlight the regional heterogeneity of roles for GABAergic inhibition in modulating behavioral states.

Glycinergic Systems

Glycine, another inhibitory neurotransmitter, has also been shown to influence the neural mechanisms underlying the sleep-wake and REM-NREM cycles (8 ,56 ,78). As in the case of GABA, glycine may provide tonic inhibition of LC and DR neurons during all behavioral states; however, increased GABAergic inhibition of the LC may be more important in progressive deactivation during NREM (56). Like GABA, glycine may regulate specific physiologic manifestations of REM. For example, medullary glycinergic cells are responsible for the postsynaptic inhibition of somatic motor neurons during REM atonia (33 ,79) (Fig. 128.4C-f), and glycinergic inhibition may play a regulatory role in the premotor functions of the pons (79).

Glutamatergic Systems

Glutamate, the most ubiquitous central nervous system excitatory neurotransmitter, has also been shown to influence the sleep-wake and the REM-NREM cycles (3 ,11 ,24 ,33). Semba (6) summarized evidence for excitatory glutamatergic input to the mesopontine tegmentum from the medial prefrontal cortex, from co-release by mesopontine cholinergic neurons themselves, from the PRF, and from the subthalamic nucleus (Fig. 128.4C-a). As described earlier in the section on modification of the original reciprocal interaction hypothesis, such glutamatergic excitation (Fig. 128.2B , Fig. 128.3 , and Fig. 128.4C-a) may widely interact with cholinergic and cholinoceptive neurons to generate the exponential increase of mesopontine and pontine reticular activity associated with REM sleep activation (2 ,6 ,11). Pontine glutamatergic cells may transmit REM sleep atonia-related signals to medullary sites (Fig. 128.4C-g), where they may then activate inhibitory glycinergic as well as GABAergic cells, which, in turn, suppress somatic motor neurons (3 ,33 ,79).

Nitroxergic Systems

The diffusible gaseous transmitter nitric oxide (NO) has been widely implicated in sleep-cycle modulation (for review, see ref. 80). NO is hypothesized to function primarily as an intercellular messenger capable of diffusing into and producing a wide variety of physiologic effects on neighboring cells including enhanced synaptic transmission through enhanced release of neurotransmitter (80). One such neurotransmitter

is ACh (80). NO is co-produced by all cholinergic neurons of the LDT and PPT (80), and inhibition of the NO synthesizing enzyme in the PRF both decreased ACh release and attenuated both natural and ACh-agonist–induced REM sleep (80). Leonard and Lydic (80) suggested that pontine NO thereby serves an important role in maintaining the cholinergically mediated REM sleep state (Fig. 128.4C-h).

Both natural and experimentally increased activity of mesopontine cholinergic neurons results in increased NO release in the thalamus (81) (Fig. 128.4C-i). NO has been shown to enhance capillary vasodilation (80), and, given its co-release with ACh, this has important implications for cholinergically mediated changes in regional blood flow during REM (see later).

Histaminergic Systems

The histaminergic arousal system, originating in neurons of the posterior hypothalamus, is fully discussed in terms of its role in the hypothalamic mediation of the sleep-wake cycle in the following neuroanatomic section. In brief, projections of this system innervate the entire forebrain (Fig. 128.4A-b) and have reciprocal projections to brainstem regions known to be involved in behavioral state control such as the mesopontine tegmentum (82). Histamine is a wake-promoting substance, and experimental lesions of histaminergic nuclei in the posterior hypothalamus such as the tuberomamilary nucleus (TMN) result in hypersomnolence (83). Moreover, microinjection of histamine or histamine agonist into the mesopontine tegmentum results in an increase in waking (82). In turn, there exists strong evidence that, during sleep, wake-active posterior hypothalamic histaminergic neurons are themselves tonically inhibited by GABAergic and galaninergic projections from the anterior hypothalamus and adjacent BF (75 ,76 ,83) (for reviews on histaminergic influences in sleep, see refs. 76 and 83) (Fig. 128.4A-b).

Adenosinergic Systems

Adenosine has received much experimental attention as a probable physiologically important endogenous somnogen (84). Adenosine has been shown to exert multiple effects on behavioral state (18 ,84). For example, adenosine may exert tonic inhibition over the glutamatergic excitatory inputs to the cholinergic cells of the LDT and PPT (18 ,85) (Fig. 128.4B-d), and it may contribute to the REM-related suppression of serotonergic raphe neurons (18) (Fig. 128.4C-j). In addition, adenosine may exert tonic inhibition over BF cholinergic neurons (18). Finally, as noted, extracellular buildup of adenosine may constitute the sleep-promoting factor associated with prolonged wakefulness (18 ,84) (Fig. 128.4B-c).

Neuropeptidergic Systems

Many different neuropeptides have been implicated in regulation of the sleep-wake and REM-NREM cycles. These include galanin (75 ,76) (Fig. 128.4B-b), orexin (86 ,87) (Fig. 128.4A-b), vasoactive intestinal polypeptide (VIP), and numerous hormones (33 ,48 ,88) (Fig. 128.4C-g) (for reviews, see ref. 89 for the cytokines, ref. 90 for hormonal influences, and ref. 88 for VIP).

Many such neuropeptides have, like adenosine, been proposed to be endogenous sleep substances whose accumulation over prolonged waking promotes sleep onset (for review, see ref. 89). As noted by Kreuger and Fang (89), there is especially strong evidence for specific promotion of NREM by tumor necrosis factor, interleukin 1, growth hormone–releasing hormone, and prostaglandin D_2 as well as for promotion of REM by VIP and prolactin. Other neuropeptides such as atriopeptin, bombesin, corticotropinreleasing factor, and substance P are released from mesopontine cholinergic terminals, and these may, in turn, modify the sleep regulatory effects of co-released ACh (reviewed in ref. 6).

Three findings highlight the importance of neuropeptides in the regulation of sleep. The first is that inhibitory neurons in the VLPO of the hypothalamus, a specifically sleep-active area (74), use galanin as well as GABA to inhibit ascending arousal systems such as the LC (75 ,76) (Fig. 128.4B-a). As noted by Shiromani et al. (83), the association of galanin with NREM is particularly significant because this neuropeptide also stimulates release of growth hormone from the pituitary and may therefore play an important role in the pulsatile release of growth hormone specifically associated with NREM SWS. The second is the demonstration that corticotropin-releasing factor participates in the pontomedullary control of REM atonia (33) (Fig. 128.4C-g). The third finding has come from studies on the genetic basis of narcolepsy using animal models. The neuropeptide orexin (or hypocretin), produced only by neurons in the lateral hypothalamus, may play a key role in sleep regulation by its modulation of ascending cholinergic and monoaminergic arousal systems (86 ,87) (Fig. 128.4A-b).

Second Messengers and Intranuclear Events

As in much of neuroscience, research on behavioral state control is now beginning to extend its inquiry beyond the neurotransmitter and its receptors to the roles of intracellular second messengers (7), as well as intranuclear events (91). Results of the molecular genetic approach to sleep research include the discovery of the role of orexin in sleep regulation (see earlier) and the discovery that choline acetyltransferase mRNA levels show behavioral state dependency in the rat (92). As noted by Shiromani et al. (83), the experimental observation of behavioral state dependent immediate early

gene expression in hypothalamic cells indicates that intranuclear events must participate in the control of sleep-wake cycles.

NEUROANATOMY OF REM-NREM CONTROL SYSTEMS

Part of "128 - Basic Mechanisms of Sleep: New Evidence on the Neuroanatomy and Neuromodulation of the NREM-REM Cycle "

At the same time that sleep studies have probed more and more deeply, to the cell membrane and beyond, efforts to understand the mechanisms of state control at a more global and regional level have continued. The picture that we are attempting to sketch is designed to provide an integrating framework for the analytic studies just reviewed. A major breakthrough in the regional brain research effort has been provided by the explosive growth of imaging studies of the human brain. This work provides still another bridge for integration: Now for the first time, we can begin to relate the cellular and molecular-level data from animals to the regional data in humans. This comparison is of particular relevance to dream theory and psychopathology.

Brainstem Executive Control of the REM-NREM Cycle

In the pons, cholinergic, serotonergic, noradrenergic, glutamatergic, and GABAergic neurotransmission among the LDT and PPT (containing cholinergic cells), the DRN (mainly serotonergic), and the LC (mainly noradrenergic) forms the core circuits for the executive control of the REM-NREM cycle (for reviews, see refs. 1 ,2 ,3 and 4 , 6 , 8 , and 17) (Fig. 128.4B and C).

In the EEG-aroused states of REM and waking, the LDT and PPT provide a large proportion of the excitatory cholinergic input to both the thalamus and the BF (Fig. 128.4A-c), which then activate limbic structures and the cortex (6) (Fig. 128.4A-d). Expression of the physiologic signs of REM is, however, modifiable at diverse sites both rostral and caudal to these executive networks, as detailed later. For example, in addition to inputs from the DRN and LC, the LDT-PPT also receives projections from limbic forebrain structures (6) and the basal ganglia (93) (Fig. 128.4C-k). Moreover, in the control of sleep onset, the physiologic features of NREM sleep, and circadian control of the sleep-wake cycle, diencephalic structures come to be of primary importance (Fig. 128.4A and B).

Other Brainstem Structures

Many different brainstem structures in addition to LDT-PPT cholinergic cells and the LC and raphe nuclei are crucially involved in the modulation of REM sleep and its distinctive physiologic signs. These include noncholinergic areas within the PPT-LDT as well as peribrachial areas caudal to the LDT and PPT in the cat (8), diverse cholinoceptive areas in the medial pontine reticular system such as the nucleus pontis oralis of the rat (33), and the gigantocellular tegmental field and LC α of the cat (7 ,14). Adding to the functional complexity of mesopontine cholinergic areas are its important roles in functions other than behavioral state control such as motor control (106).

Structures rostral and caudal to the pons such as the ventrolateral periaqueductal gray (56 ,94) and the medulla (33 ,79) also play key roles in the transmission and modulation of the physiologic signs of REM sleep. For example, sleep-associated inhibition of LC and DR by glycinergic input originates in the periaqueductal gray, the midbrain reticular formation, and various medullary sites (e.g., raphe magnus, gigantocellular α, paragigantocellular nuclei) (Fig. 128.4B-e), whereas GABAergic inhibition originates in an even more diverse set of brainstem and diencephalic regions (56) (Fig. 128.4B-a). Similarly, REM-associated postural atonia involves brainstem networks extending rostrally from midbrain areas such as the periaqueductal gray, through the mesopontine PPT, through the pontine inhibitory area and peri-LC areas to the medullary nucleus magnocellularis, and thence to spinal motor neurons caudally (33) (Fig. 128.4C-f,g). Brainstem structures rostral to the pons could facilitate brainstem-limbic interactions in REM sleep (see later) such as those also hypothesized to contribute to REM-related atonia (33).

Figure 128.4A to C schematizes major neuromodulatory systems involved in the generation and maintenance of the three cardinal behavioral states at different levels of the central nervous system. (For reviews on the functional neuroanatomy of brainstem control of sleep cycles, see refs. 1 ,2 ,3 and 4 , 6 , 8 , 14 , 17 , 56 , and 79 .)

Forebrain Structures

Other important contemporary research now extends the study of sleep-wake and REM sleep control mechanisms rostrally from the pontine brainstem to diencephalic structures in a manner consistent with connectivity studies (48). In addition to the well-described brainstem-thalamus-cortex axis, subcortical sleep control mechanisms intercommunicate with each other and with the cortex through an interconnected network of structures extending rostrally from the brainstem RAS to the hypothalamus, BF, and limbic system.

Saper et al. (76) classified three ascending arousal systems: the brainstem cortical projection system (including DR serotonergic, LC noradrenergic, and LDT-PPT cholinergic elements), the cholinergic BF projection system, and the histaminergic hypothalamic cortical projection system. The BF system projects to topographically specific cortical areas, whereas the other systems project diffusely to widespread areas of the forebrain.

We now briefly summarize recent findings on this extended

subcortical system that are pertinent to sleep-wake and REM sleep control. We focus here on findings in the hypothalamus, BF nuclei, amygdala, thalamus, and basal ganglia.

Hypothalamus

Histaminergic neurons originating in the posterior hypothalamus innervate virtually the entire brain including brainstem structures such as the mesopontine tegmentum (82) (Fig. 128.4A-b). These brainstem regions, in turn, innervate both anterior and posterior hypothalamus (82 ,83) (Fig. 128.4A-e).

Anterior portions of the hypothalamus (preoptic area and adjacent BF) are known to be essential to promoting sleep (73 ,76 ,82 ,83). In contrast, tonic firing of histaminergic neurons in the posterior hypothalamus play an important role in cortical arousal and the maintenance of wakefulness (for reviews, see refs. 76 and 83) (Fig. 128.4A-b). The TMN plays a particularly important role in this posterior hypothalamic histaminergic arousal system, and these neurons have been shown to decrease their firing during sleep (74 ,76 ,83).

Sherin et al. (74) proposed that a monosynaptic pathway in the hypothalamus may constitute a “switch” for the alternation of sleep and wakefulness. Using immediate early gene (*c-fos*) techniques, these workers identified a group of GABAergic and galaninergic neurons in the VLPO, which are specifically activated by sleep (74) and have since been shown to be selectively sleep active using electrophysiology (83). These GABAergic and galaninergic VLPO neurons constitute the main source of innervation to the histaminergic neurons of the TMN (Fig. 128.4B-b) and may therefore specifically inhibit histaminergic neurons of the TMN to preserve sleep (74 ,75 and 76).

One study demonstrated extensive histaminergic innervation of the mesopontine tegmentum including the LDT (82) (Fig. 128.4A-b). Suppression of slow-wave activity and an increase in waking followed microinjection of histamine and histamine agonist into these areas (82). Histaminergic projections from the TMN to areas of the BF involved in sleep-wake control (Fig. 128.4A-b) were also demonstrated in the cat (95). VLPO neurons were also shown to innervate other components of ascending arousal systems such as the monoaminergic nuclei of the brainstem (Fig. 128.4B-b), and there they may also exert a sleep-promoting inhibitory influence (75 ,95). An important finding is that the orexinergic cells of the lateral hypothalamus also innervate most of the brainstem and diencephalic ascending arousal systems (Fig. 128.4A-b), and, therefore, these may also play a modulatory role in the sleep-wake cycle (86 ,87).

Combining the foregoing findings with evidence of the somnogenic effects of adenosine, Shiromani et al. (83) proposed the following model for the control of the wake-NREM transition and its link to the ultradian REM-NREM cycle, detailed here and illustrated in Fig. 128.5 :

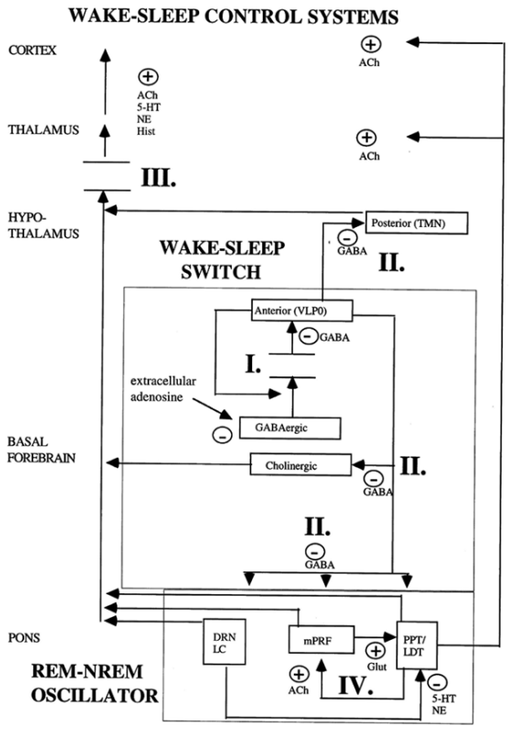


FIGURE 128.5. Integrated model of sleep onset and REM-NREM oscillation proposed by Shiromani et al. (83). *I*, During prolonged wakefulness, accumulating adenosine inhibits specific GABAergic anterior hypothalamic and basal forebrain neurons that have been inhibiting the sleep-active ventrolateral preoptic area (VLPO) neurons during waking. *II*, Disinhibited sleep-active GABAergic neurons of the VLPO and adjacent structures then inhibit the wake-active histaminergic neurons of the tuberomamillary nucleus (TMN), as well as those of the pontine aminergic (DRN and LC) and cholinergic (LDT-PPT) ascending arousal systems thereby initiating NREM sleep. *III*, Forebrain activation by ascending aminergic and orexinergic arousal systems is dysfacilitated. *IV*, Once NREM sleep is thus established, the executive networks of the pons initiate and maintain the ultradian REM-NREM cycle. ACh, acetylcholine; DRN, dorsal raphe nucleus; 5-HT, serotonin; LC, locus ceruleus; LDT, laterodorsal tegmental nucleus; NE, norepinephrine; PRF, pontine reticular formation.

Figure 128.5I : During prolonged wakefulness, accumulating adenosine inhibits specific GABAergic anterior hypothalamic and BF neurons that have been inhibiting the sleep-active VLPO neurons during waking (Fig. 128.4A-h). During waking, TMN neurons may also reciprocally inhibit sleep-active neurons of the anterior hypothalamus and BF (Fig. 128.4A-f).

Figure 128.5II : The thus disinhibited sleep-active GABAergic neurons of the VLPO and adjacent structures then inhibit the wake-active histaminergic neurons of the TMN as well as those of the pontine aminergic (DR and LC) and cholinergic (LDT-PPT) ascending arousal systems, thereby initiating NREM sleep.

Figure 128.5III : Forebrain activation by these ascending arousal systems (Fig. 128.4A-b,c,g) is thus dysfacilitated.

Figure 128.5IV : Once NREM sleep is thus established, the executive networks of the pons initiate and maintain the ultradian REM-NREM cycle described by reciprocal interaction.

Tying the hypothalamus to the pons in this dynamic manner may provide a critical link between the circadian clock in the suprachiasmatic nucleus and the NREM-REM sleep cycle oscillator (83). The circadian control of sleep by the suprachiasmatic nucleus is beyond the scope of this review but constitutes a key influence not only on the sleep wake alternation but also on shorter-period oscillations such as the ultradian REM-NREM alternation (96). (For a comprehensive review of circadian rhythms and the suprachiasmatic nucleus in sleep-wake regulation, see ref. 97 .)

Basal Forebrain

BF nuclei have anatomic connections with the LC, raphe, and pontine nuclei (69 ,73) and, in turn, project to more rostral structures such as the cortex, thalamus, and limbic systems (69 ,98). In addition to its brainstem and cortical connectivity, the BF also has anatomic connections with the anterior and posterior hypothalamus, the amygdala, and the thalamus (73). (For reviews of BF connectivity see refs. 69 and 73 .)

Neurochemically, ACh plays a major role in BF control of behavioral state chiefly by its activating effects on the cortex and other forebrain structures (69). Magnocellular cholinergic cells of the BF nuclei promote the activation of those cortical and limbic structures to which they project (69 ,73 ,98) (Fig. 128.4A-d). For example, cholinergic cells of the nucleus basalis of Meynert activate topographically distinct areas of the cortex (69 ,73 ,98), and cholinergic activity of the medium septum and vertical limb of the diagonal band of Broca mediates hippocampal theta rhythm (42).

There are extensive functional interactions between the brainstem structures (LC, raphe nuclei, LDT-PPT) and the

BF in sleep-wake control (6 ,69 ,73). The PPT-LDT provides a major source of excitatory input to the BF (6) (Fig. 128.4C-l). For example, Shiromani et al. (83) showed that PPT stimulation induces *c-fos* expression in a variety of BF nuclei. Similarly, PPT-LDT neurons were shown to release ACh into the BF of rats (99).

Glutamatergic activation is an important transmitter of reciprocal activation between the brainstem and the BF (Fig. 128.4C-m). For example, BF structures that activate the cortex can be activated by brainstem glutamatergic cells (100), the cortical EEG activation produced by cholinergic mesopontine stimulation of the BF is transmitted within the BF by glutamatergic (versus cholinergic) mechanisms (6) (Fig. 128.4C-n), and glutamatergic systems of the BF can, in turn, affect brainstem behavioral state modulation through projections to the mesopontine tegmentum (101).

In concert with closely related anterior hypothalamic cells such as those of the VLPO, specifically sleep-active BF cells (anatomically and neurochemically distinct from the wake-REM active cholinergic magnocellular neurons) function as NREM sleep-promoting elements (73 ,83). This may occur by GABAergic inhibition of posterior hypothalamic (Fig. 128.4B-b) and brainstem (Fig. 128.4B-a) arousal systems (73 ,74 and 75) as well as of more rostral structures (69) (Fig. 128.4B-f).

In addition to their role in cortical activation and desynchronization (Fig. 128.4A-d and Fig. 128.4C-d), BF neurons have been associated with the specific physiologic features of NREM sleep. For example, although BF lesions can cause the emergence of slow waves (102), magnocellular cholinergic neurons may also participate in the generation of cortical slow waves (103).

As in the brainstem, neuromodulatory systems interact within the BF itself (e.g., the foregoing BF cholinergic-glutamatergic interactions). For example, BF cholinergic neurons may be under tonic inhibition by adenosine (18) (Fig. 128.4B-g). The BF nuclei therefore both directly participate in behavioral state-related functions and modify the activity of other areas involved in sleep such as the pontine REM generator.

Amygdala

Of particular interest in view of recent findings from human neuroimaging studies (see ref. 3 for review), the amygdala has reciprocal connections with pontine regions involved in the control of REM sleep (for reviews, see refs. 48 and 104).

Physiologic signs of REM have been shown to occur spontaneously in the amygdala and in other limbic structures (104). For example, in the cat, PGO-like EEG activity has been detected in the basolateral amygdala, and single-unit activity in the central nucleus (CN) of the amygdala increases during NREM sleep immediately preceding REM (variously termed the “transitional stage,” “SP,” or “SPHOL”) and then increases further during actual REM (104). In addition, amygdalar lesions in the cat have been shown to decrease PGO frequency (104).

Physiologic signs of REM are also modifiable from the amygdala. For example, electrical stimulation of the cat amygdala significantly increased PGO number, spike density, and burst density (104). In addition, cholinergic stimulation of the cat CN enhanced REM sleep for several days by an increased number of REM episodes, an effect akin to the long-term REM enhancement by cholinergic stimulation of the peribrachial pons (104). Moreover, cholinergic stimulation of the CN concomitantly increased PGO density (104). Amygdalar stimulation also increased the amplitude and rate of acoustically elicited pontine PGO waves in the waking rat and burst firing of pontine cells in the rabbit (48).

Aminergic stimulation of the amygdala has also been shown to modify sleep in the direction predicted by reciprocal interaction for pontine sites. For example, serotonergic stimulation of the amygdala in the cat caused short latency changes of state from either NREM or REM (48). Similarly, noradrenergic stimulation of the amygdala suppressed sleep relative to wakefulness (48).

It has therefore been proposed, by Morrison et al. (48) and by Calvo and Simon-Arceo (104), that the amygdala (and, in particular, the CN) stimulates lateral pontine areas involved in REM sleep initiation (Fig. 128.4C-k) that, in turn, initiate REM sleep. However, the continuation of REM-NREM cycling in the pontine cat (2) obviates an obligatory role for amygdalar stimulation of the pons in the initiation of REM, at least in the cat.

Thalamic Structures and Intrinsic Thalamocortical Oscillatory Rhythms

Dysfacilitation of thalamocortical relay neurons after sleep onset allows the emergence of underlying thalamocortical oscillatory rhythms (for reviews, see refs. 127 and 128) (Fig. 128.4B-h). GABAergic neurons of the thalamic reticular nucleus hyperpolarize and dysfacilitate thalamic relay neurons as NREM deepens (127,128). In this hyperpolarized condition, thalamic neurons become constrained to burst-firing patterns first in spindle (12 to 14 Hz) and later in delta (1 to 4 Hz) frequencies as NREM deepens from stage 2 to delta sleep (127,128). The cortex may further constrain these spindle and delta wave–generating thalamocortical bursts within a newly described slow (less than 1 Hz) oscillation seen in cats (127,128) and humans (106).

Mesopontine cholinergic neurons provide a major excitatory input to the thalamus during the EEG activated states of wakefulness and NREM sleep and, during REM, cholinergic excitation alone may prevent the emergence of intrinsic thalamocortical oscillations (6). Findings that ascending cholinergic input from the mesopontine tegmentum influences thalamic blood flow (109) and thalamic NO release (81) (Fig. 128.4C-i) are of particular interest in light of positron emission tomography neuroimaging studies showing a REM-associated increased thalamic blood flow (reviewed in ref. 3). Other thalamic structures such as centralis lateralis nucleus possibly participate directly in the modulation of REM sleep (110).

Basal Ganglia

Although receiving less focus in past experimental studies of sleep in animals, the extensive activation of basal ganglia and related structures in positron emission tomography neuroimaging studies of human REM sleep suggests their involvement in the neural networks subserving this behavioral state (reviewed in ref. 3). There are extensive reciprocal projections between the basal ganglia and the PPT region (3 ,193). This connectivity has been postulated to subserve several specific roles. For example, Allan Braun and colleagues at the National Institutes of Health in Bethesda, Maryland suggested that the basal ganglia may play a role in the rostral transmission of PGO waves (reviewed in ref. 2), whereas Rye (3) suggested that striatal projections to the pedunculopontine region may serve to modulate movement to accord with behavioral state. Additional possibilities are suggested by the extensive activation of the ventral striatum and surrounding BF in REM (reviewed in ref. 2). For example,

Eric Nofzinger and colleagues at the University of Pittsburgh suggested that an important function of REM sleep is the integration of neocortical function with BF and hypothalamic motivational and reward mechanisms (reviewed in ref. 2).

CONCLUSIONS

Part of "128 - Basic Mechanisms of Sleep: New Evidence on the Neuroanatomy and Neuromodulation of the NREM-REM Cycle "

Interactions of diverse neuromodulatory systems operate in widespread subcortical areas to amplify or suppress REM sleep generation as well as to facilitate onset and offset of the control of behavioral state by the pontine REM-NREM oscillator. An ascending medial brainstem and diencephalic system of multiple nuclei with extensive reciprocal interconnections and system-wide sensitivity to neuromodulation controls the regular alternation and integration of the sleep-wake and REM-NREM cycles. The existence an executive aminergic-cholinergic reciprocal interaction system controlling REM-NREM alternation in the pontine brainstem has been strongly confirmed by recent findings. The interaction of pontine structures with other brain structures can now begin to be studied in ways that will enrich our understanding of how the distinctive features of each conscious state are mediated and how their stereotyped sequencing is controlled.

ACKNOWLEDGMENTS

Part of "128 - Basic Mechanisms of Sleep: New Evidence on the Neuroanatomy and Neuromodulation of the NREM-REM Cycle "

This project was funded by National Institute of Drug Abuse grant RO1-DA11744-01A1, the MacArthur Foundation Mind Body Network, and National Institutes of Health grants MH-48,832, MH13923, and MH01287. We wish to thank James Quattrochi, Bernat Kocsis, Robert Stickgold, Rosalia Silvestri-Hobson, Subimal Datta, Matthew Walker, Roar Fosse, and Dawn Opstad.

REFERENCES

1. Hobson JA, Steriade M. The neuronal basis of behavioral state control. In: Bloom FE, ed. *Handbook of physiology: the nervous system*, vol 4. Bethesda, MD: American Physiological Society, 1986:701–823.

2. Hobson JA, Stickgold R, Pace-Schott EF. The neuropsychology of REM sleep dreaming. *Neuroreport* 1998;9:R1–R14.

3. Rye DB. Contributions of the peduculopontine region to normal and altered REM sleep. *Sleep* 1997;20:757–788.

4. Steriade M, McCarley RW. *Brainstem control of wakefulness and sleep*. New York: Plenum, 1990.

5. McCarley RW, Hobson JA. Neuronal excitability modulation over the sleep cycle: a structural and mathematical model. *Science* 1975;189:58–60.

6. Semba K. The mesopontine cholinergic system: a dual role in REM sleep and wakefulness. In: Lydic R, Baghdoyan HA, eds. *Handbook of behavioral state control: molecular and cellular mechanisms*. Boca Raton, FL: CRC, 1999:161–180.

7. Capece MC, Baghdoyan HA, Lydic R. New directions for the study of cholinergic REM sleep generation: specifying pre- and post-synaptic mechanisms. In: Mallick BN, Inoue S, eds. *Rapid eye movement sleep*. New York: Marcel Dekker, 1999:123–141.

8. Datta S. Cellular basis of pontine ponto-geniculo-occipital wave generation and modulation. *Cell Mol Neurobiol* 1997;17:341–365.

9. Leonard CS, Llinas RR. Electrophysiology of mammalian peduculopontine and laterodorsal tegmental neurons *in vitro:* implications for the control of REM sleep. In: Steriade M, Biesold D, eds. *Brain cholinergic mechanisms*. Oxford: Oxford Science Publications, 1990:205–223.

10. Leonard CS, Llinas R. Serotonergic and cholinergic inhibition of mesopontine cholinergic neurons controlling REM sleep: an *in vitro* electrophysiological study. *Neuroscience* 1994;59:309–330.

11. Sakai K, Koyama Y. Are there cholinergic and non-cholinergic paradoxical sleep-on neurons in the pons? *Neuroreport* 1996;7:2449–2453.

12. Steriade M, Pare D, Parent A, et al. Projections of cholinergic and non-cholinergic neurons of the brainstem core to relay and associational thalamic nuclei in the cat and macaque monkey. *Neuroscience* 1988;25:47–67.

13. Jasper AH, Tessier J. Acetylcholine liberation from cerebral cortex during paradoxical sleep (REM). *Science* 1971;172:601–602.

14. Sakai K. Executive mechanisms of paradoxical sleep. *Arch Ital Biol* 1988;126:239–257.

15. Egan TM, North RA. Acetylcholine hyperpolarizes central neurons by acting on an M2 muscarinic receptor. *Nature* 1986;319:405–407.

16. Luebke JL, McCarley RW, Greene RW. Inhibitory action of the acetylcholine agonist carbachol on neurons of the rat laterodorsal tegmental nucleus in the *in vitro* brainstem slice. *J Neurosci* 1993;70:2128–2135.

17. Jones BE. Paradoxical sleep and its chemical/structural substrate in the brain. *Neuroscience* 1991;40:637–656.

18. McCarley RW, Strecker RE, Porkka-Hieskanen T, et al. Modulation of cholinergic neurons by serotonin and adenosine in the control of REM and NREM sleep. In: Hayaishi O, Inoue S, eds. *Sleep and sleep disorders: from molecule to behavior*. Tokyo: Academic:49–63.

19. Calvo J, Datta S, Quattrochi JJ, et al. Cholinergic microstimulation of the peribrachial nucleus in the cat: delayed and prolonged increases in REM sleep. *Arch Ital Biol* 1992;130:285–301.

20. Quattrochi JJ, Hobson JA. Carbachol microinjection into the caudal peribrachial area induces long-term enhancement of PGO wave activity but not REM sleep. *J Sleep Res* 1999;8:281–290.

21. Datta S, Siwek DF, Patterson EH, et al. Localization of pontine PGO wave generation sites and their anatomical projections in the rat. *Synapse* 1998;30:409–423.

22. Capece ML, Efange SMN, Lydic R. Vesicular acetylcholine transport inhibitor surpresses REM sleep. *Neuroreport* 1997;8:481–484.

23. Baghdoyan HA, Lydic R. M2 muscarinic receptor subtype in the feline medial pontine reticular formation modulates the amount of rapid eye movement sleep. *Sleep* 1999;22:835–847.

24. Datta S, Siwek DF Excitation of the brain stem pedunculopontine tegmentum cholinergic cells induces wakefulness and REM sleep. *J Neurophysiol* 1997;77:2975–2988.

25. Kodama T, Takahashi Y, Honda Y. Enhancement of acetylcholine release during paradoxical sleep in the dorsal tegmental field of the cat brain stem. *Neurosci Lett* 1990;114:277–282.

26. Lydic R, Baghdoyan HA. Pedunculopontine stimulation alters respiration and increases ACh release in the pontine reticular formation. *Am J Physiol* 1993;264:R544–R554.

27. Williams JA, Comisarow J, Day J, et al. State-dependent release of acetylcholine in rat thalamus measured by microdialysis. *J Neurosci* 1994;14:5236–5242.

28. Kodama T, Honda Y. Acetylcholine releases of mesopontine PGO-on cells in the lateral geniculate nucleus in sleep-waking cycle and serotonergic regulation. *Prog Neuropsychopharmacol Biol Psychiatry* 1996;20:1213–1227.

29. Curro-Dossi R, Pare D, Steriade M. Short-lasting nicotinic and long-lasting muscarinic depolarizing response of thalamocortical neurons to stimulation of mesopontine cholinergic nuclei. *J Neurophysiol* 1991;65:393–406.

30. Sakai K, Jouvet M. Brain stem PGO-on cells projecting directly to the cat dorsal lateral geniculate nucleus. *Brain Res* 1980;194:500–505.

31. Oakman SA, Faris PL, Cozzari C, et al. Characterization of the extent of pontomesencephalic cholinergic neurons' projections to the thalamus: comparison with projections to midbrain dopaminergic groups. *Neuroscience* 1999;94:529–547.

32. Webster HH, Jones BE. Neurotoxic lesions of the dorsolateral pontomesencephalic tegmentum-cholinergic cell area in the cat. II. Effects on sleep-waking states. *Brain Res* 1988;458:285–302.

33. Lai YY, Siegel JM. Muscle atonia and REM sleep. In: Mallick BN, Inoue S, eds. *Rapid eye movement sleep*. New York: Marcel Dekker, 1999:69–90.

34. Jacobs BL, Fornal CA. An integrative role for serotonin in the central nervous system. In: Lydic R, Baghdoyan HA, eds. *Handbook of behavioral state control: molecular and cellular mechanisms*. Boca Raton, FL: CRC, 1999:181–193.

35. Monti JM, Monti D. Functional role of serotonin 5-HT1 and 5-HT2 receptor in the regulation of REM sleep. In: Mallick BN, Inoue S, eds. *Rapid eye movement sleep*. New York: Marcel Dekker, 1999:142–151.

36. Luppi P-H, Gervasoni D, Peyron C, et al. Norepinephrine and REM sleep. In: Mallick BN, Inoue S, eds. *Rapid eye movement sleep*. New York: Marcel Dekker, 1999:107–121.

37. Steininger TL, Wainer BH, Blakely RD, et al. Serotonergic dorsal raphe nucleus projections to the cholinergic and noncholinergic neurons of the pedunculopontine tegmental region: a light and electron microscopic anterograde tracing and immunohistochemical study. *J Comp Neurol* 1997;382:302–322.

38. Portas CM, Bjorvatn B, Fagerland S, et al. On-line detection of extracellular levels of serotonin in dorsal raphe nucleus and frontal cortex over the sleep/wake cycle in the freely moving rat. *Neuroscience* 1998;83:807–814.

39. Portas CM, McCarley RW. Behavioral state-related changes of extracellular serotonin concentration in the dorsal raphe nucleus: a microdialysis study in the freely moving cat. *Brain Res* 1994;648:306–312.

40. Iwakiri H, Matsuyama K, Mori S. Extracellular levels of serotonin in the medial pontine reticular formation in relation to sleep-wake cycle in cats: a microdialysis study. *Neurosci Res* 1993;18:157–170.

41. Wilson C, James ML, Behnke EJ, et al. Direct measures of extracellular serotonin change in the human forebrain during waking.*Soc Neurosci Abstr* 1997;23:2130(abst).

42. Vertes R, Kocsis B. Brainstem-diencephalo-septohippocampal systems controlling the theta rhythm of the hippocampus. Neuroscience 1997;81:893–926.

43. Horner RL, Sanford LD, Annis D, et al. Serotonin at the laterodorsal tegmental nucleus suppresses rapid eye movement sleep in freely behaving rats. *J Neurosci* 1997;17:7541–7552.

44. Okabe S, Kubin L. Role of 5HT1 receptors in the control of hypoglossal motorneurons in vivo.*Sleep* 1997;19[Suppl. 10]:S150–S153.

45. Sanford, LD, Ross RJ, Seggos AE, et al. Central administration of two 5-HT receptor agonists: effect on REM sleep and PGO waves. *Pharmacol Biochem Behav* 1994;49:93–100.

46. Thakkar M, Strecker RE, McCarley RW. Behavioral state control through differential serotonergic inhibition in the mesopontine cholinergic nuclei: a simultaneous unit recording and microdialysis study. *J Neurosci* 1998;18:5490–5497.

47. Portas CM, Thakkar M, Rainnie D, et al. Microdialysis perfusion of 8-hydroxy-2-(di-N-propylamino) tetralin (8-OH-DPAT) in the dorsal raphe nucleus decreases serotonin release and increases rapid eye movement sleep in the freely moving cat.*J Neurosci* 1996;16:2820–2828.

48. Morrison AR, Sanford LD, Ross RJ. Initiation of rapid eye movement sleep: beyond the brainstem. In: Mallick BN, Inoue S, eds. *Rapid eye movement sleep*. New York: Marcel Dekker, 1999:51–68.

49. Rajkowski J, Silakov V, Ivanova S, et al. Locus coeruleus (LC) neurons in monkey are quiescent in paradoxical sleep (PS). *Soc Neurosci Abstr* 1997;23:828.2(abst).

50. Singh S, Mallick BN. Mild electrical stimulation of pontine tegmentum around locus coeruleus reduces rapid eye movement sleep in rats. *Neurosci Res* 1996;24:227–235.

51. Bier MJ, McCarley RW. REM-enhancing effects of the adrenergic antagonist idazoxan infused into the medial pontine reticular formation of the freely moving cat. *Brain Res* 1994;634:333–338.

52. Nicholson AN, Belyavin A, Pascoe PA. Modulation of rapid eye movement sleep in humans by drugs that modify monoaminergic and purinergic transmission. *Neuropsychopharmacology* 1989;2:131–143.

53. Sanford LD, Kearney K, McInerney B, et al. Rapid eye movement sleep (REM) is not regulated by 5HT2 receptor mechanisms in the laterodorsal tegmental nucleus. *Sleep Res* 1997;26:127(abst).

54. Ross RJ, Gresch PJ, Ball WA, et al. REM sleep inhibition by desipramine: evidence for an alpha-1 adrenergic mechanism. *Brain Res* 1995;701:129–134.

55. Williams JA, Reiner PB. Noradrenaline hyperpolarizes identified rat mesopontine cholinergic neurons *in vitro. J Neurosci* 1993;13:3878–3883.

56. Luppi P-H, Peyron C, Rampon C, et al. Inhibitory mechanisms in the dorsal raphe nucleus and locus coeruleus during sleep. In: Lydic R, Baghdoyan HA, eds. *Handbook of behavioral state control: molecular and cellular mechanisms*. Boca Raton, FL: CRC, 1999:195–211.

57. Nitz D, Siegel JM. GABA release in the locus coeruleus as a function of sleep/wake state. *Neuroscience* 1997;78:795–801.

58. Li XY, Greene RW, Rainnie DG, et al. Dual modulation of nicotine in DR neurons. *Sleep Res* 1997;26:22(abst).

59. Roth MT, Fleegal MA, Lydic R, et al. Pontine acetylcholine release is regulated by muscarinic autoreceptors. *Neuroreport* 1996;7:3069–3072.

60. Hobson JA, Pace-Schott EF. Reply to Solms, Braun and Reiser. *Neuropsychoanalysis* 1999;1:206–224.

61. Miller JD, Farber J, Gatz, P. Activity of mesencephalic dopamine and non-dopamine neurons across stages of sleep and waking in the rat.*Brain Res* 1983;273:133–141.

62. Trulson ME, Preussler DW, Howell AG. Activity of the substantia nigra across the sleep-wake cycle in freely moving cats. *Neurosci Lett* 1981;26:183–188.

63. Nunes GP Jr, Tufik S, Nobrega JN. Autoradiographic analysis of D1 and D2 dopaminergic receptors in rat brain after paradoxical sleep deprivation. *Brain Res Bull* 1994;34:453–456.

64. Python A, de Saint Hilaire Z, Gaillard JM. Effects of a D2 receptor agonist RO 41-9067 alone and with clonidine on sleep parameters in the rat. *Pharmacol Biochem Behav* 1996;53:291–296.

65. de Saint Hilaire Z, Python A, Blanc G, et al. Effects of WIN 35,428 a potent antagonist of dopamine transporter on sleep and locomotion. *Neuroreport* 1995;6:2182–2186.

66. Gillin JC, Post R, Wyatt RJ, et al. REM inhibitory effect of L-DOPA infusion during human sleep. *Electroencephalogr Clin Neurophysiol* 1973;35:181–186.

67. Nofzinger EA., Reynolds CF III, et al. REM sleep enhancement by bupropion in depressed men. *Am J Psychiatry* 1995;152:274–276.

68. Moore H, Fadel J, Sarter M, et al. Role of accumbens and cortical dopamine receptors in the regulation of cortical acetylcholine release. *Neuroscience* 1999;88:811–822.

69. Jones BE, Muhlethaler M. Cholinergic and GABAergic neurons of the basal forebrain: role in cortical activation. In: Lydic R, Baghdoyan HA, eds. *Handbook of behavioral state control: molecular and cellular mechanisms*. ed. Boca Raton, FL: CRC, 1999:213–233.

70. Steriade M, Pare D, Datta S, et al. Different cellular types in mesopontine cholinergic nuclei related to ponto-geniculooccipital waves. *J Neurosci* 1990;10:2560–2579.

71. Mallick BN, Kaur S, Jha SK, et al. Possible role of GABA in the regulation of REM sleep with special reference to REM-OFF neurons. In: Mallick BN, Inoue S, eds. *Rapid eye movement sleep*. New York: Marcel Dekker, 1999:153–165.

72. Nitz D, Siegel JM. GABA release in the dorsal raphe nucleus: role in the control of REM sleep. *Am J Physiol* 1997;273:R451–R455.

73. Szymusiak R. Magnocellular nuclei of the basal forebrain: substrates of sleep and arousal regulation. *Sleep* 1995;18:478–500.

74. Sherin JE, Shiromani PJ, McCarley RW, et al. Activation of ventrolateral preoptic neurons during sleep. *Science* 1996;271:216–219.

75. Sherin JE, Elmquist JK, Torrealba F, et al. Innervation of histaminergic tuberomammilary neurons by GABAergic and galininergic neurons in the ventrolateral preoptic nucleus of the rat. *J Neurosci* 1998;18:4705–4721.

76. Saper CB, Sherin JE, Elmquist JK. Role of the ventrolateral preoptic area in sleep induction. In: Hayaishi O, Inoue S, eds. *Sleep and sleep disorders: from molecule to behavior*. Tokyo: Academic, 1997:281–294.

77. Xi MC, Morales FR., Chase MH. Evidence that wakefulness and REM sleep are controlled by a GABAergic pontine mechanism. *J Neurophysiol* 1999;82:2015–2019.

78. Chase MH, Soja PJ, Morales FR. Evidence that glycine mediates the post synaptic potentials that inhibit lumbar motorneurons during the atonia of active sleep. *J Neurosci* 1989;9:743–751.

79. Gottesmann C. Introduction to the neurophysiological study of sleep: central regulation of skeletal and ocular activities. *Arch Ital Biol* 1997;135:279–314.

80. Leonard TO, Lydic, R. Nitric oxide: a diffusible modulator of physiological traits and behavioral states. In: Mallick BN, Inoue S, eds. *Rapid eye movement sleep*. New York: Marcel Dekker, 1999:167–193.

81. Williams JA, Vincent SR, Reiner PB. Nitric oxide production in rat thalamus changes with behavioral state, local depolarization and brainstem stimulation. *J Neurosci* 1997;17:420–427.

82. Lin JS, Hou Y, Sakai K, et al. Histaminergic descending inputs to the mesopontine tegmentum and their role in the control of cortical activation and wakefulness in the cat. *J Neurosci* 1996;16:1523–1537.

83. Shiromani PJ, Scammell T, Sherin JE, et al. Hypothalamic regulation of sleep. In: Lydic R, Baghdoyan HA, eds. *Handbook of behavioral state control: molecular and cellular mechanisms*. Boca Raton, FL: CRC, 1999:311–325.

84. Porkka-Heiskanen T, Strecker RE, Thakkar M, et al. Adenosine: a mediator of the sleep-inducing effects of prolonged wakefulness. *Science* 1997;276:1265–1268.

85. Rannie DG, Grunze HC, McCarley RW, et al. Adenosine inhibition of mesopontine cholinergic neurons: implications for EEG arousal. *Science* 1994;263:689–692.

86. Chemelli RM, Willie JT, Sinton CM, et al. Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell* 1999;98:437–451.

87. Lin L, Faraco J, Li R, et al. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* 1999;98:365–376.

88. Steiger A, Holsboer F. Neuropeptides and human sleep. *Sleep* 1997;20:1038–1052.

89. Krueger JM, Fang J. Cytokines and sleep regulation. In: Lydic R, Baghdoyan HA, eds. *Handbook of behavioral state control: molecular and cellular mechanisms*. Boca Raton: CRC, 1999:609–622.

90. Obal F, Krueger JM. Hormones and REM sleep. In: Mallick BN, Inoue S, eds. *Rapid eye movement sleep*. New York: Marcel Dekker, 1999:233–247.

91. Bentivoglio M, Grassi-Zucconi G. Immediate early gene expression in sleep and wakefulness. In: Lydic R, Baghdoyan HA, eds. *Handbook of behavioral state control: molecular and cellular mechanisms*. Boca Raton, FL: CRC, 1999:235–253.

92. Greco MA, McCarley RW, Shiromani PJ. Choline acetyltransferase expression during periods of behavioral activity and across natural sleep-wake states in the basal forebrain. *Neuroscience* 1999;93:1369–1374.

93. Inglis WL, Winn P. The pedunculopontine tegmental nucleus: where the striatum meets the reticular formation. *Prog Neurobiol* 1995;47:1–29.

94. Sastre JP, Buda CP, Kitahama K, et al. Importance of the ventrolateral region of the periaqueductal gray and adjacent tegmentum as studied by muscimol microinjection in the cat. *Neuroscience* 1996;74:415–426.

95. Lin JS, Vanni-Mercier G, Jouvet M. Histaminergic ascending and descending projections in the cat, a double immunocytochemical study focused on basal forebrain cholinergic cells and dorsal raphe nucleus serotoninergic neurons. *Sleep Res* 1997;26:24(abst).

96. Czeisler CA, Zimmerman JC, Ronda JM, et al. Timing of REM sleep is coupled to the circadian rhythm of body temperature in man. *Sleep* 1980;2:329–346.

97. Miller JD, Morin LP, Schwartz WJ, et al. New insights into the mamallian circadian clock. *Sleep* 1996;19:641–667.

98. Metherate R, Cox CL, Ashe JH. Cellular bases of neocortical activation: modulation of neural oscillations by the nucleus basalis and endogenous acetylcholine. *J Neurosci* 1992;12:4701–4711.

99. Consolo S, Bertorelli R, Forloni GL, et al. Cholinergic neurons of the pontomesencephalic tegmentum release acetylcholine in the basal nuclear complex of freely moving rats. *Neuroscience* 1990;37:717–723.

100. Rasmussen DD, Clow K, Szerb JC. Modification of neocortical acetylcholine release and electroencephalogram desynchronization due to brainstem stimulation by drugs applied to the basal forebrain. *Neuroscience* 1994;60:665–677.

101. Manfridi A, Mancia M. Desynchronized (REM) sleep inhibition induced by carbachol microinjections into the nucleus basalis of Meynert is mediated by the glutamatergic system. *Exptl Brain Res* 1996;109:174–178.

102. Kleiner S, Bringmann A. Nucleus basalis magnocellularis and pedunculopontine tegmental nucleus: control of the slow EEG waves in rats. *Arch Ital Biol* 1996;134:153–167.

103. Nunez A. Unit activity of rat basal forebrain neurons-relationship to cortical activity. *Neuroscience* 1996;72:757–766.

104. Calvo JM, Simon-Arceo K. Cholinergic enhancement of REM sleep from sites in the pons and amygdala. In: Lydic R, Baghdoyan HA, eds. *Handbook of behavioral state control: molecular and cellular mechanisms*. Boca Raton, FL: CRC, 1999:391–406.

105. Steriade M, Contreras D, Dossi C, et al. The slow (<1 Hz) oscillation in reticular thalamic and thalamocortical neurons: scenario of sleep rhythm generation in interacting thalamic and neocortical networks. *J Neurosci* 1993;13:3284–3299.

106. Achermann P, Borbely AA. Low frequency (<1 Hz) oscillations in the human sleep electroencephalogram. *Neuroscience* 1997;81:213–222.

107. Steriade M. Synchronized activities of coupled oscillators in the cerebral cortex and thalamus at different levels of vigilance. *Cereb Cortex* 1997;7:583–604.

108. Steriade M. Cellular substrates of oscillations in corticothalamic systems during states of vigilance. In: Lydic R, Baghdoyan HA, eds. *Handbook of behavioral state control: molecular and cellular mechanisms*. Boca Raton, FL: CRC, 1999:327–347.

109. Koyama Y, Toga T, Kayama Y, et al. Regulation of regional blood flow in the laterodorsal thalamus by ascending cholinergic fibers from the laterodorsal tegmental nucleus. *Neurosci Res* 1994;20:79–84.

110. Mancia M, Marini G. Thalamic mechanisms in sleep control. In: Hayaishi O, Inoue S, eds. *Sleep and sleep disorders: from molecule to behavior*. Tokyo: Academic, 1997:377–393.

129

Circadian Phase Sleep and Mood Disorders

Alfred J. Lewy

Alfred J. Lewy: Department of Psychiatry, School of Medicine, Oregon Health Science University, Portland, Oregon.

CIRCADIAN ANATOMY AND PHYSIOLOGY

Part of "129 - Circadian Phase Sleep and Mood Disorders "

Anatomy

The Suprachiasmatic Nucleus: Locus of the Biological Clock

Much is known about the neuroanatomic connections of the circadian system. In vertebrates, the locus of the biological clock (the endogenous circadian pacemaker, or ECP) that drives all circadian rhythms is in the hypothalamus, specifically, the suprachiasmatic nucleus (SCN) (1 ,2). This paired structure derives its name because it lies just above the optic chiasm. It contains about 10,000 neurons. The molecular mechanisms of the SCN are an active area of research. There is also a great deal of interest in clock genes and clock components of cells in general, not just in the SCN. The journal *Science* designated clock genes as the second most important breakthrough for the recent year; the year before it was also on the runner-up list. Although clock genes may endow most, if not all, cells with the capacity for circadian time keeping, only the SCN is an endogenous pacemaker. That is, only the SCN receives environmental time cues. In turn, the SCN passes on temporal information to the cells of the rest of the organism.

The Retinohypothalamic Tract: Pathway for Photic Effects

The SCN is sensitive to environmental time cues in the form of photic impulses that are conveyed from the retina via a specific neural pathway, the retinohypothalamic tract (RHT). The RHT is a separate pathway from that which mediates vision (3); however, bilateral enucleation causes circadian blindness, just as it causes visual blindness.

SCN Efferent Pathways

Not much is known about how the SCN entrains overt circadian rhythms. We know that the SCN is the master pacemaker, but regarding its regulation of the rest/activity cycle, core body temperature rhythm and cortisol rhythm, among others, it is not clear if there is a humoral factor or neural connection that transmits the SCN's efferent signal; however, a great deal is known about the efferent neural pathway between the SCN and pineal gland.

The Pineal Gland

In mammals, the pineal gland is located in the center of the brain; however, it lies outside the blood–brain barrier. Postganglionic sympathetic nerves (called the nervi conarii) from the superior cervical ganglion innervate the pineal (4). The preganglionic neurons originate in the spinal cord, specifically in the thoracic intermediolateral column. The pathway between the SCN and the spinal cord synapses in the paraventricular nucleus (PVN) and traverses through the medial forebrain bundle.

Light has two effects on melatonin production. In common with all other circadian rhythms controlled by the SCN, the 24-hour cycle of ambient light and darkness synchronizes (entrains) the melatonin rhythm to a period of precisely 24 hours. The circadian rhythm of melatonin production is the only one in which basal levels are confined to the day alternating with increased levels at night; that is, melatonin production occurs within the margins of the scotoperiod (dark period).

Melatonin is unique in another way: Light acutely stops melatonin production (5). This effect can occur only between dusk and dawn, because melatonin levels remain low during the day given that the SCN turns off melatonin production for about 12 hours each day. That is, darkness during the day cannot increase melatonin production. These two effects of light combine to restrict the duration of melatonin production to about 12 hours or less during the night. In many animals, the changing duration of the scotoperiod across the year results in a corresponding change

in the duration of melatonin production, which is used as a seasonal time cue for regulating hibernation, migration, and estrous (6).

Sympathetic stimulation of the pineal results in the synthesis and secretion of melatonin into the venous circulation, as well as into the cerebrospinal fluid. A major target of melatonin is the SCN. Melatonin is synthesized from serotonin in two steps, one of which (*N*-acetylation) is rate-limiting (7). After conversion of serotonin to *N*-acetylserotonin, hydroxyindole-O-methyltransferase synthesizes melatonin (8). Although this enzyme is not rate-limiting, it may ultimately control the maximum amount of melatonin that can be produced each night.

Pharmacology and Physiology

Pineal Adrenergic Receptors

Norepinephrine (NE) is the neurotransmitter released by postganglionic sympathetic neurons. NE stimulates ß-1-adrenergic receptors on the pinealocytes, resulting in activation of *N*-acetyltransferase (NAT) (9). NAT activation is potentiated by stimulation of α1-adrenergic stimulation of the pinealocytes (10). There are also α2-adrenergic autoreceptors on the postganglionic sympathetic neurons, which decrease melatonin production when stimulated by NE (11). Perhaps unique in the autonomic nervous system, there is no dual parasympathetic innervation opposing sympathetic control of pineal melatonin production; however, melatonin levels do not increase with general "fight or flight" sympathetic stimulation. Regulation of melatonin is tightly controlled by the SCN, which appears to be active during the day (12), thereby exerting inhibition of an otherwise always "on" signal in the PVN for increasing melatonin production (13). Nevertheless, melatonin production is a reliable measure of sympathetic and noradrenergic activity in the pineal gland. Consistent with the catecholamine hypothesis of affective disorders, melatonin production is greater in manic than in depressive states in bipolar patients (14 ,15), although the jury is out as to whether or not melatonin production is decreased in unipolar depression (16).

Pineal Pharmacology

Drugs that affect NE and its receptors can change melatonin levels. ß-Blockers reduce melatonin production (17). In humans, changes in circulating levels in NE must be extreme if melatonin levels are to be increased, such as resulting from high-altitude marathon races; however, tricyclic antidepressants reliably increase melatonin production. Drugs that stimulate and block α2-adrenergic receptors also have the predictable effects on melatonin production. For example, clonidine decreases melatonin production in humans (11), whereas yohimbine increases it (18). Benzodiazepines are also reported to decrease melatonin production. Changes in melatonin duration probably are not as important in humans, who lack seasonal rhythms. Drugs that affect melatonin production generally lower or raise the entire nighttime profile symmetrically; therefore, a circadian phase shift is less likely to result. This concept is best understood after reviewing how melatonin feeds back onto the SCN in order to shift circadian phase (see the following).

SHIFTING CIRCADIAN PHASE USING BRIGHT LIGHT

Part of "129 - Circadian Phase Sleep and Mood Disorders "

Intensity of Light

Acute Suppression of Nighttime Melatonin Production

One of the most remarkable effects of light is suppression of nighttime melatonin production (19). However, scientists in the late 1970s had concluded that humans lacked this and other chronobiologic responses to light. This erroneous thinking was based on temporal isolation studies that showed that social cues were more effective than light in entraining human circadian rhythms (20) and on studies in which light failed to suppress melatonin production in humans (17 ,21 ,22). In both types of cases, ordinary-intensity room light was used. These negative findings caused some scientists to speculate that humans lacked the neural pathways for mediating chronobiologic effects of light.

However, in 1980 we reported that bright light could suppress melatonin production in humans (the brighter the light, the greater the effect) (23) (Fig. 129.1). One implication of this finding was that humans might have biological rhythms that were cued to sunlight and were relatively unperturbed by ordinary-intensity indoor light. A second implication was that bright artificial light might be substituted for sunlight, in order to experimentally, and perhaps therapeutically manipulate biological rhythms in humans.

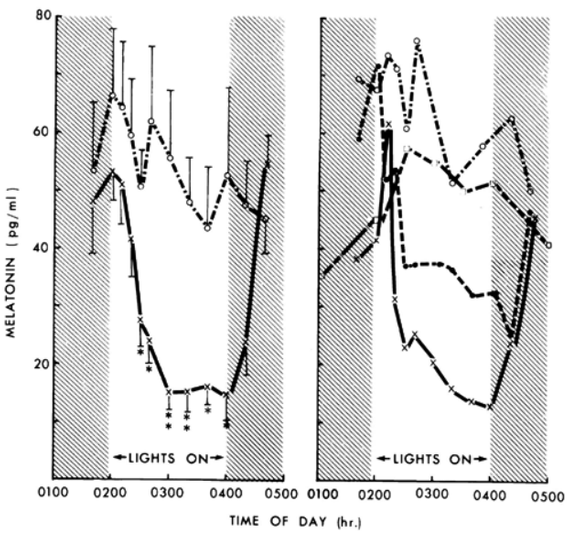


FIGURE 129.1. Left: Effect of light on melatonin secretion. Each point represents the mean concentration of melatonin (±standard error) for six subjects. A paired *t*-test, comparing exposure to 500 lux with exposure to 2,500 lux, was performed for each data point. A two-way analysis of variance with repeated measures and the Newman-Keuls statistic for the comparison of means showed significant differences between 0230 and 0400 (*, $P < .05$; **, $P < .01$). **Right:** Effect of different light intensities on melatonin secretion. The averaged values for two subjects are shown. Symbols: (o) 500 lux; (**x**) 2,500 lux; (●) 1,500 lux; and (□) asleep in the dark. From Lewy AJ, Wehr TA, Goodwin FK, et al. Light suppresses melatonin secretion in humans. *Science* 1980;210:1267–1269, with permission.

Winter Depression

Following our discovery of bright light suppression of melatonin production in humans (23), researchers began to treat depressed patients with bright light (24). We have been concentrating on treating winter depression, or seasonal affective disorder (SAD). We were unaware of the existence of SAD until Herb Kern contacted us, because of a history of an annual rhythm in mood changes: During the winter he became depressed, which he was able to document using a diary he had kept over for over 13 years (25). Since 2,000-light was effective in suppressing melatonin production, we assumed that this intensity would have other chronobiologic effects as well. Mr. Kern responded to a regimen of 2,000-lux light given between 6 and 9 AM and 4 and 7 PM (26).

We exposed him to bright light at these times, because animals tell time of year by the interval between the twilight transitions (6): Our thinking was that we could increase the length of his biological day by scheduling bright light as soon as he awakened, followed by another exposure ending 13 hours later. (Mr. Kern's depression would usually spontaneously remit in the spring.) These two studies (23 ,26) were critical in starting the field of bright light treatment, not only for SAD but for using bright light in other chronobiologic disorders as well (discussed in the following).

Wavelength of Light

With George Brainard (27), we have shown that the peak wavelength for suppressing melatonin production is 509 nm (blue-green light). Most white light sources have this wavelength, which is in the middle of the scotopic spectral distribution. Generally, we prefer to use regular fluorescent light. Rods, not cones, are most sensitive at this wavelength; however, the precise retinal photoreceptors that mediate chronobiologic effects of light have not yet been identified.

Timing of Light

The Light Phase Response Curve

The next step in the course of our work was to address the circadian phase-shifting effects of light. As mentioned in the preceding, Mr. Kern (and most of the other early SAD patients) was treated with bright light based on the idea of lengthening their winter photoperiod to one more typical of spring, the time of year when they spontaneously remitted (26 ,28). It occurred to us that we should have the same phase response curve (PRC) to light commonly found in other animals, except that humans would need bright light to most reliably demonstrate these effects (29). Hence, we postulated the human light PRC, which can be described as follows. A phase advance (shift to an earlier time) results from light exposure between the middle of the night and morning. A phase delay (shift to a later time) results from light exposure between the evening and middle of the night. These phase shifts are greatest in the middle of the night. During the day there are decreased responses to light. Rutger Wever, who previously was the driving force behind the importance of social cues, published a study in 1983, demonstrating that continuous bright light during the day had a more potent effect on the human circadian system than did ordinary-intensity light (30).

Phase-Typing Circadian Rhythm Disorders

Our much less elegant anecdotal report the same year indicated that bright light could be used to treat circadian phase disorders in humans (29); however, we specified that bright light should be confined either to the morning or evening. We proposed that there were two types of circadian disorders, the phase-delayed and phase-advanced types (Table 129.1). We also showed how bright light scheduled in the morning could be used to treat delayed sleep phase syndrome (DSPS) in order to provide a corrective phase advance and that evening bright light could be used to treat at least the early morning awakening of nonseasonal depressives in order to provide a corrective phase delay. However, the first person to treat this latter group of patients was Dan Kripke (24), who used a different approach, giving them bright light in the morning, not to cause a phase shift but

rather to illuminate a “critical interval.” Suffice it to say that no one in those early days, except Wever and our group, was thinking about the circadian phase-shifting effects of light. The following year we showed that—holding the sleep/wake cycle constant—we could shift the melatonin rhythm (a biological marker that we had proposed would be ideal for assessing circadian phase position in humans) by shifting the light/dark cycle (31 ,32).

Phase-Advanced Type	Phase-Delayed Type
ASPS	DSPS
East to west jet lag	West to east jet lag
Adjusting to night work	Readjusting to off-work
	Winter depression

TABLE 129.1. PHASE TYPING SLEEP AND MOOD DISORDERS

The Phase Shift Hypothesis for Winter Depression

When we proposed “phase typing” circadian disorders (29), we hypothesized that most people with SAD were of the phase-delayed type (33). It was our thinking that circadian rhythms drift later with the later dawn of winter and that this is the cue for some people to get depressed at this time of year. Accordingly, we hypothesized that the optimum time for bright light exposure was in the morning, which would provide a corrective phase advance. We further hypothesized that morning light would advance the circadian rhythms that were tightly coupled to the ECP, such as the melatonin rhythm, with respect to the sleep/wake cycle and its evoked rhythms, and therefore that any shift to an earlier sleep time should be held to a minimum. We also expressed concern about the possibility that too much morning light could overly phase advance these rhythms (34). Finally, we pointed out that there might be a small group of SAD patients who cued to dusk rather than dawn, and for whom evening bright light would be most antidepressant, producing a corrective phase delay (34). We also proposed an elaboration of the phase shift hypothesis (PSH): Typical patients are phase delayed, but not necessarily compared to normal controls, in that the phase delay could be ipsative (35); that is, we expected typical patients to be delayed when depressed in the winter compared to when they were euthymic. According to the nomenclature, if SAD patients are thought to have relatively long intrinsic circadian periods, they should have a relatively late dim light melatonin onset (DLMO) given either as clock time or as zeitgeber time (ZT) (that is, delayed relative to wake time). Indeed, as mentioned, it is thought that the DLMO of most SAD patients is delayed relative to the sleep/wake cycle (and therefore the ambient light/dark cycle).

Although some investigators were quick to embrace the PSH (such as David Avery, who underscored the significance of morning hypersomnia as a predictor of response to morning light) (36), many distinguished experts in the field advanced other ways of conceptualizing SAD. For example, the NIMH group published a paper in 1985 recommending 5 to 6 hours of 2,000 to 2,500 lux light in the evening (37). According to their “photon counting” hypothesis, light at any time of day should be antidepressant, as long as light of sufficient duration and intensity was used (38). Because people with SAD do not like to get up any earlier than they have to in the morning, evening was proposed to be the most convenient time for light treatment.

Both hypotheses received help from a review by the Terman group (39). Morning light was shown to be more antidepressant than evening light; however, evening bright light was shown to be more effective than evening dim light. It should be noted, though, that there was no control for the dim evening light condition; therefore, evening dim light could have been eliciting purely a placebo response. Significantly, the Terman group made the important suggestion that 10,000 lux could be used for a shorter duration than the 2,000- to 2,500-lux light that had been the previous standard.

Morning Versus Evening Light

Support for the PSH depends on the superiority of morning light for most patients. Some studies have shown that morning light is more effective than evening light (40 ,41); whereas other studies showed that they are equally effective (42). The former studies used a crossover design, whereas the latter studies used parallel groups. Critics of the PSH pointed out the advantages of the latter type of study design, whereas advocates of the PSH pointed out the advantages of the former compared to the latter. Some critics of the PSH also proposed that exposure to morning light prevented an antidepressant response to subsequent treatment with evening light (43). The Terman group also proposed a corollary to the PSH that could be construed as quite different from the original hypothesis: Light at any time of day should be antidepressant for SAD, as long as it does not produce a phase delay (44).

With the publication of three large studies by independent groups in 1998 (45 ,46 and 47), there is now general agreement that morning light is more antidepressant than evening light in the treatment of SAD. An “order” effect (43) does not seem to confound these studies, and morning light has been shown to be more effective than evening light in both parallel (45 ,46 and 47) and crossover comparisons (40 ,41 ,46). Morning light does not seem to prevent an antidepressant response to evening light (47). Moreover, evening light does not seem to be more antidepressant than a credible placebo control (45).

However, the superiority of morning light does not prove the PSH, because people could simply have a greater overall sensitivity to light at that time of day. This seemed to us to be an unlikely explanation, given that patients become more depressed when switched from morning to evening light, even after they have responded. Recently, however, the Terman group (48) as well as our group (49) found that the antidepressant response to morning light correlates with the amount of phase advance. We had previously shown this relationship with patients exposed to 30 minutes versus 120 minutes of morning light (50), which has the obvious disadvantage that patients would expect light of a

greater duration to be more antidepressant. The PSH continues to remain the most viable hypothesis for explaining why patients with SAD become depressed in the winter and for explaining how light is antidepressant in these patients.

It should be mentioned that even if the PSH is ultimately shown to be correct, it might explain only part of the response to light. As Charmane Eastman has shown, the placebo response is a major component to light treatment (51). Whether or not a specific mechanism for this can be found (e.g., an energizing effect) (52) remains to be determined.

Other Chronobiologic Hypotheses

Finally, four other chronobiologic hypotheses for SAD should be mentioned, none of which are mutually exclusive with the PSH. Martin Teicher has proposed that SAD patients are not stably entrained to the light/dark cycle (53), which in many respects is similar to the PSH. Domien Beersma has proposed that people with SAD are supersensitive to light (54). Accordingly, these people might delay in the winter in response to ordinary-intensity room light in the evening that would not be sufficiently bright to phase delay normal controls. The jury remains out on this hypothesis, which, in any event, is not inconsistent with the PSH. Also, Thomas Wehr has found that the melatonin duration (the time interval between the melatonin onset and the melatonin synthesis offset, or SynOff) expands in the winter in SAD patients but not normal controls (55); however, he also finds an overall delay in the melatonin rhythm in SAD patients in the winter compared to the summer, particularly in the SynOff. A third hypothesis, suggested by Charles Czeisler, is that SAD patients have diminished circadian amplitude when depressed in the winter (56). Czeisler has not done much testing of his hypothesis, and other investigators have not found much support for it (57 ,58).

The DLMO as a Marker for Circadian Phase Position

History

In the early 1980s, we thought that plasma melatonin sampled every 30 to 60 minutes might be able to show differences in circadian phase position between individuals and to monitor the phase-shifting effects of bright light (31 ,32). We further proposed that only one night of sampling is needed, indeed, only during the evening, so as to determine the time when melatonin levels begin to increase. We recommended that subjects be studied under dim light, so as not to suppress the rise in melatonin levels. The DLMO continues to be a useful marker for circadian phase position (59 ,60 and 61). We have also recommended that the time when melatonin levels begin to fall, the synthesis offset, or SynOff, is another useful marker for the ECP (Fig. 129.2) (61); however, the SynOff requires sample collection during sleep.

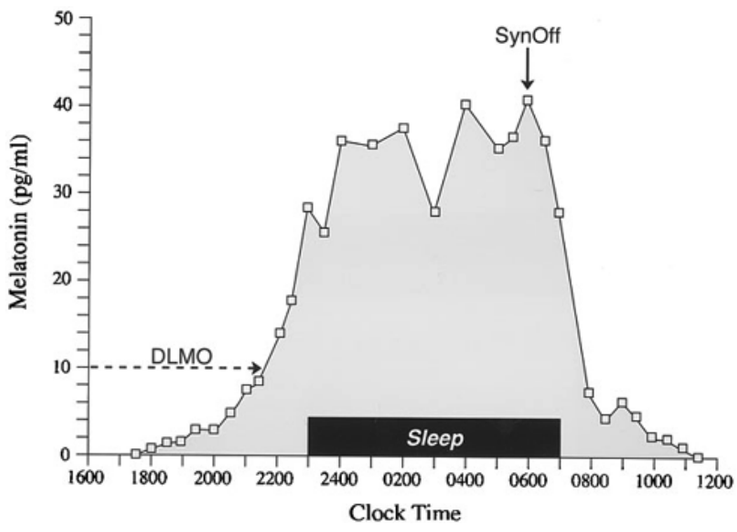


FIGURE 129.2. Representative overnight endogenous melatonin profile, showing the dim light melatonin onset (DLMO), the threshold when melatonin levels cross 10 pg/mL and continues to rise, and the melatonin synthesis offset (SynOff), indicating the beginning of the end of melatonin production.

Current Recommendations

Currently, we are suggesting that the light be so dim that subjects are not able to read without the aid of small book lamps. We further suggest that dim light begin 1 hour before blood sampling (61). The same recommendations apply to salivary collections. It is important not to put the subject in absolute darkness too early in the day, because this might cause a significant phase advance in the ECP (62).

The DLMO is now being employed extensively in research labs, and the salivary DLMO is beginning to be used in clinical settings as well (63 ,64). The DLMO has several advantages. It appears to be relatively free of noise and is easily obtained compared to other techniques for assessing circadian phase position. Outpatient determination of the DLMO usually does not interfere very much with sleep. Saliva can be collected at home.

Low Melatonin Producers

There are a couple of caveats to keep in mind. Very low or very high producers of melatonin may have an artifactually late or early DLMO, respectively (61). When there is a large range in melatonin amplitude between individuals, the DLMO can be adjusted algorithmically, or a lower threshold than the usual 10 pg/mL can be used (the 2 pg/mL DLMO is minimally affected by amplitude). Thomas Wehr has recently shown that the melatonin duration does not change across the year in normal subjects (55); however, in

SAD patients he has reported a greater delay in the SynOff than in the melatonin onset in the winter compared to the summer. Thus, we may be underestimating the amount of delay in SAD patients when just using the DLMO; therefore, for studies of circadian phase, measuring the DLMO may miss a finding that can be demonstrated using the rest of the melatonin profile. However, in most situations, measurement of the DLMO should suffice.

Circadian Amplitude

It is not clear if the overnight melatonin profile is a good marker for the amplitude of its endogenous circadian pacemaker. Furthermore, it is not clear if circadian amplitude is as important as circadian phase, in that an amplitude disturbance has yet to be shown. Moreover, no technique has been shown to enhance circadian amplitude or to reliably diminish it. The jury is out over whether or not suppressing amplitude is important for bright light to cause phase shifts (65 ,66).

The Melatonin Phase Response Curve

Blindness and Constant Dark Conditions

In 1983, Redmond, Armstrong, and Ng showed that a free-running mammal (in this case, rats) could be entrained to a daily dose of melatonin (67). We had been interested in the possibility of using melatonin to entrain the free-running rhythms of totally blind people. There appear to be at least three types of blind people: normally entrained, entrained at an abnormal phase, and free-running [blind free-runners (BFRs)] (68 ,69). BRFs are usually without any light perception, subjective or objective. Of the million or so legally blind in the United States, about 200,000 are totally blind. At least half of them are probably BFRs (70). BFRs generally have a recurrent sleep disorder that occurs when their sleep propensity rhythm drifts 12 hours out of phase with their preferred sleep time. On these days, they have insomnia and have to fight hard to resist urge to sleep during the day.

The technique we use to determine the free-running period of BFRs is to determine the melatonin onset on multiple occasions. We call this the multiple melatonin onset test (MMOT). In a typical BFR, the melatonin onset will be several minutes later each day and several hours later over a few weeks.

Although we were the first group to give melatonin to the blind (71), other investigators preceded us in giving melatonin to sighted people. Jo Arendt was the first to use melatonin in the treatment of jet lag (72); she also was the first to study its phase-shifting effects on the endogenous melatonin rhythm (73). Bruno Claustrat's group (74) also did some early work this area.

Another reason why we started with blind people is that we were concerned that the light/dark cycle would prevent melatonin's phase-shifting effects. The data on sighted subjects were intriguing, but were neither robust nor consistent (75). However, in the blind we were able to demonstrate quite reliable phase shifts using a dose of 5 mg of melatonin (71 ,76 ,77); thus encouraged, we were then ready to study sighted people, even though we had not yet achieved our goal of reliably entraining blind people to a daily dose of melatonin.

Sighted People

In sighted people, we reduced the dose to .5 mg, which produces melatonin levels of the same order of magnitude that occur physiologically. As opposed to previous studies of melatonin (which used higher doses and gave melatonin in the late afternoon or evening), we administered melatonin at different times. In each trial we gave melatonin on four consecutive days, and the results were the first unequivocal demonstration of both phase delays and phase advances, as well as the first description of the melatonin phase response curve (PRC) in humans (78 ,79 ,80 ,81 and 82). The melatonin PRC has now been replicated, by us (83) and two other groups (84 ,85); the earlier melatonin is given in the advance zone, the greater is the phase advance (however, less is known about the delay zone).

The melatonin PRC appears to have an advance zone of about 12 hours' duration, a delay zone of about 12 hours' duration, and it appears to be about 12 hours out of phase with the light PRC. That melatonin causes phase shifts opposite to those of light should not be surprising, because melatonin appears to be a chemical signal for darkness.

Circadian Time

Internal body clock time can be assessed using a variety of marker rhythms. We prefer the DLMO. On average, the DLMO occurs 14 hours after wake time. Because "lights on" is, by tradition, designated circadian time (CT) 0, we use the DLMO as CT 14. That is, no matter what clock time the DLMO occurs, its circadian time is CT 14.

On average the crossover times of the melatonin PRC appear to be at about 1 PM and about 1 AM. In circadian time, these are CT 6 and CT 18, respectively (83). That is, advance responses are obtained when melatonin is given between CT 6 and CT 18, and delay responses are obtained when melatonin is given between CT 18 and CT 6. When comparing the phase of the melatonin PRC to the light PRC, we use the beginning of the stimulus. Accordingly, the Czeisler PRC (which uses 5 hours of bright light) places the crossover times in the middle of the day at CT 6 and in the middle of the night at CT 18, according to these same conventions recommended by the PRC Atlas (86 ,87). This can be confusing, because the light PRC's crossover time in the middle of the night is also thought to occur at

the temperature minimum (86 ,88 ,89 ,90 and 91), which is usually just a few hours before wake time (CT 0).

Phase Relationship between the Light and Melatonin PRCs

In any event, light exposure during the interval between CT 6 and CT 18 causes phase delays, and during the interval between CT 18 and CT 6 causes phase advances; therefore, the light and melatonin PRCs are 12 hours out of phase with each other (Fig. 129.3 and Fig. 129.4). If the midpoint of the stimulus is used as the phase reference instead, the melatonin and light PRCs are still about 12 hours out of phase with each other. The times are given both as average clock times and as CTs. CT is more accurate if it can be calculated using an internal phase marker, such as the DLMO, which is designated CT 14. CT can be more easily but more roughly estimated using sleep onset as CT 16, or preferably, sleep offset as CT 0. The clock times in the figures below are averages, based on the presumption of a habitual wake time of 7 AM. The crossover times could be a few hours later or earlier, depending on how much the habitual wake time (not the wake time on a particular day) differs from 7 AM. For example, a person who habitually awakens at 6 AM will likely have crossover times at noon and midnight, instead of 1 PM and 1 AM.

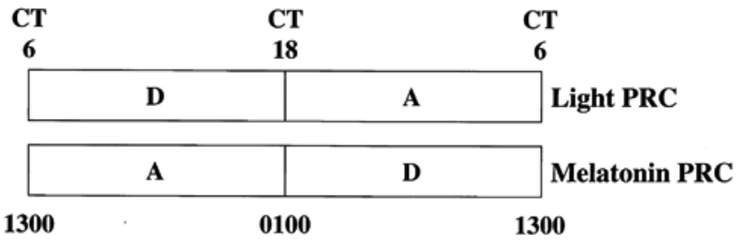


FIGURE 129.3. Schematic diagram of the light and melatonin phase response curves (PRCs). The melatonin PRC appears to be about 12 hours out of phase with the light PRC. The crossover times separating the advance (A) and delay (D) zones appear to be at about CT 6 and CT 18, on average about 1 PM and 1 AM, respectively, for people who awaken at 7 AM. The endogenous melatonin profile usually extends from about CT 14 to about CT 1. Adapted with permission from Lewy AJ, Bauer VK, Ahmed S, et al. The human phase response curve (PRC) to melatonin is about 12 hours out of phase with the PRC to light. *Chronobiol Int* 1998;15:71–83.

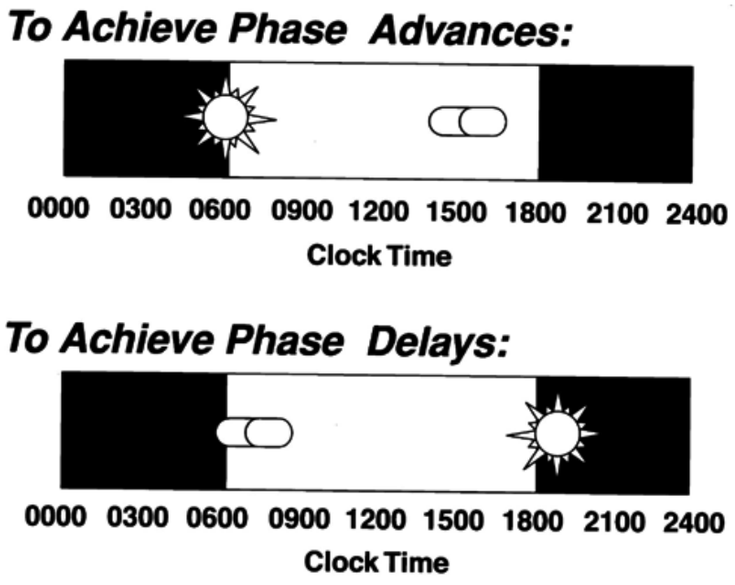


FIGURE 129.4. Phase-shifting effects of light and exogenous melatonin, assuming a normal circadian phase. To cause a phase advance, light should be administered in the morning and melatonin in the afternoon/evening. To cause a phase delay, light should be given in the evening and melatonin should be taken in the morning. The timing of light and melatonin administration is best done with reference to circadian time rather than clock time. Adapted with permission from Lewy A, Sack R. The role of melatonin and light in the human circadian system. In: Buijs R, Kalsbeek A, Romijn H, et al, eds. *Progress in brain research, vol. 111. Hypothalamic integration of circadian rhythms*. Amsterdam: Elsevier, 1996:205–216.

Circadian Time and Zeitgeber Time

A somewhat confusing issue is how to use the DLMO as a marker for internal body clock time. When scheduling bright light and melatonin using the light and melatonin PRCs, we use the CT, which as mentioned in the preceding, is designated CT 14. Therefore, the advance zone of the melatonin PRC begins 8 hours before the DLMO and ends 4 hours after the DLMO. Also mentioned, the DLMO occurs on average 14 hours after wake time. However, if a person has a relatively short intrinsic circadian period, he or she will be entrained by the light/dark cycle in a relatively advanced phase position, and if a person has a relatively long intrinsic circadian period, he or she will be entrained by the light/dark cycle in a relatively delayed phase position. When using the DLMO to provide information about the intrinsic period of the ECP, we speak of the zeitgeber time (ZT) of the DLMO. For example, if a person has a DLMO that is 13 hours after wake time (or the time of first light exposure, the main entraining agent of the ECP), the phase angle of entrainment, or zeitgeber time of the DLMO, is ZT 13. A person with a DLMO of ZT 13 probably has a relatively short intrinsic period. However, the CT of the DLMO (when operationally defined as the crossing of the 10 pg/mL plasma melatonin threshold) is always designated CT 14. In other words, the same DLMO clock time may be ZT 13 and CT 14. This is discussed further in the section on advanced and delayed sleep phase syndromes .

SAD UPDATE

Part of "129 - Circadian Phase Sleep and Mood Disorders "

Confirmation of the PSH for SAD?

The use of melatonin has allowed us to conduct a further test of the PSH for winter depression in order to answer

the following question: Does the correlation between the antidepressant response and the phase-advancing effect of morning light indicate a causal relationship? First, we did a pilot study. Five patients were given placebo in the afternoon and five were given melatonin, which was administered in two divided doses of .125 mg at CT 8 and CT 12. After 2 weeks, there was a significant improvement in the melatonin group compared to the placebo group (92).

These findings led to the study of 81 patients, divided into three groups. One group received low doses of melatonin in the afternoon and evening for 3 weeks. A second group received the same dosing regimen in the morning. There is also a placebo group. Interestingly, the antidepressant response appears to be related to the amount of phase advance (49). Therefore, the phase advance in the ECP appears to be the best-defined mechanism of action for the antidepressant effect of morning light in winter depression.

Clinical Implications of the PSH

We have cautioned (34) that it is possible to overly phase advance the body clock. A phase advance of about 1.5 hours relative to the sleep/wake cycle seems to be the optimal amount (48). Although the duration can be reduced after the patient has responded in 1 or 2 weeks, some patients will not comply with the 1 to 2 hours of bright light exposure immediately on awakening, because they do not want to get up too early. Furthermore, any advance in sleep time should be minimized, since this will retard the antidepressant response to advancing the ECP, because—according to the PSH—the ECP needs to be advanced with respect to the timing of the sleep bout (93 ,94).

Administration of melatonin (.5 mg) in the afternoon will cause a phase advance (82). This can be achieved while minimizing its soporific side effect, by using very low doses (.75 to 1.25 mg) given three or four times every 2 to 3 hours beginning 7 to 8 hours after habitual wake time. The addition of melatonin will reduce the need for an inconveniently long duration of morning bright light exposure by providing some additional phase advance. Those patients who do not get sleepy on melatonin may be able to take a sufficient dose of melatonin so that they do not require any morning bright light. Many patients with SAD seem to be unusually sensitive to the soporific effect of melatonin and will require bright light in combination with melatonin to achieve a therapeutic phase advance.

It takes less of a phase-resetting agent to maintain a certain circadian phase position than to initiate it. Hence, once a patient has responded, the duration of bright light and/or the dose of melatonin can be reduced. Patients will need to be treated until the photoperiod lengthens in the spring. When the patient's habitual wake time is occurring about 30 minutes past dawn, obtaining 15 to 30 minutes of outdoor light immediately on awakening is sufficiently therapeutic. At this time of the year, he or she is probably beginning to spontaneously remit. By the way, another way to cause a phase advance with respect to sleep time is to use a dawn simulator set to start slowly increasing light intensity a few hours before awakening (95 ,96). According to the light PRC, even relatively low intensity of light (diminished further because of closed eyelids) can cause a phase advance if given in the middle of the night.

In order to compare the phase-shifting effects of light and melatonin, ideally each treatment should be optimized. Absent this, we compared 2 weeks of 2,500 lux light at 6 AM to 8 AM or 7 PM to 9 PM in our largest light treatment of winter depression study (46). Phase shifts owing to light were of the same order of magnitude as phase shifts after 3 weeks of a divided dose of .225 to .3 mg of melatonin (which produces high but physiologic levels) used in our ongoing melatonin-treatment study described in the preceding (Table 129.2).

	Phase Advances (Hours ±)	Phase Delays (Hours ±)
Bright light	1.40 (± 0.21)	1.12 (± 0.16)
Melatonin	1.03 (± 0.14)	0.20 (± 0.12)

TABLE 129.2. PHASE SHIFTS WITH BRIGHT LIGHT AND MELATONIN ADMINISTRATION

We think it unlikely that a third week of treatment substantially increases the phase shift. This table indicates the phase-shifting effects of light and melatonin are of the same order of magnitude, although each treatment should be optimized for the most useful comparison.

OTHER CIRCADIAN PHASE DISORDERS

Part of "129 - Circadian Phase Sleep and Mood Disorders "

Shift Workers

The magnitude of the phase shifts in the melatonin PRC are smaller than those of the light PRC. However, subjects in the light PRCs studies had their sleep/wake cycles either shifted 12 hours before bright light exposure was scheduled or free to shift at will (86,88,89 and 90). In the melatonin PRC study, the sleep/wake cycle was held constant, and hence the ambient light/dark cycle also was held constant (82,83).

When we gave .5 mg of melatonin just before bedtime to night workers whose sleep/wake cycles were shifted about 12 hours, we found phase shifts of the same order of magnitude as those obtained with bright light (97). Apparently, holding the light/dark cycle constant diminishes the phase-shifting effects of melatonin (as it does light). Nevertheless, bright light seems to be a somewhat more robust phase-resetting agent than melatonin, although melatonin is of course much more convenient. In all likelihood, greater

phase shifts can be achieved in sighted people when melatonin and bright light are combined.

Jet Lag

The first use of melatonin in humans was to treat jet lag (98). There are surprisingly few studies in this area, however. Sunlight exposure at destination can be scheduled according to the nomogram we published in 1984 (99) (Fig. 129.5). This applies to obtaining (and in some cases avoiding) bright light for the day of arrival. The schedule for subsequent days can be arrived at by assuming a 2- to 3-hour phase shift per day and looking up the instructions for a crossing of the number of time zones. For example, if you have flown to Israel from Portland, OR, follow the directions for 10 time zones on the first day in Israel. For the second day in Israel, follow the directions for crossing eight time zones. Melatonin can also be helpful in the treatment of jet lag. We are currently developing guidelines for the optimal use of melatonin in the treatment of jet lag.

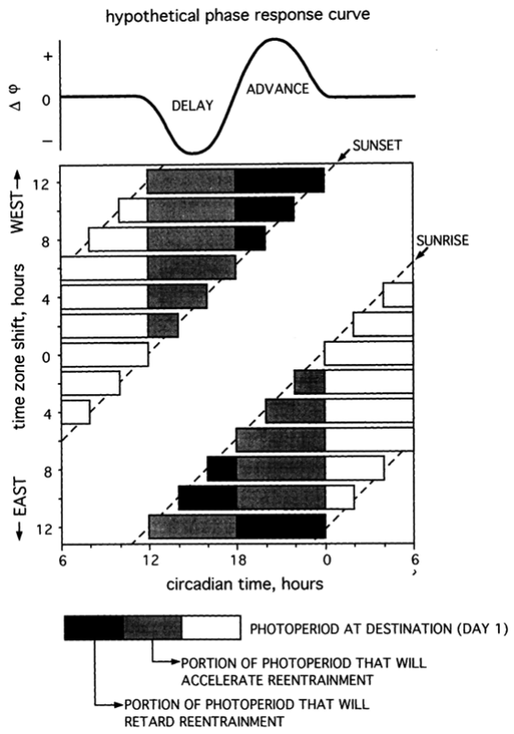


FIGURE 129.5. Proposed times for when bright light exposure should occur and when bright light exposure should be avoided the first few days after transmeridional flight. For example, after a 2-hour west-to-east trip, bright light exposure should begin at dawn and should be (optimally) 2 hours in duration. In another example, after a 10-hour west-to-east trip, however, bright light exposure should be *avoided* until 4 hours after sunrise and should occur (optimally) for 6 hours. From Daan S, Lewy AJ. Scheduled exposure to daylight: a potential strategy to reduce “jet lag” following transmeridian flight. *Psychopharmacol Bull* 1984;20:566–568, with permission.

Entrainment of BFRs

Following the first demonstration of phase shifting the ECP of a BFR (71), others and our group have found cases of varying degrees of certainty in which a daily dose of melatonin appeared to cause entrainment (100,101 and 102). In some instances, only the sleep/wake cycle appeared to be entrained (103,104 and 105).

We estimate that there are at least 100,000 totally blind people in the United States who have periodic insomnia. We have recently discovered a way to entrain most of these people (101,106). A dose of 10 mg given within an hour of preferred bedtime should in all likelihood eventually entrain most of them. Entrainment occurs in just a few days or a few weeks by initiating treatment when the MO is occurring around bedtime. Once the free-running clock of the blind person has been “captured,” the maintenance dose can be decreased to as low as .5 mg. Ongoing work in our laboratory is investigating the possibility of moving the melatonin dose earlier, to 7 to 13 hours after habitual wake time, so as to provide a typical phase angle of entrainment. In three BFRs who had pretreatment circadian periods less than 24.4 hours, we were able to capture their circadian rhythms with a *de novo* bedtime dose of .5 mg. Therefore, it may be possible to entrain people initially with .5 mg, particularly if their free-running periods are not much greater than 24 hours. Indeed, the only person who failed to entrain to the 10-mg dose had the longest free-running period of our group (24.9 hours).

BFRs are perhaps ideally suited for phenotyping people according to their intrinsic circadian period for clock gene studies, particularly bilaterally enucleated people in whom there is no chance for ocularly mediated effects of light on the circadian system. Of course, this presupposes minimal, if any influence of behaviorally related zeitgebers (BRZs) (107,108) or nonocular light (109). If these other possible modes of entrainment are shown to be negligible, then entrained BFRs with one or two eyes are probably still sensitive to the ocularly mediated light zeitgeber, even though they have no conscious light perception or any objective sign of light response, such as the melatonin suppression test, which we developed in sighted people (23,110,111,112,113 and 114) and has also been recommended for blind people (107). However, until we have ruled out entrainment by ocularly mediated light in what are thought to be totally blind people, we do not recommend the use of the melatonin suppression test, which we recently have come to think may risk desensitizing the few remaining photoreceptors that may have been sufficiently

sensitive to mediate entrainment. Furthermore, the melatonin suppression test does not seem to be very useful clinically, because it does not discriminate BFRs from entrained blind people very well. These issues cannot be resolved until the entrainment effects of BRZs (107,108,115) or nonocular light (109) are established or ruled out.

Are Nonseasonal Affective Disorders Chronobiologic?

Daniel Kripke has done more work in this area than anyone else (24). We do not think that bright light has the same robust antidepressant effect in nonseasonal depression as it has in SAD, and the jury is out as to whether it works better than placebo. However, if a patient has a circadian rhythm component to his or her affective disorder, such as early morning awakening or morning hypersomnia, then melatonin and/or bright light can be used to shift sleep to a more desirable time. Whether or not correcting the phase disturbance improves the remaining symptoms is not known at the present time.

Melatonin in Young and Elderly People

There are anecdotal and testimonial reports that melatonin improves the sleep of children with ADHD and adults with Alzheimer disease. A study has been done in elderly people indicating that benzodiazepines can be reduced or eliminated with concomitant melatonin administration (116). Although we might want to use lower doses of melatonin in young and elderly people, melatonin appears to be reasonably safe in these populations as long as a physician is monitoring them. Indeed, to date I know of no reports published in the scientific literature of serious irreversible side effects as the result of taking melatonin, and certainly millions of people have been doing so for the past several years. Nevertheless, there is a continued need for physician monitoring of melatonin usage, particularly when taken every day. However, we do not expect this to be a very common effect of melatonin, given the fact that melatonin can have either progondal or antigonadal effects, depending on whether the species is a fall or spring breeder (6,117) and the fact that there is only a very slight seasonal rhythm in human fertility (118,119).

Advanced and Delayed Sleep Phase Syndromes

Appropriately timed bright light exposure and melatonin administration can be used to treat other circadian phase disorders. These include advanced sleep phase syndrome (ASPS) and delayed sleep phase syndrome (DSPS). Light treatment of these disorders has already been summarized (120), and we are currently developing recommendations for melatonin treatment. A remarkable report was published recently concerning a family with advanced sleep phase syndrome (121). The clock time of the DLMO was quite early, even with respect to sleep; that is, the ZT of the DLMO (the number of hours after wake time) can be calculated as 13.3. The intrinsic period of one of these subjects was studied in temporal isolation and was found to be 23.3 hours, one of the shortest, if not the shortest, ever recorded. Therefore, if the wake time and DLMO time are known, their interval should predict intrinsic period, and perhaps obviate the need for an arduous study under temporal isolation conditions. This should be of interest to those interested in phenotyping sighted people for clock gene studies.

MELATONIN AND CIRCADIAN PHASE DISORDERS: PAST, PRESENT, AND FUTURE

Part of "129 - Circadian Phase Sleep and Mood Disorders "

Past Skepticism about Melatonin

The melatonin fad of a few years ago has stimulated a number of scientists to make skeptical comments. Many of these concerns are well taken. There is no clinical evidence that melatonin is useful for anything other than phase shifting and sleep; however, some investigators have expressed skepticism even for these well-documented uses.

Charles Czeisler has perhaps most comprehensively articulated these criticisms (122). Czeisler contends that our melatonin PRC was not conducted under sufficiently controlled conditions, namely, that subjects were studied at home and under a variety of light intensities. However, we view this as a strength of the methodology: because the subjects lived mainly at home, the findings can be more directly applied to real-life situations. Furthermore, despite “noise” owing to uncontrolled light intensities, the data clearly describe a well-defined PRC, perhaps because of the large number of data points. Moreover, it is difficult to imagine a systematic confound in the study owing to melatonin’s soporific side effect (particularly because naps in the middle of the day do not cause phase shifts) (123); waking up at night to take a placebo capsule causes phase shifts, if any, opposite to those of melatonin (Lewy, in preparation).

Czeisler is also concerned that the Claustrat replication PRC is slightly different from ours. Claustrat used a 3-hour intravenous infusion of melatonin, whereas our .5-mg oral dose kept blood levels elevated for several hours (84). We have speculated that melatonin’s phase-shifting effects are optimal if the exogenous dose overlaps with the endogenous melatonin profile (75). This might explain why the intravenous dose given in the evening produced more of a phase advance than the one given in the afternoon, in that the afternoon dose did not overlap with the endogenous melatonin profile. Our PRC shows that melatonin’s phase-advancing effects increase as it is given earlier in the afternoon: even at this time, the .5-mg oral dose raises blood levels through the time of the melatonin onset.

However, Czeisler's main problem with melatonin as a useful phase-resetting agent has been the difficulty demonstrating its ability to entrain BFRs. This issue is now moot, given the definitive findings of two independent groups (102 ,106).

We agree with Czeisler that light is the most powerful phase-resetting agent. However, Czeisler thinks that light is an order of magnitude more powerful than light, whereas as indicated in the preceding (Table 129.2) we think that there is not that much difference between the two zeitgebers. In any event, melatonin is much more convenient than using light as a phase-resetting agent. Although our group has concentrated on melatonin as a secondary zeitgeber in humans, Czeisler and his co-workers have continued to pursue a longstanding interest in the activity/rest cycle, first as a primary (124 ,125), and then as a secondary (107 ,108), zeitgeber. The jury is out as to the strength of the activity/rest zeitgeber in humans.

The Function of Endogenous Melatonin Production

In many animals, the duration of nighttime melatonin production appears to be critically involved in the regulation of seasonal rhythms. However, this does not appear to be the case in SAD, although Tom Wehr has some intriguing data on this point (55). Although seasonal rhythms are not very robust in humans (118 ,119), we have most of the circadian rhythms found in other animals, and melatonin may play a role, however humble, in helping the light/dark cycle entrain the ECP. This function for melatonin is critically dependent on suppression of melatonin by bright light.

Light entrains the ECP (located in the SCN), which regulates all overt circadian rhythms, including the nightly increase in pineal melatonin production. As mentioned in the beginning of this chapter, melatonin feeds back onto the SCN and stimulates receptors causing phase shifts opposite to those of light (126 ,127 and 128). Sufficiently bright light at the twilight transitions suppresses melatonin production, causing the endogenous melatonin onset to occur later and the endogenous melatonin offset to occur earlier. Thus, melatonin is prevented from stimulating parts of the melatonin PRC that might counteract the phase shift resulting from light (Fig. 129.6). In this way, the phase-shifting effect of light is augmented by an indirect effect of light acting on suppressing melatonin production. For example, if a person who normally gets up at 7 AM goes outdoors at 5 AM during the summer, earlier sunlight exposure will stimulate more of the advance zone of the light PRC, so as to cause a phase advance. Simultaneously, the melatonin offset will occur 2 hours earlier, reducing stimulation of the delay zone of the melatonin PRC. The same thinking can be applied to changes in bright exposure in the evening.

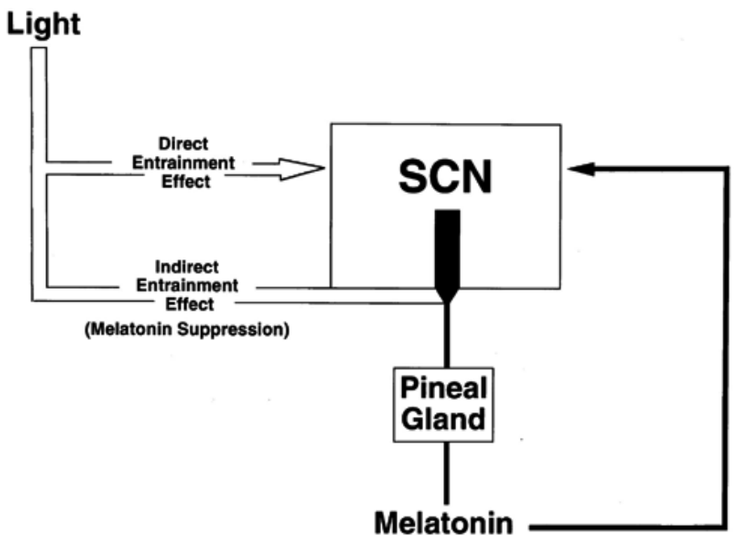


FIGURE 129.6. Schematic diagram of some of the relationships between nighttime melatonin production by the pineal gland, the light/dark cycle and the endogenous circadian pacemaker thought to be located in the hypothalamic suprachiasmatic nuclei (SCN). Acting on the SCN as described by the melatonin phase response curve (PRC) at any given time of the day or night, melatonin causes phase shifts opposite to those that light would cause *(opposing arrows)*. However, the suppressant effect of light pares the margins of the nighttime melatonin profile *(tapered vertical arrow)* and reduces endogenous melatonin's stimulation of the melatonin PRC at the day–night transitions. This second pathway for entrainment by light is particularly significant during shifts of the light/dark cycle. From Lewy AJ, Ahmed S, Jackson JML et al. Melatonin shifts circadian rhythms according to a phase-response curve. *Chronobiol Int* 1992;9:380–392, with permission.

Clearly, light is the major zeitgeber for entraining circadian rhythms. In most mammals, melatonin is used for conveying the time of the year (primarily for seasonal breeding and other photoperiodic effects) and not necessarily the time of day (6). The acute suppressant effect of light (23) is important in truncating the endogenous melatonin profile, and humans have retained the suppressant effect of light but are not really seasonal breeders. It may be that melatonin is primarily used for either circadian or seasonal time keeping, and that this difference might distinguish humans from nonhuman primates. In other words, perhaps primates use melatonin either for telling the time of the year or the time of the day, but not both. For example, in a species of primates that has a seasonal breeding cycle, melatonin has not been shown to have circadian phase-shifting effects (129). Interestingly, activity is an effective zeitgeber in at least one species of primates, which has a seasonal breeding cycle (130). A corollary of this hypothesis is that a species that does not use melatonin for telling time of the day uses activity as the secondary circadian zeitgeber. Humans, who do not appear to use activity as the secondary zeitgeber, use melatonin as one. In other words, light is the primary zeitgeber, and either melatonin or activity is the secondary zeitgeber, depending on whether or not melatonin is used for cueing seasonal rhythms.

Once again, there is no question that light is the primary zeitgeber. Light regulates the melatonin circadian rhythm

in two ways: Light entrains the SCN and acutely suppresses melatonin production. However, in humans melatonin appears to feed back onto the SCN to act as a secondary zeitgeber.

FUTURE DIRECTIONS

Part of "129 - Circadian Phase Sleep and Mood Disorders "

Optimal dosing of melatonin will depend on minimizing its soporific side effect, while maximizing its phase-shifting effects. This may entail using a low-dose sustained-release formulation to smooth out any sharp spikes in melatonin levels that appear to cause sleepiness in some people. The sustained-release formulation also has the advantage of providing continuity between exogenous levels from a low dose and the endogenous melatonin profile.

Another useful product that we might look forward to is a delayed-release sustained-release formulation that can be taken at bedtime to conveniently produce increases in melatonin throughout the night, beginning shortly after CT 18, so as to selectively stimulate the delay zone of the melatonin PRC. It may also be desirable to continue a low level of melatonin until the early afternoon in order to enhance a phase delay.

Clearly, shifting the sleep/wake (and consequently the ambient light/dark) cycle enhances the phase-shifting effects of melatonin. We have also speculated that placing a person in darkness when exogenous melatonin levels are increased may also enhance melatonin's phase-shifting effects. This needs to be more extensively tested.

Finally, we need more field studies of melatonin, not just in jet lag but with shift workers as well. One attempt at this was star-crossed (131). Our advice was not taken in its entirety, and the study design had several serious flaws in experimental design, for example, a failure to induce much jet lag in the placebo group after the first night's sleep at destination (most subjects expected to be an active treatment) and an absence of a circadian phase marker.

In some people, melatonin may be used as a mild sleep-promoting agent. The prospects for using melatonin and bright light to treat circadian phase disorders are even better. Most circadian phase disorders are relatively straightforward, and their treatment is based on the light and melatonin PRCs.

SAD is more complicated, although it is far and away the disorder most often treated with portable bright light fixtures. There appear to be two types of SAD patients (31): The typical patient with SAD is phase delayed (and complains of morning hypersomnia); the atypical patient is phase advanced (these people often report a history of getting up early year round, even on the week-ends, and often wanting to go to bed much earlier in the winter than summer). The phase disturbance in SAD is with respect to sleep as well as with respect to clock time. Therefore, there is an internal phase-angle disturbance in SAD. Sleep time should be held constant in these patients while their other circadian rhythms are shifted into the correct relative phase position with sleep. Even a confirmed PSH, however, raises new questions:

1. Which circadian rhythms tightly coupled to the ECP must be out of phase with which processes tied to sleep in order to trigger a depression each winter?
2. Does light have a specific antidepressant effect other than phase shifting?
3. How can we apply what we have learned about SAD to other affective disorders?

ACKNOWLEDGMENTS

Part of "129 - Circadian Phase Sleep and Mood Disorders "

We wish to thank the nursing staff of the OHSU General Clinical Research Center and to acknowledge the assistance of Vance K. Bauer, Hillary A. Bish, Victoria Chamberlin, and Neil R. Anderson. Supported by Public Health Service research grants MH40161 (AJL), MH00703 (AJL), MH55703 (AJL), M01 RR00334 (OHSU GCRC), and a National Alliance for Research on Schizophrenia and Affective Disorder Established Investigator Award (AJL). Dr. Lewy is co-inventor on several US melatonin use patents held by Oregon Health Sciences University that are currently not licensed to any company.

REFERENCES

1. Moore RY, Eichler VB. Loss of circadian adrenal corticosterone rhythm following suprachiasmatic lesions in the rat. *Brain Res* 1972;42:201–206.

2. Stephan FK, Zucker I. Circadian rhythms in drinking behavior and locomotor activity of rats are eliminated by hypothalamic lesions. *Proc Natl Acad Sci USA* 1972;69:1583–1586.

3. Moore RY, Lenn NJ. A retinohypothalamic projection in the rat. *J Comp Neurol* 1972;146:1–14.

4. Ariëns-Kappers J. The development, topographical relations and innervation of the epiphysis cerebri in the albino rat. *Z. Zellforsch Mikrosk Anat* 1960;52:163–215.

5. Illnerová H. Melatonin in rat pineal gland and serum; rapid parallel decline after light exposure at night. *Neurosci Lett* 1978:189–193.

6. Goldman BD, Darrow JM. The pineal gland and mammalian photoperiodism. *Neuroendocrinol* 1983;37:386–396.

7. Klein DC, Moore RY. Indole metabolism in the pineal gland: a circadian rhythm in N-acetyltransferase. *Science* 1970;169:1093–1095.

8. Axelrod J, Weissbach H. Enzymatic O-methylation of N-acetylserotonin to melatonin. *Science* 1960;131:1312.

9. Axelrod J, Zatz M. The B-adrenergic receptor and the regulation of circadian rhythms in the pineal gland. In: Litwack G, ed. *Biochemical actions of hormones*. New York: Academic Press; 1977:249–268.

10. Klein DC, Sugden D, Weller JL. Postsynaptic alpha-adrenergic receptors potentiate the B-adrenergic stimulation of pineal serotonin N-acetyltransferase. *Proc Natl Acad Sci USA* 1983;80:599–603.

11. Lewy AJ, Siever LJ, Uhde TW, et al. Clonidine reduces plasma melatonin levels. *J Pharm Pharmacol* 1986;38:555–556.

12. Schwartz WJ. Understanding circadian clocks: from c-Fos to fly balls [review]. *Ann Neurol* 1997;41:289–297.

13. Pickard GE, Turek FW. The hypothalamic paraventricular nucleus (PVN) mediates the photoperiodic control of reproduction but not the effects of light on the circadian rhythm activity. *Neurosci Lett* 1983;43.

14. Lewy AJ, Wehr TA, Gold P, et al. Plasma melatonin in manic-depressive illness. In: Usdin E, Kopin IJ, Barchas J, eds. *Catecholamines: basic and clinical frontiers*. New York: Pergamon Press, 1979:1173–1175.

15. Lewy AJ, Wehr TA, Goodwin FK. Melatonin secretion in manic-depressive illness. In: Obiols J, Ballus E, Pujol J, eds. *Biological psychiatry today*. Amsterdam: Elsevier, 1979:563–565.

16. Rubin RT, Heist EK, McGeoy SS, et al. Neuroendocrine aspects of primary endogenous depression. XI. Serum melatonin measures in patients and matched control subjects. *Arch Gen Psychiatry* 1992;49:558–567.

17. Wetterberg L. Melatonin in humans: physiological and clinical studies [review]. *J Neural Trans Suppl* 1978;13:289–294.

18. Kennedy SH, Gnam W, Ralevski E, et al. Melatonin responses to clonidine and yohimbine challenges. *J Psychiatry Neurosci* 1995;20:297–304.

19. Illnerová H, Vanêcek J. Response of rat pineal serotonin N-acetyltransferase to one min light pulse at different night times. *Brain Res* 1979;167:431–434.

20. Wever RA. *The circadian system of man. Results of experiments under temporal isolation*. New York: Springer-Verlag, 1979.

21. Arendt J. Melatonin assays in body fluids. *J Neural Trans Suppl* 1978;13:265–278.

22. Akerstedt T, Fröberg JE, Friberg Y, et al. Melatonin excretion, body temperature, and subjective arousal during 64 hours of sleep deprivation. *Psychoneuroendocrinology* 1979;4:219–225.

23. Lewy AJ, Wehr TA, Goodwin FK, et al. Light suppresses melatonin secretion in humans. *Science* 1980;210:1267–1269.

24. Kripke DF. Photoperiodic mechanisms for depression and its treatment. In: Perris C, Struwe G, Jansson B, eds. *Biological psychiatry*. Amsterdam: Elsevier, 1981:1249–1252.

25. Kern HE, Lewy AJ. Corrections and additions to the history of light therapy and seasonal affective disorder [letter]. *Arch Gen Psychiatry* 1990;47:90–91.

26. Lewy AJ, Kern HA, Rosenthal NE, et al. Bright artificial light treatment of a manic-depressive patient with a seasonal mood cycle. *Am J Psychiatry* 1982;139:1496–1498.

27. Brainard GC, Lewy AJ, Menaker M, et al. Effect of light wavelength on the suppression of nocturnal plasma melatonin in normal volunteers. *Ann NY Acad Sci* 1985;453:376–378.

28. Rosenthal NE, Sack DA, Gillin JC, et al. Seasonal affective disorder: a description of the syndrome and preliminary findings with light therapy. *Arch Gen Psychiatry* 1984;41:72–80.

29. Lewy AJ, Sack RL, Fredrickson RH, et al. The use of bright light in the treatment of chronobiologic sleep and mood disorders: the phase-response curve. *Psychopharmacol Bull* 1983;19:523–525.

30. Wever R, Polasek J, Wildgruber C. Bright light affects human circadian rhythms. *Eur J Physiol* 1983;396:85–87.

31. Lewy AJ, Sack RL, Singer CM. Assessment and treatment of chronobiologic disorders using plasma melatonin levels and bright light exposure: the clock-gate model and the phase response curve. *Psychopharmacol Bull* 1984;20:561–565.

32. Lewy AJ, Sack RL, Singer CM. Immediate and delayed effects of bright light on human melatonin production: shifting “dawn” and “dusk” shifts the dim light melatonin onset (DLMO). *Ann NY Acad Sci* 1985;453:253–259.

33. Lewy AJ, Sack RL. Phase typing and bright light therapy of chronobiologic sleep and mood disorders. In: Halaris A, ed. *Chronobiology and psychiatric disorders*. New York: Elsevier, 1987:181–206.

34. Lewy AJ. Treating chronobiologic sleep and mood disorders with bright light. *Psychiatry Ann* 1987;17:664–669.

35. Lewy AJ, Sack RL, Singer CM, et al. The phase shift hypothesis for bright light’s therapeutic mechanism of action: theoretical considerations and experimental evidence. *Psychopharmacol Bull* 1987;23:349–353.

36. Avery DH, Dahl K, Savage M, et al. Phase-typing seasonal affective disorder using a constant routine. *Soc Light Treatment Biol Rhythms Abst* 1989;1:14.

37. James SP, Wehr TA, Sack DA, et al. Treatment of seasonal affective disorder with light in the evening. *Br J Psychiatry* 1985;147:424–428.

38. Rosenthal NE, Sack DA, Carpenter CJ, et al. Antidepressant effects of light in seasonal affective disorder. *Am J Psychiatry* 1985;142:163–170.

39. Terman M, Terman JS, Quitkin FM, et al. Light therapy for seasonal affective disorder: a review of efficacy. *Neuropsychopharmacology* 1989;2:1–22.

40. Lewy AJ, Sack RL, Miller S, et al. Antidepressant and circadian phase-shifting effects of light. *Science* 1987;235:352–354.

41. Avery D, Khan A, Dager S, et al. Morning or evening bright light treatment of winter depression? The significance of hypersomnia. *Biol Psychiatry* 1991;29:117–126.

42. Wirz-Justice A, Graw P, Krauchi K, et al. Light therapy in seasonal affective disorder is independent of time of day or circadian phase. *Arch Gen Psychiatry* 1993;50:929–937.

43. Rafferty B, Terman M, Terman JS, et al. Does morning light therapy prevent evening light effect? *Soc Light Treatment Biol Rhythms Abst* 1990;2:18.

44. Terman M. Overview: light treatment and future directions of research. In: Wetterberg L, ed. *Light and biological rhythms in man*. New York: Pergamon Press, 1993:421–436.

45. Eastman CI, Young MA, Fogg LF, et al. Bright light treatment of winter depression. *Arch Gen Psychiatry* 1998;55:883–889.

46. Lewy AJ, Bauer VK, Cutler NL, et al. Morning versus evening light treatment of winter depressive patients. *Arch Gen Psychiatry* 1998;55:890–896.

47. Terman M, Terman JS, Ross DC. Timed bright light and negative air ionization for treatment of winter depression. *Arch Gen Psychiatry* 1998;55:875–882.

48. Terman M, Terman JS. Morning vs. evening light: effects on the melatonin rhythm and antidepressant response in winter depression. *Soc Light Treatment Bio Rhythms* 2000;12:1.

49. Lewy A, Bauer V, Bish H, et al. Antidepressant response correlates with the phase advance in winter depressives. *Soc Light Treatment Bio Rhythms Abstracts* 2000;12:22.

50. Lewy AJ, Sack RL, Singer CM, et al. Winter depression and the phase shift hypothesis for bright light’s therapeutic effects: history, theory and experimental evidence. *J Biol Rhythms* 1988;3:121–134.

51. Eastman CI. The placebo problem in phototherapy for winter SAD. *Soc Light Treatment Biol Rhythms Abst* 1989;1:36.

52. Lewy AJ, Sack RL. Melatonin physiology and light therapy. *Clin Neuropharmacol Suppl* 1986;9:196–198.

53. Teicher MH, Glod CA, Magnus E, et al. Circadian rest-activity disturbances in seasonal affective disorder. *Arch Gen Psychiatry* 1997;54:124–130.

54. Beersma DGM. Do winter depressives experience summer nights in winter? [letter]. *Arch Gen Psychiatry* 1990;47:879–880.

55. Wehr TA, Duncan WC Jr, Sher L, et al. SCN signal of change of season in seasonal affective disorder. *Soc Light Treatment Bio Rhythms* 2000;12:2.

56. Czeisler C, Kronauer R, Mooney J, et al. Biologic rhythm disorders, depression, and phototherapy: a new hypothesis. *Psychiat Clin N Am* 1987;10:687–709.

57. Avery D, Dahl K, Savage M, et al. Rectal temperature and TSH during a constant routine in winter depression. *Sleep Res* 1990;19:385.

58. Wirz-Justice A, Krüuchi K, Graw P, et al. Testing circadian rhythm hypotheses of winter depression in the constant routine protocol. *Neuropsychopharmacology* 1994;10:868S.

59. Lewy AJ, Sack RL, Singer CM, et al. Melatonin and Evolution. In: Hiroshige T, Honma K, eds. *Evolution of the circadian clock*. Sapporo: Hokkaido University Press, 1994:291–310.

60. Lewy AJ, Sack RL. The dim light melatonin onset (DLMO) as a marker for circadian phase position. *Chronobiol Int* 1989;6:93–102.

61. Lewy AJ, Cutler NL, Sack RL. The endogenous melatonin profile as a marker for circadian phase position. *J Biol Rhythms* 1999;14:227–236.

62. Van Cauter E, Moreno-Reyes R, Akseki E, et al. Rapid phase advance of the 24-h melatonin profile in response to afternoon dark exposure. *Am J Physiol* 1998;275:E48–E54.

63. Carskadon MA, Acebo C, Richardson GS, et al. An approach to studying circadian rhythms of adolescent humans. *J Biol Rhythms* 1997;12:278–289.

64. Sharkey KM, Eastman CI. Phase-advancing human circadian rhythms with melatonin. *Soc Light Treatment Bio Rhythms* 2000;12:9.

65. Jewett ME, Kronauer RE, Czeisler CA. Phase-amplitude resetting of the human circadian pacemaker via bright light: a further analysis. *J Biol Rhythms* 1994;9:295–314.

66. Shanahan TL, Czeisler CA. Melatonin rhythm observed during forced desynchrony: circadian and forced components. *Sleep Res* 1995;24A:544.

67. Redman J, Armstrong S, Ng KT. Free-running activity rhythms in the rat: entrainment by melatonin. *Science* 1983;219:1089–1091.

68. Lewy AJ. Human plasma melatonin studies: effects of light and implications for biological rhythm research. In: Birau N, Schloot W, eds. *Melatonin: current status and perspectives*. Oxford: Pergamon Press, 1981:397–400.

69. Lewy AJ, Newsome DA. Different types of melatonin circadian secretory rhythms in some blind subjects. *J Clin Endocr Metab* 1983;56:1103–1107.

70. Sack RL, Blood ML, Hughes RJ, et al. Circadian rhythm sleep disorders in the totally blind. *J Vis Impair Blind* 1998;92:145–161.

71. Sack RL, Lewy AJ, Hoban TM. Free-running melatonin rhythms in blind people: phase shifts with melatonin and triazolam administration. In: Rensing L, an der Heiden U, Mackey MC, eds. *Temporal disorder in human oscillatory systems*. Heidelberg: Springer-Verlag, 1987:219–224.

72. Arendt J, Aldhous M, English J, et al. Some effects of jet-lag and their alleviation by melatonin. *Ergonomics* 1987;30:1379–1393.

73. Arendt J, Bojkowski C, Folkard S, et al. Some effects of melatonin and the control of its secretion in humans. In: Evered D, Clark S, eds. *Photoperiodism, melatonin and the pineal*. London: Pitman, 1985:266–283.

74. Mallo C, Zaidan R, Faure A, et al. Effects of a four-day nocturnal melatonin treatment on the 24 h plasma melatonin, cortisol and prolactin profiles in humans. *Acta Endocrinol (Copenhagen)* 1988;119:474–480.

75. Lewy AJ, Sack RL. Exogenous melatonin's phase shifting effects on the endogenous melatonin profile in sighted humans: a brief review and critique of the literature. *J Biol Rhythms* 1997;12:595–603.

76. Sack RL, Lewy AJ. Melatonin administration phase advances endogenous rhythms in humans. *Sleep Res* 1988;17:396.

77. Sack RL, Lewy AJ, Blood ML, et al. Melatonin administration to blind people: phase advances and entrainment. *J Biol Rhythms* 1991;6:249–261.

78. Lewy AJ, Sack RL, Latham JM. Exogenous melatonin administration shifts circadian rhythms according to a phase response curve [Abstract 021]. *The Vth Colloquium of the European Pineal Study Group*. England: Guildford, 1990.

79. Lewy AJ, Sack RL, Latham JM. Circadian phase shifting of blind and sighted people with exogenous melatonin administration: evidence for a phase response curve. *Soc Light Treatment Biol Rhythms Abst* 1990;2:22.

80. Lewy AJ, Sack RL, Latham J. A phase response curve for melatonin administration in humans. *Sleep Res* 1991;20:461.

81. Lewy AJ, Sack RL, Latham JM. Melatonin and the acute suppressant effect of light may help regulate circadian rhythms in humans. In: Arendt J, Pevét P, eds. *Advances in pineal research*. London: John Libbey, 1991:285–293.

82. Lewy AJ, Ahmed S, Jackson JML, et al. Melatonin shifts circadian rhythms according to a phase-response curve. *Chronobiol Int* 1992;9:380–392.

83. Lewy AJ, Bauer VK, Ahmed S, et al. The human phase response curve (PRC) to melatonin is about 12 hours out of phase with the PRC to light. *Chronobiol Int* 1998;15:71–83.

84. Zaidan R, Geoffriau M, Brun J, et al. Melatonin is able to influence its secretion in humans: description of a phase-response curve. *Neuroendocrinology* 1994;60:105–112.

85. Middleton B, Arendt J, Stone BM. Complex effects of melatonin on human circadian rhythms in constant dim light. *J Biol Rhythms* 1997;12:467–477.

86. Czeisler CA, Kronauer RE, Allan JS, et al. Bright light induction of strong (Type O) resetting of the human circadian pacemaker. *Science* 1989;244:1328–1333.

87. Johnson CH. *An atlas of phase response curves for circadian and circatidal rhythms*. Nashville, Tennessee: Department of Biology, Vanderbilt University, 1990.

88. Honma K, Honma S. A human phase response curve for bright light pulses. *Jap J Psychiatry* 1988;42:167–168.

89. Minors DS, Waterhouse JM, Wirz-Justice A. A human phase-response curve to light. *Neurosci Lett* 1991;133:36–40.

90. Wever RA. Light effects on human circadian rhythms. A review of recent Andechs experiments. *J Biol Rhythms* 1989;4:161–186.

91. Eastman CI, Liu L, Fogg LF. Circadian rhythm adaptation to simulated night shift work: effect of nocturnal bright-light duration. *Sleep* 1995;18:399–407.

92. Lewy AJ, Bauer VK, Cutler NL, et al. Melatonin treatment of winter depression: a preliminary study. *Psychiatry Res* 1998;77:57–61.

93. Lewy AJ, Sack RL, Singer CM, et al. Winter depression: the phase angle between sleep and other circadian rhythms may be critical. In: Thompson C, Silverstone T, eds. *Seasonal affective disorder*. London: Clinical Neuroscience, 1989:205–221.

94. Lewy AJ, Sack RL, Singer CM. Bright light, melatonin, and winter depression: the phase-shift hypothesis. In: Shafii MA, Shafii SL, eds. *Biological rhythms, mood disorders, light therapy, and the pineal gland*. Washington, DC: American Psychiatric Press, 1990:143–173.

95. Terman M, Schlager D, Fairhurst S, et al. Dawn and dusk simulation as a therapeutic intervention. *Biol Psychiatry* 1989;25:966–970.

96. Avery DH, Bolte MA, Dager SR, et al. Dawn simulation treatment of winter depression: a controlled study. *Am J Psychiatry* 1993;150:113–117.

97. Sack RL, Blood ML, Lewy AJ. Melatonin administration to night-shift workers: an update. *Sleep Res* 1995;24:539.

98. Arendt J, Aldhous M, Marks V. Alleviation of “jet lag” by melatonin: preliminary results of controlled double blind trial. *BMJ* 1986;292:1170.

99. Daan S, Lewy AJ. Scheduled exposure to daylight: a potential strategy to reduce “jet lag” following transmeridian flight. *Psychopharmacol Bull* 1984;20:566–568.

100. Sack RL, Stevenson J, Lewy AJ. Entrainment of a previously free-running blind human with melatonin administration. *Sleep Res* 1990;19:404.

101. Sack RL, Brandes RW, Lewy AJ. Totally blind people with free-running circadian rhythms can be normally entrained with melatonin. *Sleep Res Online* 1999;2(Suppl 1):624.

102. Lockley SW, Skene DJ, James K, et al. Melatonin administration can entrain the free-running circadian system of blind subjects. *J Endocrinol* 2000;164:R1–R6.

103. Arendt J, Aldhous M, Wright J. Synchronisation of a disturbed sleep-wake cycle in a blind man by melatonin treatment [letter]. *Lancet* 1988;i:772–773.

104. Palm L, Blennow G, Wetterberg L. Correction of non-24-hour sleep/wake cycle by melatonin in a blind retarded boy. *Ann Neurol* 1991;29:336–339.

105. Lapierre O, Dumont M, Lespérance P, et al. Entrainment of a free-running sleep-wake cycle with melatonin in a blind retarded child. *Sleep Res* 1993;22:627.

106. Sack RL, Brandes RW, Kendall AR, et al. Entrainment of free-running circadian rhythms by melatonin in blind people. *N Engl J Med* 2000; 343:1070–1077.

107. Czeisler CA, Shanahan TL, Klerman EB, et al. Suppression of melatonin secretion in some blind patients by exposure to bright light. *N Engl J Med* 1995;332:6–11.

108. Klerman EB, Rimmer DW, Dijk D, et al. Nonphotic entrainment of the human circadian pacemaker. *Am Physiol Soc* 1998;43:R991–996.

109. Campbell SS, Murphy PJ. Extraocular circadian phototransduction in humans. *Science* 1998;279:396–399.

110. Lewy AJ, Wehr TA, Goodwin FK, et al. Manic-depressive patients may be supersensitive to light. *Lancet* 1981;i:383–384.

111. Lewy AJ, Nurnberger JI, Wehr TA, et al. Supersensitivity to light: possible trait marker for manic-depressive illness. *Am J Psychiatry* 1985;142:725–727.

112. Lewy AJ, Cutler NL, Hughes RJ, et al. Women are more sensitive than men to 500 lux light suppression of melatonin (LSM) in fall and winter. *Sleep Res* 1997;26:731.

113. Nurnberger JJ, Berrettini W, Tamarkin L, et al. Supersensitivity to melatonin suppression by light in young people at high risk for affective disorder. A preliminary report. *Neuropsychopharmacology* 1988;1:217–223.

114. Nurnberger JI, S. A, D.K. L, et al. Melatonin suppression by light in euthymic bipolar and unipolar patients. *Arch Gen Psychiatry* 2000;57:572–579.

115. Lockley SW, Skene DJ, English J, et al. Entrainment of totally blind subjects: photic or non-photic? *7th Meeting Society for Research on Biological Rhythms* 2000:28.

116. Garfinkle D, Laudon M, Nof D, et al. Treatment of elderly benzodiazepine-users with controlled-release melatonin: improvement of sleep quality. *Sleep Res* 1995;24A:303.

117. Matthews C, Guerin M, Deed J. Melatonin and photoperiodic time measurement: seasonal breeding in the sheep. *J Pineal Res* 1993;14:105–116.

118. Roenneberg T, Aschoff J. Annual rhythm of human reproduction: I. Biology, sociology, or both? *J Biol Rhythms* 1990;5:195–216.

119. Roenneberg T, Aschoff J. Annual rhythm of human reproduction: II. Environmental correlations. *J Biol Rhythms* 1990;5:217–239.

120. Terman M, Lewy AJ, Dijk DJ, et al. Light treatment for sleep disorders: consensus report. IV. Sleep phase and duration disturbances. *J Biol Rhythms* 1995;10:135–150.

121. Jones CR, Campbell SS, Zone SE, et al. Familial advanced sleep-phase syndrome: a short-period circadian rhythm variant in humans. *Nat Med* 1999;5:1062–1065.

122. Czeisler CA. Commentary: evidence for melatonin as a circadian phase-shifting agent. *J Biol Rhythms* 1997;12:618–623.

123. Buxton OM, L’Hermite-Balériaux M, Turek FW, et al. Daytime naps in darkness phase shift the human circadian rhythms of melatonin and thyrotropin secretion. *Am J Physiol Regul Integrat Comp Physiol* 2000;278:R373–R382.

124. Kronauer RE, Czeisler CA, Pilato SF, et al. Mathematical model of the human circadian system with two interacting oscillators. *Am J Physiol* 1982;242:R3–R17.

125. Czeisler C, Richardson G, Coleman R, et al. Chronotherapy: resetting the circadian clocks of patients with delayed sleep phase insomnia. *Sleep* 1981;4:1–21.

126. Reppert SM, Weaver DR, Rivkees SA, et al. Putative melatonin receptors are located in a human biological clock. *Science* 1988;242:78–81.

127. Weaver DR, Rivkees SA, Reppert SM. Localization and characterization of melatonin receptors in rodent brain by in vitro autoradiography. *J Neurosci* 1989;9:2581–2590.

128. Dubocovich ML, Benloucif S, Masana MI. Melatonin receptors in the mammalian suprachiasmatic nucleus. *Behav Brain Res* 1996;73:141–147.

129. Urbanski HF, Garyfallou VT, Kohama SG, et al. The circadian activity-rest rhythm of aged rhesus macaques: influence of melatonin. *Society for Research on Biological Rhythms*. Jacksonville, FL: Amelia Island Plantation, 2000:160.

130. Glass JD, Tardiff SD, Clemens R, et al. Photic and non-photic circadian phase-shifting responses in a diurnal monkey, the common marmoset. *Society for Research on Biological Rhythms*. Jacksonville, FL: Amelia Island Plantation, 2000:124.

131. Spitzer RL, Terman M, Williams JBW, et al. Jet lag: clinical features, validation of a new syndrome-specific scale, and lack of response to melatonin in a randomized, double-blind trial. *Am J Psychiatry* 1999;156:1392–1396.

130

Sleep Loss and Sleepiness: Physiological and Neurobehavioral Effects

Jacqueline D. Kloss

Martin P. Szuba

David F. Dinges

Jacqueline D. Kloss: Department of Psychology, Sociology, and Anthropology, Drexel University; Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania.

Martin P. Szuba and David F. Dinges: Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania.

The effects of sleep loss and sleepiness encompass a variety of neurobehavioral and physiologic alterations. This chapter reviews the causes, consequences, and mechanisms of sleep disruption and concomitant daytime sequelae, namely sleepiness and neurobehavioral performance decrements. Given the personal distress, quality of life issues, public health concerns, and economic costs of sleep loss and sleepiness, it is imperative that researchers and practitioners strive to obtain a solid understanding of these consequences and mechanisms. Several advances in the psychopharmacologic and behavioral treatments of the causes and consequences of sleep loss have recently evolved. Technologies are rapidly developing and showing promise for effective evaluation of these highly prevalent problems.

Advances and online monitoring and mathematical modeling of sleepiness and associated neurobehavioral forms are rapidly evolving novel behavioral and psychopharmacologic treatments effective for the causes and consequences of sleep loss.

PATHOPHYSIOLOGY OF DIFFICULTY INITIATING OR MAINTAINING SLEEP

Part of "130 - Sleep Loss and Sleepiness: Physiological and Neurobehavioral Effects "

Insomnia is characterized by difficulty initiating or maintaining sleep that results in psychological distress and impaired social or occupational functioning (1). Individuals with insomnia report a myriad of interpersonal, cognitive, affective, behavioral, and physical symptoms. Not only are there consequences for the individual, but also are there substantial costs to society; the direct economic costs owing to insomnia are estimated at $13.9 billion (2). The etiology of these symptoms has not been clearly delineated, however. The causal explanation that sleep deprivation accounts for the impairment of daytime functioning in insomniacs has been challenged and needs re-evaluation. This section provides a review of the daytime sequelae of the insomnia and a discussion of alternative mechanisms that may account for the daytime symptoms experienced.

Consequences

Persons with insomnia report various somatic complaints and demonstrate increased health-seeking behaviors. The primary complaints among insomniacs include drowsiness and tiredness on awakening, as well as sleepiness throughout the day (3 ,4). Insomniacs complain of physical ailments such as headache, diarrhea, stomach discomfort, heart palpitations, pain, tiredness, and weakness more frequently than do controls (5). Health-seeking behaviors such as hospital and physician visits are more frequent among a clinical sample of insomniacs compared to controls (5). Cardiovascular disease (6) and decreased immune functioning (7) may also be exacerbated in chronic insomnia.

The quality of life among insomniacs also appears to be diminished (8). Absenteeism, and work and social limitations are significantly more prevalent among insomniacs compared to normal sleepers (5 ,8). Insomniacs report restricted physical activities, poorer health, less vitality, and a decreased amount of time spent reading and engaging in recreational activities (8). Insomniacs report more time watching television, relaxing, and shopping than do noninsomniacs, whereas non-insomniacs work more, study more, and socialize more than do insomniacs. Insomnia is also associated with dissatisfaction in interpersonal relationships (4). These data suggest that insomniacs avoid or are unable

to participate in activities that require higher levels of concentration or social engagement.

Insomnia and mood disturbances often coexist. Sleep disruption is the single most common complaint of patients in a major depressive episode (9). Likewise, 30% of patients who complain of insomnia have a concurrent depressive disorder. Some have speculated that chronic insomnia may contribute to the development of major depressions (10 ,11); however, prospective controlled studies are needed to test these speculations. Kales and associates (5) found that insomniacs also exhibited symptoms mood changes, such as dysphoric mood, worry, tension, anxiety, and irritability. Even in nonclinical samples, insomniacs reported decreased mood (4), increased anxiety and depression, and less optimism (8). In contrast, Marchini and associates (12) reported that insomniacs were not particularly ruminative, tense, or physiologically aroused, but rather passive and calm. Marchini and colleagues' (12) unexpected finding led them to hypothesize that there may be different types of insomniacs: (a) hypoactive, as described; and (b) hyperactive, who are more anxious. They also suggested that insomniacs might be hypoactive during the day and hyperactive at night.

The causal direction of the relationship between insomnia and mood disorders is not clearly established. We cannot readily assume that psychiatric symptoms are merely sequelae of insomnia, nor can we definitively assume that insomnia is always a consequence of psychopathology. Clearly, we need to attend to the relationship between mood and insomnia. Even if criteria for a diagnostic disorder are not met, the interplay between moods and insomnia need to be examined in order to increase our knowledge of the etiology and guide treatment efforts.

Insomniacs' primary cognitive symptoms are impaired concentration and memory difficulties (4). Compared to noninsomniacs, insomniacs also rate their attention, memory, reasoning, problem-solving, and reaction time more poorly. Although there is some evidence that insomniacs have difficulty with semantic memory (13), reaction time, and digit span (14), objective verification of performance deficits have not consistently corroborated these subjective performance complaints. Interestingly, Sugerman and associates (15) showed that subjective insomniacs (no PSG corroboration), in contrast to objective insomniacs (PSG corroboration), displayed cognitive deficits. Thus, there may be factors other than sleep loss that account for these reported decrements. First, these data lead one to question whether or not insomniacs are indeed sleep deprived; and second, to hypothesize what could account for these reported symptoms if not sleep deprivation.

Are Insomniacs Sleep Deprived?

Daytime symptoms may not solely be attributable to sleep loss. First, for the majority of insomniacs it is questionable whether they suffer significant sleep loss compared to "good sleepers." Insomniacs have a tendency to overestimate their sleep latency, that is, the time from lights out to the onset of electrophysiologically defined sleep, and underestimate total sleep time (16). Although some insomniacs, particularly those with corroborating PSGs, demonstrate compromised sleep efficiencies and intermittent waking time (13), it is not clear that sleep is significantly disparate from that of noncomplaining sleepers in the majority of insomniacs.

Second, studies consistently find that insomniacs do not demonstrate daytime sleepiness, as measured by the Multiple Sleep Latency Test (MSLT) which measures sleep onset time when an individual is given an opportunity to sleep during the day (17 ,18). In fact, Stepanski and associates (18) found that insomniacs were less sleepy than good sleepers, based on the results of the MSLT. These results must be interpreted with strong caution, however. The MSLT measures sleepiness and ability to fall asleep. Inability to initiate sleep may be a characteristic of insomniacs both during the day and night. Thus, an insomniac may feel sleepy, but not be able to sleep during an MSLT; the results would then artificially underestimate the level of sleepiness. This measure is of dubious utility in the evaluation of sleepiness in those who cannot initiate sleep.

To circumvent this measurement difficulty, Lichstein and colleagues have used an index of sleepiness that does not depend on sleep ability, but rather diameter of the pupil as a measure of sleepiness. Although there is some evidence to suggest that insomniacs differ from noninsomniacs on sleepiness as measured by pupillometry (19 ,20), the effects were marginal. The technique may be promising, but the results are inconclusive.

Third, neither the quality nor quantity of nighttime sleep predicts the next day's functioning. One would expect that a worsening of nighttime sleep would exacerbate daytime impairment. Measures of sleep efficiency (17), total sleep time (TST), and polysomnographic (PSG) recordings (13 ,15 ,20) do not always directly relate to measures of daytime functioning. In fact, TST was correlated with increased tendency for drowsiness (17) and better nighttime sleep was correlated with increased sleep tendency during the day, for both insomniacs and noninsomniacs. Bonnet and Arand (21) also demonstrated that a worsening in sleep was not related to worsened daytime functioning. What, then, could account for the decrements in daytime functioning?

Hypothesized Mechanisms

Several studies support the notion that insomniacs are not necessarily sleep deprived; rather, they are hyperaroused and thus unable to fall asleep (17 ,18 ,21 ,22). This chronic activation may account for the inability to fall asleep at night and during the day, as measured by the MSLT (18). Bonnet and Arand (22) "yoked" the sleep of controls to that of insomniacs. Despite sleeping similar amounts, normal

sleepers exhibited a pattern resembling a sleep-deprived state (decreased tension and vigor, body temperature, and MSLT latencies). Insomniacs demonstrated a pattern of hyperarousal inconsistent with sleep deprivation (increased metabolic rate, body temperature, tension, and decreased vigor); therefore, daytime symptoms may be the result of hyperarousal, not sleep deprivation.

As alluded to in the preceding, perhaps psychopathology, either at the clinical or subclinical level, may account for both insomnia and daytime symptoms. For example, anxiety could account for sleep onset difficulties at night and symptoms of fatigue during the day (23). Likewise, Coyle (24) found that insomniacs with negative affect perceived impaired daytime cognitive functioning and motivation, whereas insomniacs with positive affect perceived better cognitive and motivational functioning.

How an insomniac reacts to his or her sleep disruption may also predict his or her experience of daytime functioning. Several hypotheses related to this notion are offered. Insomniacs may also be “short sleepers,” believe that their sleep is insufficient, and consequently become distressed about it during the day (20 ,25). Insomniacs simply may need more sleep than they are getting or be hypersensitive to small amounts of sleep loss (17). Consistent with these hypotheses, Dorsey and Bootzin (26) examined subgroups of insomniacs classified as objective insomniacs (OI) and subjective insomniacs (SI), with or without corroborating objective sleep disturbances, respectively. Like other studies, differences in performance, alertness, and night sleep parameters were not evidenced. SIs inaccurately estimated sleep/wake state in comparison to objective measures on the MSLT. OIs were more introverted, more withdrawn, and more able to accurately describe the amount of sleep that they had, but were perhaps too internally focused; SIs seemed to be more neurotic and unaware of their internal conscious state. These data suggest that the complaints of insomniacs may be differentiated and better understood by way of personality subtypes.

Studies of physiologic changes accompanying insomnia have produced inconsistent and generally unreplicable findings. The inconsistency may well derive from the heterogeneous samples form the different studies. In fact, the heterogeneity of all those diagnosed with “insomnia” confounds most studies in this field (27).

PATHOPHYSIOLOGY OF DISORDERS OF EXCESSIVE SOMNOLENCE

Part of "130 - Sleep Loss and Sleepiness: Physiological and Neurobehavioral Effects "

Obstructive Sleep Apnea

Obstructive sleep apnea (OSA) afflicts up to 4% of middle-aged adults and is characterized by respiratory pauses owing to upper airway closure during sleep, which result in acute hypoxemic events and transient arousals from sleep. Individuals with OSA primarily complain of daytime somnolence, yet are often completely unaware of their loud snoring, apneic and hypopneic events, and sleep fragmentation. Daytime performance is markedly impaired. Studies demonstrate impaired vigilance, reduced reaction times, daytime microsleeps, memory impairment, and depressive symptoms (28 ,29 and 30). Findley and associates (31) found that persons with apnea had poorer performance on “Steer Clear”—a 30-minute computer-generated visual vigilance driving simulation task that measures sustained attention and simple reaction times—than did age-matched control participants. The outcomes of these problems include impairment in work efficiency, increased automobile accident rates, and decrements in quality of life. In the case of truck drivers, pilots, and other operators of heavy machinery, these consequences of sleep-disordered breathing can be catastrophic.

Maintenance of respiration during sleep is dependent on intact central nervous system centers controlling respiration, respiratory reflexes, respiratory musculature, and innervation of the diaphragm and intercostal muscles, as well as a patent upper airway (32). The emphasis of clinical and research approaches to sleep-disordered breathing has focused on this last point—loss of airway patency. Two classes make up most of the obstructive sleep apnea patients. The majority of patients with OSA are obese and have short, thick necks with reduced upper airway diameters owing to excess tissue. A smaller proportion of patients are of normal weight, but are also prone to closure of the upper respiratory tract. These latter individuals typically have oral or maxillofacial structural changes resulting in congenitally small airways. Sleep produces a loss of postural tone of upper airway structures (32) that contributes to the onset of apneic events. Arousals from sleep, owing to combined hypoxemia and hypercapnia, are initiated to restore airway patency (33). Depression of arousal from acute or chronic sleep loss, alcohol, or other sedating drugs increases the propensity for apneic events. Thus, as apneics undergo chronic sleep loss from the repeated arousals, the daytime sequelae progress over time. It remains unclear which elements of OSA are responsible for the hypersomnolence and impaired neurobehavioral performance—repeated arousals from sleep or sleep hypoxemia (34).

Narcolepsy

Narcolepsy, with an estimated prevalence of about 50/100,000 (35), is a disabling disorder that may result in more daytime somnolence than any other major sleep disorder. In addition to the daytime somnolence, it is characterized by cataplexy, rapid switches from waking to sleeping states, sleep paralysis, and hypnagogic hallucinations (36). Not only do patients feel drowsy, but also they rapidly switch from waking to sleeping or cataplectic states—the sudden loss of all skeletal muscle control. This combination of severe excessive daytime somnolence coupled with sleep attacks

makes them particularly prone to impaired daytime neurobehavioral performance. Decrements in performance, including vigilance, and attentional and complex cognitive functioning problems, may be direct consequences of sleepiness (37). In the Findley and associates study, persons with narcolepsy demonstrated marked impairments as the duration of the task increased, with shorter latencies on the MSLT predicting greater performance errors. In addition, 20% to 30% of narcoleptics have comorbid major depressive episodes (38 ,39), although the reasons for this relationship remain largely unexplored.

Until recently, altered cholinergic and monoaminergic systems have been implicated in the pathophysiology of this disorder (40); however, new work from canine strains with narcolepsy, and a strain of Orexin knockout mice (manifesting a phenotype highly similar to narcolepsy) reveals that a genetic defect, disruption of a hypocretin (Orexin) gene may be the primary cause of narcolepsy (41 ,42). Orexins are a class of hypothalamically derived peptides with known effects on appetite and feeding behavior; however, until recently they were not recognized as sleep-modulating neurotransmitters.

Restless Legs Syndrome and Periodic Limb Movements in Sleep

Restless legs syndrome (RLS), a sleep-related disorder with an estimated prevalence of 1% to 5%, is characterized by unpleasant sensations experienced predominantly in the legs, which occur only at rest and become more pronounced in the evening or at night. Patients suffer from an urge to move their legs, often counteracted by walking, which leads to partial, temporary relief of the sensations. Most patients with RLS have periodic limb movements during sleep (PLMS) characterized by repetitive abrupt, involuntary flexion of the extremities that results in brief arousals and repeated complete awakenings. PMLS can occur as an isolated phenomenon, but often occurs with other sleep disorders, including RLS, narcolepsy, sleep apnea syndrome, or REM sleep behavior disorder.

The etiology of RLS and PLMS is unknown. It is hypothesized that PLMS results from a disinhibition of descending inhibitory pathways. Disturbances in dopaminergic, adrenergic, and opiate systems may contribute to RLS/PLMS (43); however, the evidence that these systems are responsible for the pathophysiology is inferentially derived from the fact that pharmacologic agents modulating these systems confer some clinical benefit. Daytime performance impairments that appear secondary to the sleep disruption have been poorly studied.

NEUROBEHAVIORAL AND PHYSIOLOGIC EFFECTS OF SLEEP LOSS

Part of "130 - Sleep Loss and Sleepiness: Physiological and Neurobehavioral Effects "

The evaluation of sleepiness is critical not only to minimize impairment in social or occupational settings, but also to ensure safety. Consequences of sleepiness and fatigue can lead to a myriad of neurobehavioral performance decrements and potentially dangerous situations, such as traffic and work accidents. The following section outlines the primary public health concerns and the populations that are most vulnerable to the consequences of sleepiness and fatigue, such as travelers, truck drivers, and shift workers. In addition, neurobehavioral consequences of experimentally induced sleep loss are reviewed.

Traffic Accidents

According to the National Highway Traffic Safety Administration, 56,000 automobile accidents per year are caused by drivers falling asleep at the wheel (44). Accidents involving truck drivers result in approximately 4,800 fatalities per year (45) and fatigue is the most common cause (46). Sleep-related automobile accidents are associated with fatalities (1.4%) comparable to those of alcohol-related crashes (2.1%) (47). Drowsiness can lead to rapid and frequent uncontrolled sleep or microsleeps, frequent prolonged eyelid closures, and inattention in the form of behavioral lapses that involve a failure to detect a monitored stimulus or a failure to respond in a normal timely manner (48).

Several factors are predictive of fatigue-related traffic accidents. Having an untreated sleep disorder such as insomnia, narcolepsy, or sleep apnea significantly increases the risk of having a motor vehicle accident (MVA) (49 ,50). The amount of sleep truck drivers obtain during the sleep episode prior to the accident inversely predicts the likelihood of an MVA (46). The homeostatic need for sleep and the circadian pacemaker interact in predicting performance in a dynamic, nonlinear fashion (51). Consistent with our knowledge of sleepiness and circadian neurobiology, midafternoon (approximately 3 PM) and nighttime hours (12 AM to 7 AM) are times when both sleepiness is increased and accidents are most likely to occur. Young drivers (under 45 years) are more likely to be involved in an accident during the night, individuals aged 45 to 65 are more likely to be involved in an accident around 7 AM; elderly drivers' peak accident time is at 3 PM.

Shift Workers

As many as 25% of employed individuals engage in shift work—employment outside the typical 7 AM to 7 PM workday—that can have severe personal and public health consequences. Sleepiness is reported by 70% of shift workers (52). Although it is difficult to discern the exact effects of sleepiness on daytime functioning, problem sleepiness among shift workers is associated with decreased quality of life (53), decreased productivity (54), and gastrointestinal and cardiovascular disease (55). Shift workers have a higher incidence of traffic accidents as a result of sleepiness while commuting, compared to non-shift working individuals (56). Shift

workers are also at an increased risk for injury and accidents (51). Three Mile Island, Exxon Valdez, and the Space Shuttle Challenger represent disasters where fatigue among nighttime workers has been implicated.

Both intrinsic biological and environmental factors contribute to the problem sleepiness of shift workers. Compared to individuals engaged in regular hour employment, shift workers sleep approximately 2 hours less per 24-hour sleep cycle as measured by EEG studies (57). Shift workers exist in states of chronic sleep debt because of insufficient sleep during each 24-hour period. Human entrainment to the natural 24-hour light/dark cycle establishes a fixed neurobiologic propensity to be active, alert, and performing during the daylight hours, and to sleep during the nocturne (58). Shift work requires maximum psychomotor and cognitive performance at night, that time when virtually all zeitgebers are cueing the endogenous circadian pacemaker to reduce arousal, activity, and sleep. Thus, not only must shift workers compensate for societal disruptions to their sleep, such as noise and pressures to socialize and perform domestic chores, but they must also overcome daylight and darkness time cues to work and sleep, respectively (53).

Jet Lag

Jet lag is a condition following transmeridian travel that involves a myriad of problems. Symptoms include daytime sleepiness and fatigue, impaired daytime cognitive performance, poor psychomotor coordination, dysphoric mood, and difficulty falling asleep according to the new schedule. The time needed to resynchronize to the new local light/dark cycle increases with the number of time zones crossed.

Like those of shift work, the adverse consequences of jet lag are mediated by disruptions of the sleep and circadian systems. Both the homeostatic mechanism for sleep (sleep drive that increases as duration of wakefulness increases) and circadian neurobiology interact to determine neurobehavioral alertness and performance (59). Jet lag-induced neurobehavioral performance decrements are primarily accounted for by the phase discrepancy between the organism's endogenous circadian rhythms and the new, local 24-hour light/dark cycle, although sleep loss incurred by travel can also serve to exacerbate the condition. The endogenous circadian pacemaker does not immediately adapt to the new light/dark cues, but rather requires a period for resynchronization or re-entrainment occurs during which individuals are likely to experience fatigue and performance deficits. An individual traveling eastward to a destination with a 9-hour time difference may feel compelled to sleep at 9 PM (home time) because of circadian propensity and increased homeostatic sleep drive. However, zeitgebers in the new destination (6 AM) associated with wakefulness, such as sunlight, are discrepant with the individual's endogenous pacemaker and the homeostatic sleep drive associated with sleepiness.

Sleep Deprivation: Experimentally Induced

Sleep loss results in compromised neurobehavioral performance and neurophysiologic functioning (60). Various performance assessments probe the functional capability of the CNS and offer meaning to the physiologic changes that occur as a result of sleep loss (61). Numerous studies show that as sleeplessness increases, so do subjective and objective measures of sleepiness and neurobehavioral problems. Psychomotor vigilance and probed memory impairment as well as somatic complaints appear to increase during acute total and repeated partial sleep deprivation (62 ,63 ,64 and 65). Some studies have been unable to show cognitive impairment during sleep deprivation (66), leading to speculation that chronic partial sleep deprivation does not result in cumulative decreases in performance (67 ,68). A number of factors may have contributed to the disparate outcomes among studies of waking performance after chronic sleep restriction. Many of the negative studies were limited by the fact that the primary outcome measures were performance assessments with robust practice effects (62). Learning curves confound cumulative performance deficit measurements; therefore, they compromise the validity of conclusions concerning the lack of such effects. In other words, repeated testing on a measure with a learning curve will lead to improved performance scores. Thus, if cumulative sleep loss does impair performance on this measure, the decrement will be masked by the learning-derived improvement.

Demonstration of cumulative performance deficits requires utilizing measures that are both sensitive to the effects of sleep loss and have no learning curve. Performance vigilance tasks, cognitive throughput tasks, and tasks requiring rapid response shifts incorporate both of these criteria. Studies utilizing such measures show increased lapses and heightened variability of performance during sustained vigilance tasks (62), all of which show deterioration after acute, total sleep deprivation, and after chronic partial sleep deprivation. During sleep loss, increased rates of slowing in response time result in accelerated decline in average performance with increasing task duration, independent of lapsing. Reduction in speed of response, although not a function of lapses or failure to respond, appear attributable to a decline in the ability to continuously allocate attention to the task and to respond motorically as rapidly as possible. The increase in false response rates or errors of commission, increase during chronic partial sleep deprivation, demonstrating that increased compensatory effort and a loss of motivation cannot account for these neurobehavioral performance decrements.

The magnitude of sleepiness on performance is a result of the dynamic influences of duration awake and underlying circadian rhythms. Motivation and incentive can contribute to, or override, the sleep-induced impairments, but only for a limited time. Sleepiness, fatigue, stress, and impaired

vigilance during sustained sleep restriction accumulate over time (62). Studies suggest that performance degrades in a dose–response manner (69). Kuo found that during chronic partial sleep deprivation, subjective sleepiness increased during the first week, but decreased during the second week (70), suggesting that subjects believed they were adapting to the effects of sleep loss, whereas performance measures indicated that they were not. Subjects were unaware of their neurocognitive dysfunction, because they “felt fine.”

Neurophysiologic functioning is altered during total sleep deprivation of 24 to 48 hours (TSD). (See ref. 60 for review.) Cumulative sleep loss produces decreased latency from wake to sleep onset, microsleep intrusions into wakeful periods, and involuntary sleep onsets. Constricted pupil size, difficulty with balance and coordination, and undulating slow eye movements are also observed as result of TSD. Prolonged sleep loss produces a modest dopaminergic and adrenergic activation, elevated levels of TSH, T3, and T4, hyperactivity of some immune parameters, and hypothermia. Remarkably, the hypothalamic-pituitary-adrenal axis (the “stress” axis) remains largely unaffected by sleep loss. Although recovery from TSD is marked by increased sleep intensity, sleep loss does not produce irreparable harm; changes can be reversed with recovery sleep.

Sleep deprivation in healthy individuals tends to produce little, if any, worsening of mood, anxiety, or anger, but does produce worsening self-reports of fatigue, vigor, and confusion. In contrast, depressed patients demonstrate increased locomotor activity, increased self-ratings of vigor, reduced fatigue, and improved mood after approximately 30 hours of sleep deprivation (71 ,72). This seemingly paradoxic effect in depressed individuals may reflect an underlying heightened sensitivity to the sleep deprivation-induced increases in dopamine, hypothalamic-pituitary-thyroid axis activity (73). Studies aimed at understanding these opposite effects in depressed and healthy persons to elucidate mechanisms are needed.

TREATMENT FOR SLEEP LOSS AND SLEEPINESS

Part of "130 - Sleep Loss and Sleepiness: Physiological and Neurobehavioral Effects "

Insomnia

Results from recent metaanalyses indicate that nonpharmacologic treatments for chronic insomnia are effective for the majority (70% to 80%) of patients (74) in reducing latency to sleep onset and wake after sleep onset by approximately 50% (e.g., to approximately 30 minutes). Effective treatments for insomnia include stimulus control (75), progressive muscle relaxation (76), paradoxic intention, sleep restriction (77), biofeedback (78), and cognitive therapy (79). (See ref. 27 for a further description of nonpharmacologic treatments.) Although the benefits of using combined pharmacologic and nonpharmacologic treatments has not been investigated extensively, some data suggest that behavior therapy alone, pharmacotherapy alone, and the two in combination provide comparable efficacy in the short term, but behavioral approaches may excel in the long term. (See ref. 74 for a review.)

Although patients have reported sleep quality improvement by using these strategies, the degree to which daytime sequelae, such as self-reported cognitive impairment, mood disturbance, and quality of life, remit has not yet clearly been determined. This is particularly relevant given the hypothesis that it may not be sleep loss per se that accounts for daytime impairment, but rather concomitants such as hyperarousal, cognitive distortions, and distressed mood that account for daytime performance and functioning. Strategies such as cognitive-behavioral therapy and progressive relaxation hold promise for managing these noted daytime sequelae.

The classic benzodiazepines and the newer, more selective benzodiazepine agonists zaleplon and zolpidem are extremely effective at inducing and sustaining sleep. The role of these compounds in treating insomnia is described in greater detail elsewhere in this text. Many of them unfortunately also produce residual daytime somnolence and impaired neurobehavioral performance. Typically, the longer-acting agents are more likely associated with these adverse effects. The newer agents, zaleplon and zolpidem, appear to produce less daytime problems than the older agents (80), however, whether any of these compounds reverse the daytime impairments to which insomniacs are prone remains to be seen.

Additional research is needed to assess the effects of pharmacologic and nonpharmacologic treatments not only for sleep quality (total sleep time, wake after sleep onset, sleep efficiency), but also for daytime performance, function, and distress.

Excessive Somnolence

Daytime napping is a behavioral strategy commonly used to alleviate excessive somnolence and enhance alertness in everyday life. The efficacy of napping, however, is contingent on the causes of the sleepiness and performance deficits. For whom and under what conditions is napping effective at alleviating sleep and enhancing alertness? The propensity for adults to nap in the midafternoon is relatively consistent across all cultures and appears tied to the endogenous circadian system. Some cultures, such as those in Mexico, China, or Greece, endorse taking afternoon siestas, consistent with the chronobiologic tendency. Perhaps owing to industrialization or occupational demands, other countries (e.g., the United States and Japan) do not endorse this practice, despite the endogenous drive for sleep in the midafternoon. Thus, napping is a behavior that is consistent with the circadian rhythm dip in the midafternoon and can be used to enhance functioning, even for individuals who do not exhibit sleep disorders (81).

For individuals with sleep disorders, however, the usefulness of napping in alleviating symptoms depends on the nature of the dysfunction (i.e., the underlying mechanism that contributes to the symptoms) (81). One might assume that napping is a healthy way of managing excessive somnolence regardless of the underlying mechanism. Many persons with narcolepsy find brief daytime napping to be helpful, whereas persons with untreated sleep apnea derive no benefit from napping (81). Napping improves reaction time performance in individuals with narcolepsy-cataplexy (82). Likewise, the strategy of “prophylactic napping” in advance to prevent anticipated sleepiness is quite helpful for individuals (e.g., truck drivers or shift workers) who need to work for prolonged hours (83). Appropriately timed napping can be beneficial for treatment of jet lag in some circumstances (84).

Two caveats are described regarding the use of napping for managing excessive somnolence. First, side effects of napping can include sleep inertia, which is characterized by sleepiness, diminished alertness, and reduced performance that occurs immediately on waking from sleep but that dissipates within 1 to 4 hours of awakening (85 ,86 and 87). Sleep inertia can be especially problematic for those who need to perform immediately on awakening. Second, if a nap is too long, it can interfere with nighttime sleep. Hence, napping is not recommended for individuals whose primary presenting problems directly involve difficulty initiating or maintaining nocturnal sleep.

Wake-Promoting Compounds

Caffeine is the most widely used wake-promoting compound in the world, most often consumed in high, intermittent dosages (150 to 300 mg) and usually in the hours just after awakening. Caffeine is most often used to counter the effects of morning sleep inertia. However, some also use it throughout the day to maintain wakefulness. This may be a natural countermeasure to daytime sleepiness caused by insufficient sleep the prior night. Research is needed in this area. Caffeine is a safe and simple wake-promoter that has been “staring us in the face,” but little research has focused on how to use caffeine as a practical and safe wake-promoter in the context of daytime sleepiness.

The mechanisms by which caffeine is able to promote wakefulness have not been fully elucidated (88). Most studies indicate that, at the levels reached during normal consumption, caffeine exerts its action through antagonism of central adenosine receptors (89 ,90). It reduces physiologic sleepiness (91—93) and enhances vigilance and cognitive performance (94 ,95). These beneficial effects have also been reported for caffeine taken during sleep deprivation (91 ,93 ,94).

Classical psychomotor stimulants such as methamphetamine and methylphenidate are potent centrally active compounds with central and peripheral sympathomimetic activity. In contrast to caffeine, methamphetamine and methylphenidate produce neurobehavioral activation and promote wakefulness by increasing dopaminergic and noradrenergic neurotransmission. These compounds have a number of potentially undesirable side effects, including anxiety, appetite suppression, tolerance, dependence, and abuse potential (96).

Modafinil is the first of a new class of wake promoting therapeutics (97 ,98). The mechanism(s) by which it improves alertness and vigilance and reduces sleepiness remains obscure. Some work suggests that modafinil may promote activity at α1- and ß-adrenergic receptors (99) and 5-HT2 receptors (100). Its ability to stimulate dopaminergic activity remains controversial. New work has demonstrated that it actually stimulates Orexin-containing neurons in the hypothalamus of mice (42). Unlike amphetamines, modafinil does not appear to produce dependence or have addictive potential (98 ,101). The novel wake-promoting compounds hold potential for enhancing understanding of the mechanisms of pathologic somnolence and for the treatment of the disorders of excessive sleepiness.

Obstructive Sleep Apnea

Treatments for OSA are directed at maintaining airway patency and thereby preventing the apneic events. The most effective methods developed to date include continuous positive airway pressure (CPAP), weight loss, dental appliances that reposition the jaw and/or tongue, and surgical procedures. These treatments have been demonstrated to improve the daytime somnolence, impaired vigilance, depression, and overall quality of life (28 ,29 and 30). Few randomized, well-controlled trials have been published that evaluate pharmacologic agents in the treatment of obstructive sleep apnea. Respiratory stimulants (theophylline), psychostimulants, adrenergic agonists, opioid antagonists, and nicotinic agents, have been studied with mixed results. Non-OSA sleep-related breathing disorders such as hypercapnic obesity-hypoventilation, myxedema, central apnea, and periodic breathing in congestive heart failure respond to specific pharmacologic measures. Future research including the use of the newer wake-promoting compounds, such as modafinil, is warranted.

Narcolepsy

Until recently, standard treatments for narcolepsy often included a combination of amphetamine-like stimulants for sleepiness and antidepressant therapy for abnormal rapid eye movement sleep events (cataplexy, sleep paralysis, and hypnagogic hallucinations). These treatments are purely symptomatically directed and involve activation of central dopaminergic and adrenergic systems (36). Modafinil is the first specific treatment approved in the United States for treatment of narcolepsy. With the discovery of the genetic

markers for narcolepsy, even more novel approaches appear conceivable. Gene therapy or compounds affecting Orexin, systems one of which is modafinil, are likely directions for future research.

RLS/PLMS

Treatment of RLS/PLMS is targeted toward dopaminergic, adrenergic, GABA, and opiate systems. L-Dopa, dopamine agonists, benzodiazepines, opioids, clonidine, and carbamazepine appear effective. With no obvious cause, treatment has been aimed at symptom control to date (43).

Shift Work and Jet Lag

The disturbances in circadian neurobiology associated with shift work and jet lag appear to be responsive to interventions that alter the underlying circadian system. Bright light therapy and exogenously administered melatonin are potent zeitgebers capable of inducing phase shifts in humans. Regulation of exposure to sunlight and artificial light (102 ,103), napping (104), caffeine to promote alertness at night and hypnotics to help daytime sleep (105), and melatonin to adjust circadian rhythms (106 ,107) are all helpful in limited studies. This evidence is in need of replication and application to other real-world situations.

RECENT ADVANCES IN ASSESSMENT AND PREVENTION TECHNOLOGIES

Part of "130 - Sleep Loss and Sleepiness: Physiological and Neurobehavioral Effects "

As discussed, the MSLT and pupillometry aid in the assessment of sleepiness. Wrist-worn actigraphic devices that monitor locomotion have demonstrated utility in monitoring sleep-wake patterns and sleep quality as well as assessing sleep disorders. (See refs. 108 and 109 for review.) A newly introduced technology, the sleep switch, is a handheld instrument that effectively detects latency to sleep onset (110). The patient presses and holds a button. When the patient lapses into sleep, voluntary motor tone is lost, the button is released, and an event marker notes the time. Unlike actigraphy, it cannot measure total sleep time; however, it has the distinct advantage as an objective estimate of sleep onset latencies for measuring insomnia, compared to actigraphs and compared to the subjective estimates of sleep logs that have traditionally been used (110).

The development and validation of technologies to detect and monitor fatigue is essential (111). As discussed, fatigue-related motor vehicle crashes and performance errors owing to sleep loss are pervasive and individuals are unreliable predictors of their own level of impairment (70 ,112). Moreover, current standards of proscriptive hours are not sufficient at preventing crashes, even when compliance is 100%. Thus, technology offers advantages of both objective verification of sleepiness levels and a viable alternative to enhance and improve safety while facilitating occupational and economic goals.

Four major categories comprise operator-centered fatigue monitoring technologies. First, readiness-to-perform and fitness-for-duty technologies for drowsiness—aim to measure the functional capacity for work to be performed. Some measure fatigue by physiologic fitness (pupil or ocular scanning), whereas others measure performance via a battery of simple performance tests (113). Second, mathematical models of alertness are combined with ambulatory technologies to predict fatigue (114 ,115 and 116). These typically involve a device, such an actigraph, which measures fatigue in combination with a formula (mathematical model) that predicts performance capacity for a given period of time when sleepiness is likely to occur. Third, vehicle based performance technologies focus on the vehicle, in contrast to the driver (117 ,118 ,119 and 120). They are designed to monitor the vehicle hardware systems that are subject to the alterations of the driver's performance, such as steering or speed variability or lane swaying. Fourth, in-vehicle, on-line, operator status monitoring technologies aim to monitor biobehavioral features of the operator (e.g., eyes, face, head, heart, brain electrical activity) on-line. Example of devices include: (a) video of the face, which monitors the eyelid position, blinks, movements, head nodding, direction of gaze; (b) eye trackers; (c) wearable eyelid monitors; (d) head movement detectors; (e) EEG algorithms; and (f) ECG algorithms (111). All these systems have relative merits and drawbacks. Clearly the status of these technologies is promising.

REFERENCES

1. *Diagnostic and statistical manual of mental disorders*, fourth ed. Washington, DC: American Psychiatric Association, 1994.

2. Walsh J, Engelhardt CL. The direct economic costs of insomnia in the United States for 1995. *Sleep* 1999;22(2):S386–S393.

3. Ancoli-Israel S, Roth T. Characteristics of insomnia in the United States: results of 1991 National Sleep Foundation Survey. I. *Sleep* 1999;22(2):S347–S353.

4. Roth T, Ancoli-Israel S. Daytime consequences and correlates of insomnia in the United States: results of the 1991 National Sleep Foundation Survey II. *Sleep* 1999;22(2):S354–S358.

5. Kales J, Kales A, Bixler EO, et al. Biopsychobehavioral correlates of insomnia. V: Clinical characteristics and behavioral correlates. *Am J Psychiatry* 1984;141(11):1371–1376.

6. Schwartz S, McDowell Anderson W, et al. Insomnia and heart disease: a review of epidemiologic studies. *J Psychosom Res* 1999;47(4):313–333.

7. Irwin M, Fortner M, Clarl C, et al. Reduction of natural killer cell activity in primary insomnia and in major depression. *Sleep Res* 1995;24:256.

8. Zammit GK, Weiner J, Damato N, et al. Quality of life in people with insomnia. *Sleep* 1999;22(2):S379–S385.

9. Goodwin FK, Jamison KR. *Manic-depressive illness*. New York: Oxford, 1990.

10. Breslau N, Roth T, Rosenthal L, et al. Sleep disorders and psychiatric disorders: a longitudinal epidemiologic study of young adults. *Biol Psychiatry* 1996;39(6):411–4118.

11. Ford DE, Kamerow DB. Epidemiologic study of sleep disturbances and psychiatric disorders: An opportunity for prevention? *JAMA* 1989;162:1479–1484.

12. Marchini E, Coates TJ, Magistad JG, et al. What do insomniacs do, think, and feel during the day? A preliminary study. *Sleep* 1983;6:147–155.

13. Mendelson WB, Garnett D, Gillin C, et al. The experience of insomnia and daytime and nighttime functioning. *Psychiatry Res* 1984;12:235–250.

14. Hauri PJ. Cognitive deficits in insomnia patients. *Acta Neurol (Belgium)* 1997;97:113–117.

15. Sugerman JL, Stern JA, Walsh JK. Daytime alertness in subjective and objective insomnia: some preliminary findings. *Biol Psychiatry* 1985;20:741–750.

16. Carskadon MA, Dement WC, Mitler MM, et al. Self-reports versus sleep laboratory findings in 122 drug-free subjects with complaints of chronic insomnia. *Am J Psychiatry* 1976;133(12):1382–1388.

17. Seidel W, Dement W. Sleepiness in insomnia: evaluation and treatment. *Sleep* 1982;5:S182–190.

18. Stepanski E, Zorick F, Roehrs T, et al. Daytime alertness in patients with chronic insomnia compared with asymptomatic control subjects. *Sleep* 1988;11(1):54–60.

19. Lichstein K, Johnson RS, Gupta SS, et al. Are insomniacs sleepy during the day? *Behav Res Ther* 1992;30(3):283–292.

20. Lichstein, KL, Wilson NM, Noe SL, et al. Daytime sleepiness in insomnia: behavioral, biological, and subjective indices. *Sleep* 1994;17(8):693–702.

21. Bonnet MH, Arand DL. The consequences of a week of insomnia. II: Patients with insomnia. *Sleep* 1998;21(4):359–368.

22. Bonnet MH, Arand DL. Consequences of a week of insomnia. *Sleep* 1996;19(6):453–461.

23. Chambers MJ, Kim JY. The role of state-trait anxiety in insomnia and daytime restedness. *Behav Med* 1993;19:42–46.

24. Coyle K. The subjective assessment of daytime cognitive dysfunction and motivation by insomniacs. *Br J Health Psychol* 1998;3:171–182.

25. Seidel WF, Ball S, Cohen S, et al. Daytime alertness in relation to mood, performance, and nocturnal sleep in chronic insomniacs and noncomplaining sleepers. *Sleep* 1984;7:230–238.

26. Dorsey C, Bootzin R. Subjective and psychophysiologic insomnia: an examination of sleep tendency and personality. *Biol Psychiatry* 1997;41:209–216.

27. Buyssee D. *Neuropsychopharmacology: the fifth generation of progress*. New York: Lippincott Williams & Wilkins, 2000.

28. Kribbs NB, Pack AI, Kline LR, et al. Effects of one night without nasal CPAP treatment on sleep and sleepiness in patients with obstructive sleep apnea. *Am Rev Respir Dis* 1993;147:1162–1168.

29. Flemins WW, Tsai W. Quality of life consequences of sleep-disordered breathing. *J Allergy Clin Immunol* 1997;99:S750–756.

30. Engelman HM, Martin SE, Deary IJ, et al. Effect of CPAP therapy on daytime function in patients with sleep apnea/hypopnea syndrome. *Thorax* 1997;52:114–119.

31. Findley LJ, Suratt PM, Dinges DF. Time-on-task decrements in “steer clear” performance of patients with sleep apnea and narcolepsy. *Sleep* 1999;22(6):804–809.

32. Cistulli PA, Sullivan CE. Pathophysiology of sleep apnea. In: Saunders NA, Sullivan CE, eds. *Sleep and breathing*. New York: Marcel Dekker, 1994:411–447.

33. Olson LG, Hensley MJ, Saunders NA. Effects of slow wave sleep on the ventilatory responses to hypoxia in healthy subjects. *Proc Aust Physiol Pharmacol* 1981;12:219P,34–40.

34. Kimoff RJ. Sleep fragmentation in obstructive sleep apnea. *Sleep* 1996;19:S61–66.

35. Hublin C, Partinnen M, Kaprio J, et al. Epidemiology of narcolepsy. *Sleep* 1994;17(8):S7–12.

36. Mignot E. Perspectives in narcolepsy research and therapy. *Curr Opin Pulmon Med* 1996;2:482–487.

37. Hood B, Bruck D. Sleeepiness and performance in narcolepsy. *J Sleep Res* 1996;5(2):128–134.

38. Roth B, Nevsimalova S. Depression in narcolepsy and hypersomnia. *Schweizer Archiv fur Neurologie, Neurochirurgie Psychiatrie* 1981;116(2):291–300.

39. Reynolds CD, Christiansen CL, Taska LS, et al. Sleep in narcolepsy and depression. Does it all look alike? *J Nerv Ment Dis* 1983;171(5):290–295.

40. Mefford IN, Foutz A, Baker TL, et al. Regional distribution of biogenic amines in CNS of normal and narcoleptic canines. *Sleep Res* 1981;10:69.

41. Lin L, Faraco J, Li R, et al. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (Orexin) receptor 2 gene. *Cell* 1999;98(3):365–376.

42. Chemelli RM, Willie JT, Sinton CM, et al. Narcolepsy in Orexin knockout mice: molecular genetics of sleep regulation. *Cell* 1999;98:437–451.

43. Trenkwalder C, Walters AS, Hening W. Periodic limb movements and restless legs syndrome. *Neurol Clin* 1996;3:629–650.

44. Knipling RR, Wang JS. Crashes and fatalities related to driver drowsiness/fatigue. Research note: National Highway Traffic Safety Administration. U.S. Department of Transportation, November 1994.

45. US Congress Office of Technology Assessment. *Gearing up for safety: motor carrier safety in a competitive environment*. OTA-SET-382. Washington, DC: US Government Printing Office, 1988.

46. National Transportation Safety Study. Factors that affect fatigue in heavy truck accidents (V1) report no. NTSB//SS-95/01 (PB95-9170012), January 18, 1995.

47. Pack A, Pack AM, Rodgman E, et al. Characteristics of crashes attributed to the driver having fallen asleep. *Accid Anal Prev* 1995;27:769–775.

48. Dinges DF, Kribbs NB. Performing while sleeping: effects of experimentally induced slepiness. In: Monk TH, ed. *Sleep, sleepiness, and performance*. Chichester, UK: John Wiley & Sons, 1991:97–128.

49. Findley L, Unverzagt M, Guchu R, et al. Vigilance and automobile accidents in patients with sleep apnea or narcolepsy. *Chest* 1995;108:619–624.

50. Costa e Silva JA, Chase M, et al. Special report from a symposium held by the World Health Organization and the World Federation of Sleep Research Societies: an overview of insomnias and related disorders—recognition, epidemiology, and rational management, 1996.

51. Dinges DF. An overview of sleepiness and accidents. *J Sleep Res* 1995;4(2):4–14.

52. Rutenfranz J, Haider M, Koller M. Occupational health measures for night workers and shift workers. In: Folkard S, Monk T, eds. *Hours of work: temporal factors in work scheduling*. New York: John Wiley & Sons, 1985:199–210.

53. Walker J. Social problems of shift work. In: Folkard S, Monk T, eds. *Hours of work: temporal factors in work scheduling*. New York: John Wiley & Sons, 1985:211–226.

54. Wojtczak-Jaroszowa J, Pawlowska-Skyba K. Work at night and shiftwork: I. Day andnight oscillations of working capacity and the work efficiency. *Med Pract* 1967;18(1):1–10.

55. Scott AJ, La Dou J. Shiftwork: effects of sleep and health with recommendations for medical surveillance ad screening. *Occup Med* 1990;5:273–299.

56. Richardson GS, Minor JD, Czeisler CA. Impaired driving performance in shift workers: the role of circadian system in a multifactorial model. *Alcohol Drugs Driv* 1990;5(1,4):265–273.

57. Torsval L, Arkestedt T, Gillander K, et al. Sleep on the night shift: 24 hour EEG monitoring of spontaneous sleep/wake behavior. *Psychophysiology* 1989;26:352–358.

58. Arkestedt T. Adjustment of physiologic circadian rhythms and the sleep wake cycle to shift work. In: Folkard S, Monk T, eds. *Hours of work: temporal factors in work scheduling*. New York: John Wiley & Sons, 1985:185–198.

59. Folkard S, Akerstedt T. Towards the prediction of alertness on abnormal sleep/wake schedules. In: Coblentz A, ed. *Vigilance and performance in automated systems*. Dordrecht, Netherlands: Kluwer, 1989:287–296.

60. Dinges DF, Chugh D. Physiologic correlates of sleep deprivation. In: Kinney JM, Tucker HN, eds. *Physiology, stress, and malnutrition: functional correlates, nutritional intervention*. New York: Lippincott-Raven, 1997:1–27.

61. Dinges DF. Probing the limits of functional capability: the effects of sleep loss on short-duration tasks. In: Broughton RJ, Ogilvie RD, eds. *Sleep, arousal, and performance*. Boston: Birkhauser, 1992:176–188.

62. Dinges DF, Pack F, Williams K, et al. Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 hours per night. *Sleep* 1997;20(4):267–277.

63. Blagrove M, Alexander C, Horne JA. The effects of chronic sleep reduction on the performance of cognitive tasks sensitive to sleep deprivation. *Appl Cog Psychol* 1994;9:21–40.

64. Gillberg M, Akerstedt T. Sleep restriction and SWS suppression: effects on daytime alertness ad nighttime recovery. *J Sleep Res* 1994;3:144–151.

65. Wittersheim G, Libert JP, Muzet A. Effects of partial sleep deprivation and heat exposure on performance. In: Broughton RJ, Ogilvie RD, eds. *Sleep arousal and performance*. Boston: Bikhauser, 1992:189–197.

66. Webb WB, Agnew HW. The effects of a chronic limitation of sleep length. *Psychophysiology* 1974;11:265–274.

67. Naitoh P. Minimal sleep to maintain performance: the search for sleep quantum in sustained operations. In: Stampi C, ed. *Why we nap: evolution, chronobiology, and functions of polyphasic and ultra-short sleep*. Boston: Birkhauser, 1992:199–216.

68. Horne JA, Wilkinson S. Chronic sleep reduction: daytime vigilance performance and EEG measures of sleepiness, with particular reference to “practice effects.” *Psychophysiology* 1985;22:69–78.

69. Dinges DF, Maislin GM, Kuo, et al. Chronic Sleep restriction: neurobehavioral effects of 4 hr, 6 hr, 8 hr TIB. *Sleep* 1999;22:S115.

70. Kuo A, Carlin MM, Powell JJ, et al. Chronic restriction of sleep to 4 hours per night for 14 nights changes performance linearly but not subjective sleepiness. *Sleep* 1998;21(3):241.

71. Szuba MP, Baxter LR, Fairbanks LA, et al. Effects of partial sleep deprivation on the diurnal variation of mood and motor activity in major depression. *Biol Psychiatry* 1991;30:817–829.

72. McManus KS, Szuba MP, Carlin MM, et al. Total sleep deprivation induces opposite effects on mood states in depressed and control subjects. *Sleep* 2000;23(Suppl. 2):A37.

73. Szuba MP, Baxter LR. Letter: sleep deprivation and thyroid hormones. *Psychiatry Res* 1994;53:107–108.

74. Morin CM, Hauri P, Espie C, et al. Nonpharmacologic treatment of chronic insomnia. Sleep 1999;22(8):1013–1174.

75. Bootzin RR, Perlis ML. Nonpharmacologic treatments for insomnia. *J Clin Psychiatry* 1992;53:37–41.

76. Borkovec TD, Fowles DC. Controlled investigation of the effects of progressive and hypnotic relaxation on insomnia. *J Abnorm Psychol* 1973;82(1):153–158.

77. Spielman AJ, Saskin P, Thorpy MJ. Treatment of chronic insomnia by restriction of time in bed. *Sleep* 1987;10:45–56.

78. Hauri P. Treating psychophysiologic insomnia with biofeedback. *Arch Gen Psychiatry* 1981;39(7):752–758.

79. Morin CM. *Insomnia: psychological assessment and management*. New York: Guilford Press, 1993.

80. Danjou P, Paty I, Fruncillo R, et al. A comparison of the residual effects of zaleplon and zolpidem following administration 5 to 2 h before awakening. *Br J Clin Pharmacol* 1999;48:367–374.

81. Dinges DF. Napping. *Encylopedia of sleep and dreaming*. New York: Macmillan Publishing Co., 1990:1–8.

82. Mullington J, Broughton R. Scheduled naps in the management of daytime sleepiness in narcolepsy-cataplexy. *Sleep* 1993;16(5):444–456.

83. Dinges DF, Orne MT, Whitehouse WG, et al. Temporal placement of a nap for alertness: Contributions of circadian phase and prior wakefulness. *Sleep* 1987;10: 313–329.

84. Graeber RC. Disorders of chronobiology. In: Kryger M, Dement W, Roth T, eds. *Principles and practice of sleep medicine*. Philadelphia: W.B. Saunders, 1994:463–470.

85. Dinges DF. Are you awake? Cognitive performance and reverie during the hypnopompic state. In: Bootzin RR, Kihlstrom JF, Schacter D, eds. *Sleep and cognition*. Washington, DC: American Psychological Association, 1990:159–175.

86. Van Dongen HPA, Price NJ, Mullington JM, et al. Caffeine eliminates psychomotor vigilance deficits from sleep inertia. *Sleep* 2001;24(7):1–7.

87. Jewett ME, Wyatt JK, Ritz-De Cecco A, et al. Time course of sleep inertia dissipation in human performance and alertness. *J Sleep Res* 1999;1:1–8.

88. Lieberman HR. Caffeine. In: Smith AP, Jones DM, eds. *Handbook of human performance, vol 2. Health and performance*. San Diego: Academic Press, 1992:49–72.

89. Seidel WF, et al. NREM sleep homeostasis in rats after caffeine or methamphetamine. *Sleep Res* 1994;23:82.

90. Benington JH, et al. A1 adenosine receptor stimulation mimics changes in non-REM sleep EEG slow-wave activity. *Sleep Res* 1993;22:4.

91. Walsh JK, et al. Effect of caffeine on physiologic sleep tendency and ability to sustain wakefulness at night. *Psychopharmacology* 1990;101:271–273.

92. Rosenthal L, et al. Alerting effects of caffeine after normal and restricted sleep. *Neuropsychopharmacology* 1991;4(2):103–108.

93. Bonnet MH, Arand DL. Impact of naps and caffeine on extended nocturnal performance. *Physiol Behav* 1994a;56(1):103–109.

94. Lorist MM, et al. Influence of caffeine on information processing stages in well rested and fatigued subjects. *Psychopharmacology* 1994;113(3 and 4):411–421.

95. Bonnet MH, Arand DL. The use of prophylactic naps and caffeine to maintain performance during a continuous operation. *Ergonomics* 1994b;37(6):1009–1020.

96. Gillman AG, Goodman LS, et al, eds. *The pharmacological basis of therapeutics*. New York: MacMillan, 1985:1839.

97. Lyons TJ, French J. Modafinil: the unique properties of a new stimulant. *Aviation Space Environ Med* 1991;62:432–435.

98. Hermant JL, et al. Awakening properties of modafinil: effect on nocturnal activity in monkeys (Macaca mulatta) after acute and repeated administration. *Psychopharmacology* 1991;103:28–32.

99. Lin JS, et al. Role of catecholamine in the modafinil and amphetamine induced wakefulness, a comparative pharmacological study in the cat. *Brain Res* 1992;591:319–326.

100. Tanganelli S, et al. Inhibitory effects of the psychoactive drug modafinil on gamma-aminobutyric acid outflow from the cerebral cortex of the awake freely moving guinea-pig. Possible involvement of 5-hydroxytryptamine mechanisms. *Naunyn-Schmiedebergs Arch Pharmacol* 1992;345:462–465.

101. Duteil J, et al. Central alpha 1-adrenergic stimulation in relation to the behavior stimulating effect of modafinil: studies with experimental animals. *Eur J Pharmacol* 1990;180:49–58.

102. Boulos Z, Campbell SS, Lewy AJ, et al. Light treatment for sleep disorders: consensus report:VII. Jet lag. *J Biol Rhythms* 1995;10:167–176.

103. Czeisler CA, Dijk DJ. Use of bright light to treat maladaptation to night shift work and circadian rhythm sleep disorders. *J Sleep Res* 1995;4(S2):70–73.

104. Rosekind MR, Smith RM, Miller DL, et al. Alertness management: strategic naps in operational settings. *J Sleep Res* 1995;4(S2):62–66.

105. Walsh JK, Muehlbach MJ, Schweitzer PK. Hypnotics and caffeine as countermeasures for shift work related sleepiness and sleep disturbance. *Journal of Sleep Res* 1995;4(S2):80–83.

106. Arendt J, Aldhous M, English J, et al. Some effects of jet lag and their alleviation by melatonin. *Ergonomics* 1987;30:1379–1393.

107. Arendt J, Deacon S, English J, et al. Melatonin and adjustment to phase shift. *J Sleep Res* 1995;4(S2):74–79.

108. Sadeh A, Hauri PJ, Kripke DF, et al. The role of actigraphy in evaluation of sleep disorders. *Sleep* 1995;18(4):288–302.

109. American Sleep Disorders Association. Practice parameters for the use of actigraphy in the clinical assessment of sleep disorders. *Sleep* 1995;18(4):285–287.

110. Hauri PJ. Evaluation of a sleep switch device. *Sleep* 1999;22(8):1110–1117.

111. Dinges DF, Mallis MM. Managing fatigue by drowsiness detection: can technological promises be realized? In: Hartley L, ed. *Managing fatigue in transportation*. Oxford, UK: Pergamon, 1998:209–229.

112. Dinges DF. The nature of sleepiness: causes, contexts and consequences. In: Stunkard A, Baum A, eds. *Perspectives in behavioral medicine: eating, sleeping and sex*. Hillsdale, NJ: LEA, 1989:147–149.

113. Rosekind MR, et al. Crew factors in flight operations IX: effects of planned cockpit rest on crew performance and alertness in long-haul operations. NASA Technical Memorandum 108839, 1994.

114. Akerstedt T, Folkard S. Prediction of intentional and unintentional sleep onset. In: Ogilvie RD, Harsh JR, eds. *Sleep onset: normal and abnormal*. Washington, DC: American Psychological Association, 1994:73–87.

115. Belenky G. Sustaining performance during continuous operation: The U.S. Army's sleep management system. International Conference Proceedings on Managing Fatigue in Transportation. Tampa, FL, 1997.

116. Jewett ME. Models of circadian and homeostatic regulation of human performance and alertness. Doctoral thesis. Harvard University, Cambridge, MA, 1997.

117. Wylie CD, et al. Commercial Motor Vehicle Driver Fatigue and Alertness Study. Project Report, U.S. Department of Transportation Report No. FHWA-MC-97-002, 1996.

118. Grace R. Field testing a prototype drowsy driver detection and warning system. In: Proceedings of the Technological Conference on Enhancing Commercial Motor Vehicle Driver Vigilance, American Trucking Associations Foundation, FHWA, NHTSA. McLean, VA, 1996:27–30.

119. King D. Correlation of steering behavior with heavy truck driver fatigue. Proceedings of the Technological Conference on Enhancing Commercial Motor Vehicle Driver Vigilance, American Trucking Associations Foundation, FHWA, NHTSA. McLean, VA, 1996:69.

120. Schwenk S. TRW Driver Alertness Warning System (DAWS). Proceedings of the Technological Conference on Enhancing Commercial Motor Vehicle Driver Vigilance, American Trucking Associations Foundation, FHWA, NHTSA. McLean, VA, 1996:76–77.

131

Pathophysiological and Pharmacologic Aspects of the Sleep Disorder Narcolepsy

Emmanuel Mignot

Seiji Nishino

Emmanuel Mignot and Seiji Nishino: Center for Narcolepsy, Department of Psychiatry and Behavioral Sciences, Stanford University Medical Center, Stanford, California.

Narcolepsy is frequently both over- and under-diagnosed. The condition is not rare and has population prevalence similar to that of multiple sclerosis (57 ,78 ,141). Studies have demonstrated a large psychosocial impact of the disease (22 ,23). Narcolepsy is also a unique disease model for basic sleep researchers with the availability of validated animal models and as the only known disorder with a complete disorganization of sleep and REM sleep. Our understanding of the pathophysiology of the disorder is rapidly emerging, thanks to the discovery that narcolepsy-cataplexy is associated with a deficiency in the hypocretin (Orexin) neuropeptide system (30 ,68 ,106). The fact that human narcolepsy is HLA-associated (59 ,82) also suggests a possible autoimmune mediation in many cases. In this chapter, we briefly outline how narcolepsy is diagnosed and treated, as well as discuss future directions for this rapidly evolving area.

CLINICAL AND EPIDEMIOLOGIC ASPECTS OF HUMAN NARCOLEPSY

Part of "131 - Pathophysiological and Pharmacologic Aspects of the Sleep Disorder Narcolepsy "

Cataplexy: A Pathognomonic Symptom of the Narcolepsy Syndrome

Patients with narcolepsy experience brief episodes of muscle weakness when laughing, angry, or elated, a symptom referred to as cataplexy (3 ,7 ,46 ,55 ,114). These episodes most often affect the legs or the face, leading to knee buckling, sagging of the jaw, slurring of speech, and/or dropping of the head (7 ,46 ,55). Episodes are brief (a few seconds to several minutes at most), bilateral, and rarely lead to body collapse and/or long-lasting episodes of complete paralysis. Consciousness is preserved during cataplexy (7 ,46 ,55).

The importance of carefully defining cataplexy should be emphasized. Epidemiologic studies have shown that up to 30% of the general population experiences “muscle weakness episodes in reaction to emotions” (4 ,7 ,57). Clearly, this definition is not sufficient to establish cataplexy. Genuine cataplectic episodes in narcolepsy are triggered by very specific emotions. Joking, laughing, and anger are the most reliable triggering events (7). Rare episodes of muscle weakness occurring exclusively in the context of unusual emotional triggers, for example while tense or stressed, or during sexual or athletic activities, should not be considered as cataplexy (7).

The presence of cataplexy being critical to the diagnosis, it is clinically useful to differentiate definite/clear-cut cataplexy from doubtful, possible cataplexy (very rare events, long duration, and unusual triggers). For many clinicians, the presence of clear-cut cataplexy is sufficient to diagnose narcolepsy and narcolepsy-cataplexy is etiologically homogenous. In favor of this hypothesis, almost all (85% to 100%) patients with definite cataplexy share a specific genetic marker, HLA-DQB1*0602, across various ethnic groups (81). This high association contrasts with DQB1*0602 control frequencies ranging from 12% in Japanese, to 20% to 25% in most white populations, and 38% in African Americans (69 ,81).

Other Narcolepsy Symptoms

Although cataplexy is the most specific symptom of narcolepsy, it is frequently mild and rarely the most significant problem clinically for narcoleptic patients. Rather, persistent daytime sleepiness is the most disabling symptom in most patients. Patients with narcolepsy experience a permanent background of daytime sleepiness culminating in overwhelming sleep attacks (3 ,55 ,104 ,114). Narcoleptic subjects may continue their activity during these sleep attacks in a semiautomatic manner, without any memory of the event (automatic behavior). Daytime napping usually relieves daytime sleepiness temporarily (55). Other symptoms

frequently reported include sleep paralysis, an inability to move occurring at sleep/wake transitions, and hypnagogic hallucinations, dreamlike experiences at sleep onset (3 ,55 ,104 ,114). These symptoms are also not specific to narcolepsy and are frequently observed in the general population and patients with other sleep disorders (4 ,7 ,109 ,110). Patients with narcolepsy also have frequently disturbed nocturnal sleep (3 ,55 ,104 ,114). Nightmares, REM behavior disorder, and periodic leg movements during sleep are commonly observed (56 ,104 ,135). Typically, patients with narcolepsy fall asleep easily and wake up after a few hours, unable to fall asleep again at night (3 ,55 ,104 ,114).

Diagnosis of Narcolepsy

The diagnosis of narcolepsy is primarily clinical but polysomnographic studies are useful to document a sleep abnormality and to exclude confounding and/or associated sleep disorders. These tests are also useful to justify future treatment using amphetamine-like stimulants. Most commonly, nocturnal polysomnography with monitoring of breathing and oxygen saturation is carried out to exclude sleep apnea syndrome or other problems potentially disrupting nocturnal sleep. This is followed by a four- to five-nap multiple sleep latency test (MSLT) (4 ,29). These tests must be carried out without any psychotropic treatment and after adequate washout periods (at least 2 weeks for antidepressants, because of their strong REM sleep effects). Sleep logs are used to document adequate nocturnal sleep amount prior to testing. Nocturnal polysomnography in patients with narcolepsy usually reveals a short REM latency in less than 20 minutes (50% of the cases), low sleep efficiency (associated insomnia), and frequently associated periodic leg movements (56). MSLT data indicates short mean sleep latency (SL ≤ 8 minutes) and REM episodes (2 or more sleep onset REM periods, or SOREMPs), a result generally considered diagnostic of narcolepsy (4) (Table 131.1).

Diagnostic Criteria: Narcolepsy

A. A complaint of excessive sleepiness or sudden muscle weakness
B. Recurrent daytime naps or lapses into sleep that occur almost daily for at least 3 months
C. Sudden bilateral loss of postural muscle tone in association with intense emotion (cataplexy)
D. Associated features include:
   1. Sleep paralysis
   2. Hypnagogic hallucinations
   3. Automatic behaviors
   4. Disrupted major sleep episode
E. Polysomnography demonstrates one or more of the following:
   1. Sleep latency less than 10 minutes
   2. REM sleep latency less than 20 minutes and
   3. An MSLT that demonstrates a mean sleep latency of less than 5 minutes
   4. Two or more sleep onset REM periods
F. HLA typing demonstrates DR2 positively
G. Absence of any medical or psychiatric disorder that could account for the symptoms
H. Other sleep disorders may be present but are not the primary cause of the symptoms (e.g., periodic limb movement disorder or central sleep apnea syndrome)

Minimal Criteria: B + C, or A + D + E + G

From American Sleep Disorders Association. ICSD-International Classification of Sleep Disorders diagnostic and coding manual. Rochestv, MN. American Sleep Disorders Association, 1991.

TABLE 131.1. INTERNATIONAL CLASSIFICATION OF SLEEP DISORDERS (ICSD) DIAGNOSTIC CRITERIA FOR NARCOLEPSY

Epidemiologic Studies of Narcolepsy-Cataplexy

Epidemiologic studies have only been performed for narcolepsy-cataplexy. Hublin and associates performed the best-designed study in Finland (57). Using a twin registry and systematic evaluation of 10,000 twin individuals, this study led to a prevalence of .023%. Other studies have led to very similar prevalence values (.02% to .05%) in North American and various other Western European countries (78). Population-based studies suggest a higher prevalence for narcolepsy-cataplexy in Japan (.16% to .17%), but diagnostic criteria did not require polysomnography to verify the diagnosis (54 ,150). Studies comparing sleep disorder populations in a sleep disorder center in Israel suggest a low prevalence (.002%) for the syndrome in this population (65).

Narcolepsy without Cataplexy and Disease Spectrum

In current patient populations, only 50% to 80% of narcoleptic patients have cataplexy. The presence of cataplexy is not necessary to diagnose narcolepsy based on current international diagnostic criteria. Rather, narcolepsy is either diagnosed: (a) in the presence of cataplexy with or without results from associated sleep tests; or (b) without cataplexy but with abnormal MSLT results, associated sleep paralysis or hypnagogic hallucinations and after excluding other sleep disorders (e.g., abnormal breathing during sleep) (4). The rationale for a broader definition of narcolepsy stems from the observation that sleep paralysis, hypnagogic hallucinations, and SOREMPs are all pathologic manifestations of abnormal REM sleep in narcolepsy (52 ,123).

Whereas the prevalence of narcolepsy-cataplexy is well established, the population prevalence of narcolepsy without cataplexy is unknown and could be as high as several percent of the population. Genetic studies indicate a higher HLA association in narcolepsy-cataplexy (85% to 100% DQB1*0602 positive) versus narcolepsy without cataplexy (40% DQB1*0602 positive) (80), suggesting increased etiologic heterogeneity in narcolepsy without cataplexy.

GENETIC PREDISPOSITION IN HUMAN NARCOLEPSY

The genetic aspects of human narcolepsy are complex. Since 1983–1984, human narcolepsy has been known to be associated with HLA-DR2 (59); more recent results suggest a primary association with HLA-DQ (76 ,82 ,96 ,117), with HLA DQB1*0602 playing a primary role in disease predisposition and other HLA alleles having secondary effects (88). The importance of environmental factors is indicated by the low degree of concordance of monozygotic twins (25% to 31%) (78). Genuine multiplex families are rare; most narcoleptic patients do not have a family history. Only 1% to 2% of first-degree relatives of narcolepsy patients ever develop narcolepsy-cataplexy (20 ,44 ,78). One to two percent affected in first-degree relatives indicates a 20- to 40-fold increased risk that cannot be explained by HLA-associated genetic factors alone (78). Further, multiplex families are more frequently HLA-DQB1*0602 negative than sporadic cases, suggesting the importance of non-HLA genetic factors (83).

The HLA association observed in narcolepsy suggests a primary involvement of the immune system in the pathophysiology of the disorder, yet all studies aiming at demonstrating an autoimmune mediation have failed (28 ,77 ,86). The recent discovery that narcolepsy is associated with undetectable CSF hypocretin-1 levels (106) suggests that this hypothesis should be revisited now that a potential target cell population has been identified.

ANIMALS MODELS OF NARCOLEPSY AND HYPOCRETIN (OREXIN)

Part of "131 - Pathophysiological and Pharmacologic Aspects of the Sleep Disorder Narcolepsy "

In 1973 and 1974, narcolepsy was first reported in a dachshund and a poodle (60 ,91). Autosomal recessive occurrence of narcolepsy in doberman pinschers and Labrador retrievers was subsequently discovered, and a colony of genetically narcoleptic Dobermans and Labradors was established at Stanford University (13). As with human patients, narcoleptic animals exhibit muscle weakness (cataplexy) when emotionally stimulated. Polygraphic studies in narcoleptic dogs have also demonstrated that narcoleptic canines have a shorter latency to drowsiness, light sleep, and REM sleep than do control animals (104). Sleep paralysis and hypnagogic hallucinations may also exist in narcoleptic dogs, but are impossible to document owing to their subjective nature.

In narcoleptic Dobermans and Labradors, the major susceptibility gene, canarc-1, is unlinked to dog leukocyte antigen (DLA) (87). Linkage analysis with various genetic markers, including minisatellite probes and functional candidate gene probes, revealed that the canine narcolepsy gene cosegregated with a polymorphic band cross-reacting with the switch region of the human immunoglobulin μ heavy-chain gene (87). After 10 years of chromosome walking in dogs, the canine narcolepsy gene was finally identified as the hypocretin receptor 2 gene (*Hcrtr2*) (68). Three mutations causing loss of function of Hcrtr 2 and impaired postsynaptic hypocretin neurotransmission were identified in Labradors, dachshunds, and Dobermans, respectively (68). The discovery of *canarc-1* (*Hcrtr-2*) was followed by the report that preprohypocretin (prepro-Orexin) knockout mice also exhibit a narcolepsy-like phenotype, shorter REM sleep onset, and episodes of behavioral arrest similar to cataplexy in canine narcolepsy (30). Deficits of either the hypocretin ligand or its receptor-2–mediated transmission thus generate narcolepsy in animal models.

Following up on this discovery, hypocretin 1-peptide (Orexin A) levels were measured in the cerebrospinal fluid (CSF) of narcoleptic humans. Strikingly, Hcrt 1 was below detectable levels in seven out of nine patients, in contrast to eight control subjects who all had normal Hcrt 1 levels (106). These results suggest that a deficit in hypocretin transmission is also involved in cases of human narcolepsy, although disease heterogeneity may exist (106). (See more details in monoaminergic cholinergic imbalance and deficit in hypocretin neurotransmission sections.)

TREATMENT OF HUMAN NARCOLEPSY

Part of "131 - Pathophysiological and Pharmacologic Aspects of the Sleep Disorder Narcolepsy "

Pharmacologic Treatment of Daytime Sleepiness with Amphetamine-like Compounds

Nonpharmacologic treatments (i.e., behavioral modification such as regular napping and work accommodations) are often helpful (128 ,129), but are rarely sufficient to control the symptoms. Referral to patient support groups (e.g., narcolepsy network) and giving directives regarding driving and other potentially dangerous activities is critical until a better understanding and control of the disorder by the patient is achieved. In a recent survey by a patient group organization (8), 94% of all patients reported using pharmacologic therapies, mostly stimulant medications. Sleepiness is usually treated using amphetamine-like CNS stimulants or modafinil, a novel wake-promoting compound unrelated to the amphetamines (Table 131.2). The most commonly used amphetamine-like compounds are methamphetamine, *d*-amphetamine, methylphenidate, pemoline, and mazindol (Table 131.2). The most important pharmacologic property of amphetamine-like stimulants is to release catecholamines, mostly dopamine and norepinephrine (62 ,147). Monoamine reuptake blockade and MAO inhibition also occur at high doses and the importance of these secondary pharmacologic effects varies from one amphetamine derivative to another (73). Pharmacologic studies using the canine narcolepsy model strongly suggest that presynaptic enhancements of dopamine transmission contribute to the EEG arousal effects of amphetamine-like CNS stimulants and modafinil (103). (See also the Canine Narcolepsy section.)

Of importance is the report that in animals, amphetamine-like stimulants are neurotoxic at high doses for catecholaminergic (e.g., dopaminergic) neurons (127).

Stimulant Compound	Usual Daily Doses[a]	Half-Life (Hours)	Side Effects/Notes
Sympathomimetic stimulants			
D-Amphetamine sulfate[b]	5–60 mg (15–100 mg)	10–12	Irritability, mood changes, (urinary pH-dependent) headaches, palpitations, tremors, excessive sweating, insomnia
Methamphetamine HCl[b]	5–60 mg (15–80 mg)	4–5	Same as *d*-amphetamine (urinary pH-dependent) may have a greater central over peripheral effect[c]
Methylphenidate HCl[b]	10–60 mg (30–100 mg)	3–4	Same as amphetamines; better therapeutic index than *d*-amphetamine with less reduction of appetite or increase in blood pressure; short duration of action
Pemoline	20–115 mg (37.5–150 mg)	16–18	Less sympathomimetic effect; milder stimulant; slower onset of action; a tendency for drug build-up; occasionaly produces liver toxicity; not a controlled substance
Mazindol[d]	2–6 mg (NA)	33–55	Weaker CNS stimulant effects; anorexia, dry mouth, irritability, headaches, gastrointestinal symptoms; reported to have less potential for abuse
Other Agents for treatment of EDS			
Modafinil	100–400 mg (NA)	8–14	No peripheral sympathomimetic action; headaches; nausea; reported to have less potential for abuse
Caffeine[e]	100–200 mg (NA)	3.5–5	Palpitations, hypertension; weak stimulant effect; 100 mg of caffeine roughly equivalent to one cup of coffee
MAO inhibitors with alerting effect			
Selegiline[b]	5–40 mg (NA)	0.15[f]	Low abuse potential; partial (10–40%) interconversion to amphetamine
Brofaromine	150 mg (NA)	12–19	Reversible MAOA selective inhibitor

[a]Dosages recommended by the ASDA (90) are listed in parentheses (usual starting dose and maximal dose recommended).
[b]Demonstrated anticataplectic effects in narcoleptic dogs.
[c]Methamphetamine is reported to have more central effects (38) and may predispose more to amphetamine psychosis (139). The widespread misuse of methamphetamine had led to severe legal restriction on its manufacture, sale; and prescription in many countries (112). Note that the molecular weight of this compound is about half that of *d*- and *l*-amphetamine; thus, methamphetamine contains twice as much active molecules as *d*- or *l*-amphetamine per mg dose. *L*-amphetaine (dose range 20–60 mg) is not available in the US, but probably has no advantage over *d*-amphetamine in the treatment of narcolepsy (slightly weaker stimulant).
[d]Demonstrated anticataplectic effects in humans.
[e]Caffeine can be bought without prescription in the form of tablets (NoDoz, 100 mg; Vivarin 200 mg caffeine) and is used by many patients with narcolepsy prior to diagnosis.
[f]Half-lives of metabolites (amphetamine and methamphetamine) are long.

TABLE 131.2. COMMONLY USED TREATMENTS FOR EXCESSIVE DAYTIME SLEEPINESS (EDS)

The clinical use of stimulants in narcolepsy has been the subject of a recent American Sleep Disorders Association (ASDA, now American Academy of Sleep Medicine) Standards of Practice publication (9). Typically, the patient is started at a low dose, which is then increased progressively to obtain satisfactory results (Table 131.2). Studies have shown that daytime sleepiness can be greatly improved subjectively, but that sleep variables are never completely normalized by stimulant treatment (92). Low efficacy compounds/milder stimulants (e.g., modafinil, or more rarely, pemoline) are usually tried first. More effective amphetamine-like stimulants (i.e., methylphenidate, *d*-amphetamine, and methamphetamine) are then used if needed. The final dose of stimulant medication used varies widely from patient to patient, depending on tolerance, personality, efficacy, and lifestyle (from no stimulant treatment to very high doses). Patient input and work environment is very important. Some patients prefer to use high doses of long-acting, slow-release preparations to stay awake all day long, whereas others combine lower doses and short half-life derivatives (e.g., methylphenidate) with scheduled napping. Stimulant compounds are generally well tolerated. Minor side effects such as headaches, irritability, nervousness, tremors, anorexia, palpitations, sweating, and gastric discomfort are common (Table 131.2). Cardiovascular impact such as increased blood pressure is possible considering established sympathomimetic effects in animals, but has been

remarkably difficult to document in human studies (145). Surprisingly, tolerance rarely occurs in this patient population and “drug holidays” are not recommended by the American Academy of Sleep Medicine (90). Rather, a slight increase in dosage is preferable. Exceptionally, psychotic complications are observed, most often when the medications are used at high doses and chronically disrupt nocturnal sleep.

Amphetamine was first used to treat narcolepsy in 1935 (120), only 8 years after Alles initially synthesized it (5). Both the *l*- and *d*-isomers have been used for the treatment of narcolepsy, either in isolation or as a racemic mixture (available in the United States). The *d*-isomer is a slightly more potent stimulant (113 ,115) and is most generally used. *L*-Amphetamine is occasionally used in some European countries (dose range 20 to 60 mg) (112). *D*-Amphetamine is the second most frequently prescribed narcolepsy treatment after methylphenidate (8). It is well absorbed by the gastrointestinal tract and partially metabolized in the liver using aromatic and aliphatic hydroxylation. This process yields parahydroxyamphetamine and norephedrine, respectively, both of which are biologically active (158). Amphetamine is metabolized into benzoic acid (23%), which is subsequently converted to hippuric acid or parahydroxyamphetamine (2%). This in turn is converted to parahydroxynorephedrine (.4%). Thirty-three percent of the oral dose is excreted unchanged in the urine. Importantly, urinary excretion of amphetamine and many amphetamine-like stimulants is greatly influenced by urinary pH. Amphetamine is a weak base and at a physiologic pH, it exists mainly as a charged amine [RNH3]+, which is poorly reabsorbed in the renal tubules. Acidifying the urine thus favors the excretion of the charged form of the amine (16), increases urinary excretion versus liver catabolism, and reduces the half-life. At urinary pH 5.0, the elimination half-life of amphetamine is very short (about 3 to 5 hours) but at pH 7.3 it increases to 21 hours (16). Sodium bicarbonate delays excretion of amphetamine and prolong its clinical effects, whereas ammonium chloride shortens amphetamine toxicity. Finally, dextroamphetamine is available as a sulfate-base derivative or as spansule (slow-release) capsules.

Methamphetamine is the most efficacious and most potent amphetamine derivative available. This compound is extremely useful in subjects with severe sleepiness who need high doses. The addition of a methyl group makes this derivative more lipophilic, thus increasing CNS penetration and providing a better central over peripheral profile. The widespread misuse of methamphetamine has led to severe legal restriction on its manufacture, sale, and prescription in many countries (112), but it is available in the United States. It should also be noted that the molecular weight of the most commonly used form of methamphetamine (hydrochloride) is about half that of *d*- and *l*-amphetamine salt (sulfate). Methamphetamine preparations thus contain twice as many active amphetamine molecules when compared to *d*- or *l*-amphetamine per mg dose. The simple chlorate to sulfate formulation difference largely explains the higher potency of methamphetamine.

Yoss and Daly introduced methylphenidate for the treatment of narcolepsy almost 50 years ago (160). It is now the most commonly prescribed stimulant medication in the United States, with 46% of narcoleptic patients using the compound on a regular basis (8). Part of its popularity is owing to its relatively short duration of action (approximately 3 to 4 hours). This property allows narcoleptic patients to use the compound on an as-needed basis while keeping open the possibility of napping. The compound is also reported to produce fewer psychotic complications at high doses (116). A slow release formulation is available but less frequently used.

Pemoline is generally better tolerated than methamphetamine or *d*-amphetamine but it is also less efficacious and less potent, and occasionally produces liver toxicity. After taking a therapeutic dose of pemoline (40 mg), peak levels in serum are reached within 4 to 6 hours. The half-life is 16 to 18 hours. Pemoline is partially metabolized by the liver. Metabolites include pemoline conjugates, pemoline dine, and mandelic acid. After oral administration of 40 mg of pemoline, 35% to 50% of the dose is excreted in the urine within 32 hours, and only a minor fraction is present as metabolites (41). The long duration of action of pemoline may be associated with a better compliance in narcoleptic patients (130). Pemoline most selectively blocks dopamine reuptake and only weakly stimulates dopamine release. Fatal hepatotoxicity has been reported and may be dose related (17 ,142). Pemoline should not be prescribed to patients with impaired hepatic function, and hepatic function should be carefully monitored during chronic drug administration. The recent introduction of modafinil, a novel wake-promoting agent with a similar profile and fewer side effects, has greatly diminished the use of this compound in narcolepsy.

Mazindol is less frequently used because of its weaker stimulant activity (58). It is a weak releasing agent for dopamine, but it also blocks dopamine and norepinephrine reuptake with high affinity (103). Mazindol is effective for both excessive daytime sleepiness and cataplexy (58). Mazindol is absorbed quantitatively at a medium rate from the gastrointestinal tract, and the peak blood concentration is reached after 2 to 4 hours. The half-life of clearance from blood was estimated at 33 to 55 hours (47).

Modafinil and Other Wake-Promoting Agents

Modafinil, a compound structurally distinct from amphetamines, has recently been approved in the United States for the treatment of narcolepsy and essential hypersomnia. This compound is also increasingly explored to treat other conditions, such as residual sleepiness in treated obstructive sleep

apnea or fatigue in multiple sclerosis. Modafinil has been available in France since 1986, and long-term follow-up suggests no remarkable side-effect profile and low abuse potential. Clinical trials in France and Canada have shown that 100 to 300 mg of modafinil is effective for improving daytime sleepiness in narcoleptic and hypersomnolent subjects without interfering with nocturnal sleep. It has limited efficacy on cataplexy and the symptoms of abnormal REM sleep (15 ,19 ,21). Recent double-blind trials on 283 narcoleptic subjects in 18 centers in the United States and on 75 narcoleptic subjects in 11 centers in Canada revealed that 200 and 400 mg of modafinil significantly reduced sleepiness and improved patients' overall clinical condition (1 ,26). However, it is also reported that patients who have been previously treated with methylphenidate may respond more poorly to modafinil (26). Modafinil is well tolerated by these subjects, and adverse experiences with modafinil use occur at rates comparable to placebo (1 ,26). In humans, modafinil exhibits a linear pharmacokinetic profile for doses ranging from 50 to 400 mg, with a terminal elimination half-life ($t_{1/2}$) of 9 to 14 hours (159). Modafinil is extensively metabolized to two major pharmacologically inactive metabolites, modafinil acid and modafinil sulfone, which are renally excreted. Less than 10% of the oral dose of modafinil is excreted unchanged, and 40% to 60% is excreted as unconjugated acid in urine (159).

The exact mode of action of modafinil is still uncertain. The wake promoting effects of the compounds have been suggested to involve α1-adrenergic stimulation (67) and/or serotonergic–GABAergic interactions (37). The compound interacts with the dopaminergic system at high doses and is neuroprotective in the MTPP model (37 ,39). Recent work by our group rather suggests that selective, but low-potency, dopamine reuptake inhibition mediates the wake-promoting effects of modafinil (84 ,103). In rats, modafinil acutely decreased both REM and non-REM sleep in rats for up to 5 to 6 hours without inducing a secondary rebound hypersomnolence (34). This contrasts with the intense recovery sleep seen after amphetamine administration (34). This unique feature of modafinil (wakefulness without rebound hypersomnia) may be explained by the pharmacokinetics profile of the compound (modafinil has a significantly longer half-life than amphetamine or methylphenidate) (159). Alternatively, this important difference may be owing to its unique pharmacodynamic profile, for example, dopamine uptake inhibition versus dopamine release effects for amphetamine (34).

Several factors make modafinil an attractive alternative to amphetamine-like stimulants. First, animal studies suggest that the compound does not affect blood pressure as much as amphetamines do (50) (potentially the result of its lack of effects on adrenergic release or reuptake). This suggests that modafinil might be useful for patients with a heart condition or high blood pressure. Second, animal data suggest no neurotoxic effects and no or less rebound hypersomnolence on withdrawal. Third, data obtained to date suggest that tolerance and dependence are limited with this compound (15), although a recent animal study reports a cocaine-like discriminative stimulus and reinforcing effects of modafinil in rats and monkeys, respectively (42). Finally, clinical studies suggest that the alerting effect of modafinil might be qualitatively different from that observed with amphetamine (15). In general, patients feel less irritable and/or agitated with modafinil than the amphetamines (15). In animal experiments, modafinil did not induce behavioral excitation, as measured by lack of locomotor activation (35). Considering the many advantages of modafinil over amphetamine treatment (fewer cardiovascular side effects, low abuse potential, lower levels of tolerance, and less rebound sleep), modafinil may replace amphetamine-like stimulants as a first-line treatment for excessive daytime sleepiness.

Caffeine, a xanthine derivative, may be the most popular and widely consumed stimulant in the world. The average cup of coffee contains about 50 to 150 mg of caffeine. Tea (25 to 90 mg/5 oz), cola drinks (35 to 55 mg/12 oz), chocolate (15 to 30 mg/1 oz), and cocoa (2 to 20 mg/5 oz) also contain significant amounts of caffeine. Taken orally, caffeine is rapidly absorbed, taking 47 minutes to reach maximum plasma concentration. The half-life of caffeine is about 3.5 to 5 hours (143). A slow-release soft gelatin caffeine capsule is also available with a mean delay to peak plasma concentration of 4 hours (143). The behavioral effects of caffeine include increased mental alertness, faster and clearer flow of thought, increased wakefulness, and restlessness (121). Fatigue is reduced, and the need for sleep is delayed (121). Physical effects of caffeine include palpitations, hypertension, and increased secretion of gastric acid and increased urine output (121). Heavy consumption (12 or more cups a day, or 1.5 g of caffeine) can cause agitation, anxiety, tremors, rapid breathing, and insomnia (121). The mechanism of action of caffeine involves antagonism of an adenosine (nonspecific) receptor and of adenosine-induced neuronal inhibition (121). Considering the fact that 100 mg of caffeine is roughly equivalent to one cup of coffee, caffeine does not possess the efficacy to counteract the pathologic sleepiness seen in narcolepsy. Nevertheless, caffeine in the form of tablets can be bought without a prescription (NoDoz, 100 mg caffeine; Vivarin, 200 mg caffeine), and is used by many patients with narcolepsy prior to diagnosis.

Antidepressants and the Pharmacologic Treatment of Cataplexy

Amphetamine stimulants have little effect on cataplexy, and additional compounds are most often needed to control cataplexy if the symptom is severe enough to warrant treatment. Since the 1960s, it has been known that imipramine is very effective in reducing cataplexy (2). Together with protriptyline and clomipramine, these tricyclic antidepressants

are now the most commonly used anticataplectic agents (8) (Table 131.3). Other antidepressant compounds of the tricyclic family have also been used with some success (Table 131.3). The use of tricyclic antidepressants in the treatment of cataplexy, however, is hampered by a number of problems. The first is the relatively poor side-effect profile of most tricyclic compounds. These are mostly owing to their anticholinergic properties, leading to dry mouth (and associated dental problems), tachycardia, urinary retention, constipation, and blurred vision (Table 131.3). Additional side effects are weight gain, sexual dysfunction (impotence and/or delayed orgasm), tremors, antihistaminergic effects leading to sedation and occasionally orthostatic hypotension owing to the α1-adrenergic blockade effects of some compounds. In this respect, protriptyline is often preferred because of its previously reported mild stimulant effect (49). Nighttime sleep might also become more disturbed because of increased muscle tone and leg movements (122 ,152). The cardinal pharmacologic property of tricyclic antidepressants is their ability to inhibit the reuptake of norepinephrine (and epinephrine) and serotonin (14). The degree of uptake inhibition of norepinephrine and serotonin is quite variable depending on the compound and the existence of active metabolites (mostly active on adrenergic uptake) (14). Additionally, some tricyclic compounds, such as protriptyline, are also weak dopamine reuptake inhibitors (14).

Antidepressant Compounds	Usual Daily Doses	Side Effects/Note
Commonly used compounds		
Imipramine	10–100 mg	Dry mouth, anorexia, sweating, constipation, drowsiness (51).
Desipramine	25–200 mg	Effects and side effects similar to those of imipramine demethylated metabolite of imipramine (51).
Protryptiline	5–60 mg	Some reports suggest improvement in vigilance measures (49), whereas other reports are negative (no improvement in performance or daytime sleepiness) (93).
Clomipramine	10–150 mg	Digestive problem, dry mouth, sweating, tiredness, impotence (45,140). Its active metabolite, desmethylclomipramine, is shown to be more potent in the canine model (97).
Fluoxetine[a]	20–60 mg	Nausea, dry mouth, fewer side effects, long half-life (60 hours); no anticholinergic or antihistaminergic effects; good anticataplectic effect but less potent than clomipramine (63).
Some clinical trials, but less commonly used		
Zimelidine[a]	100 mg	Less sedative effect; no anticholinergic or antihistaminergic effects; potent anticataplectic compound (94). Its active metabolite, norzimelidine, is shown to be more potent than zimelidine in the canine model (97).
Femoxitine[a]	600 mg	Fewer side effects than clomipramine but less potent (136) pharmacologic effects similar to fluoxetine.
Fluvoxamine[a]	50–300 mg	Gastrointestinal side effects (134). No active metabolities; pharmacologic profile similar to fluoxetine; less active than clomipramine.
Paroxetine[a]	20–60 mg	Less anticholinergic effects; cardiovascular side effects; effective on cataplexy (with yohimbine) (119).
Viloxazine	100–300 mg	Few side effects; selective but low potency adrenergic uptake inhibitor, active on human cataplexy and possibly sleepiness, but not available in most countries (40,43); effective anticataplectic agent in canines (85).
Venlafaxine	150–375 mg	New serotonergic and adrenergic uptake blocker; no anticholinergic effects, effective on cataplexy sleepiness in a small pilot study (146), low potency.

[a]Clinical trial results using these compounds (63,94,134,136) suggest that SSRIs are effective for the treatment of cataplexy or REM sleep related symptoms while inducing fewer side effects than classical tricyclic antidepressants. It is, however, still not conclusive whether SSRIs can be recommended as a first line of treatment, because SSRIs are usually less potent than tricyclic antidepressants (136).
SSRIs, selective serotonin reuptake inhibtors.

TABLE 131.3. ANTIDEPRESSANTS CURRENTLY USED AS ANTICATAPLECTIC AGENTS

The introduction of newer antidepressants with selective serotonergic uptake inhibition properties (e.g., SSRIs) and no anticholinergic effects, such as fluoxetine, fluvoxamine, paroxetine, sertraline, femoxitine, zimelidine, and trazodone has raised hope that the control of cataplexy could be achieved with fewer side effects. In general, however, clinicians have been less impressed with the potency of the serotonergic compounds on cataplexy (63 ,94 ,136). This experience parallels experiments in canine narcolepsy suggesting that adrenergic, not serotonergic, uptake inhibition mediates

the anticataplectic effects of most antidepressant medications (85 ,97). Among the SSRIs, fluoxetine is a viable alternative to tricyclic compounds (63). Fluoxetine has a good side-effect profile and may induce less weight gain, a significant advantage for some patients. Venlafaxine, a novel serotonergic and adrenergic reuptake blocker, also has been used recently with good success. Finally, the introduction of reboxetine, a specific adrenergic reuptake blocker, may offer a novel and more effective alternative to SSRIs and tricyclic antidepressants based on animal data.

In addition to the antidepressants listed in Table 131.3 , γ-hydroxybutyrate (GHB), a hypnotic compound discussed in more detail in the section on disrupted nocturnal sleep , has been shown to alleviate cataplexy during long-term administration. GHB is an endogenous constituent of mammalian brains, synthesized locally from GABA, which may play a role as an inhibitory neuromodulator (18). Monoamine oxidase inhibitors (MAOIs) are known to potently reduce REM sleep, and are therefore excellent candidate anticataplectic agents; however, these compounds are less often used owing to their poor safety profile. Selective or reversible MAOIs have recently become available, but clinical trials of these compounds at a large scale are still not available (104).

Treatment of Sleep Paralysis and Hypnagogic Hallucinations

The treatment of these two symptoms is not well codified. Hypnagogic hallucinations can be quite bothersome, and often occur in patients who also suffer from frequent nightmares. As they are a manifestation of sleep onset REM sleep, the compounds that suppress REM sleep are usually helpful in alleviating this symptom, and tricyclic antidepressant treatment has been reported to have some beneficial effects (149). Sleep paralysis only rarely requires treatment, but tricyclic antidepressants are also very effective for preventing this symptom. Recently, high doses (60 mg QD) of fluoxetine have been advocated as a very active treatment for isolated sleep paralysis (61). GHB is also effective in suppressing hypnagogic hallucinations, sleep paralysis, and cataplexy (72).

GHB and Treatment of Disturbed Nocturnal Sleep

Insomnia is a major complaint in narcoleptic subjects. Several studies reported that benzodiazepine hypnotics are effective in consolidating nighttime sleep in patients with narcolepsy (151). GHB, a compound with remarkable REM- and SWS-inducing properties, has also been used for consolidating nighttime sleep, an effect that leads to decreased sleepiness and cataplexy the following day (24 ,25 ,137 ,138). Large-scale double-blind placebo controlled clinical trials are in progress in the United States to re-establish GHB as a first line treatment for narcolepsy-cataplexy (104). The compound is especially useful in patients with severe insomnia and cataplexy who do not tolerate well the side effects of antidepressant medication on sexual potency. The mode of action of GHB on sleep and sleep-related symptoms is unknown, but may involve decreased dopaminergic tone after GHB (18). Because of its positive effects on mood and libido, its SWS-enhancing properties, and a subsequent increase in growth hormone release, the drug is widely abused by athletes and other populations (31 ,70); therefore, several states have passed legislation restricting access to GHB requesting by prescriptions for its use. The compound has also been reported to increase periodic leg movements in narcoleptic patients (27).

GHB is absorbed 15 to 20 minutes after oral ingestion, and peak plasma concentration occurs at 60 to 120 minutes. The elimination half-life is 20 minutes (111 ,157). Exogenous GHB is almost completely eliminated by oxidative biotransformation to carbon dioxide and water, less than 5% is detected unmetabolized in the urine (111 ,157). At low doses, GHB is anxiolytic and myorelaxant. At intermediate doses, GHB increases slow wave sleep and REM sleep (64). However, because of the short half-life of the compound, its effects on sleep architecture are short-lasting (about 3 to 4 hours) and administration thus has to be repeated two to three times during the night (20 to 40 mg/kg per night). Overdoses (a single dose of 60 to 100 mg/kg) induce dizziness, nausea, vomiting, confusion, agitation, epileptic seizures, hallucinations, and coma with bradycardia and respiratory depression (66). Death has been reported and the therapeutic window is narrow (LD_{50} = 5- to 15-fold the dose inducing coma). Although the compound is structurally related to GABA and is a natural metabolite of the neurotransmitter, its mode of action involves specific non-GABAergic binding sites (75). GHB and GABAB receptors may interplay functionally (71). GHB is also known to inhibit firing of dopaminergic neurons, dopamine release, and dopamine synthesis (156).

PHARMACOLOGY AND NEUROCHEMISTRY OF CANINE NARCOLEPSY

Part of "131 - Pathophysiological and Pharmacologic Aspects of the Sleep Disorder Narcolepsy "

Pharmacologic Control of Canine Cataplexy

The canine model of the disorder has been used to pharmacologically dissect the mechanisms involved in the control of cataplexy and alertness. These experiments led us to conclude that the control of cataplexy and REM sleep are very similar, although in several aspects, cataplexy is not identical to natural REM sleep muscle atonia (102). Activation of cholinergic systems using the acetylcholinesterase inhibitor physostigmine, for example, greatly exacerbates cataplexy (12 ,33). This cholinergic effect is mediated via muscarinic receptors because muscarinic stimulation aggravates cataplexy,

whereas its blockade suppresses it, and nicotinic stimulation or blockade has no effect (12 ,33). These results parallel data obtained on cholinergic systems and REM sleep control. (See refs. 144 and 148 for review.)

Monoaminergic transmission is also critical for the control of cataplexy. All therapeutic agents currently used to treat cataplexy (i.e., antidepressants or MAOIs) are known to act on these systems. Furthermore, whereas cholinergic neurons are activated during REM sleep, the firing rate of monoaminergic neurons in the brainstem, such as in the locus ceruleus (LC) and raphe magnus, are well known to be dramatically depressed during this sleep stage (10 ,153). In contrast, dopamine neurons of the ventral tegmental area (VTA) and substantia nigra (SN) do not significantly change their activity during the sleep cycle (89 ,154).

Although antidepressants and MAOIs enhance monoaminergic transmission, these compounds generally lack specificity and globally enhance serotonergic, adrenergic, and dopaminergic transmission. Using newer uptake inhibitors and releasing agents with selective monoamine effects, we have demonstrated that the presynaptic activation of adrenergic, but not dopaminergic or serotonergic, systems mediates the anticataplectic effects of currently available antidepressive treatments (85 ,97). This suggests that cataplexy, and possibly REM sleep atonia are more selectively modulated by adrenergic systems. Interestingly, presynaptic activation of dopamine transmission with dopamine uptake inhibitors had potent alerting effects (103) but no effect on cataplexy (85).

Receptor-Specific Regulation of Cataplexy: A Pathologic Model of REM Sleep Atonia

More than 200 compounds with various pharmacologic properties (cholinergic, adrenergic, dopaminergic, serotonergic, prostaglandins, opioids, benzodiazepines, GABAergics, and adenosinergics) have been studied in the narcoleptic canine model. (See ref. 104 for a recent review.) Although many compounds, such as M2 antagonists, α1-agonists, α2-antagonists, dopamine D2(3) antagonists, 5-HT1a agonists, TRH analogues, prostaglandin E2, and L type Ca^{2+} channel blockers reduce cataplexy, very few compounds significantly aggravate cataplexy. Because REM sleep can be easily disturbed nonspecifically in pharmacologic studies, aggravations in cataplexy are considered to be the most specific effect. The stimulation of muscarinic M2 (non-M1) receptors significantly aggravates cataplexy. Among monoaminergic receptors, the postsynaptic adrenergic α1b receptors (79 ,99) and presynaptic α2 receptors (100) were also found to aggravate cataplexy, a result consistent with a primary adrenergic control of cataplexy. It was also found that dopamine D2(3) agonists significantly aggravated cataplexy and induced drowsiness in these animals. To date, no other receptor ligands (e.g., adenosinergic, histaminergic or GABAergic) have been found to aggravate cataplexy (104).

The cataplexy-inducing effects of D2(3) compounds on cataplexy, however, are difficult to reconcile with the fact that dopaminergic uptake blockers and releasing agents have absolutely no effect on cataplexy (85). Interestingly, the aggravation of cataplexy by D2(3) agonists is blocked by adrenergic, but not dopaminenergic, uptake inhibitors (105), suggesting some functional interaction between the dopaminergic and adrenergic systems for the regulation of cataplexy.

Presynaptic Stimulation of Dopamine Transmission Mediates the EEG Arousal Effects of Amphetamine-like Stimulants

Amphetamine-like CNS stimulants currently used clinically for the management of sleepiness in narcolepsy presynaptically enhance monoaminergic transmission; however, these compounds also lack pharmacologic specificity. In order to study the mode of action of these compounds on daytime sleepiness, the stimulant properties of several dopaminergic and adrenergic uptake inhibitors were quantified and compared to the effects of amphetamine and modafinil using 6-hour daytime polygraphic recordings in the canine narcolepsy model (103). In spite of their lack of effects on cataplexy, all dopaminergic uptake inhibitors induced significant EEG arousal. In contrast, nisoxetine and desipramine, two potent adrenergic uptake inhibitors, had little effect on EEG arousal but significantly suppressed REM sleep, a finding that is consistent with their potent anticataplectic effects. Furthermore, the *in vivo* potency of dopamine uptake inhibitors on EEG arousal correlates well with the *in vitro* dopamine transporter (DAT), but not with the noradrenaline transporter (NET), binding affinities for individual compounds (103). These results are consistent with the hypothesis that presynaptic modulation of dopamine mediates the EEG arousal effects of these compounds. Interestingly, it was also found that modafinil binds to the DAT site with low affinity (84 ,103), similar to the affinity range for amineptine (a dopamine uptake inhibitor that also enhances EEG arousal in our model). Thus, DAT binding may also contribute to the stimulant properties of modafinil.

Midbrain Dopaminergic Systems Are Involved in the Regulation of Cataplexy and Excessive Daytime Sleepiness in Narcolepsy

The site of action for dopaminergic modulation of cataplexy recently has been identified using microdialysis experiments. D2(3) agonist injections into the ventral tegmental area (VTA) (126) and substantia nigra (SN) (53), two regions

where dopaminergic cell body autoreceptors are densely packed, was found to reproduce the effects of small intravenous doses of dopamine agonists or antagonists on cataplexy (98). An injection of the same compounds in the pontine reticular formation (PRF), where dopamine autoreceptors are less densely packed, has no effect (126). This suggests that the D2(3) effect is genuinely mediated by autoreceptor stimulation and that cataplexy may be modulated by changes in dopaminergic activity originating from the VTA and SN. The perfusion of dopamine uptake inhibitors in these midbrain dopaminergic nuclei does not modify cataplexy (53). Because various electrophysiologic and pharmacologic studies have demonstrated that dopamine reuptake is of physiologic importance in the limbic forebrain, striatum, and cortical hemispheres, but not in midbrain dopaminergic neurons (108), acting sites for the dopaminergic regulation of cataplexy and sleepiness may be anatomically different. This may explain why D2(3) agonists induce both cataplexy and sleepiness (53 ,126), whereas dopamine uptake inhibitors only induce significant EEG arousal but have no effect on cataplexy (105). In this model, enhancement of DA transmission at the terminal region may be sufficient to induce a wake-promoting effect; however, reduction in the activity of DA neurons and interactions with adrenergic systems may be required for the modulation of cataplexy (101 ,105).

Cholinergic Hypersensitivity and the Regulation of Cataplexy

The effects of cholinergic stimulation in various brain areas were also examined in narcoleptic and control canines. Local injection or perfusion of carbachol, a predominantly muscarinic agonist, into the PRF was found to aggravate canine cataplexy in a dose-dependent fashion (125). Acetylcholine release in the PRF was significantly elevated during the FECT when narcoleptic animals had multiple cataplectic attacks, while no increase in acetylcholine levels was observed in control animals (124). The results obtained in the PRF with cholinergic agonists and acetylcholine release were expected considering the well-established role of pontine cholinergic systems in the regulation of REM sleep. In this experiment, however, narcoleptic dogs were found to be consistently more sensitive to cholinergic stimulation than control animals (125). More surprisingly, however, it was also found that the local injection of carbachol unilaterally or bilaterally (2 to 10 nmol per site) into the BF (rostral to the preoptic area, in the vertical or horizontal limbs of the diagonal band of Broca and medial septum) also dose-dependently aggravated cataplexy (107). This manipulation induced long-lasting muscle atonia episodes with desynchronized EEG in narcoleptic canines (107). The same pharmacologic manipulation (10 nmol of carbachol) did not induce cataplexy in normal animals, but rather induced wakefulness, as previously reported in rats and cats (11). The BF is anatomically connected with the limbic system, which is regarded as a critical circuit for integrating emotions. Furthermore, BF neurons are known to respond to the arousing property of appetitive stimuli (131), which potently induce cataplexy in narcoleptic dogs. Considering the fact that emotional excitation is an alerting stimulus in normal animals, but induces cataplexy in narcoleptic animals, the BF may be involved in triggering a paradoxic reaction to emotions—atonia rather than wakefulness—in narcoleptic animals.

Monoaminergic/Cholinergic Imbalance and Hypocretin Deficiency

As detailed, cholinergic and monoaminergic imbalances are central to the pathophysiology of narcolepsy and the control of natural REM sleep. The fact that impaired hypocretin transmission is involved in both animal and human narcolepsy indicates that the hypocretin peptides must be tightly connected functionally with monoaminergic and cholinergic systems. Deficit in the production of the hypocretin ligand, as well as mutations in one of the receptors (Hcrtr 2), induces narcolepsy in mice and dogs, respectively (30 ,68). In humans, abnormalities in the production of the ligands are most likely to be the etiology of the disease in most cases (106).

De Lecea and associates first identified hypocretins using a subtraction technique aimed at the isolation of hypothalamic-specific transcripts (32). The same neuropeptide system was independently identified by Sakurai and associates and was named Orexin (133). These authors isolated two new neuropeptides, Orexin A and B, as endogenous ligands for previously poorly characterized orphan G-protein–coupled receptors (GPCRs). The hypocretin receptors are closely related to other neuropeptide GPCRs, such as the Y_2 neuropeptide Y receptor (26% identity) and the TRH receptor (25%). Hypocretin-1 and -2 correspond to Orexin A and B, respectively (132). These molecules are processed from the same precursor peptide (preprohypocretin). Hypocretins bind and activate two closely related GPCRs, the Hcrtr 1 (Ox1R) and Hcrtr 2 (Ox2 R) receptors (133). Hcrtr 1 is selective for hypocretin-1 (20 to 100 × higher affinity) whereas Hcrtr 2 exhibits similar affinity for both hypocretin-1 and -2 (133). Preprohypocretin mRNA and hypocretin-1 or -2 immunoreactivity colocalize in a small group of neurons located within and around the lateral hypothalamic area (LHA) in adult rat brains (32 ,118 ,133).

Hypocretins were initially believed to control appetite and food intake (thus the name Orexin, appetite in Greek) because of their discrete localization within the lateral hypothalamus. Similar to neuropeptide Y and leptin, hypocretins play a role in metabolic and endocrine regulation and have effects on food intake (95 ,133). The finding that hypocretin-containing neurons diffusely innervate numerous brain regions in addition to the hypothalamus (cerebral cortex,

limbic system, brainstem, and the spinal cord) (118), however, suggested that hypocretins might have other functions. Regulatory effects on blood pressure, body temperature, and the sleep-wake cycle were suggested (118). Hypocretin neurons in the LHA project to brain regions responsible for the regulation of vigilance (e.g., LC, VTA and the tuberomamillary histaminergic nucleus) and REM sleep (LC, dorsal raphe, and pontine cholinergic nuclei and PRF) (118). These anatomic and physiologic findings, taken together with the fact that deficits in hypocretin neurotransmission induce the narcolepsy phenotype, suggest that that hypocretins are the major neuromodulators for monoaminergic and cholinergic systems and hypocretins modulate sleep and sleep-related phenomena by interaction with these classical neurotransmitters.

PERSPECTIVES AND FUTURE DIRECTIONS

The discovery that a deficit in hypocretin neurotransmission, as revealed by the CSF hypocretin studies (106), frequently causes human narcolepsy opens the door to new diagnostic and therapeutic strategies. Measuring hypocretin levels in the CSF or other biological fluids may soon be used as a diagnostic test for narcolepsy. Early diagnosis may be critical for a disorder with peripubertal onset and a dramatic psychosocial impact. New therapeutic strategies should also be developed. All compounds currently used for the treatment of narcolepsy act symptomatically by enhancing monoaminergic transmission, likely downstream of the hypocretin neurotransmitter system (see section for monoaminergic/cholinergic imbalance and hypocretin deficiency). If reduced neurotransmission of hypocretin is a primary deficit in human narcolepsy and hypocretin receptors are still functional, supplementing transmission with hypocretins (and analogues) may have significant therapeutic effects in human narcolepsy (both on EDS and cataplexy).

Hypocretins are also likely to join acetylcholine and monoamines as critical sleep neurotransmitters. Basic research work relevant to the issue of sleep control will proceed at a rapid pace. Which hypocretin projections are most important for sleep control? How is the hypocretin system regulated across the sleep cycle? The preferential localization of Hcrtr2 on tuberomamillary and SN/VTA neurons suggests a preferential role for histaminergic and dopaminergic systems, but functional studies are lacking. Hypocretins are strongly excitatory in most cells studied, including monoaminergic cells (48 ,155). Removing an excitatory signal on these target cells could contribute to a monoaminergic hypoactivity. Similarly, the dense neuroanatomic distribution of Hcrtr2 in limbic structures such as the amygdala and the nucleus accumbens may explain cataplexy, a symptom triggered by emotions. Recent neuroimaging studies have shown that these regions are activated during natural REM sleep (74).

More work in the area of narcolepsy pathophysiology is also needed. Mutation screening studies of hypocretin genes indicate in narcolepsy-cataplexy very rare hypocretin system gene disease causing mutations, even in familial and non–HLA-DQB1*-0602 positive narcoleptic subjects (36). This indicates some degree of disease heterogeneity and the possibility that other neuronal systems closely related to the hypocretin/narcolepsy system still remain to be discovered. Does the observation that most cases of narcolepsy have undetectable CSF hypocretin levels indicate that hypocretin cells are destroyed in human narcoleptic brains? The finding that narcolepsy is tightly associated with HLA suggests a possible autoimmune process directed against these LHA cells. Autoantibodies against hypocretins or a substance closely related to the hypocretin system/LHA may secondarily cause narcolepsy. The relationship between feeding regulation, energy metabolism, and narcolepsy also should be explored. The role of the hypocretin system in the pathophysiology of narcolepsy without cataplexy remains to be investigated. Taking into account a recent, rapid gain in narcolepsy research, further progress in identifying the cause of human narcolepsy is likely to proceed at a rapid pace.

REFERENCES

1. Randomized trial of modafinil as a treatment for the excessive daytime somnolence of narcolepsy: US Modafinil in Narcolepsy Multicenter Study Group [in process citation]. *Neurology* 2000;54:1166–1175.

2. Akimoto H, Honda Y, Takahashi Y. Pharmacotherapy in narcolepsy. *Dis Nerv Sys* 1960;21:704–706.

3. Aldrich MS. The neurobiology of narcolepsy-cataplexy. *Prog Neurobiol* 1993;41:533–541.

4. Aldrich MS, Chervin RD, Malow BA. Value of the multiple sleep latency test (MSLT) for the diagnosis of narcolepsy. *Sleep* 1997;20:620–629.

5. Alles GA. The comparative physiological actions of d 1-beta-phenylisopropylamines: pressor effects and toxicity. *J Pharmacol Exp Ther* 1933;47:339–354.

6. *ICSD-international classification of sleep disorders: diagnostic and coding manual*. Rochester, MN: American Sleep Disorders Association, 1991.

7. Anic-Labat S, Guilleminault C, Kraemer H, et al. Validation of a cataplexy questionnaire in 983 sleep disorder patients. *Sleep* 1999;22:77–87.

8. Association AN. Stimulant medication survey. *The Eye Opener* 1992;5:1–3.

9. Association ASD. Practice parameters for the use of stimulants in the treatment of narcolepsy. *Sleep* 1994;17:348–351.

10. Aston-Jones G, Bloom FE. Activity of norepinephrine-containing locus coeruleus neurons in behaving rats anticipates fluctuations in the sleep-waking cycle. *J Neurosci* 1981;1:876–886.

11. Baghdoyan HA, Spotts JL, Snyder SG. Simultaneous pontine and basal forebrain microinjection of carbachol suppresses REM sleep. *J Neurosci* 1993;13:229–242.

12. Baker TL, Dement WC. Canine narcolepsy-cataplexy syndrome: evidence for an inherited monoaminergic-cholinergic imbalance. In: McGinty DJ, Drucker-Colin R, Morrison A, et al, eds. *Brain mechanisms of sleep*. New York: Raven Press, 1985:199–233.

13. Baker TL, Foutz AS, McNerney V, et al. Canine model of narcolepsy: genetic and developmental determinants. *Exp Neurol* 1982;75:729–742.

14. Baldessarini RJ. How do antidepressants work? In: Davis JM, Mass JW, eds. *The affective disorders*. Washington, DC: American Psychiatric Press, 1983:243–260.

15. Bastuji H, Jouvet M. Successful treatment of idiopathic hypersomnia and narcolepsy with modafinil. *Prog Neuropsychopharmacol Biol Psychiatry* 1988;12:695–700.

16. Beckett AH, Rowland M, Turner P. Influence of urinary pH on excretion of amphetamine. *Lancet* 1965;1:303.

17. Berkovitch M, Pope E, Phillips J, et al. Pemoline-associated fulminant liver failure: testing the evidence for causation. *Clin Pharmacol Ther* 1995;57:696–698.

18. Bernasconi R, Mathivet P, Bischoff S, et al. Gamma-hydroxybutyric acid: an endogenous neuromodulator with abuse potential? *Trends Pharmacol Sci* 1999;20:135–141.

19. Besset A, Tafti M, Villemine E, et al. Effect du modafinil (300 mg) sur le sommeil, la somnolence et la vigilance du narcoleptique. *Neurophysiol Clin* 1993;23:47–60.

20. Billiard M, Pasquie-Magnetto V, Heckman M, et al. Family studies in narcolepsy. *Sleep* 1994,17:S-54–S-9.

21. Boivin DB, Montplaisir J, Petit D, et al. Effect of modafinil on symptomatology of human narcolepsy. *Clin Neuropharmacol* 1993;16:46–53.

22. Broughton R, Ghanem Q, Hishikawa Y, et al. Life effects of narcolepsy. Relationships to geographic origin (North American, Asian or European) and to other patient and illness variables. *Can J Neurol Sci* 1983;10:100–104.

23. Broughton R, Guberman A, Roberts J. Comparison of psychosocial effects of epilepsy and of narcolepsy-cataplexy: a controlled study. *Epilepsia* 1984;25:423–433.

24. Broughton R, Mamelak M. Gamma-hydroxybutyrate in the treatment of compound narcolepsy: a preliminary report. In: Guilleminault C, Dement WC, Passouant P, eds. *Narcolepsy*. New York: Spectrum, 1976:59–67.

25. Broughton R, Mamelak M. The treatment of narcolepsy-cataplexy with nocturnal gamma-hydroxybutyrate. *Can J Neurol Sci* 1979;6:1.

26. Broughton RJ, Fleming JA, George CF, et al. Randomized, double-blind, placebo-controlled crossover trial of modafinil in the treatment of excessive daytime sleepiness in narcolepsy. *Neurology* 1997;49:444–451.

27. Bédard M, Jacques M, Godbout R, et al. Nocturnal g-hydroxybutyrate. *Clin Neuropharmacol* 1989;12:29–36.

28. Carlander B, Eliaou JF, Billiard M. Autoimmune hypothesis in narcolepsy. *Neurophysiol Clin* 1993;23:15–22.

29. Carskadon MA, Dement WC, Mitler MM, et al. Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. *Sleep* 1986;9:519–524.

30. Chemelli RM, Willie JT, Sinton CM, et al. Narcolepsy in Orexin knockout mice: molecular genetics of sleep regulation. *Cell* 1999;98:437–451.

31. Chin M, Kreutzer RA, Dyer JL. Acute poisoning from g-hydroxybutyrate in California. *West J Med* 1992;156:380–384.

32. De Lecea L, Kilduff TS, Peyron C, et al. The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity. *Proc Natl Acad Sci USA* 1998;95:322–327.

33. Delashaw JB, Foutz AS, Guilleminault C, et al. Cholinergic mechanisms and cataplexy in dogs. *Exp Neurol* 1979;66:745–757.

34. Edgar DM, Seidel WF. Modafinil induces wakefulness without intensifying motor activity or subsequent rebound hypersomnolence in the rat. *J Pharmacol Exp Ther* 1997;283:757–769.

35. Edgar DM, Seidel WF, Contreras P, et al. Modafinil promotes EEG wake without intensifying motor activity in the rat. *Can J Physiol Pharmacol* 1994;72:S1–362.

36. Faraco J, Rogers W, Overseem S, et al. Mutation screening of hypocretin system genes in human narcoleptics. *Sleep* 2000;23:A90.

37. Ferraro L, Tananelli S, O'Connor WT, et al. The vigilance promoting drug modafinil increases dopamine release in the rat nucleus accumbens via the involvement of a local GABAergic mechanism. *Eur J Pharmacol* 1996;306:33–39.

38. Fujimori M, Himwich HE. Electroencephalographic analyses of amphetamine and its methoxy derivatives with reference to their sites of EEG alerting in the rabbit brain. *Int J Neuropharmacol* 1969;8:601–613.

39. Fuxe K, Janson AM, Rosen L, et al. Evidence for a protective action of the vigilance promoting drug modafinil on the MPTP-induced degeneration of the nigrostriatal dopamine neurons in the black mouse: an immunocytochemical and biochemical analysis. *Brain Res* 1992;88:117–130.

40. Godbout R, Poirier G, Montplaisir J. New treatment for narcolepsy (viloxazine). In: Burton SA, Dement WC, Ristanovic R, eds. *Narcolepsy 3rd International Symposium*. San Diego: ICI Pharma, 1988:79–81.

41. Goenechea S, Wagner GM. Quantitative determination of pemoline in serum and urine after ingestion of therapeutic doses. *Arzneimittel-Forschung* 1977;27:1604–1605.

42. Gold LH, Balster RH. Evaluation of the cocaine-like discriminative stimulus effects and reinforcing effects of modafinil. *Psychopharmacology* 1996;126:286–292.

43. Guilleminault C, Mancuso J, Salva MAQ, et al. Viloxazine hydrochloride in narcolepsy: a preliminary report. *Sleep* 1986;9:275–279.

44. Guilleminault C, Mignot E, Grumet FC. Familial patterns of narcolepsy. *Lancet* 1989;2:1376–1379.

45. Guilleminault C, Raynal D, Takahashi S, et al. Evaluation of short-term and long-term treatment of the narcolepsy syndrome with clomipramine hydrochloride. *Acta Neurol Scand* 1976;54:71–87.

46. Guilleminault C, Wilson RA, Dement WC. A study on cataplexy. *Arch Neurol* 1974;31:255–261.

47. Hadler AJ. Mazindol, a new non-amphetamine anorexigenic agent. *J Clin Pharmacol* 1972;12:453–458.

48. Hagan JJ, Leslie RA, Patel S, et al. Orexin A activates locus coeruleus cell firing and increases arousal in the rat [in process citation]. *Proc Natl Acad Sci USA* 1999,96:10911–10916.

49. Henry GK, Hart RP, Kwentus JA, et al. Effects of protriptyline on vigilance and information processing in narcolepsy. *Psychopharmacology* 1988;95:109–112.

50. Hermant JF, Rambert FA, Deuteil J. Lack of cardiovascular effects after administration of modafinil in conscious monkeys. In: French Association des Pharmacologistes Tours. *Fundam Clin Pharmacol* 1991;5:825.

51. Hishikawa Y, Ida H, Nakai K, et al. Treatment of narcolepsy with imipramine (Tofranil) and desmethylimipramine (Pertofrane). *J Neurol Sci* 1965;3:453–461.

52. Hishikawa Y, Koida H, Yoshino K, et al. Characteristics of REM sleep accompanied by sleep paralysis and hypnagogic hallucinations in narcoleptic patients. *Waking Sleeping* 1978;2:113–123.

53. Honda K, Riehl J, Mignot E, et al. Dopamine D3 agonists into the substantia nigra aggravates cataplexy but does not modify sleep [in process citation]. *Neuroreport* 1999;10:3111–3118.

54. Honda Y. Census of narcolepsy, cataplexy and sleep life among teen-agers in Fujisawa city. *Sleep Res* 1979;8:191.

55. Honda Y. Clinical features of narcolepsy. In: Honda Y, Juji T, ed. *HLA in narcolepsy*. Berlin: Springer-Verlag, 1988:24–57.

56. Hong S, Hayduk R, Lim J, et al. Clinical and polysomnographic features in DQB1*0602 positive and negative narcolepsy patients: results from the modafinil clinical trial. *Sleep Med* 2000;1:33–39.

57. Hublin C, Kaprio J, Partinene M, et al. The prevalence of narcolepsy: an epidemiologic study of the Finnish twin cohort. *Ann Neurol* 1994;35:709–716.

58. Iijima S, Sugita Y, Teshima Y, et al. Therapeutic effects of mazindol on narcolepsy. *Sleep* 1986;9:265–268.

59. Juji T, Satake M, Honda Y, et al. HLA antigens in Japanese patients with narcolepsy. All the patients were DR2 positive. *Tissue Ant* 1984;24:316–319.

60. Knecht CD, Oliver JE, Redding R, et al. Narcolepsy in a dog and a cat. *J Am Vet Med Assoc* 1973;162:1052–1053.

61. Koran L, Raghavan S. Fluoxetine for isolated sleep paralysis. *Psychomatics* 1993;34:184–187.

62. Kuczenski R, Segal DS, Cho A, et al. Hippocampus norepinephrine, caudate dopamine and serotonin and behavioral responses to the stereoisomers of amphetamine and methamphetamine. *J Neurosci* 1995;15:1308–1317.

63. Langdon N, Shindler J, Parkes JD, et al. Fluoxetine in the treatment of cataplexy. *Sleep* 1986;9:371–373.

64. Lapierre O, Montplaisir J, Lamarre M, et al. The effect of gamma-hydroxybutyrate on nocturnal and diurnal sleep of normal subjects: further consideration on REM sleep-triggering mechanisms. *Sleep* 1990;13:24–30.

65. Lavie P, Peled R. Narcolepsy is a rare disease in Israel. *Sleep* 1987;10:608–609.

66. Li J, Stokes SA, Woeckener A. A tale of novel intoxication: seven cases of gamma-hydroxybutyric acid overdose [see comments]. *Ann Emerg Med* 1998;31:723–728.

67. Lin JS, Roussel B, Akaoka H, et al. Role of catecholamines in modafinil and amphetamine induced wakefulness, a comparative pharmacologic study. *Brain Res* 1992;591:319–326.

68. Lin L, Faraco J, Li R, et al. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (Orexin) receptor 2 gene. *Cell* 1999;98:365–376.

69. Lin L, Jin L, Kimura A, et al. DQ microsatellite association studies in three ethnic groups. *Tissue Ant* 1997;50:507–520.

70. Mack RB. Love potion number 8 1/2. Gamma-hydroxybutyrate poisoning. *NC Med J* 1993;54:232–233.

71. Maitre M. The gamma-hydroxybutyrate signalling system in brain: organization and functional implications. *Prog Neurobiol* 1997;51:337–361.

72. Mamelak M, Scharf MB, Woods M. Treatment of narcolepsy with γ-hydroxybutyrate. A review of clinical and sleep laboratory findings. *Sleep* 1986;9:285–289.

73. Mantle TJ, Tipton KF, Garrett NJ. Inhibition of monoamine oxidase by amphetamine and related compounds. *Biochem Pharmacol* 1976;25:2073–2077.

74. Maquet P, Peters J, Aerts J, et al. Functional neuroanatomy of human rapid-eye-movement sleep and dreaming. *Nature* 1996;383:163–166.

75. Mathivet P, Bernasconi R, De Barry J, et al. Binding characteristics of gamma-hydroxybutyric acid as a weak but selective GABAB receptor agonist. *Eur J Pharmacol* 1997;321:67–75.

76. Matsuki K, Grumet FC, Lin X, et al. DQ rather than DR gene marks susceptibility to narcolepsy. *Lancet* 1992;339:1052.

77. Matsuki K, Juji T, Honda Y. Immunological features of narcolepsy in Japan. In: Honda Y, Juji T, eds. *HLA in narcolepsy*. New York: Springer-Verlag, 1988:150–157.

78. Mignot E. Genetic and familial aspects of narcolepsy. *Neurology* 1998;50:S16–S22.

79. Mignot E, Guilleminault C, Bowersox S, et al. Central alpha-1 adrenoceptor subtypes in narcolepsy-cataplexy: a disorder of REM sleep. *Brain Res* 1989;490:186–191.

80. Mignot E, Hayduk R, Black J, et al. HLA Class II studies in 509 narcoleptic patients. *Sleep Res* 1997;26:433.

81. Mignot E, Hayduk R, Grumet FC, et al. HLA DQB1*0602 is associated with cataplexy in 509 narcoleptic patients. *Sleep* 1997;20:1012–1020.

82. Mignot E, Lin X, Arrigoni J, et al. DQB1*0602 and DQA1*0102 (DQ1) are better markers than DR2 for narcolepsy in Caucasian and black Americans. *Sleep* 1994;17:S6–S7.

83. Mignot E, Meehan J, Grumet FC, et al. HLA class II and narcolepsy in thirty-three multiplex families. *Sleep Res* 1996;25:303.

84. Mignot E, Nishino S, Guilleminault C, et al. Modafinil binds to the dopamine uptake carrier site with low affinity. *Sleep* 1994;17:436–437.

85. Mignot E, Renaud A, Nishino S, et al. Canine cataplexy is preferentially controlled by adrenergic mechanisms: evidence using monoamine selective uptake inhibitors and release enhancers. *Psychopharmacology* 1993;113:76–82.

86. Mignot E, Tafti M, Dement WC, et al. Narcolepsy and immunity. *Adv Neuroimmunol* 1995;5:23–37.

87. Mignot E, Wang C, Rattazzi C, et al. Genetic linkage of autosomal recessive canine narcolepsy with an immunoglobulin heavy-chain switch-like segment. *Proc Natl Acad Sci USA* 1991;88:3475–3478.

88. Mignot E, Lin L, Rogers W, et al. Complex HLA-DR and -DQ interactions confer risk of narcolepsy-cataplexy in three ethnic groups. *Am J Hum Genet* 2001;68:686–699.

89. Miller JD, Farber J, Gatz P, et al. Activity of mesencephalic dopamine and non-dopamine neurons across stages of sleep and waking in the rat. *Brain Res* 1983;273:133–141.

90. Mitler MM, Aldrich MS, Koob GF, et al. Narcolepsy and its treatment with stimulants. *Sleep* 1994;17:352–371.

91. Mitler MM, Boysen BG, Campbell L, et al. Narcolepsy-cataplexy in a female dog. *Exp Neurol* 1974;45:332–340.

92. Mitler MM, Hajdukovic R. Relative efficacy of drugs for the treatment of sleepiness in narcolepsy. *Sleep* 1991;14:218–220.

93. Mitler MM, Shafor R, Hajdukovich R, et al. Treatment of narcolepsy: objective studies on methylphenidate, pemoline, and protriptyline. *Sleep* 1986;9:260–264.

94. Montplaisir J, Godbout R. Serotonergic reuptake mechanisms in the control of cataplexy. *Sleep* 1986;9:280–284.

95. Moriguchi T, Sakurai T, Nambu T, et al. Neurons containing Orexin in the lateral hypothalamic area of the adult rat brain are activated by insulin-induced acute hypoglycemia. *Neurosci Lett* 1999;264:101–104.

96. Neely S, Rosenberg R, Spire J, et al. HLA antigens in narcolepsy. *Neurology* 1987;37:1858–1860.

97. Nishino S, Arrigoni J, Shelton J, et al. Desmethyl metabolites of serotonergic uptake inhibitors are more potent for suppressing canine cataplexy than their parent compounds. *Sleep* 1993;16:706–712.

98. Nishino S, Arrigoni J, Valtier D, et al. Dopamine D2 mechanisms in canine narcolepsy. *J Neurosci* 1991;11:2666–26671.

99. Nishino S, Fruhstorfer B, Arrigoni J, et al. Further characterization of the alpha-1 receptor subtype involved in the control of cataplexy in canine narcolepsy. *J Pharmacol Exp Ther* 1993;264:1079–1084.

100. Nishino S, Haak L, Shepherd H, et al. Effects of central alpha-2 adrenergic compounds on canine narcolepsy, a disorder of rapid eye movement sleep. *J Pharmacol Exp Ther* 1990;253:1145–1152.

101. Nishino S, Honda K, Riehl J, et al. Extracellular single unit recordings of dopaminergic neurons in the ventral tegmental area in narcoleptic Dobermans. *Sleep (APSS abstract book)*. 2000;23:A1.

102. Nishino S, Mao J, Sampathkumaran R, et al. Adrenergic, but not dopaminergic, uptake inhibition reduces REM sleep and cataplexy concomitantly. *Sleep Res* 1997;26:445.

103. Nishino S, Mao J, Sampathkumaran R, et al. Increased dopaminergic transmission mediates the wake-promoting effects of CNS stimulants. *Sleep Res Online* 1998;1:49–61. http://www.sro.org/1998/Nishino/49/

104. Nishino S, Mignot E. Pharmacologic aspects of human and canine narcolepsy. *Prog Neurobiol* 1997;52:27–78.

105. Nishino S, Riehl J, Mignot E. Effects of dopaminergic and noradrenergic uptake inhibitors on dopaminergic D3 agonist-included cataplexy. *Sleep* 1998;21:181 (APSS abstract).

106. Nishino S, Ripley B, Overeem S, et al. Hypocretin (Orexin) deficiency in human narcolepsy. *Lancet* 2000;355:39–40.

107. Nishino S, Tafti M, Reid MS, et al. Muscle atonia is triggered by cholinergic stimulation of the basal forebrain: implication for the pathophysiology of canine narcolepsy. *J Neurosci* 1995;15:4806–4814.

108. Nissbrandt N, Engberg G, Pileblad E. The effects of GBR 12909, a dopamine re-uptake inhibitor, on monoaminergic neurotransmission in rat striatum, limbic forebrain, cortical hemispheres and substantia nigra. *Naunyn-Schmiedeberg's Arch Pharmacol* 1991;344:16–28.

109. Ohayon M, Zulley J, Guilleminault C, et al. Prevalence and pathologic associations of sleep paralysis in the general population. *Neurology* 1999;52:1194–1200.

110. Ohayon MM, Priest RG, Caulet M, et al. Hypnagogic and hypnopompic hallucinations: pathologic phenomena? *Br J Psychiatry* 1996;169:459–467.

111. Palatini P, Tedeschi L, Frison G, et al. Dose-dependent absorption and elimination of gamma-hydroxybutyric acid in healthy volunteers. *Eur J Clin Pharmacol* 1993;45:353–356.

112. Parkes D. Amphetamines and alertness. In: Guilleminault C, Dement WC, Passouant P, eds. *Narcolepsy*. New York: Spectrum, 1976:643–658.

113. Parkes D. Amphetamines and other drugs. In: *Sleep and its disorders*. London: WB Saunders, 1985:459–482.

114. Parkes JD, Baraitser M, Marsden CD, et al. Natural history, symptoms and treatment of the narcoleptic syndrome. *Acta Neurol Scand* 1975;52:337–353.

115. Parkes JD, Fenton GW. Levo(-) amphetamine and dextro (+) amphetamine in the treatment of narcolepsy. *J Neurol Neurosurg Psychiatry* 1973;36:1076–1081.

116. Pawluk LK, Hurwitz TD, Schulter JL, et al. Psychiatric morbidity in narcoleptics on chronic high dose methylphenidate therapy. *J Nerv Ment Dis* 1995;183:45–48.

117. Pelin Z, Guilleminault C, Rish N, et al. HLA-DQB1*0602 homozygosity increases relative risk for narcolepsy but not disease severity in two ethnic groups. *Tissue Ant* 1998;51:96–100.

118. Peyron C, Tighe DK, van den Pol AN, et al. Neurons containing hypocretin (Orexin) project to multiple neuronal systems. *J Neurosci* 1998;18:9996–10015.

119. Poceta JS, Hajdukovic R, Mitler MM. Improvement in cataplexy with yohimbine and paroxetine: case report. *Sleep Res* 1994;23:304.

120. Prinzmetal M, Bloomberg W. The use of Benzedrine for the treatment of narcolepsy. *J Am Med Assoc* 1935;105:2051–2054.

121. Rall TR. Central nervous system stimulants. In: Gilman AG, Goodman LS, Rall TW, et al, eds. *The pharmacologic basis of therapeutics*, 7th ed. New York: Pergamon Press, 1985:345–382.

122. Raynal D. Polygraphic aspects of narcolepsy. In: Guilemminault C, Passouant P, ed. *Narcolepsy*. New York: Spectrum, 1976:669–684.

123. Rechschaffen A, Dement W. Studies on the relation of narcolepsy, cataplexy and sleep with low voltage random EEG activity. In: Kety S, Evarts E, Williams H, eds. *Sleep and altered states of consciousness*. Baltimore: Williams & Wilkins, 1967:488–505.

124. Reid MS, Siegel JM, Dement WC, et al. Cholinergic mechanisms in canine narcolepsy: II. Acetylcholine release in the pontine reticular formation is enhanced during cataplexy. *Neuroscience* 1994;59:523–530.

125. Reid MS, Tafti M, Geary J, et al. Cholinergic mechanisms in canine narcolepsy: I. Modulation of cataplexy via local drug administration into pontine reticular formation. *Neuroscience* 1994;59:511–522.

126. Reid MS, Tafti M, Nishino S, et al. Local administration of dopaminergic drugs into the ventral tegmental area modulate cataplexy in the narcoleptic canine. *Brain Res* 1996;733:83–100.

127. Ricaurte GA, McCann UD. Neurotoxic amphetamine analogues: effects in monkeys and implications for humans. *Ann NY Acad Sci* 1992;648:371–382.

128. Roehrs T, Zorick F, Wittig R, et al. Alerting effects of naps in patients with narcolepsy. *Sleep* 1986;9:194–199.

129. Rogers AE. Problems and coping strategies identified by narcoleptic patients. *J Neurosurg Nursing* 1984;16:326–334.

130. Rogers AE, Aldrich MS, Berrios AM, et al. Compliance with stimulant medications in patients with narcolepsy. *Sleep* 1997;20:28–33.

131. Rolls ET, Sanghera MK, Roper-Hall A. The latency of activation of neurones in the lateral hypothalamus and substantia innominata during feeding in the monkey. *Brain Res* 1979;164:121–135.

132. Sakurai T, Amemiya A, Ishii M, et al. Orexins and Orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior [comment]. *Cell* 1998,92:696.

133. Sakurai T, Amemiya A, Ishil M, et al. Orexins and Orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* 1998,92:573–585.

134. Schachter M, Parkes JD. Fluvoxamine and clomipramine in the treatment of cataplexy. *J Neurol Neurosurg Psychiatry* 1980;43:171–174.

135. Schenck CH, Mahowald MW. Motor dyscontrol in narcolepsy: rapid-eye-movement (REM) sleep without atonia and REM sleep behavior disorder. *Ann Neurol* 1992;32:3–10.

136. Schrader H, Bendixen-Markset AC, Treidene HE. The treatment of accessory symptoms in narcolepsy: a double-blind cross-over study of a selective serotonin re-uptake inhibitor (femoxetine) versus placebo. *Acta Neurol Scand* 1986;74:297–303.

137. Scrima L, Hartman PG, Johnson FH, et al. Efficacy of gamma-hydroxybutyrate versus placebo in treating narcolepsy-cataplexy: double-blind subjective measures. *Biol Psychiatry* 1989;26:331–343.

138. Scrima L, Johnson FH, Thomas EG, et al. The effects of gamma-hydroxybutyrate (GHB) on multiple sleep latency test (MSLT) in narcolepsy patients, va long term study. *Sleep Res* 1990;19:288.

139. Seijun T. Methamphetamine psychosis. In: Ellinwood EH, Cohen åAS, ed. *Current concept on amphetamine abuse*. Rockville, MD: National Institute of Mental Health, 1970:159–161.

140. Shapiro WR. Treatment of cataplexy with clomipramine. *Arch Neurol* 1975;32:653–656.

141. Shepherd DI, Summers A. Prevalence of multiple sclerosis in Rochdale. *J Neurol Neurosurg Psychiatry* 1996;61:415–417.

142. Shevell M, Schreiber R. Pemoline-associated hepatic failure: a critical analysis of the literature. *Pediatr Neurol* 1997;16:14–16.

143. Sicard BA, Perault MC, Enslen M, et al. The effects of 600 mg of slow release caffeine on mood and alertness. *Aviat Space Environ Med* 1996;67:859–862.

144. Siegel JM. Brainstem mechanisms generating REM sleep. In: Kryger MH, Roth T, Dement WC, eds. *Principles and practice of sleep medicine*. Philadelphia: WB Saunders, 1994:125–144.

145. Simpson LL. Blood pressure and heart rate responses produced by d-amphetamine: correlation with blood levels of drug. *J Pharmcol Exp Ther* 1978;205:366–373.

146. Smith M, Parkes J, Dahlitz M. Venlafaxine in the treatment of the narcoleptic syndrome. *J Sleep Res* 1996;5:217.

147. Snyder SH, Taylor KM, Coyle JT, et al. The role of brain dopamine in behavioral regulation and the actions of psychotropic drugs. *Am J Psychiatry* 1970;127:199–207.

148. Steriade M, McCarley RW. *Brainstem control of wakefulness and sleep*. New York: Plenum, 1990.

149. Takahashi S. The action of tricyclics (alone or in combination with methylphenidate) upon several symptoms of narcolepsy. In: Guilleminault C, Dement WC, Passouant P, eds. *Narcolepsy*. New York: Spectrum, 1976:625–638.

150. Tashiro T, Kambayashi T, Hishikawa Y. An epidemiologic study of narcolepsy in Japanese. In: The 4th International Symposium on Narcolepsy. Tokyo, Japan. 1994, p. 13.

151. Thorpy M, Snyder M, Ledereich P, Starz K. Short term triazolam improves nocturnal sleep of narcoleptics. *Sleep* 1992;15:212–216.

152. Thorpy MJ, Goswami M. Treatment of narcolepsy. In: Thorpy MJ, ed. *Handbook of sleep disorders*. New York: Marcel Dekker, 1990:235–258.

153. Trulson ME, Jacobs BL. Raphe unit activity in freely moving cats: correlation with level of behavioral arousal. *Brain Res* 1979;163:135–150.

154. Trulson ME, Preussler DW, Howell GA. Activity of substantia nigra units across the sleep-waking cycle in freely moving cats. *Neurosci Lett* 1981;26:183–188.

155. Uramura K, Yata T, Muraya S, et al. Neurostimulation and signal transduction by Orexin on A10 dopaminergic cells in rats. In: Orexin Workshop II. Vol Showa University School of Medicine, Tokyo, Japan. 2000, p. 4.

156. Vayer P, Mandel P, Maitre M. Gamma-hydroxybutyrate, a possible neurotransmitter. *Life Sci* 1987;41:1547–1557.

157. Vickers M. Gamma-hydroxybutyric acid. *Int Anesthesiol Clin* 1969;7:75–89.

158. Williams RT, Caldwell RJ, Dreng LG. Comparative metabolism of some amphetamine in various species. In: Schneider SH, Esdin E, eds. *Frontiers of catecholamine research*. Oxford, England: Pergamon, 1973:927–932.

159. Wong YN, King SP, Laughton WB, et al. Single-dose pharmacokinetics of modafinil and methylphenidate given alone or in combination in healthy male volunteers. *J Clin Parmacol* 1998;38:276–282.

160. Yoss RE, Daly DD. Treatment of narcolepsy with Ritalin. *Neurology* 1959;9:171–173.

132

Basic Mechanisms of Sedative/Hypnotics

Wallace B. Mendelson

Wallace B. Mendelson: Department of Psychiatry, The University of Chicago, Chicago, Illinois.

THEORIES OF HYPNOTIC ACTION

Part of "132 - Basic Mechanisms of Sedative/Hypnotics "

How compounds of a wide variety of chemical classes can have relatively similar effects in inducing sleep is an intriguing pharmacologic question. Probably the dominant theory—that these compounds exert their actions by altering the physical properties of lipids in neuronal membranes—comes from the alcohol and anesthetic literature (1). Although there are many viewpoints about which specific aspect of lipid properties is affected (i.e., fluidity, thickness, or surface tension), it is largely a physicochemical approach. Ultimately, it was found inadequate for a number of reasons; perhaps the most important is that there are very little or no detectable changes in lipid bilayers at the concentrations at which these compounds induce sleep or anesthesia (1). Ultimately, interest has turned to more specific mechanisms; by far the most satisfying one (and the focus of this chapter) is the notion that altering neurotransmitter-gated receptor channels induces sleep and anesthesia (2).

Another approach to understanding sedative/hypnotics has been to hypothesize that sleep results from drug-induced reduction in energy metabolism; barbiturates, for instance, decrease cerebral glucose metabolic rate in human positron emission tomography (PET) scan studies (3). On the other hand, the results of animal studies have been more variable, such that barbiturates may (4) and benzodiazepines may not (5) decrease cerebral metabolic rate of oxygen (CMRO2). A more cogent argument against the notion that hypnotics induce sleep by lowering metabolic rate is that it stems from a view of sleep as being a very passive process, which seems to contradict the more contemporary understanding of sleep as a multifaceted, actively regulated process (6). Indeed, at doses that induce sleep (and prior to achieving anesthetic doses), patients receiving most hypnotic medications demonstrate the alternating ultradian rhythm of NREM and REM sleep, which indicates an active regulatory mechanism. It seems more parsimonious, then, to hypothesize that sedative/hypnotics act at specific sites involved in sleep regulation, rather than producing a nonspecific “slowing” of the nervous system.

THE MOLECULAR LEVEL: THE CENTRAL $GABA_A$–BENZODIAZEPINE RECEPTOR COMPLEX

Part of "132 - Basic Mechanisms of Sedative/Hypnotics "

Background and Description

A major insight into the action of hypnotics began in the 1970s with the discovery of high affinity, stereospecific receptors for benzodiazepines in the central nervous system (7 ,8). $GABA_A$-benzodiazepine receptors are the most abundant inhibitory receptor system in the central nervous system (CNS) (9). They can be viewed as representatives of a large family of hetero-oligomeric ligand-gated ion channels (e.g., nicotinic acetylcholine, glycine, and serotonin-3 receptors) (10). Functionally, they contain three distinct but interacting moieties: recognition sites for GABA and benzodiazepines and a chloride ionophore (11). Binding of a benzodiazepine agonist to its recognition site results in increased chloride ion flux, which in turn hyperpolarizes the postsynaptic membrane at a level below that at which spike generation is possible.

Molecular cloning data indicate that the $GABA_A$-benzodiazepine receptor complex is comprised of at least five subunits; these in turn may have various isoforms (9). Each subunit is comprised of four membrane-spanning regions. The intracellular loops of some subunits contain phosphorylation sites, which have been hypothesized to be a locus of receptor modulation. It is possible, for instance, that the receptor system might respond to GABA differently in the phosphorylated versus dephosphorylated state. α- and γ-Subunits need to be present in order to be responsive to benzodiazepines, and a complete system of α, ß, and γ are

needed for a fully responsive receptor (12). Recent data using murine gene targeting techniques indicate that the α-1 isoform is not required for the anxiolytic actions of benzodiazepines, but may mediate sedation and ataxic effects; in principle, this might make it possible to develop anxiolytics with a more benign side effect profile than those currently available (13). The γ-1 subunit, which in some senses may be viewed as a marker of the subgroup of GABA receptors representing the $GABA_A$–benzodiazepine receptor complex, is found in 60% to 90% of GABA receptors (14). $GABA_A$-benzodiazepine receptors have also been divided into Type I and II (sometimes referred to as o-1 and -2), differing on the particular α subunit isoform; whether this distinction has pharmacologic significance in terms of effects of drugs that selectively bind to the Type I receptor is still under investigation. There are peripheral and "Valium-insensitive" receptors that appear not to be involved in sleep-related processes in addition to the central receptors (the focus here); hence, they are beyond the scope of this chapter.

Presynaptic Effects: Alterations in Calcium Channel Function

Although most work on the effector mechanism of benzodiazepine receptor agonists focuses on the postsynaptic mechanism of alterations in chloride ionophore function, it should be noted that there are also presynaptic actions involving calcium ion flux that have been less fully explored but that may have relevance to sedative/hypnotic properties. In rats, for instance, the dihydropyridine calcium channel blocker nifedipine given intraventricularly blocks the sleep-inducing property of systemically administered flurazepam, whereas BAY-K-8644, which facilitates calcium ion flux in dihydropyridine-sensitive channels, greatly augments the hypnotic effects of flurazepam (15).

The Role of GABA Agonists

One intriguing set of issues that has not been clearly resolved is that more specific $GABA_A$ agonists are relatively weak in their hypnotic effects, do not necessarily augment the effects of benzodiazepines, and may indeed have very different actions. The $GABA_A$ agonist muscimol when given IP to rats, for instance, has mild effects on reducing sleep latency and does not alter total sleep time (16). Analogously, the $GABA_A$ antagonist bicuculline slightly increases sleep latency without altering total sleep in the rat; neither interact with intraperitoneally administered triazolam (16). Muscimol has also been found to have effects on sleep different from midazolam in rats; the former increased both NREM and REM sleep, whereas midazolam increased NREM, decreased REM, and produced opposite effects on low-frequency EEG activity (17). Again, in contrast to triazolam, microinjection of muscimol into the medial preoptic area has been reported to have no effect on sleep in rats (18). The reasons for these differences are not clear; among the possibilities are that in the intraperitoneal administration studies muscimol produced nonspecific actions at GABA receptors throughout the brain, or that muscimol enters the CNS poorly (19); the lack of effects on sleep after microinjection directly into the MPA would seem to rule these possibilities out. It also appears that muscimol may alter chloride channel function in a manner different from GABA, activating channels for greater durations (20). It also may be that muscimol indiscriminately activates GABA receptors, whereas in contrast benzodiazepines may only alter function of receptors with specific compositions (21). Although this is still being assessed, the differences in effects of muscimol and benzodiazepines provide a cautionary note that it may be important not to equate the hypnotic actions of benzodiazepines with simple GABAergic effects.

The $GABA_A$–Benzodiazepine Receptor Complex as a Common Site for Diverse Pharmacologic Classes of Hypnotics

The affinities of various benzodiazepines were found to correlate well with their anxiolytic, anticonvulsant, and muscle relaxant properties in the early reports on central $GABA_A$–benzodiazepine receptors (7). Later evidence indicated that the receptor complex mediates the hypnotic actions of benzodiazepines as well. This role was clarified by studies of the inverse agonist 3-hydroxymethyl-ß-carboline (3-HMC), which induced awakening and decreased sleep in rats, an effect prevented by the benzodiazepine receptor blocker CGS 8216 (15). Similarly, the hypnotic properties of the clinically used benzodiazepine flurazepam were blocked by a low dose of 3-HMC that had minimal effects when given alone (22). The stereospecificity of the site was demonstrated by studies of the benzodiazepine B-10 enantiomers, in which the (+) compound induced sleep, whereas the (-) compound increased wakefulness (15). It became clear, then, that it was indeed interaction with this receptor complex that mediates the effects of benzodiazepines on sleep and waking.

The $GABA_A$-benzodiazepine receptor complex contains modulatory sites not only for benzodiazepines, but also for a number of other types of sedating compounds, including barbiturates, neurosteroids, and ethanol (23 ,24), etomidate (25), and the anesthetic propofol (26). The newer nonbenzodiazepine hypnotics zolpidem, zopiclone, and zaleplon bind to the type I benzodiazepine recognition site as well. The end result is thought to be an enhancement of chloride ion flux, as described. These compounds may achieve this effect by different means, nonetheless; hence, benzodiazepines may increase the frequency of channel opening (27), whereas barbiturates, for instance, may increase the duration of opening (28). Ethanol facilitates GABA-stimulated chloride flux at low concentrations, and directly enhances flux

at higher concentrations (29). Similarly, some cytokines such as interleukin-1 enhance GABA-dependent chloride influx into synaptoneurosomes (30).

We are beginning to gain insight into the problem we posed originally: how administration of sedative/hypnotic compounds from such diverse pharmacologic classes can result in sleep induction. It appears that most or all of them produce pharmacologic effects by altering the function of various moieties of the $GABA_A$-benzodiazepine receptor complex.

NEUROANATOMIC STUDIES

Part of "132 - Basic Mechanisms of Sedative/Hypnotics "

Microinjection of Hypnotics into the CNS

In some ways, it has been more challenging to determine the neuroanatomic sites of action of hypnotics than to assess their actions at a molecular level. The most parsimonious neuroanatomic approach would seem to be that hypnotics act at sites thought to be involved in physiologic sleep regulation, on the basis of lesion or stimulation studies. Using this reasoning, the author's laboratory set out some years ago to administer benzodiazepines into such sites, many of which were chosen on the basis of the classical studies of Hernandez-Peon (31). The compound chosen for these studies was the benzodiazepine triazolam, which was the most widely administered clinical hypnotic. Perhaps the most surprising finding was the absence of effects on sleep following triazolam microinjection into many loci thought to be involved in sleep regulation, including the locus ceruleus, horizontal limb of the diagonal band of Broca, lateral preoptic area, basomedial nucleus of the amygdala (32), and ventrolateral preoptic area (33). In contrast, microinjection at several sites—notably the dorsal raphe nuclei and medial preoptic area— profoundly altered sleep and waking. We examine these in turn.

The Dorsal Raphe Nuclei

The dorsal raphe nuclei, tubular structures in the upper pons and lower midbrain that contain most of the forebrain serotonin (34) send fibers via the medial forebrain bundle to a various areas including the hypothalamus, basal forebrain, septal area, striatum, hippocampus, and cerebral cortex (35 ,36). The heaviest innervation in the hypothalamus is in the medial preoptic area, suprachiasmatic nucleus, and dorsal and ventral premammillary nuclei (37), as well as in histamine-containing neurons of the tuberomamillary nucleus (38). Stimulation of the dorsal raphe has been demonstrated by the 2-deoxyglucose method to increase glucose utilization in the somatosensory cortex, thalamus, hypothalamus, and extrapyramidal system (39). Ascending fibers innervate the suprachiasmatic nucleus (34), raising the possibility that they influence circadian rhythm function. The dorsal raphe nuclei in turn receive descending fibers originating from the preoptic area (40), suggesting a kind of reciprocal innervation discussed in the following. The ventrolateral preoptic area (VLPO), which has been postulated to play a role in sleep initiation (41), sends GABAergic and galaninergic descending fibers to the dorsal and median raphe nuclei as well as the locus ceruleus, which might result in inhibitory action on ascending monoaminergic systems (42). In addition, dorsal raphe cells receive inhibitory serotonergic input from other local dorsal raphe neurons (43 ,44 ,45 and 46) as well as from the median raphe (43). Interestingly, microinjections of triazolam into the dorsal raphe *deceased* sleep in the rat—sleep latency was significantly increased and total sleep time was significantly reduced (47).

The Medial Preoptic Area

There has been considerable work on the role of the hypothalamus and adjacent structures in the regulation of sleep and waking, ever since pathological studies following the epidemic of encephalitis lethargica in the 1920s (48). Later lesion studies by Hess (49) and Nauta (50) suggested that the anterior hypothalamus is involved in sleep maintenance, whereas a more caudal area might promote wakefulness. Stimulation of a basal forebrain area, including the medial preoptic area (MPA), enhances (51) and lesions decrease (52) sleep in cats. Lesions of the medial preoptic area acutely decrease sleep in the rat, albeit subject to effects of ambient temperature (53). The MPA receives visual and auditory inputs, and contains neurons sensitive to glucose and steroids (54), osmotic and cardiovascular measures (55), and temperature (56). It is also a thermoregulatory site, and it may coordinate many processes in homeostatic and reproductive functions (36).

Cell bodies and fibers in the MPA cross react in immunohistochemical studies with a wide variety of neurotransmitters such as substance P and neuropeptide Y (57). Push-pull cannula studies of the MPA have documented release of catecholamines, GABA, and glutamate (58). GABA is uniformly found throughout the hypothalamus, and its synthetic enzyme GAD has very high concentrations in the preoptic area in particular (59). The MPA is also rich in some forms of the $GABA_A$-benzodiazepine receptor complex, suggesting that benzodiazepine hypnotic compounds might bind there (14). Neurons that become more active during NREM and REM sleep are inside the MPA (60). Its projections travel throughout the forebrain and brainstem (61) to the median and dorsal raphe nuclei, and possibly to the locus ceruleus (40). Stimulation of the MPA influences firing rates in the midbrain reticular formation (62 ,63). In summary, the preoptic area/basal forebrain appears to play an important role in the regulation of sleep and its integration with other processes; thus, it seems a likely candidate to be influenced by compounds that alter sleep and waking. Studies in the author's laboratory indicate that microinjections of pentobarbital (64), triazolam (65),

and propofol (66) into the MPA result in enhanced sleep in rats. Other groups have found similar results after microinjection of ethanol (67), adenosine (68), and prostaglandin D2 (69). Interestingly, the ability of adenosine to induce sleep when infused into the MPA is prevented by the benzodiazepine receptor blocker flumazenil (70).

Lesion Studies

One question that arises as a result of these studies is whether the MPA is the only site at which these compounds act to induce sleep; phrased somewhat differently, we know from the microinjection studies that the MPA is *sufficient*, but not whether it is *necessary*, for the hypnotic action of benzodiazepines. One way to approach this is to lesion the MPA, and then determine whether peripherally administered triazolam will still induce sleep. In order to assess this, we induced ibotenic acid lesions of the preoptic area in rats, and allowed 7 to 9 days of recovery in order for sleep to return to prelesion levels. We then administered triazolam .8 mg per kg intraperitoneally, and found that it still potently reduced sleep latency and increased total sleep (71). To put this finding in context, many of the studies of sleep physiology have suggested the presence of redundant mechanisms, which perhaps reflects the importance to the CNS of maintaining sleep/wake processes. One classic example, for instance, is that after anatomic or pharmacologic lesions of the dorsal raphe nuclei, sleep is initially greatly reduced, but then slowly returns to normal amounts (72 ,73). It seems likely, then, that although the MPA represents a common area at which a wide range of hypnotic compounds may act, there remain additional redundant mechanisms for sleep regulation, at which these agents can alter sleep in the absence of an intact MPA.

Do Hypnotics Induce Sleep by Altering Brain Temperature?

As we have described, the MPA contains warm- and cold-sensitive neurons, and is involved in thermoeffector activities (56). Typically, brain temperature decreases in NREM sleep compared to waking (74), and preoptic temperature drops in behaviorally defined sleep (75). It is known that peripherally administered benzodiazepines reduce rodent core (76) and brain (77) temperature. Thus, one interesting issue is whether the hypnotic effects of drugs such as triazolam are owing to “direct” effects on sleep regulatory mechanisms, or alternatively, whether the observed effects on sleep are secondary to drug-induced changes in temperature. In general, both rectal (78) and peritoneal (65) temperatures rise briefly following microinjection of triazolam into the MPA, but they do so equally in animals injected with vehicle. This transient rise in peripheral temperature, then, appears to be a nonspecific response related to the mechanics of injecting a fluid into the MPA. It is not clear whether brain temperature itself is altered by this microinjection procedure. In a recent study, cerebral brain temperature was measured following microinjection of triazolam .25 µg or vehicle into the MPA. Although sleep latency and waking time were significantly reduced, and total sleep increased in the 2 hours after injection, there were no changes in mean temperature in the first or second hour, temperature at sleep onset, or mean change from preinjection baseline (79). This suggests that the effects of triazolam microinjections into the MPA on sleep are not secondary to alterations in brain temperature, and raise the interesting possibility that pharmacologic and physiologic sleep initiation may differ in their relationship to temperature. Whether benzodiazepines induce subtler changes in hypothalamic temperature and play a role in sleep initiation remain to be determined.

CONCLUSION

Part of "132 - Basic Mechanisms of Sedative/Hypnotics "

Integrating Pharmacology of Hypnotics with Physiologic Sleep Regulatory Mechanisms

How should we picture the results of these microinjection and lesion studies in the context of what is known about the regulation of sleep? One of the most intriguing questions is: How does triazolam, which has a potent hypnotic effect when taken systemically, *reduce* sleep when injected into the dorsal raphe nuclei? The most likely explanation is that triazolam effectively inhibits raphe nuclei function. Benzodiazepines given systemically or iontophoretically have been reported to potentiate GABA inhibition of the dorsal raphe (80). Moreover, both depletion of serotonin by PCPA (81) or anatomic lesions of the dorsal raphe (82) greatly decrease sleep. Thus, it seems likely that the reduction in sleep seen after microinjection of triazolam mimics the effects of lesions of this structure.

As described, ascending fibers originating in the dorsal raphe travel via the medial forebrain bundle to various areas, including the hypothalamus (particularly the MPA and tuberomamillary nucleus), basal forebrain, septal area, striatum, hippocampus, and cerebral cortex (35). Descending fibers, in turn, travel to the dorsal raphe from the preoptic area. The dorsal and median raphe and the locus ceruleus also receive GABAergic and galaninergic innervation from the ventrolateral preoptic area (VLPO), which has an inhibitory influence on ascending monoaminergic systems (42).

Just as the MPA sends descending fibers to these specific monoaminergic brainstem nuclei, it also profoundly alters function of the reticular formation. This diffuse ascending system originates in the upper brainstem thalamic core, and projects largely through the intralaminar and other thalamic nuclei to the cortex, where it induces arousal (83). The main neurotransmitters involved in this process appear to be acetylcholine and glutamate (84). Stimulation of the MPA evokes inhibitory field potentials and suppresses neuronal

activity in the midbrain reticular formation, or MRF (62). Other stimulation data suggest that it induces an initial excitation followed by postexcitatory discharge suppression, suggesting that some actions of the basal forebrain on sleep might be mediated by alterations in activity in the MRF (63 ,85). In summary, we have described a system in which ascending monoaminergic pathways innervate the preoptic area, which appears to be an integrative center for sleep and a variety of physiologic processes, including thermoregulation and cardiovascular function, and at which microinjections of a wide range of classes of hypnotic compounds induce sleep. In turn, descending fibers from the preoptic area provide an inhibitory influence on both specific aminergic nuclei such as the dorsal raphe, as well as the midbrain reticular formation. We theorize that benzodiazepine hypnotics such as triazolam act by altering function of this reciprocal system of innervation between the hypothalamus/basal forebrain and brainstem structures.

Future Directions

Obviously, there are many potential areas to explore as possible sites of action of hypnotics. One possibility that has received little attention has been that classical hypnotics, such as benzodiazepines or barbiturates, might alter the ascending histaminergic arousal system, which is presumably the mechanism by which antihistamines produce sedating effects. Certainly one area of interest is the tuberomamillary nucleus, which lies adjacent to the mamillary bodies, just above the ventral surface of the hypothalamus (38). It is a histamine-producing cell group thought to be part of the ascending arousal system, with fibers going to the amygdala, hippocampus, and cortex (86). Data using the retrograde tracer cholera toxin subunit B indicate that the ventrolateral preoptic area (VLPO), which has been postulated to be involved in physiologic sleep initiation, innervates the histaminergic neurons in the tuberomamillary nucleus, via a GABAergic pathway (41 ,85 ,87). In principle, benzodiazepine or other hypnotic compounds might act by enhancing GABAergic inhibition of the tuberomamillary nucleus, decreasing its arousing effects. As described, the tuberomamillary nucleus also receives ascending serotonergic input from the dorsal raphe nuclei (38); therefore, it is conceivable that effects of hypnotics on sleep might also be mediated by their actions on the dorsal raphe, which in turn affect this structure.

ACKNOWLEDGMENT

Part of "132 - Basic Mechanisms of Sedative/Hypnotics "

This work was partially supported by NIH grants K07 HL03640 and 1RO1DA10682-01.

REFERENCES

1. Franks NP, Lieb WR. What is the molecular nature of general anaesthetic target sites [Abstract]? *TIPS* 1987;8:169–174.

2. Franks NP, Lieb WR. Which molecular targets are most relevant to general anaesthesia [Abstract]? *Toxicol Lett* 1998;Nov. 23:100–101.

3. Theodore WH, DiChiro G, Margolin R, et al. Barbiturates reduce human cerebral glucose metabolism [Abstract]. *Neurology* 1986;36:60–64.

4. Nemoto EM, Klementavicius R, Melick JA, et al. Suppression of cerebral metabolic rate for oxygen (CMR02) by mild hypothermia compared to thiopental [Abstract]. *J Neurosurg Anesthesiol* 1996;8:52–59.

5. Carlsson C, Chapman AG. The effect of diazepam on the cerebral metabolic state in rats and its interaction with nitrous oxide [Abstract]. *Anesthesiology* 1981;54:488–495.

6. Mendelson WB. *Human sleep: research and clinical care*. New York: Plenum Press, 1987.

7. Squires RF, Braestrup C. Benzodiazepine receptors in rat brain. *Nature* 1977;266:732–734.

8. Mohler H, Okada T. Benzodiazepine receptor: demonstration in the central nervous system. *Science* 1977;198:849–851.

9. Skolnick P. Is receptor heterogeneity relevant to the actions of benzodiazepine receptor ligands? In: Briles S, File S, eds. *New concepts in anxiety*. London: Macmillan, 1991:199–202.

10. Ortells MO, Lunt GG. Evolutionary history of the ligand-gated ion-channel superfamily of receptors. *Trends Neurosci* 1995;18:121–127.

11. Skolnick P, Mendelson WB, Paul SM. Benzodiazepine receptors in the central nervous system. In: Wheatley D, ed. *Psychopharmacology of sleep*. New York: Raven Press, 1981:117–134.

12. Hadingham KL, Wingrove P, le-Bourdelles BB, et al. Cloning of cDNA sequences encoding human alpha 2 and alpha 3 gamma-aminobutyric acid-A receptor subunits and characterization of the benzodiazepine pharmacology of alpha 5-containing human gamma-aminobutyric acid-A receptors. *Mol Pharmacol* 1993;43:970–975.

13. Tecott LH. Designer genes and anti-anxiety drugs. *Nat Neurosci* 2000;3:529–530.

14. Fritschy J-M, Mohler H. GABA-A receptor heterogeneity in the adult rat brain: differential regional and cellular distribution of seven major subunits. *J Comp Neurol* 1995;359:154–194.

15. Mendelson WB. Neuropharmacology of sleep induction by benzodiazepines. *Crit Rev Neurobiol* 1992;6:221–232.

16. Mendelson WB, Monti D. Do benzodiazepines induce sleep by a GABAergic mechanism? *Life Sci* 1993;53:81–87.

17. Lancel M, Cronlein AM, Faulhaber J. Role of GABAa receptors in sleep regulation: differential effects of muscimol and midazolam on sleep in rats. *Neuropsychopharmacology* 1996;15:63–74.

18. Mendelson WB. Sleep and core temperature measures following microinjection of muscimol into the medial preoptic area of rats [Abstract]. *Sleep Hypnosis* 1999;1:158–162.

19. Lancel M. Role of GABA-A receptors in the regulation of sleep: initial sleep responses to peripherally administered modulators and agonists. *Sleep* 1999;22:33–44.

20. Barker JL, Harrison NL, Mariani AP. IV. Benzodiazepine pharmacology of cultured mammalian CNS neurons. *Life Sci* 1986;39:1959–1968.

21. Seltzer AM, Donoso AO. Restraining action of GABA on estradiol-induced LH surge in the rat: GABA activity in brain nuclei and effects of GABA mimetics in the medial preoptic nucleus. *Neuroendocrinology* 1992;55:28–34.

22. Mendelson WB, Cain M, Cook JM, et al. A benzodiazepine receptor antagonist decreases sleep and reverses the hypnotic actions of flurazepam. *Science* 1983;219:414–416.

23. Olsen RW, Tobin AJ. Molecular biology of the GABA-A receptor. *FASEB J* 1990;4:1469–1480.

24. Wong G, Sci Y, Skolnick P. Stable expression of type I gamma-aminobutyric acid-A-benzodiazepine receptors in a transfected cell line. *Mol Pharmacol* 1992;42:996–1003.

25. Uchida I, Kamatchi G, Burt D, et al. Etomidate potentiation of GABA-A receptor gated current depends on the subunit composition. *Neurosci Lett* 1995;185:203–206.

26. Krasowski MD, Koltchine VV, Rick CE, et al. Propofol and other intravenous anesthetics have sites of action n the gamma-aminobutyric acid-A receptor distinct from that for isoflurane. *Mol Pharmacol* 1998, in press.

27. Twyman RE, Rogers CJ, Macdonald RL. Differential regulation of γ-aminobutyric acid receptor channels by diazepam and pentobarbital. *Ann Neurol* 1989;25:312–320.

28. Macdonald RL, Rogers CJ, Twyman RE. Barbiturate regulation of kinetic properties of the GABA-A receptor channel of mouse spinal neurones in culture. *J Physiol* 1989;417:483–500.

29. Suzdak PD, Schwartz RD, Skolnick P, et al. Ethanol stimulates gamma-aminobutyric acid receptor-mediated chloride transport in rat brain synaptoneurosomes. *Proc Natl Acad Sci USA* 1986;83:4071–4075.

30. Miller GL, Galpern RW, Dunlap K, et al. Interleukin-1 augments gamma-aminobutyric acid(a) receptor function in brain. *Amer Soc Pharmacol Exp Ther* 1993;39:105–108.

31. Hernandez Peon R, Chavez Ibarra G. Sleep induced by electrical or chemical stimulation of the forebrain. *EEG Clin Neurophysiol* 1963;24:188–198.

32. Mendelson WB. State-altering effects of benzodiazepines and barbiturates. In: Lydic R, Baghdoyan HA, eds. *Handbook of behavioral and state control*. Boca Raton, FL: CRC Press, 1998:407–419.

33. Mendelson WB. Effects of microinjections of triazolam into the ventrolateral preoptic area on sleep in the rat. *Life Sci* 2000, in press.

34. Azmitia EC, Segal M. An autoradiographic analysis of the differential ascending projections of the dorsal and median raphe nuclei in the rat. *Comp Neurol* 1978;179:641–668.

35. Halliday G, Harding A, Paxinos G. Serotonin and tachykinin systems. In: Paxinos G, ed. *The rat nervous system*. New York: Academic Press, 1995:929–974.

36. Simerly RB, Swanson LW. The organization of neural inputs to the medial preoptic nucleus of the rat. *J Comp Neurol* 1985;246:312–342.

37. Tork I. Raphe nuclei and serotonin containing systems. In: Paxinos G, ed. *The rat nervous system*. Sydney: Academic Press, 1985:43–78.

38. Ericson H, Blomqvist A, Kohler C. Brainstem afferents to the tuberomamillary nucleus in the rat brain with special reference to monoaminergic innervation. *J Comp Neurol* 1989;281:169–192.

39. Cudennec A, Duverger D, Serrano A, et al. Influence of ascending serotonergic pathway on glucose use in the conscious rat brain. II. Effects of electrical stimulation of the rostral raphe nuclei. *Brain Res* 1988;444:227–241.

40. Bjorklund S, Hokfelt T, Swanson LW, eds. *Handbook of chemical neuroanatomy*, fifth ed. New York: Elsevier 1987:1–124.

41. Sherin JE, Shiromani PJ, McCarley RW, et al. Activation of ventrolateral preoptic neurons during sleep. *Science* 1996;271:216–219.

42. Sherin JE, Elmquist JK, Torrealba F, et al. Innervation of histaminergic tuberomamillary neurons by GABAergic and galaninergic neurons in the ventrolateral preoptic nucleus of the rat. *J Neurosci* 1998;18:4705–4721.

43. Mosko SS, Haubrich D, Jacobs BL. Serotonergic afferents to the dorsal raphe nucleus: evidence from HPR and synaptosomal uptake studies. *Brain Res* 1977;119:269–290.

44. Hobson JA, Lydic R, Baghdoyan N. Evolving concepts of sleep cycle generation: From brain centers to neuronal populations. *Behav Brain Sci* 1986;9:371–448.

45. Hobson JA, McCarley RW, Wyzinski PW. Sleep cycle oscillations: reciprocal discharge by two brainstem neuronal groups. *Science* 1975;18:55–58.

46. Gallager DW, Aghajanian GK. Inhibition of firing of raphe neurones by tryptophan and 5-hydroxytryptophan: blockade by inhibition serotonin synthesis with Ro-4-4602. *Neuropharmacology* 1976;15:149–156.

47. Mendelson WB, Martin JV, Perlis M, et al. Arousal induced by injection of triazolam into the dorsal raphe nucleus of rats. *Neuropsychopharmacology* 1987;1:85–88.

48. Von Economo C. Schlaftheorie. *Ergeb Physiol* 1929;28:312–339.

49. Hess WR. Hirneizversuche uber den meshanismum der schlafes. *Psychiatr Nervenkr* 1929;86:287–292.

50. Nauta WJH. Hypothalamic regulation of sleep in rats: an experimental study. *J Neurophysiol* 1946;9:285–316.

51. Sterman MB, Clemente CD. Forebrain inhibitory mechanisms: sleep patterns induced by basal forebrain stimulation in the behaving cat. *Exp Neurol* 1962;6:103–117.

52. McGinty DJ, Sterman MB. Sleep suppression after basal forebrain lesions in the cat. *Science* 1968;160:1253–1255.

53. Szymusiak R, Satinoff E. Ambient temperature-dependence of sleep disturbances produced by basal forebrain damage in rats. *Brain Res Bull* 1984;12:295–305.

54. Boulant JA, Silva NL. Neuronal sensitivities in preoptic tissue slices: interactions among homeostatic systems. *Brain Res Bull* 1988;20:871–878.

55. Hori T, Nakashima T, Koga H, et al. Convergence of thermal, osmotic and cardiovascular signals on preoptic and anterior hypothalamic neurons in the rat. *Brain Res Bull* 1988;20:879–885.

56. McGinty DJ, Szymusiak RS. Hypothalamic thermoregulatory control of slow wave sleep. In: Mancia M, Marini G, eds. *The diencephalon and sleep*. New York: Raven Press, 1990:97–110.

57. Simerly RB, Gorski RA, Swanson LW. Neurotransmitter specificity of cell and fibers in the medial preoptic nucleus: an immunohistochemical study in the rat. *J Comp Neurol* 1986;246:343–363.

58. Demling J, Fuchs E, Baumert M, et al. Preoptic catecholamine, GABA, and glutamate release in ovariectomized estrogen-primed rats utilizing a push-pull cannula technique. *Neuroendocrinology* 1985;41:212–218.

59. Tappaz ML, Brownstein MJ, Kopin IJ. Glutamate decarboxylase (GAD) and γ-aminobutyric acid (GABA) in discrete nuclei of hypothalamus and substantia nigra. *Brain Res* 1977;125:109–121.

60. Koyama Y, Hayaishi O. Firing of neurons in the preoptic/anterior hypothalamic areas in rat: its possible involvement in slow wave sleep and paradoxical sleep. *Neurosci Res* 1994;19:31–38.

61. Simerly RB, Swanson LW. Projections of the medial preoptic nucleus: a Phaseolus vulgaris leukoagglutinin anterograde tract-tracing study in the rat. *J Comp Neurol* 1988;270:209–242.

62. Bremer F. Preoptic hypnogenic area and the reticular activating system. *Arch Ital Biol* 1973;111:85–111.

63. Szymusiak R, McGinty D. Responses evoked in midbrain reticular formation neurons by stimulation of the medial and lateral basal forebrain. *Sleep Res* 1988;17:

64. Mendelson WB. Sleep induction by microinjection of pentobarbital into the medial preoptic area in rats. *Life Sci* 1996;59:1821–1828.

65. Mendelson WB, Monti D. Effects on sleep of triazolam and nifedipine injections into the medial preoptic area. *Neuropsychopharmacology* 1993;8:227–232.

66. Tung A, Mendelson WB. Are anesthetics sleeping pills? Sleep induction by injection of propofol into the medial preoptic area of rats. *Anesth Analg* 2000;92S:S474.

67. Ticho SR, Stojanovic M, Lekovic G, et al. Effects of ethanol injection to the preoptic area on sleep and temperature in rats. *Alcohol* 1992;9:275–278.

68. Ticho SR, Radulovacki M. Role of adenosine in sleep and temperature regulation in the preoptic area of rats. *Pharmacol Biochem Behav* 1991;40:33–40.

69. Ueno R, Ishikawa Y, Wakayama T, et al. Prostaglandin D2 induces sleep when microinjected into the preoptic area of conscious rats. *Biochem Biophys Res Commun* 1982;109:576–582.

70. Mendelson WB. Sleep-inducing effects of adenosine microinjections into the medial preoptic area are blocked by flumazenil [Abstract]. *Brain Res* 2000;852:479–481.

71. Mendelson WB. Effects of parenterally administered triazolam on sleep in rats with lesions of the preoptic area. *Pharmacol Biochem Behav* 1998;61:81–86.

72. Dement W, Zarcone V, Ferguson J, et al. Some parallel findings in schizophrenic patients and serotonin-depleted cats. In: Sankar S, ed. *Schizophrenia: current concepts and research*. Hicksville, NY: PJD Publications, 1969:775–809.

73. Morgane PJ. Amine pathways and sleep regulation. *Brain Res Bull* 1982;9:743–749.

74. Gltozbach SF, Heller C. Temperature regulation. In: Kryger MH, Roth T, Dement WC, eds. *Principles and practice of sleep medicine*. Philadelphia: WB Saunders, 1989:260–275.

75. Abrams R, Hammel HT. Hypothalamic temperature in unanesthetized rats during feeding and sleeping. *Am J Physiol* 1964;206:641–646.

76. Wolffgramm J, Mikolaicczyk C, Coper H. Acute and subchronic benzodiazepine-barbiturate interactions on behaviour and physiological responses of the mouse. *Naunyn-Schmiedebergs Arch Pharmacol* 1994;349:279–286.

77. Schwartz RD, Huff RA, Yu X. Postischemic diazepam is neuroprotective in the gerbil hippocampus. *Brain Res* 1994;647:153–160.

78. Mendelson WB, Martin JV. Characterization of the hypnotic effects of triazolam microinjections into the medial preoptic area. *Life Sci* 1992;50:1117–1128.

79. Pham TT, Mendelson WB. Effects of microinjections of triazolam into medial preoptic area on sleep and brain temperature in rats. *Sleep Hypnosis* 1999;1:10–13.

80. Gallagher DW. Benzodiazepines: potentiation of a GABA inhibitor response in the dorsal raphe nucleus. *Eur J Pharmacol* 1978;49:133–143.

81. Pujol JF, Buguet A, Froment JL, et al. The central metabolism of serotonin in the cat during insomnia: a neurophysiological and biochemical study after p-chlorphenylalanine or destruction of the raphe system. *Brain Res* 1971;29:195–212.

82. Jouvet M, Renault J. Insomnie persistente apres lesions des noyaux raphe chez le chat. *Compt Rend Soc Biol (Paris)* 1966;160:1461–1465.

83. Steriade M. Arousal: revisiting the reticular activating system. *Science* 1996;272:225–226.

84. Steriade M, McCormick DA, Sejnowski TJ. Thalamocortical oscillations in the sleeping and aroused brain. *Science* 1993;262:679–682.

85. Szymusiak R, McGinty D. Effects of basal forebrain stimulation on the waking discharge of neurons in the midbrain reticular formation of cats. *Brain Res* 1989;498:355–359.

86. Wilcox BJ, Seybold VS. Localization of neuronal histamine in rat brain. *Neurosci Lett* 1982;29:105–110.

87. Krilowicz BL, Szymusiak R, McGinty D. Regulation of posterior lateral hypothalamic arousal neuronal discharge by preoptic anterior hypothalamic warming. *Brain Res* 1994;668:30–38.

133

Current and Experimental Therapeutics of Insomnia

Daniel J. Buysse

Cynthia M. Dorsey

Daniel J. Buysse: Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

Cynthia M. Dorsey: Department of Psychiatry, McLean Hospital, Belmont, Massachusetts.

INSOMNIA: DEFINITIONS, IMPACT, AND DIAGNOSIS

Part of "133 - Current and Experimental Therapeutics of Insomnia "

Definition of Insomnia Symptoms and Disorders

Insomnia can refer to either a symptom or clinical disorder. The *symptom* of insomnia is the subjective complaint of difficulty falling or staying asleep, poor quality sleep, or inadequate sleep duration, despite having an adequate opportunity for sleep. Two points in this definition deserve specific attention. First, insomnia is a subjective complaint not currently defined by laboratory test results or a specific duration of sleep or wakefulness. Second, the insomnia symptom occurs despite the individual having adequate opportunity to sleep. This distinguishes insomnia from sleep deprivation, which has different causes, consequences, and clinical presentations. As a *disorder*, insomnia is a syndrome consisting of the insomnia complaint, together with specific diagnostic features (either clinical or laboratory), significant distress or functional impairment, and the absence of specific exclusionary features.

Prevalence of Insomnia Symptoms and Disorders

Insomnia symptoms and disorders are widely prevalent. A number of recent epidemiologic studies indicate a prevalence of 30% to 45% for insomnia *symptoms* in the prior year (1 ,2 and 3). The specific prevalence depends on the definition of insomnia symptoms used. The prevalence of insomnia *disorders* is obviously much lower, in the range of 10% to 15% (4 ,5). Epidemiologic studies point to a consistent set of risk factors for insomnia. These include a previous history of insomnia, increasing age, female gender, psychiatric symptoms and disorders, medical symptoms and disorders, impaired activities of daily living, anxiolytic and hypnotic medication use, and low socioeconomic status. The increasing prevalence of insomnia with age may be explained in large part by increasing comorbidity with medical and psychiatric disorders and medication use. The incidence of insomnia also increases with age and is greater in women than men. On the other hand, remission of insomnia decreases with age and is less common in women. Together, prevalence, incidence, and remission data indicate that insomnia is often a chronic condition. Between 50% and 80% of individuals with insomnia at baseline have a persistent complaint after follow-up intervals of 1 to 3.5 years (1 ,6 ,7 and 8).

Impact of Insomnia

Studies in working populations show that individuals complaining of insomnia have more mood symptoms, gastrointestinal symptoms, headache, and pain (9). In addition, individuals with insomnia have greater self-ratings of role impairment, days of limited activity, days spent in bed, and higher total health costs (10). Health-related quality of life is significantly lower for individuals with insomnia than for those without (11). Individuals with insomnia may also have higher rates of serious accidents or injuries (12) and injurious falls (13). The economic costs of insomnia are also substantial. One recent estimate places the annual direct costs for insomnia-related problems at nearly $14 billion (including $11 billion related to nursing home care) (14). Insomnia has been identified as a significant risk factor for institutionalization in the elderly in some studies (15), but not in others (3). Despite these morbidities, insomnia does not appear to be an independent risk factor for mortality (3 ,16).

Perhaps the greatest morbidity associated with insomnia is an increased risk for psychiatric disorders. Several large,

carefully controlled studies have found that individuals with insomnia are at significantly increased risk for the development of depression, anxiety, and substance use disorders (4 ,17 ,18 ,19 ,20 and 21). These studies have included subjects from young adults to the elderly, and follow-up intervals from 1 to 35 years. Figure 133.1 shows data from the Breslau and associates study that are representative of these findings. The obvious—and unanswered—question is whether early identification and intervention in insomnia could prevent this costly outcome.

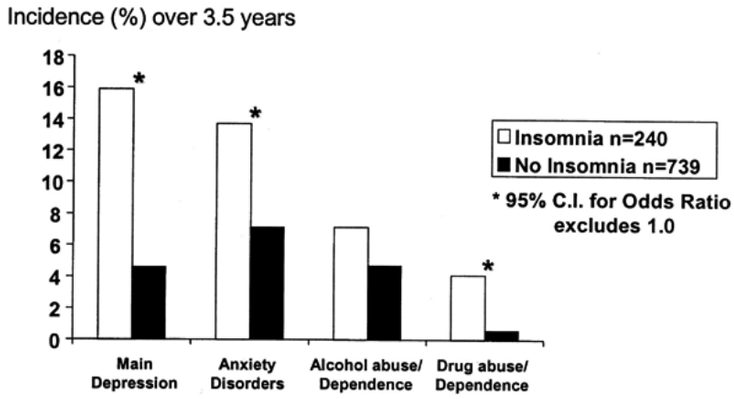


FIGURE 133.1. Insomnia is a risk factor for depression. Odds ratio's for developing new-onset psychiatric disorders among individuals with and without insomnia. Subjects were young adults in a health maintenance organization, interviewed at baseline, and at a 3.5-year follow-up. Baseline insomnia was a significant risk factor for incident depressive, anxiety, and substance use disorders. Data from ref. 19.

Differential Diagnosis

Insomnia can be the final result of many factors acting singly or in combination (Fig. 133.2). Many of these factors may lead to behaviors and conditioning that further reinforce the original problem. For instance, an individual who sleeps poorly may spend more time in bed in an effort to "catch up" on sleep. This extended time in bed occurring in an individual with impaired *ability* to sleep can further contribute to impaired sleep continuity and to the bed becoming a conditioned stimulus for wakefulness. Furthermore, most of the factors underlying insomnia contribute to increased physiologic arousal, which may constitute a final common pathway leading to insomnia complaints.

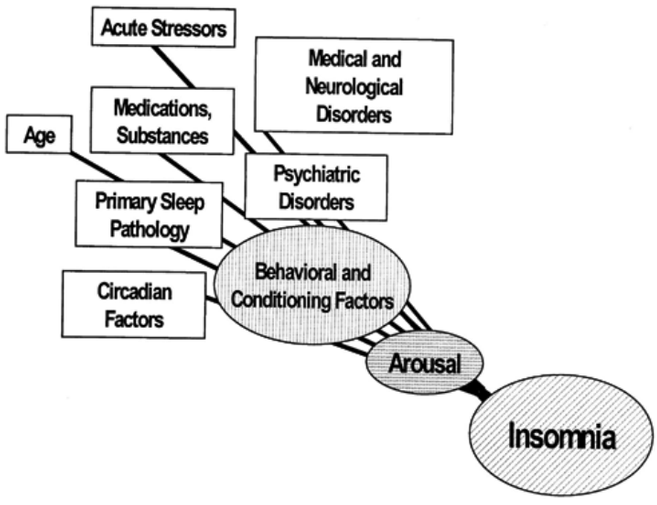


FIGURE 133.2. Causes of insomnia. Insomnia may result from many factors acting simply or in combination. Behavioral and conditioning factors often perpetuate insomnia that originally arises from other causes. Cognitive, affective, and somatic arousal may be the final common pathways leading to insomnia.

The differential diagnosis of insomnia disorders includes insomnia secondary to other medical, psychiatric, or substance use conditions; insomnia occurring during the course of primary sleep disorders; and primary insomnia. Insomnia secondary to psychiatric disorders is the most prevalent insomnia disorder, both in general population samples and clinically referred samples (5 ,22), accounting for 40% to 75% of all diagnoses. Virtually any psychiatric disorder can be associated with insomnia, although mood disorders (major depression, bipolar mood disorder, dysthymia) and anxiety disorders (generalized anxiety disorder, panic disorder, posttraumatic stress disorder) are the most common. Insomnia also may result from specific medical and neurologic disorders; those associated with pain, impaired mobility, and central nervous system (CNS) dysfunction are the most common. Common examples include arthritis, congestive heart failure, chronic obstructive pulmonary disease, and Parkinson's disease. A wide variety of drugs can also cause or contribute to insomnia, including alcohol, caffeine, decongestants, and other CNS stimulants, corticosteroids, and antidepressant medications, particularly selective serotonin reuptake inhibitors. Medications and substances can cause insomnia not only during the time they are being used, but also during withdrawal.

Insomnia also can be associated with specific sleep disorders. One common example is *restless legs syndrome* (RLS), which consists of an urge to move one's legs accompanied by uncomfortable dysesthesias, usually described as "creepy

crawly" or restless feelings. These sensations regularly increase at night, decrease during the day, and are temporarily relieved by movement. RLS is often associated with significant sleep onset insomnia. In addition, RLS is often accompanied by *periodic limb movement disorder* (PLMD), which consists of repetitive, short (.5- to 3.0-second) bilateral jerks in the toes, feet, ankles, and legs. These movements can lead to brief arousals and a complaint of nonrestorative sleep. *Sleep apnea syndromes* do not typically present with a complaint of insomnia. More often, sleep apnea presents within a syndrome of excessive daytime sleepiness, loud snoring, breathing pauses during sleep, and obesity or craniofacial abnormalities; however, a minority of patients, including older individuals, may present with insomnia complaints. *Circadian rhythm sleep disorders* often include prominent insomnia complaints. For instance, individuals with delayed sleep phase syndrome complain of difficulty falling asleep accompanied by difficulty awakening in the morning. Conversely, individuals with advanced sleep phase syndrome complain of early morning awakening and sleepiness in the evening hours. Jet lag and shift work sleep disorders are further examples of circadian sleep disorders that can present with insomnia problems.

Individuals who do not have other sleep disorders are diagnosed with primary insomnia. According to the DSM-IV, this syndrome is defined by a significant insomnia complaint; evidence of distress or impairment; and the absence of a concurrent psychiatric, medical, or sleep disorder that could explain the problem. Approximately 10% to 20% of individuals with significant insomnia are diagnosed with primary insomnia (5 ,22). This condition is broadly analogous to the term "psychophysiologic insomnia." The latter term invokes the etiologic factors of physiologic and cognitive arousal in association with the insomnia complaint.

Etiology and Neurobiology of Insomnia

Despite the prevalence and consequences associated with insomnia, relatively little is known regarding its neurobiology. One of the earliest and most enduring conceptualizations of insomnia is that of psychophysiologic arousal. Individuals with insomnia have several indicators of sympathetic and hypothalamic-pituitary-adrenal (HPA) axis activation, together with other peripheral indicators of "arousal." For instance, individuals with insomnia may have elevated temperature and muscle tone at sleep onset (23 ,24), elevated heart rate and sympathovagal tone in heart rate variability (25), and positive correlations among wake time after sleep onset and urinary norepinephrine, DOPAC, and DHPG (26). Studies of whole body metabolic rate, assessed by oxygen consumption, show elevated rates for individuals with insomnia compared to healthy controls, a difference which persists 24 hours per day (27). The psychologic arousal of insomnia is supported by higher rates of self-reported ruminations and intrusive thoughts among individuals with insomnia. It is less clear whether excess cognitive activity actually causes insomnia or is simply a byproduct of it. The cognitive hyperarousal may be the result of a ruminative, anxious personality style that also has been associated with insomnia.

Unfortunately, most studies identifying hyperarousal in insomnia are based on peripheral or "downstream" measures of arousal, rather than CNS arousal per se. Evidence for this type of arousal comes from electroencephalographic studies. Several investigators have demonstrated that individuals with insomnia have reduced sleep propensity not only at night, but also during the day. Individuals with insomnia also have lower δ EEG power (usually taken as an indicator of homeostatic sleep drive) and elevated amounts of ß EEG power (usually interpreted as evidence of EEG activation or cognitive activity) (28). In one recent investigation of depressed patients with insomnia, Nofzinger and colleagues found that ß EEG activity correlated positively with glucose metabolic rate in the medial orbitofrontal cortex, a region implicated in both behavioral and electroencephalographic activation (29). Behavioral evidence also supports the concept of increased cortical activity during sleep among individuals with insomnia. For instance, individuals with insomnia have better ability to recall auditory stimuli presented during the early sleep period relative to control subjects (30).

An integrative neurobiological model of insomnia should account for evidence of cortical activation during sleep, vulnerability to developing mood disorders, and evidence for sympathetic and HPA axis activation. It should also account for insomnia subjects' complaints of impaired concentration and memory, as well as their reduced propensity for sleep, even during daytime hours. Such a model may involve relative activation of ascending cholinergic and noradrenergic systems with diffuse projections to the cortex through thalamic and basal forebrain systems. The model may also involve reduced efficiency of processing in the frontal cortex, which may explain insomnia patients' complaints of poor concentration and attention. Another component of an integrative neurobiological model of insomnia would involve affective dysregulation. This might include amygdala activation or reduced activity in the subgenual anterior cingulate, similar to that observed in depression (31). Overactivity of ascending arousal systems, together with limbic system dysregulation, could lead to sustained activation of hypothalamic efferent systems, including activation of the sympathetic nervous system and HPA axis.

BEHAVIORAL AND NONPHARMACOLOGIC TREATMENT OF INSOMNIA

Part of "133 - Current and Experimental Therapeutics of Insomnia "

Rationale and Efficacy

Most behavioral and cognitive interventions aim to decrease or change factors that interfere with sleep, including maladaptive

sleep habits, cognitive or physiologic hyperarousal, and dysfunctional beliefs about sleep and insomnia. The ultimate goal of behavioral treatments for insomnia is to help patients manage their sleep and sleep habits more effectively. In addition to providing a safe alternative to pharmacotherapy, these nondrug treatments offer patients the potential benefit of a greater sense of control over their sleep problems. Most insomnia patients indicate that they would prefer a nonpharmacologic solution to their insomnia (32).

A comprehensive review of the efficacy of nonpharmacologic treatments for chronic primary insomnia, based on two metaanalyses and 48 individual treatment studies, showed reliable improvements in the main outcome measures of latency to sleep- and wake-time after sleep onset (33). Data consistently indicated that approximately 70% to 80% of insomniacs benefited from treatment. The magnitude of improvement was approximately 50%, with sleep latency reduced by about 30 minutes on average, from 60 to 30 minutes, and wake-time after sleep onset reduced from 70 to 38 minutes. Subjective report of sleep quantity and quality improved, based on sleep diary data. Relatively few studies have used PSG or actigraphy to document objective improvement. Improvements with behavioral treatment are well maintained over at least 6 months (34).

Good Sleep Practices (Sleep Hygiene Education)

Sleep hygiene education aims to promote environmental and lifestyle factors that are conducive to sleep, and to minimize those that affect sleep in a negative way, such as late-night caffeine consumption, sleeping with a television or radio on, or engaging in exercise in close proximity to bedtime. Many of these behaviors are not intrinsically problematic, but become detrimental to sleep if they are timed inappropriately. For example, exercise too close to bedtime can cause physiologic arousal that can impair sleep onset, whereas exercise during the late afternoon or earlier evening can have beneficial effects on sleep (35). Lifestyle factors alone are rarely the cause of chronic insomnia, but rather may complicate insomnia arising from other causes. Sleep hygiene modification is seldom considered sufficient treatment for chronic insomnia, and results from intervention studies support its limited efficacy. There is no single standard set of sleep hygiene recommendations; a sample of commonly reported elements is included in Table 133.1 .

Caffeine is a stimulant and should be discontinued 4–6 hours before bedtime.
Nicotine is a stimulant and should be avoided near bedtime and on awakening.
Alcohol is a depressant that can facilitate sleep onset, but can disrupt sleep later in the night. It should be avoided in close proximity to bedtime.
A heavy meal too close to bedtime can interfere with sleep and should be avoided. A light snack is all right.
Regular exercise in the late afternoon or early evening may deepen sleep, whereas exercise too close to bedtime may disrupt sleep.
Minimize light, noise, and excessive temperature during sleep.

Adapted from Morin (1990).

TABLE 133.1. SLEEP HYGIENE RECOMMENDATIONS

Stimulus Control Therapy

Stimulus control techniques (36) are based on the premise that insomnia is exacerbated or maintained by a maladaptive conditioned response to the bedroom environment and bedtime routine, which develops as a result of repeated difficulty sleeping. Whatever the initial cause of the insomnia, when an individual has experienced the frustration of lying in bed being unable to sleep, anxiety develops about the ability to sleep and the potential consequences of lack of sleep. Greater effort is made to lie in bed and consciously try to sleep. This behavior and effort are incompatible with sleep, cause further frustration and alertness, and result in “conditioned” insomnia.

The goal of stimulus control is to recondition cues such as the bedroom and bedtime routine to elicit relaxation and sleep as opposed to anxiety, frustration, and wakefulness. Stimulus control instructions are outlined in Table 133.2 . This is a paradoxic approach; the patient must accept the rationale of the treatment and trust the therapist enough to do something that does not make intuitive sense (i.e., get out of bed when he or she wants so desperately to sleep). Stimulus control requires effort and persistence, and often leads to initial resistance and temporary worsening before improvement. Consistency and motivation are important ingredients for a successful response. Quantitative reviews of controlled intervention trials consistently support the efficacy of stimulus control therapy.

Lie down intending to go to bed only when you are sleepy.
Use the bed and bedroom for sleep and sex only. Do not watch TV, listen to the radio, eat, or read in bed.
Get out of bed if you cannot fall asleep or go back to sleep within 10–15 minutes; return to bed only when you feel sleepy.
If you still cannot fall asleep, repeat the processing step as often as is necessary during the night.
Set your alarm and maintain a regular arising time in the morning, irrespective of how much sleep you got during the night.
Do not nap during the day.

Adapted from refs. 36 and 36a.

TABLE 133.2. STIMULUS CONTROL INSTRUCTIONS

Sleep Restriction Therapy

Patients with insomnia often try to compensate for lost sleep by getting into bed early or remaining in bed after awakening in the morning. Many individuals assume that bed rest may be restorative, even if no sleep is achieved. Unfortunately, the excess time in bed results in increased wakefulness in bed, which causes more frustration about difficulty sleeping and leads to even more pronounced insomnia. Increased time awake in bed can thus contribute to conditioned insomnia.

The goal of sleep restriction is to decrease time in bed in order to maximize the sleep efficiency of time spent in bed. Unlike stimulus control, sleep restriction addresses only the amount of time one spends in bed, rather than how the time in or out of bed is spent. This approach involves an initial curtailment of time in bed to the amount of time actually spent sleeping, based on sleep diary entries averaged over at least 1 week's time. Average sleep efficiency, which represents the proportion of time in bed spent asleep, is computed from sleep diaries. After sleep efficiency reaches desired levels (typically 90%), time allowed in bed can be increased by increments of 15 minutes until desired total sleep time at night is reached. If sleep efficiency remains low (<80%), after the initial restriction, time in bed is further curtailed by 15-minute increments until sleep continuity improves sufficiently. Time in bed is not changed if sleep efficiency is between 80% and 90%.

Relaxation and Biofeedback Therapies

Relaxation techniques target the cognitive or physiologic arousal that interferes with sleep, as discussed. A number of relaxation therapies have been used for insomnia, including progressive muscle relaxation and biofeedback to diminish physiologic arousal, and imagery techniques, autogenic training, and meditation to reduce cognitive arousal. Relaxation treatments may be most useful for sleep onset insomnia. In general, the magnitude of improvement seen with relaxation is smaller than for other behavioral approaches (37).

Cognitive and Multimodal Therapies

Insomnia often involves negative, unrealistic, or exaggerated beliefs about sleep and consequences of insomnia. These dysfunctional beliefs can cause emotional arousal and exacerbate the sleep problem (38). Cognitive restructuring has been used to help patients question the validity of automatic, maladaptive thoughts and reformulate them to make them more realistic and adaptive.

Many cognitive and behavioral techniques share common elements, and they are increasingly being used together within multimodal treatment protocols. It makes intuitive sense for some of the approaches to be combined, such as stimulus control and sleep restriction. The change in behavior advocated and the net result of each are similar, although the rationales are different. An adjustment in thinking, as can be accomplished with cognitive restructuring, can be helpful for successful completion of any behavioral or cognitive treatment. Effective, circumscribed, multicomponent therapies, such as that developed by Morin (39) combine several different treatment approaches within a limited number of treatment sessions to treat insomnia. The treatment protocol potentially can benefit a variety of patients who may respond differently to various aspects of the program. The treatments are integrated in a later session, and relapse prevention is addressed, promoting an overall focus on self-efficacy. From the existing literature, it is not clear that such combined approaches are more effective than the most effective of the individual techniques (e.g., stimulus control) used alone; however, such multifaceted therapies may have the added benefit of treating a broader range of patients without having to individualize treatment.

Other Nonpharmacologic Treatments

Phototherapy

As noted, insomnia associated with circadian rhythm sleep disorders results from problems related to the timing of sleep versus sleep itself. Because light is the most potent zeitgeber, or time cue, for the circadian timing system, phototherapy can be used as part of a treatment regimen to adjust the timing of the sleep/wake cycle and address a corresponding complaint of insomnia and/or sleepiness.

Exposure to bright light shifts circadian phase in a time-dependent manner (40). In general, bright light in the early morning hours shifts sleep and circadian rhythms to an earlier time (i.e., causes a phase advance); bright light in the evening hours shifts sleep and circadian rhythms to a later time (i.e., causes phase delays). Phototherapy can be delivered through artificial light, or by exposure to diffuse natural outdoor light. Artificial bright light has been shown to improve sleep maintenance insomnia in older adults (41) and younger adults with chronic insomnia (42); however, it is more typically used to treat circadian rhythm sleep disorders, such as delayed sleep phase syndrome. Practice parameters recently have been developed by the American Academy of Sleep Medicine (AASM) regarding use of phototherapy in the treatment of sleep disorders, including recommendations for light intensity and duration (43).

Exercise

The effects of exercise on sleep have been reviewed elsewhere (35 ,44). Exercise can increase sleep quality and slow-wave sleep, and reduce sleep latency in some individuals, including older adults (45). Individuals in good physical condition

observe the greatest effects on sleep with exercise in the late afternoon or early evening. Exercise performed in close proximity to bedtime, or by individuals who are unaccustomed to such exercise, can cause arousal. Exercise can be used as part of sleep hygiene recommendations or an overall training program potentially to improve sleep and health in general. Some of the effect of exercise on sleep may be mediated by changes in core body temperature. In particular, a rise in core body temperature with exercise may be followed by an exaggerated temperature decline during early sleep. This temperature decline may promote slow-wave sleep (46).

Passive Body Heating

Elevation of core body temperature by external body heating during the early evening also increases slow-wave sleep in both young and older individuals, and improves sleep continuity in older women with insomnia (47 ,48). Passive body heating involves immersion in hot water of at least 40°C for at least 30 minutes during afternoon or evening hours prior to bedtime. Although larger trials are needed, this procedure may constitute a relatively noninvasive, nonpharmacologic technique for treating insomnia.

PHARMACOLOGIC TREATMENTS FOR INSOMNIA

Part of "133 - Current and Experimental Therapeutics of Insomnia "

Several medication classes are used for the treatment of insomnia, although the strength of evidence regarding their efficacy and tolerability varies considerably. The major classes are benzodiazepine receptor agonists (BzRA), antidepressant drugs (AD), antihistamines, melatonin, and various herbal remedies including valerian root extracts. Of these medications, only BzRAs are formally approved for the indication of insomnia treatment in the United States. Nevertheless, physician-prescribing data show that prescriptions for BzRA hypnotics declined by 150% between 1987 and 1996, at the same time that benzodiazepine nonhypnotic prescriptions for insomnia remained stable, and antidepressant prescriptions increased by 150% (49). In particular, prescriptions for trazodone increased sixfold.

Benzodiazepine Receptor Agonists

Benzodiazepine receptor agonists (BzRAs) include the true benzodiazepines (e.g., triazolam, temazepam, estazolam, and lorazepam) as well as a structurally dissimilar group of nonbenzodiazepine agents, including an imidazopyridine (zolpidem), pyrazolopyrimidine (zaleplon), and cyclopyrolone (zopiclone). BzRAs are the only pharmacologic agents currently approved by the FDA for the treatment of insomnia, and they are labeled for short-term use (i.e., less than 4 weeks).

BzRAs share common pharmacodynamic actions, discussed in Chapter 68 . In summary, these agents bind at a specific recognition site in the benzodiazepine-γ aminobutyric acid (GABA)–chloride ion channel macromolecular complex. This binding is responsible for the hypnotic, anxiolytic, myorelaxant, and anticonvulsant actions of BzRAs. There is some evidence that the nonbenzodiazepine BzRAs, specifically zolpidem, may be relatively more specific for hypnotic effects relative to anticonvulsant and anxiolytic effects; this may be related to greater specificity for benzodiazepine type I receptors.

Specific BzRAs differ significantly in pharmacokinetic properties, including the rate of absorption, extent of distribution, and rate of elimination. BzRAs also range widely in elimination half-life, from 1 hour for zaleplon to 120 hours for flurazepam and its metabolites. Finally, these agents differ in terms of active metabolites, which may have longer half-lives than the parent compound. Table 133.3 outlines relevant pharmacokinetic properties for commonly used BzRAs.

Medication	Usual Adult Therapeutic Dose (mg)	Time to Onset (Minutes)	Terminal Elimination Half-Life (Hrs)	Active Metabolites
Marketed as hypnotics				
Estazolam	0.5–2.0	15–30	8–24	No
Flurazepam	15–30	30–60	2–5[a] 47–120[b]	Yes
Quazepam	7.5–30	20–45	15–40[a] 39–120[b]	Yes
Temazepam	7.5–30	45–60	8–20	No
Triazolam	0.125–0.25	15–30		No
Zaleplon	5–10	15	1.0	No
Zolpidem	5–10	30	1.5–4.5	No
Not marketed as hypnotics				
Clonazepam	0.25–2.0	20–60	19–60	No
Lorazepam	0.25–2.0	30–60	8–24	No

[a]Parent compound.
[b]Active metabolite.

TABLE 133.3. PHARMACOKINETIC PROPERTIES OF BENZODIAZEPINE RECEPTOR AGONISTS

BzRAs are efficacious in the short-term treatment of insomnia. Recent metaanalyses examined BzRA effects on sleep latency, sleep duration, number of awakenings, and sleep quality (50 ,50a). For each of these outcomes, the effect size d ranged between .55 and .75, indicating moderately large effect sizes and substantiating the superiority of these agents over placebo. Other data support a broader range of beneficial outcomes. For instance, treatment with zopiclone for both 14 days and 8 weeks of treatment was associated with greater improvements in quality of life measures, social activities, and professional activities compared to placebo (51). A telephone survey of patients with untreated insomnia and those receiving benzodiazepines showed that the later group reported fewer symptoms of feeling blue, down in the dumps, or depressed, and being easily upset compared to the former group (52).

BzRAs have consistent effects on PSG sleep measures. (See ref. 53 for review.) As expected, BzRAs are associated with reduced sleep latency and wakefulness during the night, and increased sleep duration and sleep quality ratings. Other specific PSG effects depend on the particular agent. For instance, zaleplon, with its very short half-life, has not been demonstrated to consistently affect sleep duration despite its effect on sleep latency. Traditional benzodiazepines reduce REM and stage 3 to 4 NREM sleep, whereas zaleplon and zolpidem are not associated with changes in sleep stages. In addition, BzRAs reduce the number of periodic limb movements and arousals associated with these movements (54). BzRAs can also lead to oxyhemoglobin desaturations during sleep and can theoretically worsen sleep apnea; however, in patients with moderate degrees of sleep apnea, the change in number of apneas and oxyhemoglobin saturation is felt to be clinically insignificant (55).

Although BzRAs have been studied primarily for short-term treatment of insomnia, insomnia is often a chronic

condition, and many patients take their hypnotics for longer periods of time. Some patients clearly developed tolerance with continued use of BzRAs, and some polysomnographic studies support this phenomenon (56); however, other PSG studies show continued efficacy over several nights of continued nightly administration. For instance, triazolam, zolpidem, and zaleplon have shown continued efficacy over a period of 4 to 5 weeks in double-blind, placebo-controlled studies (57 ,58 ,59 and 60), and single-blind studies have shown continued efficacy by PSG for as long as 6 months (61 ,62). Studies using self-ratings or observing-ratings have documented efficacy for even longer amounts of time. For instance, double-blind studies have shown continued efficacy for up to 24 weeks with no evidence of tolerance according to mean subject ratings (63 ,64) and single-blind studies have shown efficacy for up to 1 year of treatment (65 ,66); however, the role of BzRAs in long-term treatment or maintenance treatment of insomnia remains to be more clearly defined.

BzRAs can have significant adverse effects. The most common of these is a continuation of their desired therapeutic effect, sedation during the daytime. Daytime sleepiness is clearly more severe with longer-acting agents such as flurazepam, which has been documented in PSG studies (57 ,67). Similar PSG studies of short-acting hypnotics have not shown an increase in daytime sleepiness. BzRAs are also associated with dose-related anterograde amnesia that may even be partially responsible for their therapeutic affect (68 ,69). BzRAs can also impair other aspects of psychomotor performance, including reaction time, recall, and vigilance. Whether or not such deficits improve with this continuation of the drug is more controversial, with some studies noting improvement following discontinuation (70 ,71) and other studies failing to show such improvement (72).

BzRAs are significantly related to an increased risk of injurious falls and hip fractures in elderly people. In particular, risk seems to be increased with the use of long-acting agents, high doses, multiple agents, and cognitive impairment in patients (73 ,74). Data regarding automobile crashes are somewhat mixed. Self-report data have shown an increased rate of motor vehicle accidents in those taking hypnotics (12) but a case-control study in older individuals failed to show an elevated risk associated with benzodiazepines (75). Other studies have shown risk associated with long half-life drugs and recent initiation of treatment, but not with longer-term treatment (76). An analysis of data from over one million subjects in a national cancer database demonstrated an increased risk of all-cause mortality among older individuals using hypnotics, even after controlling for symptoms of insomnia, although the hypnotic drugs examined included barbiturate and other drugs as well as benzodiazepines (77). In fact, examination of two specific benzodiazepine agents in this cohort did not show an elevated mortality risk.

Several discontinuation phenomena have been examined in relation to BzRAs. *Rebound insomnia* refers to an increase in insomnia symptoms beyond their baseline level. Rebound is thought to be associated primarily with short-acting BzRAs, although recent evidence for zolpidem and zaleplon does not show this effect. Patients who demonstrate rebound insomnia tend to have worse baseline sleep and higher medication doses than patients without rebound (78 ,79). The behavioral aspect of taking a pill may contribute to rebound insomnia. Individuals who have shown a poor

response to treatment may show the greatest rebound (80). *Withdrawal* refers to the appearance of new symptoms on discontinuation of the drug. Withdrawal may occur in 40% to 100% of patients treated chronically with benzodiazepines, and can persist for days or weeks following discontinuation (81 ,82). Withdrawal symptoms can include dizziness, confusion, depression, and feelings of unreality. Cognitive and behavioral treatments can help patients discontinue chronic benzodiazepine use (83). The prevalence of true withdrawal phenomenon in any individuals treated with once-daily hypnotic doses of BzRAs is not well known. *Recurrence* is another potential discontinuance syndrome that has received little attention in insomnia. Given that insomnia tends to be chronic, it should not be surprising that many patients complained of their original symptom after discontinuation of an affected treatment. The role of recurrence in chronic BzRA treatment also remains to be well defined. Finally, *abuse* of BzRAs used for insomnia appears to be uncommon. One telephone survey showed no greater use of increased doses for BzRAs compare to antidepressants (84). Although data are difficult to obtain, benzodiazepines may be used by .5% to 3.0% of the population for nonmedical purposes in any 1 year (85). Among those who wish to discontinue chronic use of BzRAs, their pattern of use tends to suggest stability or declining doses over time as well as a tendency to intermittent rather than consistent dosing (86). Thus, among individuals with no prior substance use history, abuse of BzRAs appears to be uncommon.

Antidepressant Drugs

Although use of antidepressant drugs (AD) for insomnia has increased dramatically, evidence to support their efficacy is relatively sparse. The most commonly used ADs for insomnia include trazodone, tertiary tricyclic agents, and mirtazapine. These drugs clearly have diverse effects on neurotransmission, as reviewed in Chapter 79 . In general, the sedating properties of antidepressants are related to antagonism of serotonin 5-HT_2, histamine, and α_1-adrenergic receptors.

The effects of various antidepressant agents on sleep have been described primarily in the context of depression treatment. These effects are summarized in Table 133.4 and several recent reviews (87 ,88). As the table indicates, antidepressant drugs vary widely in their effects on sleep continuity, EEG delta activity and slow-wave sleep, and REM sleep. Sleep continuity effects are likely to be most important in the treatment of insomnia. Some antidepressant drugs also can cause or exacerbate insomnia problems. Selective serotonin reuptake inhibitors (SSRIs) bupropion, noradrenergic selective tricyclic drugs, and strongly serotonergic tricyclic drugs (e.g., clomipramine) are the most common agents to have such effects. In addition, serotonergic specific antidepressants can lead to anomalous sleep stages characterized by eye movements during NREM sleep and they can also cause or exacerbate restless leg syndrome and periodic limb movements (48). Antidepressants may also be associated with slight improvements in sleep apnea (89).

	Sleep Continuity	Slow Wave Sleep	REM Sleep
Tricyclic	↓ to ↑	→ to ↑	↓ to ↓↓↓
SSRIs, Venlafaxine	→ to ↓	→ to ↓	↓ to ↓↓
Trazodone, Nefazodone	↑	→ to ↑	→ to ↑
Bupropion	↓	→ to ↓	↑

→, No change.
↑, Increase.
↓, Decrease.

TABLE 133.4. EFFECTS OF ANTIDEPRESSANT DRUGS ON EEG SLEEP

Studies with small numbers of subjects and diverse inclusion criteria suggested the beneficial effects of trazodone 150 to 400 mg on sleep continuity measures, as well as a tendency to increase Stage 3 to 4 sleep and improve subjective sleep quality ratings, in insomnia patients (90 ,91 and 92). A more recent 2-week double-blind placebo-controlled study compared the effects of trazodone 50 mg and zolpidem 10 mg to placebo among individuals with primary insomnia (93). This study showed improvements in subjective sleep latency and sleep duration with both active drugs, although there was some evidence for superiority of zolpidem during the second treatment week. Both drugs were well tolerated. Other studies involving primary insomnia have shown beneficial effects of short-term treatment with low-dose doxepin (94) and trimipramine (95) compared to placebo. Finally, a recent open-label trial of paroxetine for primary insomnia in the elderly showed significant improvement in a multivariate measure of sleep quantity based on both diary and polysomnographic sleep measures (96). Small improvements were noted in diary-based measures of sleep quality and PSG measures of sleep efficiency; however, the greatest improvements were noted in daytime symptoms of mood and well being. Thus, it may not simply be the sedating properties of antidepressants that lead to improvements in insomnia.

Indirect evidence for the efficacy of antidepressants, and differential effects among agents, comes from studies in individuals with major depression. For instance, fluvoxamine has a relatively alerting effect relative to desipramine that in turn is more alerting than amitriptyline (97 ,98). A comparison of trimipramine and imipramine found that both drugs improve sleep quality, although trimipramine was associated with more positive effects on PSG sleep (99). A comparison of fluoxetine with trazodone showed that the later drug was associated with more improvements in insomnia

symptoms, but also with a greater percentage of sedating events during the daytime (100). A series of comparisons between fluoxetine and nefazodone has consistently shown that both drugs improve subjective sleep quality among depressed patients, although the change appears to be larger with nefazodone (101 ,102). Nefazodone also led to improvements in PSG sleep efficiency, whereas fluoxetine was associated with mild decrements.

Antihistamines

Antihistamines such as diphenhydramine and doxylamine are the most widely available over-the-counter preparations for insomnia. The mechanism of action of these drugs involves inhibition of histamine H_1 receptors. Histaminic neurons in the posterior hypothalamus promote wakefulness through interactions with ascending cholinergic nuclei and through projections through the thalamus. Inhibition of H_1 receptors leads to decreased alertness and subjective sedation. The elimination half-life of diphenhydramine ranges from 3 to 5 hours, within increases in elderly persons. In addition to their effects on histamine, these medications can also have antimuscarinic anticholinergic effects.

Despite their widespread use, a large body of well-documented research does not support the efficacy of antihistamines. Diphenhydramine 50 mg, improved subjective ratings of sleep quality, sleep time, sleep latency, and wakefulness after sleep onset in middle-aged subjects with insomnia (103). A more recent study comparing the effects of lorazepam versus a combination of lorazepam plus diphenhydramine showed a slight advantage for the combination preparation in terms of sleep latency and subjective sleep quality (104). On most sleep measures, the two drug preparations were fairly similar. Studies of antihistamines in elderly people demonstrate subjective sedative properties comparable in magnitude to those of benzodiazepines and confirmed by effects such as increased sleep time, decreased awakening, and shorter sleep latency (105 ,106).

Adverse effects of antihistamines include a range of cognitive and performance impairments (107). The anticholinergic effects of these medications may be of particular concern in elderly subjects. The relative safety and efficacy of antihistamines with more sustained use has not been examined.

Melatonin

Melatonin has been widely used as a “natural” sleep-promoting agent. Data regarding its efficacy and safety have been mixed. The study designs, doses, and outcome measures used in melatonin trials have been quite variable and may contribute to inconsistent findings (108). Melatonin is secreted by the pineal gland during hours of darkness in both diurnal and nocturnal mammals. Melatonin’s effect on sleep and wakefulness may result from interaction with specific receptors in the suprachiasmatic nucleus of the hypothalamus (109). In addition, melatonin shifts circadian rhythms according to a phase response curve (110 ,111). The half-life of endogenous melatonin is less than 1 hour. Exogenous melatonin is absorbed from the gastrointestinal tract, but a wide variety of preparations are commercially available, ranging from very short-acting to very long-acting agents, with half-lives ranging from several minutes to approximately 8 hours. Doses greater than 1 mg are likely to induce supraphysiologic concentrations. Clinical trials have employed doses ranging from .1 to 80 mg.

During daytime administration, melatonin causes sleepiness in fatigue and healthy subjects (112 ,113). When administered at night to healthy subjects, melatonin decreases sleep latency (114) and the number of awakenings, and improves sleep efficiency in an experimental insomnia paradigm (115).

Studies in insomnia patients have also yielded inconsistent findings. Single-night administration seems to produce very little effect (116). Subjective sleep ratings showed no effect in another trial of 5 mg for 1 week (117), whereas a 14-day trial of 75 mg resulted in increased subjective sleep time (118). Trials of melatonin in elderly people have ranged from 1 to 21 days. The most consistent effect is reduced sleep latency with some evidence as well for reduced nighttime wakefulness using sustained-released preparations (119 ,120 ,121 and 122). In a carefully designed 14-day crossover trial, immediate- and sustained-release melatonin were associated with shortened sleep latency, but no change in sleep time, sleep efficiency, wakefulness, or subjective sleep measures (123).

Adverse effects associated with melatonin have not been carefully evaluated. Melatonin has effects on reproductive cycles in several mammalian species, and reports have indicated the potential for worsening of sleep apnea and impaired cognitive and psychomotor performance during daytime administration. There are also some concerns regarding vasoconstriction as a potential side effect.

Valerian Extract

Valerian extract is one of the most widely used herbal remedies for insomnia. These extracts are derived from roots of the genus Valeriana, most often of the species *V. officinelis*. They contain a number of potentially active compounds, including sesquiterpenes and valepotriates. Valerian extracts show affinity for $GABA_A$ receptors, which may be related to the high amount of GABA itself that is often contained in these preparations (124 ,125). However, GABA does not cross the blood–brain barrier, so this is an unlikely mechanism of action. Other potential actions include affinity for serotonin and adenosine receptors.

Clinical studies with valerian extracts show mild sedative and anxiolytic effects. In particular, four double-blind placebo-controlled studies have examined doses of 400 to 900

mg of valerian extract over periods of time from 1 to 8 days, and in diverse subject populations ranging from healthy young adults to elderly insomniacs (126 ,127 ,128 and 129). Subjective effects include decreased sleep latency and improved sleep quality (126 ,127 ,129). One study also reported decreased subjectively rated awakenings (126). Polysomnographic studies have shown an increase in stage 3 to 4 NREM sleep and reduced stage 1 sleep (128), with no change in sleep onset time, awake time after sleep onset, or other measures of sleep continuity (128 ,129). Likewise, valerian was found not to influence the EEG power spectrum during sleep (129). Findings from these studies are hampered by small numbers of subjects, different inclusion criteria, and inconsistent findings. These studies do not demonstrate the efficacy of valerian extract in most groups of individuals with primary insomnia.

Clinical studies have suggested a generally favorable side effect profile for valerian extract; however, the sedative effects of valerian may potentiate the effects of other CNS antidepressants (125).

FUTURE DIRECTIONS

Part of "133 - Current and Experimental Therapeutics of Insomnia "

Although considerable progress has been made with regard to the epidemiology of insomnia, further work needs to be done regarding its consequences for health and role functioning. Individuals with insomnia complain not only of sleep disturbance, but daytime consequences as well. In addition, investigations into the neurobiology of insomnia are clearly needed. This will help to define the underlying pathophysiology of insomnia in the general sense, but also help to define the boundaries of specific insomnia disorders. Techniques from cognitive and affective neuroscience, as well as electrophysiology and psychophysiology, will lead to an improved understanding of this condition. Functional neuroimaging experiments will also contribute to our understanding of the circuitry involved in insomnia, and its boundaries with mood and anxiety disorders. To date, no animal model exists for insomnia that would also help to promote research in humans. Finally, genetic studies have been very useful for identifying abnormalities associated with narcolepsy and circadian rhythm sleep disorders. Similar genetic and genetic epidemiology strategies remain to be applied to a study of insomnia.

Several issues also remain with regard to treatment aspects of insomnia. First, the relative benefits and risks of treatment in terms of symptomatic relief, health-related quality of life, and morbidity remain to be defined. These issues are of considerable importance, given the potential for some insomnia treatments to cause significant adverse effects, such as cognitive impairment and injurious falls. The optimal duration of treatment and the conceptualization of potential “maintenance” treatments for insomnia is also an area open for further investigation.

With regard to behavioral treatments, one of the major challenges is designing well manualized and “exportable” treatments that can be applied more readily in a variety of treatment settings, including primary care settings. Several studies have begun to examine the optimal combination of behavioral and medication-treatment approaches. Some of the evidence suggests better durability of treatment effects with behavioral treatment alone (33); however, sequential treatments as well as concurrent treatments need to be investigated. In addition, treatment strategies for nonresponders to either behavioral or pharmacologic interventions must be developed.

Advances in the neurobiology of insomnia may come from basic neuroscience sources. For instance, recent evidence has accumulated regarding the role of adenosine as a modulator of sleep/wake states (130). Relative underactivity of adenosinergic neurotransmission could potentially result in reduced sleep drive. Another focus for dysregulation in insomnia may involve the ventrolateral preoptic area (VLPO) and its interactions with the tuberomamillary nucleus in the posterior hypothalamus (131 ,132). The GABAergic VLPO has been identified as one of the few “sleep active” areas of the brain; dysregulation in this nucleus and its efferent projections to histaminergic, cholinergic, and noradrenergic nuclei could conceivably shift the sleep/wake balance in the direction of wakefulness. Finally, recent findings regarding the role of orexin in sleep/wake regulation could have direct implications for the neurobiology and pharmacologic treatment of insomnia (133 ,134). Neuroscience and clinical studies can both inform the optimal management of insomnia disorders.

ACKNOWLEDGMENTS

Part of "133 - Current and Experimental Therapeutics of Insomnia "

This work was supported in part by AG15138, AG00972, MH24652, MH30915, AG13961. DJB has served as a consultant and on speaker’s bureaus for Searle, Wyeth-Ayerst, and Cephalon.

REFERENCES

1. Ganguli M, Reynolds CF, Gilby JE. Prevalence and persistence of sleep complaints in a rural elderly community sample: the MoVIES Project. *J Am Ger Soc* 1996;44:778–784.

2. Mellinger GD, Balter MB, Uhlenhuth EH. Insomnia and its treatment: prevalence and correlates. *Arch Gen Psychiatry* 1985;42:225–232.

3. Foley DJ, Monjan AA, Brown SL, et al. Sleep complaints among elderly persons: an epidemiologic study of three communities. *Sleep* 1995;18:425–432.

4. Ford DE, Kamerow DB. Epidemiologic study of sleep disturbances and psychiatric disorders. *JAMA* 1989;262:1479–1484.

5. Ohayon MM. Prevalence of DSM-IV diagnostic criteria of insomnia: distinguishing insomnia related to mental disorders from sleep disorders. *J Psychiatr Res* 1997;31:333–346.

6. Foley DJ, Monjan A, Simonsick EM, et al. Incidence and remission of insomnia among elderly adults: an epidemiologic study of 6,800 persons over three years. *Sleep* 1999;22:S366–S372.

7. Katz DA, McHorney CA. Clinical correlates of insomnia in patients with chronic illness. *Arch Int Med* 1998;158:1099–1107.

8. Dodge R, Cline MG, Quan SF. The natural history of insomnia and its relationship to respiratory symptoms. *Arch Int Med* 1995;155:1797–1800.

9. Kuppermann M, Lubeck DP, Mazonson PD, et al. Sleep problems and their correlates in a working population. *J Gen Int Med* 1995;10:25–32.

10. Simon GE, Von Korff M. Prevalence, burden, and treatment of insomnia in primary care. *Am J Psychiatry* 1997;154:1417–1423.

11. Zammit GK, Weiner J, Damato N, et al. Quality of life in people with insomnia. *Sleep* 1999;22:S379–S385.

12. Balter MB, Uhlenhuth EH. New epidemiologic findings about insomnia and its treatment. *J Clin Psychiatry* 1992;53:34–39.

13. Koski K, Luukinen H, Laippala P, et al. Risk factors for major injurious falls among the home-dwelling elderly by functional abilities. *Gerontology* 1998;44:232–238.

14. Walsh JK, Engelhardt CL. The direct economic costs of insomnia in the United States for 1995. *Sleep* 1999;22:S386–S393.

15. Pollak CP, Perlick D. Sleep problems and institutionalization of the elderly. *J Ger Psychiatry Neurol* 1991;4:204–210.

16. Althuis MD, Fredman L, Langenberg PW, et al. The relationship between insomnia and mortality among community-dwelling older women. *J Am Ger Soc* 1998;46:1270–1273.

17. Livingston G, Blizard B, Mann A. Does sleep disturbance predict depression in elderly people? A study in inner London. *Br J Gen Pract* 1993;43:445–448.

18. Eaton WM, Badawi M, Melton B. Prodromes and precursors: epidemiologic data for primary prevention of disorders with slow onset. *Am J Psychiatry* 1995;152:967–972.

19. Breslau N, Roth T, Rosenthal L, et al. Sleep disturbance and psychiatric disorders: a longitudinal epidemiological study of young adults. *Biol Psychiatry* 1996;39:411–418.

20. Chang PP, Ford DE, Mead LA, et al. Insomnia in young men and subsequent depression. The Johns Hopkins Precursors Study. *Am J Epidemiol* 1997;146:105–114.

21. Weissman MM, Greenwald S, Nino-Murcia G, et al. The morbidity of insomnia uncomplicated by psychiatric disorders. *Gen Hosp Psychiatry* 1997;19:245–250.

22. Buysse DJ, Reynolds CF, Hauri PJ, et al. Diagnostic concordance for DSM-IV sleep disorders: a report from the APA/NIMH DSM-IV field trial. *Am J Psychiatry* 1994;151:1351–1360.

23. Freedman RR, Sattler HL. Physiologic and psychological factors in sleep-onset insomnia. *J Abnorm Psychol* 1982;91:380–389.

24. Monroe LJ. Psychological and physiologic differences between good and poor sleepers. *J Abnorm Psychol* 1967;72:255–264.

25. Bonnet MH, Arand DL. Heart rate variability: sleep stage, time of night, and arousal influences. *Electroencephalograph Clin Neurophysiol* 1997;102:390–396.

26. Vgontzas AN, Tsigos C, Bixler EO, et al. Chronic insomnia and activity of the stress system: a preliminary study. *J Psychosom Res* 1998;45:21–31.

27. Bonnet MH, Arand DL. 24-Hour metabolic rate in insomniacs and matched normal sleepers. *Sleep* 1995;18:581–588.

28. Merica H, Gaillard JM. The EEG of the sleep onset period in insomnia: a discriminant analysis. *Physiol Behav* 1992;52:199–204.

29. Nofzinger EA, Price JC, Meltzer CC, et al. Towards a neurobiology of dysfunctional arousal in depression: the relationship between beta EEG power and regional cerebral glucose metabolism during NREM sleep. *Psychiatry Res Neuroimag* 2000;98:71–91.

30. Perlis ML, Smith MT, Orff HJ, et al. The mesograde amnesia of sleep may be attenuated in subjects with primary insomnia. *Physiol Behav* 2001;74:71–76.

31. Drevets WC, Price JL, Simpson JR, et al. Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* 1997;386:824–827.

32. Morin CM, Gaulier B, Barry T, et al. Patients' acceptance of psychological and pharmacological therapies for insomnia. *Sleep* 1992;15:302–305.

33. Morin CM, Hauri PJ, Espie CA, et al. Nonpharmacologic treatment of chronic insomnia. *Sleep* 1999;22:1134–1156.

34. Morin CM, Stone J, McDonald K. Psychological management of insomnia: a clinical replication series with 100 patients. *Behav Ther* 1994;25:291–309.

35. Horne JA. The effects of exercise upon sleep: a critical review. *Biol Psychol* 1981;12:241–290.

36. Bootzin RR, Nicassio PM. Behavioral treatments of insomnia. In: Hersen M, Eisler RE, Miller PM, eds. *Progress in behavior modification*, vol. 6. New York: Academic Press, 1978:1–45.

36a. Bootzin RR. A stimulus control treatment for insomnia. *Proceedings of the American Psychological Association* 1972;395–396.

37. Morin CM, Culbert JP, Schwartz SM. Nonpharmacological interventions for insomnia: a meta-analysis of treatment efficacy. *Am J Psychiatry* 1994;151:1172–1180.

38. Morin CM, Stone J, Trinkle D, et al. Dysfunctional beliefs and attitudes about sleep among older adults with and without insomnia complaints. *Psychol Aging* 1993;8:463–467.

39. Morin CM. A cognitive-behavioral conceptualization of insomnia. In: Barlow DH. *Insomnia-psychological assessment and management*. New York: Guilford Press, 1993:46–60.

40. Czeisler CA, Kronauer RE, Allan JS, et al. Bright light induction of strong (type 0) resetting of the human circadian pacemaker. *Science* 1989;244:1328–1333.

41. Campbell SS, Dawson D, Anderson MW. Alleviation of sleep maintenance insomnia with time exposure to bright light. *J Am Ger Soc* 1993;41:829–836.

42. Guilleminault C, Clerk A, Black J, et al. Nondrug treatment trials in psychophysiologic insomnia. *Arch Int Med* 1995;155:838–844.

43. Chesson AL Jr, Littner M, Davila D, et al. Practice parameters for the use of light therapy in the treatment of sleep disorders. Standards of Practice Committee, American Academy of Sleep Medicine. *Sleep* 1999;22:641–660.

44. Kubitz KA, Landers DM, Petruzzello SJ, et al. The effects of acute and chronic exercise on sleep. A meta-analytic review. *Sports Med* 1996;21:277–291.

45. King AC, Oman RF, Brassington GS, et al. Moderate-intensity exercise and self- rated quality of sleep in older adults. A randomized controlled trial. *JAMA* 1997;277:32–37.

46. Gillberg M, Akerstedt T. Body temperature and sleep at different times of day. *Sleep* 1982;5:378–388.

47. Horne JA, Reid AJ. Night-time sleep EEG changes following body heating in a warm bath. *Electroencephalograph Clin Neurophysiol* 1985;60:154–157.

48. Dorsey CM, Lukas SE, Teicher MH, et al. Effects of passive body heating on the sleep of older female insomniacs. *J Ger Psychiatry Neurol* 1996;9:83–90.

49. Walsh JK, Schweitzer PK. Ten-year trends in the pharmacological treatment of insomnia. *Sleep* 1999;22:371–375.

50. Nowell PD, Mazumdar S, Buysse DJ, et al. Benzodiazepines and zolpidem for chronic insomnia: a meta-analysis of treatment efficacy. *JAMA* 1997;278:2170–2177.

50a. Holbrook AM, Crowther R, Lotter A, et al. Meta-analysis of benzodiazepine use in the treatment of insomnia. *CMAJ* 2000;162:225–233.

51. Leger D, Querasalva MA, Philip P. Health-related quality of life in patients with insomnia treated with zopiclone. *Pharmacoeconomics* 1996;10:39–44.

52. Balter MB, Uhlenhuth EH. The beneficial and adverse effects of hypnotics. *J Clin Psychiatry* 1991;52:16–23.

53. Parrino L, Terzano MG. Polysomographic effects of hypnotic drugs: a review. *Psychopharmacology* 1996;126:1–16.

54. Hening W, Allen R, Earley C, et al. The treatment of restless legs syndrome and periodic limb movement disorder. *Sleep* 1999;22:970–999.

55. Berry RB, Kouchi K, Bower J. Triazolam in patients with obstructive sleep apnea. *Am J Respir Crit Care Med* 1995;151:450–454.

56. Kales A, Bixler EO, Scharf M, et al. Sleep laboratory studies of flurazepam: a model for evaluating hypnotic drugs. *Clin Pharmacol Therapeut* 1976;19:576–583.

57. Mitler MM, Seidel WF, Van Den Hoed J, et al. Comparative hypnotic effects of flurazepam, triazolam, and placebo: a long-term simultaneous nighttime and daytime study. *J Clin Psychopharmacol* 1984;4:2–13.

58. Scharf MB, Roth T, Vogel GW, et al. A multicenter, placebo-controlled study evaluating zolpidem in the treatment of chronic insomnia. *J Clin Psychiatry* 1994;55:192–199.

59. Walsh JK, Vogel GW, Scharf M, et al. A five week, polysomnographic assessment of zaleplon 10 mg for the treatment of primary insomnia. *Sleep Med* 2000;1:41–49.

60. Elie R, Ruther E, Farr I, et al. Sleep latency is shortened during 4 weeks of treatment with zaleplon, a novel nonbenzodiazepine hypnotic. Zaleplon Clinical Study Group. *J Clin Psychiatry* 1999;60:536–544.

61. Pecknold J, Wilson R, LeMorvan P. Long term efficacy and withdrawal of zopiclone: a sleep laboratory study. *Int Clin Psychopharmacol* 1990;5:57–67.

62. Kummer J, Guendel L, Linden J, et al. Long-term polysomnographic study of the efficacy and safety of zolpidem in elderly psychiatric in-patients with insomnia. *J Int Med Res* 1993;21:171–184.

63. Oswald I, French C, Adam K, et al. Benzodiazepine hypnotics remain effective for 24 weeks. *Br Med J* 1982;284:860–863.

64. Allen RP, Mendels J, Nevins DB, et al. Efficacy without tolerance or rebound insomnia for midazolam and temazepam after use for one to three months. *J Clin Pharmacol* 1987;27:768–775.

65. Schlich D, L'Heritier C, Coquelin JP, et al. Long-term treatment of insomnia with zolpidem: a multicentre general practitioner study of 107 patients [published erratum appears in *J Int Med Res* 1993;21(6):346]. *J Int Med Res* 1991;19:271–279.

66. Maarek L, Cramer P, Attali P, et al. The safety and efficacy of zolpidem in insomniac patients: a long-term open study in general practice. *J Int Med Res* 1992;20(2):162–170.

67. Carskadon MA, Seidel WF, Greenblatt DJ, et al. Daytime carryover of triazolam and flurazepam in elderly insomniacs. *Sleep* 1982;5:361–371.

68. Roth T, Roehrs R, Wittig R, et al. Benzodiazepines and memory. *Br J Clin Pharmacol* 1984;18:45S–49S.

69. Perlis ML, Giles DE, Mendelson WB, et al. Psychophysiologic insomnia: the behavioural model and a neurocognitive perspective. *J Sleep Res* 1997;6:179–188.

70. Tonne U, Hiltunen AJ, Vikander B, et al. Neuropsychological changes during steady-state drug use, withdrawal and abstinence in primary benzodiazepine-dependent patients. *Acta Psychiatr Scand* 1995;91:299–304.

71. Salzman C, Fisher J, Nobel K, et al. Cognitive improvement following benzodiazepine discontinuation in elderly nursing home residents. *Int J Ger Psychiatry* 1992;7:89–93.

72. Tata PR, Rollings J, Collins M, et al. Lack of cognitive recovery following withdrawal from long-term benzodiazepine use. *Psychol Med* 1994;24:203–213.

73. Mustard CA, Mayer T. Case-control study of exposure to medication and the risk of injurious falls requiring hospitalization among nursing home residents. *Am J Epidemiol* 1997;145:738–745.

74. Neutel CI, Hirdes JP, Maxwell CJ, et al. New evidence on benzodiazepine use and falls: the time factor. *Age Ageing* 1996;25:273–278.

75. Leveille SG, Buchner DM, Koepsell TD, et al. Psychoactive medications and injurious motor vehicle collisions involving older drivers. *Epidemiology* 1994;5:591–598.

76. Hemmelgarn B, Suissa S, Huang A, et al. Benzodiazepine use and the risk of motor vehicle crash in the elderly. *JAMA* 1997;278:27–31.

77. Kripke DF, Klauber MR, Wingard DL, et al. Mortality hazard associated with prescription hypnotics. *Biol Psychiatry* 1998;43:687–693.

78. Merlotti L, Roehrs T, Zorick F, et al. Rebound insomnia: duration of use and individual differences. *J Clin Psychopharmacol* 1991;11:368–373.

79. Roehrs TA, Zorick FJ, Wittig RM, et al. Dose determinants of rebound insomnia. *Br J Clin Pharmacol* 1986;22:143–147.

80. Hajak G, Clarenbach P, Fischer W, et al. Rebound insomnia after hypnotic withdrawal in insomniac outpatients. *Eur Arch Psychiatry Clin Neurosci* 1998;248:148–156.

81. Griffiths RR, Weerts EM. Benzodiazepine self-administration in humans and laboratory animals—implications for problems of long-term use and abuse. *Psychopharmacology* 1997;134:1–37.

82. Lader M. Anxiety or depression during withdrawal of hypnotic treatments. *J Psychosom Res* 1994;38:113–123.

83. Morin CM, Colecchi CA, Ling WD, et al. Cognitive behavior therapy to facilitate benzodiazepine discontinuation among hypnotic-dependent patients with insomnia. *Behav Ther* 1995;26:733–745.

84. Balter MB, Uhlenhuth EH. Prescribing and use of benzodiazepines: an epidemiologic perspective. *J Psychoactive Drugs* 1992;24:63–64.

85. Woods JH, Winger G. Current benzodiazepine issues. *Psychopharmacology* 1995;118:107–115, 118.

86. Romach M, Busto U, Somer G, et al. Clinical aspects of chronic use of alprazolam and lorazepam. *Am J Psychiatry* 1995;152:1161–1167.

87. Sharpley AL, Cowen PJ. Effect of pharmacologic treatments on the sleep of depressed patients. *Biol Psychiatry* 1995;37:85–98.

88. Thase ME. Depression, sleep, and antidepressants. *J Clin Psychiatry* 1998;59:55–65.

89. Buysse DJ, Reynolds CF, Hoch CC, et al. Longitudinal effects of nortriptyline on EEG sleep and the likelihood of recurrence in elderly depressed patients. *Neuropsychopharmacology* 1996;14:243–252.

90. Montgomery I, Oswald I, Morgan K, et al. Trazodone enhances sleep in subjective quality but not in objective duration. *Br J Clin Pharmacol* 1983;16:139–144.

91. Scharf MB, Sachais BA. Sleep laboratory evaluation of the effects and efficacy of trazodone in depressed insomniac patients. *J Clin Psychiatry* 1990;51:13–17.

92. Parrino L, Spaggiari MC, Boselli M, et al. Clinical and polysomnographic effects of trazodone CR in chronic insomnia associated with dysthymia. *Psychopharmacology* 1994;116:389–395.

93. Walsh JK, Erman M, Erwin CW, et al. Subjective hypnotic efficacy of trazodone and zolpidem in DSMIII-R primary insomnia. *Hum Psychopharmacol* 1998;13:191–198.

94. Hajak G, Rodenbeck A, Adler L, et al. Nocturnal melatonin secretion and sleep after doxepin administration in chronic primary insomnia. *Pharmacopsychiatry* 1996;29:187–192.

95. Hohagen F, Montero RF, Weiss E, et al. Treatment of primary insomnia with trimipramine: an alternative to benzodiazepine hypnotics? *Eur Arch Psychiatry Clin Neurosci* 1994;244:65–72.

96. Nowell PD, Reynolds CF, Buysse DJ, et al. Paroxetine in the treatment of primary insomnia: preliminary clinical and EEG sleep data. *J Clin Psychiatry* 1999;60:89–95.

97. Shipley JE, Kupfer DJ, Griffin SJ, et al. Comparison of effects of desipramine and amitriptyline on EEG sleep of depressed patients. *Psychopharmacology* 1985;85:14–22.

98. Kupfer DJ, Perel JM, Pollock BG, et al. Fluvoxamine versus desipramine: comparative polysomnographic effects. *Biol Psychiatry* 1991;29:23–40.

99. Ware JC, Brown FW, Moorad PJ, et al. Effects on sleep: a double-blind study comparing trimipramine to imipramine in depressed insomniac patients. *Sleep* 1989;12:537–549.

100. Beasley CM, Dornseif BE, Pultz JA, et al. Fluoxetine versus trazodone: efficacy and activating-sedating effects. *J Clin Psychiatry* 1991;52:294–299.

101. Armitage R, Yonkers K, Cole D, et al. A multi-center, double-blind comparison of the effects of nefazodone and fluoxetine on sleep architecture and quality of sleep in depressed outpatients. *J Clin Psychopharmacol* 1997;17:161–168.

102. Rush AJ, Armitage R, Gillin JC, et al. Comparative effects of nefazodone and fluoxetine on sleep in outpatients with major depressive disorder. *Biol Psychiatry* 1998;44:3–14.

103. Rickels K, Morris RJ, Newman H, et al. Diphenhydramine in insomniac family practice patients: a double- blind study. *J Clin Pharmacol* 1983;23:234–242.

104. Saletu B, Saletu-Zyhlarz G, Anderer P, et al. Nonorganic insomnia in generalized anxiety disorder. *Neuropsychobiology* 1997;36:130–152.

105. Mesulam M-M. Central cholinergic pathways: neuroanatomy and some behavioral implications. In: Avoli M, Reader TA, Dykes RW, et al, eds. *Neurotransmitters and cortical function*. New York: Plenum Press, 1988:237–260.

106. Meuleman JR, Nelson RC, Clark RL Jr. Evaluation of temazepam and diphenhydramine as hypnotics in a nursing-home population. *Drug Intell Clin Pharm* 1987;21:716–720.

107. Meltzer EO. Antihistamine-and decongestant-induced performance decrements. *J Occup Med* 1990;32:327–334.

108. Mendelson WB. A critical evaluation of the hypnotic efficacy of melatonin. *Sleep* 1997;20:916–919.

109. Sack RL, Hughes RJ, Edgar DM, et al. Sleep-promoting effects of melatonin: at what dose, in whom, under what conditions, and by what mechanisms? *Sleep* 1997;20:908–915.

110. Lewy AJ, Ahmed S, Jackson JM, et al. Melatonin shifts human circadian rhythms according to a phase-response curve. *Chronobiol Int* 1992;9:380–392.

111. Deacon S, Arendt J. Melatonin-induced temperature suppression and its acute phase-shifting effects correlate in a dosedependent manner in humans. *Brain Res* 1995;688:77–85.

112. Dollins AB, Zhdanova IV, Wurtman RJ, et al. Effect of inducing nocturnal serum melatonin concentrations in daytime on sleep, mood, body temperature and performance. *Proc Natl Acad Sci USA* 1994;91:1824–1828.

113. Tzischinsky O, Lavie P. Melatonin possesses time-dependent hypnotic effects. *Sleep* 1994;17:638–645.

114. Zhdanova IV, Wurtman RJ, Morabito C, et al. Effects of low oral doses of melatonin, given 2–4 hours before habitual bedtime, on sleep in normal young humans. *Sleep* 1996;19:423–431.

115. Waldhauser F, Saletu B, Trinchard-Lugan I. Sleep laboratory investigations on hypnotic properties of melatonin. *Psychopharmacology* 1990;100:222–226.

116. James SP, Sack DA, Rosenthal NE, et al. Melatonin administration in insomnia. *Neuropsychopharmacology* 1989;3:19–23.

117. Ellis CM, Lemmens G, Parkes JD. Melatonin and insomnia. *J Sleep Res* 1996;5:61–65.

118. MacFarlane JG, Cleghorn JM, Brown GM, et al. The effects of exogenous melatonin on the total sleep time and daytime alertness of chronic insomniacs: a preliminary study. *Biol Psychiatry* 1991;30:371–376.

119. Garfinkel D, Laudon M, Nof D. et al. Improvement of sleep quality in elderly people by controlled-release melatonin. *Lancet* 1995;346:541–544.

120. Haimov I, Lavie P, Laudon M, et al. Melatonin replacement therapy of elderly insomniacs. *Sleep* 1995;18:598–603.

121. Wurtman RJ, Zhdanova I. Improvement of sleep quality by melatonin. *Lancet* 1995;346:1491.

122. Lushington K, Pollard K, Lack L, et al. Daytime melatonin administration in elderly good and poor sleepers: effects on core body temperature and sleep latency. *Sleep* 1997;20:1135–1144.

123. Hughes RJ, Sack RL, Lewy AJ. The role of melatonin and circadian phase in age-related sleep-maintenance insomnia: assessment in a clinical trial of melatonin replacement. *Sleep* 1998;21:52–68.

124. Houghton PJ. The scientific basis for the reputed activity of Valerian. *J Pharm Pharmacol* 1999;51:505–512.

125. Wong AH, Smith M, Boon HS. Herbal remedies in psychiatric practice. *Arch Gen Psychiatry* 1998;55:1033–1044.

126. Leathwood PD, Chauffard F, Heck E, et al. Aqueous extract of valerian root (Valeriana officinalis L.) improves sleep quality in man. *Pharmacol Biochem Behav* 1982;17:65–71.

127. Lindahl O, Lindwall L. Double blind study of a valerian preparation. *Pharmacol Biochem Behav* 1989;32:1065–1066.

128. Schulz H, Stolz C, Muller J. The effect of valerian extract on sleep polygraphy in poor sleepers: a pilot study. *Pharmacopsychiatry* 1994;27:147–151.

129. Balderer G, Borbély AA. Effect of valerian on human sleep. *Psychopharmacology* 1985;87:406–409.

130. Porkka-Heiskanen T, Strecker RE, Thakkar M, et al. Adenosine: a mediator of the sleep-inducing effects of prolonged wakefulness. *Science* 1997;276:1265–1268.

131. Sherin JE, Shiromani PJ, McCarley RW, et al. Activation of ventrolateral preoptic neurons during sleep. *Science* 1996;271:216–219.

132. Sherin JE, Elmquist JK, Torrealba F, et al. Innervation of histaminergic tuberomamillary neurons by GABAergic and galaninergic neurons in the ventrolateral preoptic nucleus of the rat. *J Neurosci* 1998;18:4705–4721.

133. Lin L, Faraco J, Li R, et al. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* 1999;98:365–376.

134. Chemelli RM, Willie JT, Sinton CM, et al. Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. *Cell* 1999;98:437–451.

134

Sleep Disturbances Associated with Neuropsychiatric Disease

Eric A. Nofzinger

Matcheri Keshavan

Eric A. Nofzinger and Matcheri Keshavan: Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

Several observations suggest important links between sleep and mental disorders. From a psychological perspective, it has long been thought that trouble sleeping is related to a troubled mind, an intrusion of mental activity into the quiescence of sleep. Historically, this vein has fueled much of the interest in the relationship between sleep and mental disorders until the discovery of electrophysiologically defined sleep stages, particularly rapid eye movement (REM) sleep in the 1950s. Even after that time and into the 1960s and 1970s, this basic psychological tenet guided interest in uncovering relationships between REM sleep and mental disorders. Over time, interest shifted to defining the neurobiology of mental disorders. In this service, EEG sleep staging became a tool to be used in either diagnosis or validation of the biological nature of mental disorders. A second, nonpsychiatric, line of investigation during this time concerned itself with the physiology of normal sleep. This led to a significant expansion in our understanding of the brain mechanisms that govern basic sleep/wake, or behavioral state, regulation. Few attempts, however, have been made to bridge these diverging lines of investigation despite the complimentary information derived from each area. In part, this may be related to the vastly divergent levels of observation of the brain capable in preclinical and human studies. With the advent of advances in preclinical work defining the brain mechanisms underlying electrophysiologic oscillations measured at the brain surface and in human brain imaging work defining correlates of underlying functional neuroanatomy, these divergent fields are beginning to communicate in meaningful ways to enrich the discoveries made in each domain. Evidence of this evolution comes in the form of chapters in this volume devoted to neuropsychopharmacology by neuroscientists such as Allan Hobson and Emmanuel Minot and the inclusion in this clinical chapter of preclinical data that may guide interpretations of human electrophysiology and functional neuroanatomy.

As a guide, this chapter attempts to first distill aspects of preclinical work that may inform our understanding of findings in electrophysiology and functional neuroanatomy in clinical populations. We then review the subjective, EEG, and brain imaging work in the major mental disorders that help to define pathophysiology through a sleep window into brain function. Disorders that are highlighted are the major mood disorders, schizophrenia, and degenerative disorders of aging given the extensive sleep research conducted in each of these areas.

NEUROBIOLOGY OF HEALTHY SLEEP

Part of "134 - Sleep Disturbances Associated with Neuropsychiatric Disease "

Overview of the Sleep/Wake Cycle

Sleep can be conceptualized as a motivated behavior, something the organism “needs” to do in order to survive and for which there is a pressure to perform. Sleep propensity is lowest shortly after awakening, increases in mid-afternoon, plateaus across the evening, is greatest during the night and declines across sleep. Sleep has been subclassified polysomnographically into NREM and REM sleep states based on three measures: (1) electroencephalography (EEG), (2) electromyography (EMG), and (3) electrooculography (EOG). With sleep onset, the EEG frequency slows, the amplitude increases and the EMG decreases. This sleep is classified as NREM stages 1, 2, 3, or 4, which are distinguished by increasing amounts of low-frequency, high-amplitude EEG activity, also known as “delta” activity. Delta sleep decreases across the night. REM sleep follows the first NREM period and is characterized by low-amplitude, mixed high-frequency EEG, the occurrence of intermittent REMs, skeletal muscle atonia, and irregular cardiac and respiratory events. Across the night, brain function oscillates between the globally distinct states of NREM and REM sleep about three or four times, approximately every 90 minutes. Across successive sleep cycles within a night, stages 3 and 4 sleep decrease

and then disappear, whereas REM sleep and lighter NREM sleep stages increase.

Historical Development of Sleep Neurobiology

The current understanding that sleep and wakefulness are both active brain states that are generated and maintained from within the brain has its origins in the pioneering work of Berger (1930), Economo (1929), Bremer (1935), Moruzzi and Magoun (1949), and Jouvet (1962). Prior to the work of these investigators, the state of sleep was thought to represent an inactive period for the brain. Berger first described electrophysiologic patterns of brain electrical activity that correlated with changes in behavioral state. Other investigators discovered that the brain itself was responsible for generating its own intrinsic electrical activity and for switching between behavioral states. Clinicoanatomic clues came in the discovery by Economo (1929) that lesions to the posterior hypothalamus in encephalitis lethargica produced somnolence or coma, suggesting a role for the posterior hypothalamus in maintaining wakefulness. Moruzzi and Magoun defined the concept of a nonspecific ascending reticular activating system (ARAS) responsible for cortical activation following their discovery that electrical stimulation of the brainstem reticular formation suppressed high-amplitude EEG waves in the cortex. Finally, following the discovery of the behavioral state of REM sleep by Aserinsky and Kleitman (1), the French neurophysiologist Michel Jouvet (2) localized the generation of this sleep state to structures in the pontobulbar brainstem using rostropontine transections in cats. From these studies it is now recognized that behavioral state is intrinsically regulated and generated by the brain itself and manifested by electrical activity in the brain that can be recorded by EEG at the scalp.

An extensive preclinical literature has developed that describes the brain mechanisms of the distinct electrophysiologic oscillations that characterize the various behavioral states of waking, NREM, and REM sleep. Several features are important with respect to localizing the brain structures underlying slow wave sleep. First, the electrical oscillations observed at the macroscopic level are the end result of electrical oscillations involving widespread thalamocortical neurons that are synchronized in a global fashion. Second, widespread changes in these oscillations can result from state-dependent changes in modulatory systems such as the brainstem, hypothalamus, and basal forebrain. Third, slow oscillations in the 1 to 4 Hz delta range have both cortical and thalamic components. This line of research emphasizes that rhythmic oscillations are the end result, therefore, of integrated corticothalamic circuits and modulatory structures. Changes in delta sleep found in diverse neuropsychiatric disorders, therefore, may result from functional changes at one or more levels including the cortex, thalamus, and modulatory structures.

At the other end of the oscillatory spectrum are the fast oscillations in the beta and gamma frequencies (roughly > 20 Hz) that have been associated with increased vigilance and are most prevalent during waking and REM sleep (3 ,4). These fast rhythms are present in cortical and thalamic neurons and depend on neuronal depolarization characteristic of brain activation, a state in which neurons are maximally ready to respond to either external (in waking) or internal (in REM sleep) stimuli. These fast synchronous rhythms may serve a purpose of binding aspects of a stimulus into a global representation. Importantly, these fast rhythms as well as the slow rhythms are dependent on the level of excitability in local intracortical circuits as may be modulated by ascending activating modulatory systems such as the brainstem monoaminergic and cholinergic nuclei (5). Disturbances in these fast frequency oscillations in neuropsychiatric disorders may reflect pathology at one or more of these levels modulating these electrophysiologic rhythms.

Mechanisms Underlying Behavioral State Changes: Core Structures Related to Arousal

From the electrophysiologic preclinical literature, it has become apparent that the set point for electrical oscillations in widespread thalamocortical circuits can be modulated by global ascending activating influences. Identification of the structures that are important in providing this input, therefore, may provide clues as to the brain mechanisms leading to altered electrophysiologic activity in diverse neuropsychiatric disorders. Arousal and the maintenance of an aroused state is an active process requiring the integrated activity of a series of arousal systems shown diagrammatically in Fig. 134.1 . (See Robbins and Everitt, 1996 , for review and RMoore and colleagues , 2001, for a discussion of the relationship between the Orexin neurons and these arousal systems) (6 ,7).

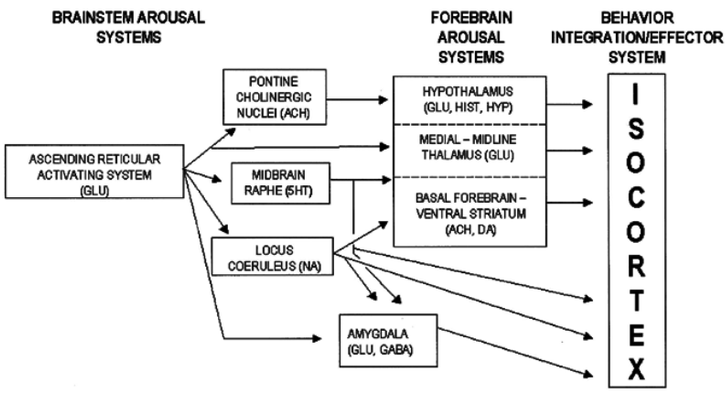


FIGURE 134.1. Arousal systems in the human brain (see text for description). The transmitters associated with each system are abbreviated. ACH, acetylcholine; DA, dopamine; GABA, γ aminobutyric acid; GLU, glutamate; HIST, histamine; HYP, hypocretin; NA, noradrenaline; 5HT, serotonin.

The central brainstem arousal system is the ascending reticular activating system (ARAS) (8 ,9). The ARAS projects into a series of specific brainstem systems including the pontine cholinergic nuclei, midbrain raphe nuclei, and the locus ceruleus, and into a series of forebrain structures involved in arousal. These include the midline and medial thalamus with widespread cortical projections and the amygdala, which has interconnections with isocortex and other areas involved in arousal, particularly hypothalamus and ventral striatum. The amygdala is particularly involved with autonomic regulation and the emotional component of arousal. There are two important components of the basal forebrain–ventral striatum system (10). One is the cholinergic neurons of the medial septum-nucleus of the diagonal band-nucleus basalis complex that innervates the entire forebrain. The second is the nucleus accumbens-ventral striatum complex that is involved in transmitting the arousing aspects of reinforcing stimuli. Until recently, the importance

of the hypothalamus was not fully recognized, but this is rapidly changing (10 ,11). First, we now appreciate that an important component of the circadian control of behavioral state is the maintenance of arousal by the circadian pacemaker, the suprachiasmatic nucleus (12). Second, there are extensive hypothalamic projections to isocortex, predominantly from posterior hypothalamus. These include a newly discovered projection from a group of neurons that produce a novel peptide hypocretin. This projection is of particular interest because the hypocretin neurons project not only over the entire isocortex but also to all of the arousal systems noted in Fig. 134.1 , including extraordinarily dense projections to locus ceruleus, raphe nuclei, pontine cholinergics, midline thalamus, nucleus basalis, and amygdala. (See ref. 7 for review.) The hypocretin projection has become of particular interest because a hypocretin gene knockout produces a narcolepsy-like syndrome in mice (13) and hypocretin is below detectable levels in CSF from narcoleptics in comparison to controls (14).

Mechanisms Underlying Behavioral State Changes: Core Structures Involved in REM Sleep

The nature of dreaming has led to extensive discussions about the relationships between REM sleep and mental disorders (9). The finding of EEG sleep staging abnormalities in REM sleep in certain mental disorders has also fueled discussions about the relationships between the biology of REM sleep and the pathophysiology of mental disorders; therefore, a brief review of brain structures that may have overlapping roles in REM sleep regulation and the pathophysiology of mental disorders may guide theoretical models of behavioral state regulation in diverse mental disorders (see Chapter 128 , Chapter 129 , Chapter 130 , Chapter 131 , Chapter 132 and Chapter 133). Evidence from a variety of approaches suggests that the laterodorsal and pedunculopontine tegmental cholinergic nuclei (LDT and PPT) in the pontine reticular formation underlie the phasic and tonic components of REM sleep (9). A reciprocal interaction hypothesis (9) claims that these cholinergic nuclei become disinhibited during the entry into REM sleep by the removal of tonic inhibition from noradrenergic and serotonergic nuclei as these monoaminergic nuclei slow or become silent in the transition from NREM to REM sleep (15). Modifications of this model now account for the influence of additional brainstem neurotransmitter systems such as GABAergic, nitroxergic, glutamatergic, glycinergic, histaminergic, adenosinergic, and dopaminergic; various peptide systems such as galanin, Orexin, vasoactive intestinal polypeptide, and nerve growth factor; and hormonal influences such as growth hormone-releasing hormone, prolactin, and corticotropin-releasing factor (9 ,11 ,16 ,17 and 18). Brainstem reticular nuclei include a dorsal pathway innervating the thalamus and a ventral pathway innervating the basal forebrain that thereby mediates widespread cortical arousal indicative of the REM sleep state. Human brain imaging studies of REM sleep show that the ventral pathway predominates during human REM sleep in activating anterior paralimbic structures (19 ,20).

Mechanisms Underlying Behavioral State Changes: Modulatory Structures

In large part, these preclinical studies have focused on the primary nuclei that may be generative centers for one behavioral state or another. Less information is available regarding structures that modulate activity in these centers and that may play a role in the abnormal modulation of behavioral states in various mental disorders (8 ,9). Recent work, for example, suggests that the amygdala has significant anatomic connections and recently established modulatory effects on the brainstem centers involved in REM sleep production (21 ,22). Similarly, other forebrain structures such as the hypothalamus (13) and basal forebrain (10 ,17) are

known to have both anatomic and functional relationships with brainstem centers thought to play a role in behavioral state regulation in addition to the primary roles they each play in cortical arousal. Although it is possible that certain disease states are associated with pathologic changes in discrete brain structures that generate discrete behavioral states, it is more likely (e.g., in the case of depression) that there are pathologic functions in brain structures that have modulating effects on these core brain structures in producing the behavioral state changes.

Mechanisms Underlying Behavioral State Changes: Evidence from Human Sleep Imaging Studies

The advent of brain imaging methods that can be used to study the functional neuroanatomy of human sleep has recently created a venue for linking preclinical work with human sleep electrophysiology. Findings from these new studies have enriched our understanding of the brain structures that are preferentially active across behavioral states and that may play a role in behavioral state regulation as well as functional roles for discrete behavioral states.

Global Changes in Brain Function across Behavioral States

In terms of global, or whole brain changes, from waking to NREM sleep, there are reductions in measures of cortical blood flow or metabolism (19). During REM sleep, global blood flow or metabolism ranges from 10% below to 41% above levels obtained during wakefulness (23 ,24). A recent report cited a positive correlation between waking global and regional cerebral blood flow and slow wave sleep measures from the subsequent night of sleep (25). The authors interpreted these findings as reflecting an energy conservation or restorative role for slow wave sleep.

Relative Regional Changes during NREM Sleep

In terms of regional relative changes during NREM sleep, blood flow has been shown to negatively correlate with the presence of NREM sleep in the anterior cingulate (19 ,26 ,27), pontine reticular formation (19 ,26 ,27), thalamus (19 ,26 ,27 and 28), basal forebrain/hypothalamus (19 ,26), amygdala (26), and orbitofrontal cortex (19 ,26 ,27). These changes are consistent with preclinical studies showing reductions in brainstem, basal forebrain, and hypothalamus sources of ascending activation. Declining function in the amygdala suggests the possibility that this structure modulates activity in ascending activating structures.

Relative Regional Changes during REM Sleep

In terms of regional relative changes during REM sleep, this state has been reliably associated with the selective activation of limbic and paralimbic structures including the amygdala, ventral striatum, anterior cingulate, and medial prefrontal cortex (19 ,20 ,23 ,29 ,30). This pattern of activation is superimposed on brainstem activation known to play a role in REM sleep generation on the basis of preclinical work.

Functional Brain Changes Associated with Sleep Deprivation

Several studies have discovered that sleep deprivation is associated with a global reduction in metabolism with some preference for the prefrontal cortex. This reduction is magnified with successive nights of sleep deprivation. These changes have recently been shown to play a role in the cognitive alterations associated with sleep deprivation (31).

THEORETICAL MODELS OF THE FUNCTION OF SLEEP

Part of "134 - Sleep Disturbances Associated with Neuropsychiatric Disease "

Homeostatic Function

Several models of sleep/wake regulation attempt to define parameters that may influence the probability at any point in the day that sleep may occur. One such model is the two-process model described by Borbely (32). One process, called process S, describes a homeostatic sleep process. Process S is thought to be dependent on the amount of prior wakefulness and is reflected by the amount of EEG slow wave (.5 to 4.5 Hz EEG) activity. As sleep deprivation increases, for example, process S increases and amplifies sleep propensity. The second process, process C, varies throughout the day in relation to a sinusoidal circadian phase across a 24-hour day. The intensity of this sleep propensity is unrelated to the amount of prior wakefulness. The sleep parameter most affected by this process is process C; nonsleep correlates of this process include core body temperature, plasma melatonin, and plasma cortisol levels. The regulatory structure for this process is the suprachiasmatic nucleus.

Sleep and Neuronal Plasticity

Several theorists have conceptualized distinct roles for the diverse sleep/wake states based on emerging knowledge regarding forebrain function in information processing. Buzsaki (33) emphasized a two-stage model of waking hippocampal memory trace formation and suggested parallels for two stages of memory processing during NREM and REM sleep. Karni and associates and Wilson and McNaughton have demonstrated direct evidence in support of memory processing during sleep for REM and NREM sleep, respectively.

Although the general states of neuronal activation have been described for NREM and REM sleep, little is known within these states about the relevance of this neuronal activation.

Llinas and Pare (34) suggest that wakefulness and REM sleep are fundamentally equivalent states of activation characterized by intrinsic oscillatory thalamo-cortical loops differing only in the degree to which external stimuli are capable of modulating the global brain state. They suggest that these oscillations serve the purpose of generating an internal representation of the world guided by innate predispositions of the brain to categorize and integrate the sensory world in certain ways. During REM sleep, these innate templates, which have been molded by experience, may be used to recreate world-analogues. During wakefulness, these templates are modulated by sensory events. Winson (35) suggested that, in humans, psychological experience essential to survival is integrated and further consolidated during REM sleep. He based this hypothesis on: (a) the presence of hippocampal theta during REM sleep; (b) the sole occurrence of hippocampal theta during species-specific survival-dependent behavior in nonprimates; and (c) the role of hippocampal theta in the induction of long-term potentiation (LTP), a model for the synaptic modulation underlying certain types of memory. Kavanau (36), reviewing data on synaptic events related to the formation of enhanced synaptic efficacy in the context of evolutionary biology, suggests that REM sleep may perform the service of repetitive activations of synapses in neural circuits that underlie essential adaptive behaviors. This “dynamic stabilization” is thought to ensure the efficacy of circuits that otherwise may suffer from a disuse atrophy. Although clearly these hypotheses require validation, the underlying premises are consistent with psychological theories that REM sleep and dreaming may play a role in affective adaptation; in the integration of recent, remote, and perhaps phyletic memory; maintaining or facilitating behavioral repertoires underlying crucial behavior; and the genetic programming of inherited behavior.

SLEEP IN MAJOR NEUROPSYCHIATRIC DISEASES

Part of "134 - Sleep Disturbances Associated with Neuropsychiatric Disease "

We turn now to characterizing the sleep disturbances in the major mental disorders. The vast majority of relevant data come from clinical and EEG sleep reports. Early findings from functional brain imaging studies are also reviewed.

DEPRESSION

Part of "134 - Sleep Disturbances Associated with Neuropsychiatric Disease "

Subjective Findings

The majority of patients with mood disorders describe difficulty falling asleep, staying asleep, and returning to sleep after early morning awakenings. Clinically, they report a paradoxic state of physical daytime fatigue, yet with persistent mental activity that makes it difficult for them to fall asleep at night. Subjective sleep quality has been measured with a validated instrument called the Pittsburgh Sleep Quality Index (PSQI) (37). In one study, subjective sleep quality was rated worse by patients with major depression than by patients presenting with a chief complaint of insomnia or other sleep disorder patients. Whereas insomnia characterizes the melancholia of middle age and elderly unipolar depression, younger patients and bipolar depressed patients often describe, atypically, difficulty getting up in the morning and hypersomnia during the daytime. This subjective hypersomnia, however, does not translate into an increased physiologic tendency to fall asleep when measured objectively by the Multiple Sleep Latency Test (MSLT).

EEG Findings

An extensive literature describes the changes in EEG sleep in patients with depression (38 ,39). Measures derived from the EEG sleep recordings that have been found to differ between healthy and depressed subjects include measures of sleep continuity, measures of visually scored EEG sleep stages, and automated measures of characteristics of the EEG waveform across the sleep period such as period amplitude or EEG spectral power measures.

The changes in subjective sleep complaints are paralleled by EEG measures of sleep. These include increases in sleep latency and decreases in sleep continuity. In terms of EEG sleep stages or “sleep architecture,” depressed patients often show reduced stage 3 and 4 NREM sleep (also known as “slow wave sleep” because of the presence of slow EEG delta activity during these stages). Several changes in REM sleep also have been noted. These include an increase in the amount of REM sleep, shortening of the time to onset of the first REM period of the night, shortened REM latency, and increase in the frequency of eye movements within a REM period.

In terms of quantitative EEG changes in sleep, many (40 ,41) but not all studies have reported reductions in the amplitude or a reduction in the number of low frequency (0 to 4 Hz) delta waves during sleep in depressed patients. Increased high frequency EEG activity has also been reported in depressed patients, including alpha (40) and beta. Importantly, sex differences have been found in these abnormalities. Depressed women appear to have relative preservation of delta sleep in relation to depressed men, despite elevations in higher frequency EEG activity in both groups (42).

Several observations regarding sleep disruption in depressed patients suggest that there may be a timing abnormality in the evolution of sleep across the night in depressed patients (38 ,41). A number of factors can influence the sleep EEG findings characteristic of depression. For instance, sleep continuity deteriorates, slow wave sleep decreases, and REM latency shortens with age, even in healthy subjects; however, age-related changes are more pronounced in patients

with depression (43). By contrast, sleep EEG measures are generally less abnormal in adolescents and prepubertal children with depression, and only appear consistently in those adolescents who are hospitalized and/or suicidal (44).

Other studies have shown that patients with psychotic depression have particularly severe EEG sleep disturbances and very short REM sleep latencies; patients with recurrent depression have more severe REM sleep disturbances than patients in their first episode; and sleep continuity and REM sleep disturbances are more prominent early in the depressive episode than later (45). Some studies suggest that patients with dysthymia and mania (surprisingly) have EEG sleep disturbances very similar to those observed in major depression.

Stressful life events also interact with EEG sleep. For instance, individuals who have severe stressful events preceding the onset of depression are less likely to have reduced REM latency than patients without such a stressor. Among older depressed patients, poor sleep is associated with shorter episode duration, older age, greater medical burden, the presence of life stressors, and a lower level of perceived social support (45).

EEG sleep findings help to inform our understanding of the neurobiology of longitudinal course and treatment outcome in depression. Although severely reduced REM latencies, phasic REM measures, and sleep continuity disturbances generally move toward control values after remission of depression, most sleep measures show high correlations across the course of an episode. Reduced REM latency is associated with increased response rates to pharmacotherapy (46) but not psychotherapy. Depressed patients with abnormal sleep profiles (reduced REM latency, increased REM density, and poor sleep continuity) are significantly less likely to respond to cognitive behavior therapy and interpersonal therapy than patients with a “normal” profile. Other studies have indicated that reduced REM latency and decreased delta EEG activity are associated with increased likelihood or decreased time until recurrence of depression in patients treated with medications or psychotherapy (47).

Sleep Neuroendocrine Findings

Cortisol secretion has been linked with the circadian cycle and growth hormone with slow wave sleep processes. In depressed subjects, studies have shown increases in cortisol secretion rates, a flattening of the circadian rhythms in cortisol, and elevated cortisol nadir. Secretion of growth hormone on the other hand, is reduced in the first half of the night in depressed subjects both in the acute phase and following remission from depression. Evidence such as this has been used to support hypotheses that a balance between the oppositional actions of CRF and GHRH may be shifted toward CRF in major depression. Given the activating qualities of CRF and its direct or indirect inhibition of GHRH, sleep would be shifted toward states of cortical activation, that is, either waking or REM sleep, and away from NREM sleep, which has been associated with the homeostatic function of sleep.

Sleep Neuropharmacologic Findings

Each of the major neurotransmitter systems shown to modulate the ascending activation of the cortex, that is, the cholinergic, noradrenergic, and serotonergic systems, have been implicated in the pathophysiology of mood disorders. The role of additional brainstem neurotransmitter systems such as GABAergic, nitroxergic, glutamatergic, glycinergic, histaminergic, adenosinergic, dopaminergic, and various peptide systems such as galanin, Orexin, vasoactive intestinal polypeptide, and nerve growth factor in the sleep disturbances in depression remain to be defined. Nearly all effective antidepressant medications show a pronounced inhibition of REM sleep including a prolongation of the first REM cycle and a reduction in the overall percent of REM sleep. (Exceptions include nefazodone [48] and bupropion, which do not suppress REM sleep [49].) Enhanced cholinergic function concurrent with reduced monoaminergic tone in the central nervous system has been proposed as a pharmacologic model for depression. In an exaggerated sense, the state of REM sleep mimics this formulation, that is, a cholinergically driven state with reduced firing of noradrenergic and serotonergic neurons. Cholinergic agents such as the muscarinic agonist RS 86, arecoline, physostigmine, and scopolamine produce exaggerated REM sleep effects in depressed patients in comparison with patients with eating disorders, personality disorders, anxiety disorders, and healthy controls (50). These studies suggest that there may be a supersensitivity of the cholinergic system driving REM sleep in mood disorders patients, although an alternative plausible hypothesis is that there may be reduced monoaminergic (5-HT and/or NE) inhibition of the brainstem cholinergic nuclei in mood disorders patients. Cholinergic activation may also play a role in the hyperactivity in the HPA axis and in the blunting of growth hormone secretion noted in depressed patients across the night, given the influence of cholinergic drugs on HPA activity and GH release.

Selective serotonin reuptake inhibitors are known to have prominent REM suppressing activity, most notably early in the night when enhances in REM sleep are most often seen in mood disorders patients (39). A tryptophan-free diet, which depletes central serotonin activity, is noted to decrease REM latency in healthy controls and in depressed patients (51) and ipsapirone, a 5-HT1a agonist, is noted to prolong REM latency in both normal controls and in depressed patients (52). Anatomically, 5-HT1a receptors have been conceptualized as the limbic receptors given their high densities in the hippocampus, septum, amygdala, and cortical paralimbic structures. The action in these structures has been shown to be largely inhibitory (hyperpolarizing).

Given the importance of limbic and paralimbic structures in REM sleep modulation, the influence of SSRI medications may be mediated by these limbic receptors. Importantly, in the brainstem LDT, a locus of cholinergic cells identified in the generation of REM sleep, bursting cholinergic neurons are inhibited by the action of 5-HT on 5-HT1a receptors. Finally, the effects of the 5-HT1a-antagonist pindolol on EEG sleep in healthy subjects was studied and noted to reduce REM sleep. This was interpreted as supportive of a reduction in raphe serotonergic autoregulation, resulting in increased serotonergic input to pontine cholinergic centers and inhibiting REM sleep.

Functional Neuroimaging Findings

Given the selective activation of limbic and paralimbic structures during REM sleep in healthy subjects, the study of the functional neuroanatomy during REM sleep in depressed patients

may provide clues as to alterations in limbic and paralimbic function related to the pathophysiology of depression. In contrast to healthy controls (4), depressed patients fail to activate anterior paralimbic structures (subgenual and pregenual anterior cingulate and medial prefrontal cortices) from waking to REM sleep. In contrast to healthy controls, depressed subjects show large activations in the dorsal tectum (superior colliculus and periaqueductal gray) during REM sleep. Finally, in contrast to controls, depressed subjects activate left sensorimotor cortex, left inferior temporal cortex, left uncal gyrus and amygdala, and left subicular complex during REM sleep. These findings suggest that depressed patients demonstrate uniquely different patterns of activation from waking to REM sleep than do healthy controls. In the context of neuroscience models relating forebrain function during REM sleep to attention, motivation, emotion, and memory, these results suggest that prior REM sleep abnormalities in mood disorders patients, therefore, likely reflect alterations in limbic and paralimbic forebrain function related to depression (Fig. 134.2).

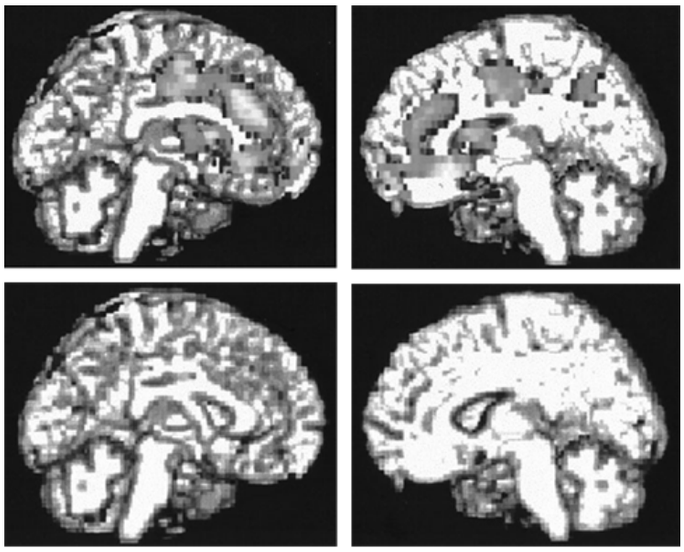


FIGURE 134.2. Healthy and depressed subjects' changes in relative glucose metabolism from waking to rapid eye movement (REM) sleep. The top two images are midsagittal sections showing areas of the brain that have greater relative glucose metabolism during REM sleep than during waking. These areas include anteriorly located paralimbic structures such as the ventral striatum, anterior cingulate cortex, and medial prefrontal cortex. The lower two images show the same comparisons, but in depressed patients. Of note is the absence of an increase in relative metabolism in any of these anteriorly located paralimbic structures in depressed patients from waking to REM sleep. See color version of figure.

Functional neuroimaging of NREM sleep in depressed subjects would be expected to provide evidence regarding the functioning of homeostatic mechanisms in mood disorders patients since this is a time of nonselective nonactivation of the cortex in which the buildup of a sleep dependent process, process S, is discharged and during which growth hormone secretion occurs. Ho and Gillin and colleagues (53) demonstrated that whole brain and regional cerebral glucose metabolism was elevated during the first NREM period of the night for depressed men in relation to healthy men. These findings are supportive of a deficiency of homeostatic mechanisms in mood disorders patients, perhaps secondary to cortical hyperarousal. Clark and associates (25) reported that reductions in delta sleep in depressed patients were associated with reductions in afternoon waking relative and global blood flow. This suggests that the elevations in glucose metabolism during NREM sleep in depressed patients are not related to a waking hypermetabolic state. Studies across waking and NREM sleep are needed in order to clarify this notion.

Nofzinger and associates (54) sought to clarify the neurobiological basis of variations in one aspect of central nervous system "arousal" in depression by characterizing the functional neuroanatomic correlates of beta EEG power density during NREM sleep. First, nine healthy (*N* = 9) subjects underwent concurrent EEG sleep studies and [^{18}F]2-fluoro-2-deoxy-D-glucose ([^{18}F]FDG) positron emission tomography (PET) scans during their first NREM period of sleep in order to generate hypotheses about specific brain structures that show a relationship between increased beta power and increased relative glucose metabolism. Second, brain structures identified in the healthy subjects were then used as *a priori* regions of interest in similar analyses from identical studies in 12 depressed subjects. Statistical parametric mapping was used to identify the relationship between beta power and relative regional cerebral glucose metabolism (rCMRglu) during NREM sleep. Regions that demonstrated significant correlations between beta power and relative cerebral glucose metabolism in both the healthy and depressed subjects included the ventromedial prefrontal cortex and the right lateral inferior occipital cortex. During a baseline night of sleep, depressed patients demonstrated a trend toward greater beta power in relation to a separate age- and gender-matched healthy control group. In both healthy and depressed subjects, beta power negatively correlated with subjective sleep quality. Finally, in the depressed group, there was a trend for beta power to correlate with an indirect measure of absolute whole brain metabolism during NREM sleep. This study demonstrated a similar relationship between electrophysiologic arousal and glucose metabolism in the ventromedial prefrontal cortex in depressed and healthy subjects. Given the increased electrophysiologic arousal in some depressed patients and the known anatomic relations between the ventromedial prefrontal cortex and brain activating structures, this study raises the possibility that the ventromedial prefrontal cortex plays a significant role in mediating one aspect of dysfunctional arousal found in more severely aroused depressed patients.

Wu and associates (55) characterized the functional neuroanatomic changes following sleep deprivation therapy in depressed patients. They found that depressed patients who demonstrated high pretreatment relative glucose metabolic rates in the medial prefrontal cortex were more likely to respond to sleep deprivation. Further, a reduction in relative metabolism in this region was found following sleep deprivation.

Depression and Alcoholism

Significant comorbidity exists between depression and alcoholism. EEG sleep studies have been used as a tool to explore the relationships between the neurobiology of the two disorders. Clark and colleagues (56) found that alcoholism played a more significant role in sleep disruption than depression. Increased REM density in primary alcoholics with and without a lifetime diagnosis of secondary depression predicted 3-month alcohol-related relapse rates. Drummond and co-workers (57) examined the relationships among sleep, natural course, and relapse in abstinent pure primary alcoholic patients. They found that the sleep was short, fragmented, and shallow early in abstinence with some incomplete improvements over the first year of abstinence. Further, the presence of sleep disruption at 5 months was shown to predict relapse by 14 months.

SCHIZOPHRENIA

Part of "134 - Sleep Disturbances Associated with Neuropsychiatric Disease "

REM Sleep

Early sleep EEG studies sought to test the intriguing hypothesis that schizophrenia is a spillover of the dream state into wakefulness. No evidence has accrued to support this prediction; however, subtle alterations in architecture of REM sleep may occur. REM latency was found decreased in seven of the 10 studies (58). It has been proposed that this may result from a deficit in SWS in the first NREM period leading to a passive advance, or early onset of the first REM period. An alternative explanation is "REM pressure." However, studies of the amounts of REM sleep have been conflicting with increases, decreases, as well as no change being found (58 ,59). Studies examining treatment naive schizophrenia patients show no increases in REM sleep (60 ,61); the increases in REM sleep observed in previously treated subjects may reflect effects of medication withdrawal and/or changes related to the acute psychotic state (61). It is unlikely that the observed decreases in REM latency in some schizophrenia patients result from primary abnormalities in REM sleep.

Slow Wave Sleep

Slow wave sleep is of particular interest to schizophrenia because of the implication of the prefrontal cortex in this disorder (62) and in generation of SWS (63). Several studies have shown a reduction of SWS in schizophrenic patients; SWS deficits have been seen in acute, chronic, as well as remitted states; and in never-medicated, neuroleptic-treated, as well as unmedicated patients (58); however, not all studies show these deficits. Studies that fail to find differences in SWS have generally used conventional visual scoring. Three studies that have quantified sleep EEG parameters have consistently shown reductions in SWS. Ganguli

and associates (60) observed no change in visually scored SWS, but instead, a significant reduction in delta wave counts in drug-naive schizophrenic patients, suggesting that visually scored SWS may not be sensitive enough, and automated counts may be a better marker of SWS deficiency in schizophrenia. Other groups have described similar reductions in delta counts. SWS deficits have been recently demonstrated in early course schizophrenia using sensitive approaches such as spectral analysis (64).

Sleep Deprivation Studies

Sleep deprivation provides a naturalistic, physiologic challenge for dynamic manipulation of sleep processes, and can help clarify the primary nature of sleep abnormalities. An intriguing reduction in REM rebound following REM sleep deprivation has been described in several studies in acute schizophrenia, but a normal or exaggerated REM rebound in remitted schizophrenic patients (58). This rebound failure in acute schizophrenia has been attributed to a possible “leakage” of phasic REM events from REM sleep into NREM sleep, although no systematic investigation has supported this hypothesis.

Sleep deprivation studies can also help clarify whether SWS abnormalities in schizophrenia are secondary to pathology in neuronal circuits in this disorder, or whether they reflect primary homeostatic disruption in sleep processes. There is evidence (65) that following total sleep deprivation, recovery of Stage 4 sleep is diminished in schizophrenia. SWS deprivation is known to consistently cause impaired attention, prolonged reaction time, verbal learning, and vigilance similar to what is seen in frontal lobe dysfunction and schizophrenia (66). A defect in SWS recovery might be consistent with impairments in critical cognitive processes such as psychomotor vigilance observed in schizophrenia. Such a defect might also suggest impairment in trophotropic or restorative processes in schizophrenia.

Relation between Sleep Abnormalities and Clinical Measures in Schizophrenia

The neurobiological correlates of the psychopathological dimensions are critical for our understanding of the pathophysiology of psychiatric disorders. Research during the past decade has focused increasingly on the positive and negative syndromes, a conceptual distinction of particular importance to pathophysiology of schizophrenia. A number of studies have examined the association between REM sleep parameters and clinical parameters. Tandon and associates (61) reported an inverse association between REM latency and negative symptoms. No association has been seen between sleep abnormalities and depressive symptoms (61), but two studies have shown that increased REM sleep may correlate with suicidal behavior in schizophrenia (67 ,68).

It is important to understand the longitudinal nature of sleep abnormalities in schizophrenia, in order to elucidate their significance for pathophysiology. Stage 4 does not appear to improve while other sleep stages change following 3 to 4 weeks of conventional antipsychotic treatment (69). In a longitudinal study, alterations of SWS appeared to be stable when polysomnographic studies were repeated at 1 year, but the REM sleep parameters appeared to change. These observations suggest that SWS deficits in schizophrenia might be trait-related (70). Consistent with this view, delta sleep abnormalities have been found to correlate with negative symptoms (60) and with impaired outcome at 1 and at 2 years (71).

An association between attentional impairment and SWS deficits also has been reported in early studies (72). The thalamus plays a crucial role in attention and gating of information because it is the major relay station receiving input from the reticular activating system and limbic and cortical association areas. A defect in this structure could explain much of the psychopathology of schizophrenia and alterations in SWS.

Relation Between Sleep Findings and Neurobiology in Schizophrenia

Some clues to the possible pathophysiologic significance of SWS deficits in schizophrenia may derive from the studies of the ontogeny of sleep during normal adolescence, in the context of a neurodevelopmental framework for schizophrenia. Converging evidence suggests a substantial reorganization of human brain function during adolescence; a marked decline in synaptic density is seen in the postmortem prefrontal cortex during adolescence. Pronounced reductions in cortical gray matter volume and regional cerebral metabolism also have been seen during adolescence. Polysomnographic studies show major changes in EEG sleep patterns in adolescence; the amounts of SWS decrease (73) across the age span from childhood to late adolescence. The time courses for maturational changes in SWS, cortical metabolic rate, and synaptic density during the postnatal phase of human development are strikingly similar; Feinberg (74) suggested that cortical neurons, after birth, go through a phase of initial overproduction and subsequent regression (pruning) of neural elements. Thus, the maturational processes in sleep EEG, cortical synaptic density and regional cerebral metabolism might reflect a common underlying biological change (i.e., a large-scale synaptic elimination).

It is instructive to examine the polysomnographic abnormalities in schizophrenia in relation to the brain maturational parameters discussed in the preceding. In addition to SWS deficits, consistent alterations in structure and function of cortical and subcortical brain regions have been observed in schizophrenia. Cross sectional studies of the correlations between such alterations and sleep are valuable to better understand the pathophysiologic substrate of schizophrenia.

Altered Neuroanatomy

Reductions in cortical gray matter are seen in schizophrenia, perhaps more prominently in the frontal and temporal cortex as well as in thalamic volume (75). SWS generation appears to be regulated by a complex neural system involving the anterior brain regions and the thalamus. The relationship between alterations in these brain structures and SWS, therefore, is interesting. SWS tends to be inversely correlated with anterior horn ratio, a measure of frontal lobe size (76), and positively correlated with lateral ventricular volume (77). These observations may reflect reductions in subcortical structures such as the thalamus, which forms a substantial part of the ventricular boundaries. A correlation between thalamic volume and SWS deficits may be predicted, and is consistent with the former's critical role in generation of SWS.

Altered Neurochemistry

What are the neurochemical mechanisms involved in SWS deficits? SWS may result from several neurochemical processes that produce neural inhibition, functional differentiation, and EEG synchrony. Activation of the cholinergic system facilitates arousal and hastens REM sleep. Cholinergic hyperfunction postulated to underlie schizophrenia, therefore, could account for SWS and REM latency reduction in schizophrenia sleep (78). Interestingly, schizophrenic patients show supersensitive REM sleep induction with the cholinergic agonist R5 86 (79), suggesting cholinergic hyperfunction. Nicotinic receptors may also be involved; sensory gating deficits as evidenced by P50 event-related potentials, which are possibly related to central nicotinergic system alterations, are reversed following sleep in schizophrenia (80). Alternatively, serotonergic abnormalities may also be involved, as indicated by an inverse correlation between serotonin metabolites in the CSF and SWS in schizophrenia (81). Disturbances in catecholaminergic mechanisms may also underlie SWS deficits in schizophrenia. Norepinephrine, which is presumed to be hyperfunctional in schizophrenia, is inhibitory to REM; therefore, it may be argued that cholinergic and monoaminergic abnormalities mediate the constellation of reduced REM latency and SWS deficit without increases in REM sleep amounts in schizophrenia (78).

The possible relation between hormonal substances and delta sleep has also received some attention. Van Kammen and associates (82) showed an association between reduction in delta sleep induction peptide, a putative endogenous sleep modulator, and SWS decrements. Adenosine, an amino acid neuromodulator has drawn increasing interest in recent years as a possible endogenous sleep-promoting agent, as it tends to accumulate during waking hours (83). Adenosine agonists have been proposed as possible therapeutic agents in schizophrenia (84).

Altered Physiology

Some evidence, albeit modest, for decreased frontal lobe metabolism has been documented in schizophrenia using a variety of techniques, including PET, single photon emission computed tomography (SPECT), 31P MRS, and xenon 133 inhalation technique. It may be instructive to examine SWS deficits in the context of such physiologic alterations. An association has been demonstrated between SWS deficits and reduced frontal lobe membrane phospholipid metabolism as examined by 31P MRS (85).

Decreased synaptic density, postulated to underlie schizophrenia, could result in reduced SWS by decreased membrane surface (fewer dendrites per neuron), causing a smaller voltage response to the synchronizing stimulus. Recent studies in cats using single cell recordings (86) have shown that slower (<1 Hz) synchronized oscillations originate mainly in the neocortex, whereas delta waves (1 to 4 Hz) arise primarily from activity of thalamocortical neurons. A finer analysis of these oscillations may clarify the nature of pathophysiology in schizophrenia. Preliminary analysis of this question using period amplitude analyses suggested more prominent deficits in the <1 Hz range in schizophrenia, pointing to a thalamocortical dysfunction (64). This finding deserves further study and replication.

Effects of Antipsychotic Drugs on Sleep

Studies of the acute effects of neuroleptics have consistently shown improvements in sleep continuity, as reflected by reduced sleep latencies, improved sleep time, greater sleep efficiency, and prolongation of REM latency (58); however, changes in SWS have been less consistent. Studies that have examined the sedative effect of conventional neuroleptics have reported either no effects or modest increases in SWS. A recent study showed a robust increase in SWS with olanzapine following acute administration (87). On the other hand clozapine increases stage 2 sleep, but may actually decrease stage 4 sleep (88). These studies have frequently used small sample sizes; few studies have examined sleep variables in relation to acute versus long-term treatment with neuroleptics in a longitudinal design.

Attempts to examine polysomnographic characteristics of schizophrenia have to consider potential effects of neuroleptic discontinuation on sleep EEG. Neylan and associates (89) reported significant worsening of REM and non-REM sleep in a series of schizophrenic patients undergoing controlled neuroleptic discontinuation. Patients experiencing relapse has larger impairments in sleep. The effects of neuroleptic discontinuation continued to worsen from 2 to 4 weeks of a neuroleptic-free condition, and did not correlate with clinical change (90). These findings highlight the importance of controlling for medication state in investigation of EEG sleep in schizophrenia.

Future Directions

New knowledge on brain mechanisms of sleep is likely to open new avenues to investigate the pathophysiology of schizophrenia. First, functional brain imaging studies suggest distinct patterns of regional brain activation in SWS and REM sleep; such studies can provide clues to the pathophysiology of schizophrenia, especially used in conjunction with physiologic perturbation paradigms such as sleep deprivation (31). Second, sleep architecture changes dramatically during development; sleep studies during development in health and disease can shed considerable light on developmentally mediated neuropsychiatric disorders (91). Finally, sleep changes are often the earliest signs of disturbance, and may even represent trait-related vulnerability markers for psychiatric disorders; sleep studies of individuals at risk for schizophrenia are likely to be fruitful (92).

GERIATRIC DISORDERS

Part of "134 - Sleep Disturbances Associated with Neuropsychiatric Disease "

Late Life Depression

Assessments of the timing and quality of NREM and REM sleep cycles through EEG sleep studies have proven to be particularly useful indicators of homeostatic and adaptive physiologic processes during successful and pathologic aging in humans. In a study of healthy “old old” and “young old,” Reynolds and colleagues (93) reported: (a) a small age-dependent decrease in slow-wave sleep, in contrast to the stability of REM sleep measures from “young old” to “old old”; (b) much better preservation of slow wave sleep among aging women than men, particularly in the first NREM period of the night, but no sex-related differences in REM sleep measures; (c) greater stability of sleep maintenance among aging men than women; and (d) longer REM sleep latencies among aging women than men. In comparison with 20 year olds, 80 year olds have significant reductions in both REM sleep percent and latency, as well as a slower recovery from the effects of acute sleep loss. Elderly people show greater rigidity in sleep patterns, with less intersubject and intrasubject variability in habitual sleep times compared to the young (94); therefore, although some loss of REM and slow wave NREM sleep characterize the aging process, successful aging is associated with a relative stability of sleep states over the later years. One explanation for these findings may be some mild losses in both homeostatic as well as adaptive physiologic mechanisms that result in NREM and REM sleep decrements, respectively. Alternatively, information processing related to both homeostatic and adaptive behavior may be more stable and efficient in the elderly, especially in those for whom the aging process has been bridged successfully.

EEG sleep studies have also provided insights into the pathophysiology of disorders in which affective adaptation is significantly stressed in late-life, that is, in response to significant losses or in acute depressive episodes. In a study of elderly volunteers who had lost a spouse, Reynolds and colleagues were able to distinguish EEG sleep changes discriminating subjects who did from those who did not develop an episode of major depression. Bereaved subjects with depression had significantly lower sleep efficiency, more early morning awakening, shorter REM latency, greater REM percent, and lower rates of delta-wave generation in the first NREM period, as compared to nondepressed bereaved volunteers. These findings are similar to those of elderly patients with recurrent unipolar depression. In a subsequent longitudinal study of bereaved elders who did not become clinically depressed, only increases in phasic REM activity and density (compared to normal controls) were observed throughout the first 2 years of bereavement (95). These findings are similar to those of Cartwright for depressed versus nondepressed divorcing women, suggesting that short REM latency and slow wave sleep are correlates of depression during stressful life events, whereas increased REM activity may correlate with successful recovery from stress in the absence of major depression. Given the selective activation of limbic and paralimbic structures during REM sleep, that is, structures related to affective adaptation, the increase in REM activity in the successful resolution of stressful events may relate to the activation of brain structures mediating adaptation to these stressful events. A failure to terminate this normal adaptive response, similar to the failure of counterregulatory neuroendocrine events in the generalized stress response, appears to result in a shift of central neurophysiologic events favoring cortical activation. This results in shifts away from homeostatic NREM sleep states, and toward more activated waking and REM sleep cortical states. In the EEG these are manifested by reductions in intrinsic, nonactivated thalamocortical synchronous EEG waves of NREM sleep, and an early entry into desynchronized REM sleep. Subjectively, this may relate to depressed patients complaints of both increased fatigue, or anergy (loss of homeostatic mechanisms) and increased internal “arousal” (shifting toward a more continuous on-line state of readiness to process salient stimuli).

Alzheimer’s Disease

Disturbances in sleep commonly accompany Alzheimer’s disease (AD) (96). These disturbances are a significant cause of distress for caregivers, often leading to institutionalization of these patients (97). The changes in sleep often parallel the changes in cognitive function in demented patients (98). Also, daytime agitation has been associated with sleep quality at night (99). A large-scale community based study of AD patients reported that sleeping more than usual and early morning awakenings were the most common sleep disturbances in noninstitutionalized patients. Nighttime awakenings, however, were more disturbing to caregivers.

Nighttime awakenings were associated with male gender, and greater memory and functional declines. Three groups of subjects were identified in association with nocturnal awakenings: (a) patients with only daytime inactivity; (b) patients with fearfulness, fidgeting, and occasional sadness; and (c) patients with multiple behavioral problems including frequent episodes of sadness, fearfulness, inactivity, fidgeting, and hallucinations.

In terms of sleep laboratory-based evaluations, sleep continuity disturbances in these patients include decreased sleep efficiencies, increased lighter stage 1 NREM sleep, and an increased frequency of arousals and awakenings. Sleep architecture abnormalities include decreases in stages 3 and 4 NREM sleep and some reports of decreases in REM sleep (100). Loss of sleep spindling and K complexes have also been noted in dementia. Sleep apnea has been observed in 33% to 53% of patients with probable AD (101). It is unclear if there is an increased prevalence of sleep apnea, however, in AD patients in relation to age- and gender-matched controls. Nocturnal behavioral disruptions, or “sundowning” are reported commonly in the clinical management of AD patients, although specific diagnostic criteria for a “sundowning” episode have been difficult to define (96 ,102). Despite extensive clinical research in this area, the pathophysiology of sundowning, including its relationship with brain mechanisms that control sleep/wake and circadian regulation remain unclear. Overall, the literature on sleep in AD suggest that the primary defect in this disease is the more general neurodegenerative changes that lead to the profound cognitive and functional declines of this disease and that the sleep changes are secondary manifestations of the disorder. If sleep is viewed as generated by core sleep systems that then require an intact neural structure throughout the rest of the brain for expression of behavioral states, then the sleep changes in AD are most likely related to end-organ failure as opposed to pathology in key sleep or circadian systems themselves.

Parkinson’s Disease

Light, fragmented sleep occurs frequently in Parkinson’s disease (PD) patients. Sleep problems have been reported in as many as 74% to 96% of patients (103). Complaints include frequent awakenings, early awakening, nocturnal craps, pains, nightmares, vivid dreams, visual hallucinations, vocalizations, somnambulisms impaired motor function during sleep, myoclonic jerks, excessive daytime sleepiness, REM sleep behavior disorder, and sleep-related violence leading to injury (103 ,104 ,105 and 106). These changes may result from the disease itself, or to complication from treatment with dopaminergic agents (103 ,104 and 105 ,107). Additionally, depression is common in PD and the sleep disruption may in part be related to this comorbid disorder (103).

Sleep architecture abnormalities include increased awakenings, reductions in stages 3 and 4 sleep, REM sleep, and sleep spindles (107 ,108). Reductions in REM latency have been observed. Increased muscular activity, contractions and periodic limb movements may prevent slow wave sleep and foster light fragmented sleep. Disorganized respiration also is found (109).

FUTURE DIRECTIONS

Part of "134 - Sleep Disturbances Associated with Neuropsychiatric Disease "

As one of the earlier tools available to psychiatric research for discovering the biological basis of mental disorders, EEG sleep recordings have been used extensively to characterize alterations in brain function across diverse mental disorders. This literature now forms the background for new research that can uncover the brain mechanisms that underlie these descriptive changes. Tools available for such purposes include refinements in electrophysiologic recordings using automated EEG and the concurrent use of electrical recordings of cognitive processes such as evoked responses to characterize changes in information processing during sleep in relation to mental disorders. Extensive advances in functional neuroimaging that are currently available, as well as in development, promise to provide us with dynamic images of brain function as it transitions throughout the sleep/wake cycle. In this manner, the functional neuroanatomic basis of the electrophysiologic abnormalities can be determined and interventions designed targeting not only specific neurotransmitter systems but systems that are specific to a discrete brain region that is responsible for the sleep/wake disturbance. The discovery of the genetic basis of narcolepsy brings the dream of uncovering the genetic basis of sleep disruption into closer view. Application of the advances in cognitive neuroscience to sleep also promise to define the importance of sleep in shaping cognitive function as well as the role of pathological sleep in disrupting waking cognitive activity.

ACKNOWLEDGMENTS

Part of "134 - Sleep Disturbances Associated with Neuropsychiatric Disease "

This work was supported in part by MH30915, MH52247, MH37869, MH00295, MH01414, MH45203, and the Theodore and Vada Stanley Foundation. The authors thank Robert Y. Moore for his contributions. EAN has received honoraria for speaking from Glaxo-Wellcome, Inc. MK has received research support from Pfizer Inc., Janssen Pharmaceutical Research Foundation, and Eli Lilly Pharmaceutical.

REFERENCES

1. Aserinsky E, Kleitman N. Regularly occurring periods of eye motility, and concomitant phenomena, during sleep. *Science* 1953;118:273–274.

2. Jouvet M. Recherches sur les structures nerveuses et les mecanismes responsables des differentes phases du sommeil physiologique. *Arch Italiennes Biol* 1962;100:125–206.

3. Steriade M, Amzica F, Contreras D. Synchronization of fast (30 to 40 Hz) spontaneous cortical rhythms during brain activation. *J Neurosci* 1996;16:392–417.

4. Nofzinger EA, Nichols TE, Meltzer CC, et al. Changes in forebrain function from waking to REM sleep in depression: preliminary analyses of [^{18}F] FDG PET studies. *Psychiatry Res Neuroimag* 1999;91:59–78.

5. Destexhe A, Contreras D, Steriade M. Cortically-induced coherence of a thalamic-generated oscillation. *Neuroscience* 1999;92:427–443.

6. Robbins TW, Everitt BJ. Arousal systems and attention. In: Gazzaniga M. *The cognitive neurosciences*. Cambridge: MIT Press, 1996:703–720.

7. Moore RY, Abrahamson EA, van den Pol A. The hypocretin neuron system: an arousal system in the human brain. *Arch Italiennes Biol* 2001;139:195–205.

8. Steriade M, McCarley RW. *Brainstem control of wakefulness and sleep*. New York: Plenum Press, 1990.

9. Hobson JA, Pace-Schott E, Stickgold R. Dreaming and the brain: towards a cognitive neuroscience of conscious states. *Behav Brain Sci* 2000;23:793–842.

10. Szymusiak R. Magnocellular nuclei of the basal forebrain: substrates of sleep and arousal regulation. *Sleep* 1995;18:478–500.

11. Shiromani PJ, Scammell T, Sherin JE, et al. Hypothalamic regulation of sleep. In: Lydic R, Baghdoyan HA. *Handbook of behavioral state control: cellular and molecular mechanisms*. Boca Raton, FL: CRC Press, 1999:311–326.

12. Edgar DM. Sleep-wake circadian rhythms and aging: potential etiologies and relevance to age-related changes in integrated physiological systems. *Neurobiol Aging* 1994;15:499–501.

13. Chemelli RM, Willie JT, Sinton CM, et al. Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation [see comments]. *Cell* 1999;98:437–451.

14. Nishino S, Ripley B, Overeem S, et al. Hypocretin (Orexin) deficiency in human narcolepsy [letter]. *Lancet* 2000;355:39–40.

15. McCarley RW, Hobson JA. Neuronal excitability modulation over the sleep cycle: a structural and mathematical model. *Science* 1975;189:58–60.

16. Krueger JM, Obal F, Fang J. Humoral regulation of physiological sleep: cytokines and GHRH. *J Sleep Res* 1999;8:53–59.

17. Semba K. The mesopontine cholinergic system: a dual role in REM sleep and wakefulness. In: Lydic R, Baghdoyan HA. *Handbook of behavioral state control: cellular and molecular mechanisms*. Boca Raton, FL: CRC Press, 1999:161–180.

18. Luppi P-H, Peyron C, Rampon C, et al. Inhibitory mechanisms in the dorsal raphe nucleus and locus ceruleus during sleep. In: Lydic R, Baghdoyan HA. *Handbook of behavioral state control: cellular and molecular mechanisms*. Boca Raton, FL: CRC Press, 1999:195–212.

19. Braun AR, Balkin TJ, Wesenten NJ, et al. Regional cerebral blood flow throughout the sleep-wake cycle. An $H_2(^{15})O$ PET study. *Brain* 1997;120:1173–1197.

20. Nofzinger EA, Mintun MA, Wiseman MB, et al. Forebrain activation in REM sleep: an FDG PET study. *Brain Res* 1997;770:192–201.

21. Calvo JM, Simon-Arceo K. Cholinergic enhancement of REM sleep from sites in the pons and amygdala. In: Lydic R, Baghdoyan HA. *Handbook of behavioral state control: cellular and molecular mechanisms*. Boca Raton, FL: CRC Press, 1999:391–406.

22. Sanford LD, Tejani-Butt SM, Ross RJ, et al. Amygdaloid control of alerting and behavioral arousal in rats: involvement of serotonergic mechanisms. *Arch Italiennes Biol* 1995;134:81–99.

23. Buchsbaum MS, Gillin JC, Wu J, et al. Regional cerebral glucose metabolic rate in human sleep assessed by positron emission tomography. *Life Sci* 1989;45:1349–1356.

24. Maquet P, Franck G. Cerebral glucose metabolism during the sleep-wake cycle in man as measured by positron emission tomography. In: Horne J. *Sleep '88*. New York: Gustav Fischer, 1989:76–78.

25. Clark C, Dupont R, Lehr P, et al. Is there a relationship between delta sleep at night and afternoon cerebral blood flow, assessed by HMPAO-SPECT in depressed patients and normal control subjects? Preliminary data. *Psychiatry Res* 1998;84:89–99.

26. Maquet P, Degueldre C, Delfiore G, et al. Functional neuroanatomy of human slow wave sleep. *J Neurosci* 1997;17:2807–2812.

27. Hofle N, Paus T, Reutens D, et al. Regional cerebral blood flow changes as a function of delta and spindle activity during slow wave sleep in humans. *J Neurosci* 1997;15:4800–4808.

28. Andersson JL, Onoe H, Hetta J, et al. Brain networks affected by synchronized sleep visualized by positron emission tomography. *J Cereb Blood Flow Metab* 1998;18:701–715.

29. Lydic R, Baghdoyan HA, Hibbard L, et al. Regional brain glucose metabolism is altered during rapid eye movement sleep in the cat: a preliminary study. *J Comp Neurol* 1991;304:517–529.

30. Maquet P, Peters JM, Aerts J, et al. Functional neuroanatomy of human rapid-eye-movement sleep and dreaming. *Nature* 1996;383:163–166.

31. Drummond SPA, Brown GC, Gillin JC, et al. Altered brain response to verbal learning following sleep deprivation. *Nature* 2000;43:655–657.

32. Borbély AA. A two process model of sleep regulation. *Hum Neurobiol* 1982;1:195–204.

33. Buzsaki G. Commentary: two-stage model of memory trace formation: a role for "noisy" brain states. *Neuroscience* 1989;31:551–70.

34. Llinas RR, Pare D. Commentary: of dreaming and wakefulness. *Neuroscience* 1991;44:521–535.

35. Winson J. The biology and function of REM sleep. *Curr Opin Neurobiol* 1993;3:243–248.

36. Kavanau JL. Sleep and dynamic stabilization of neural circuitry: a review and synthesis. *Behav Brain Res* 1994;63:111–126.

37. Buysse DJ, Reynolds CF, Monk TH, et al. The Pittsburgh Sleep Quality Index (PSQI): a new instrument for psychiatric research and practice. *Psychiatry Res* 1989;28:193–213.

38. Armitage R, Hoffmann RF, Rush AJ. Biological rhythm disturbance in depression: temporal coherence of ultradian sleep EEG rhythms. *Psychol Med* 1999;29:1435–1448.

39. Nofzinger EA, Buysse DJ, Reynolds CF, et al. Sleep disorders related to another mental disorder (nonsubstance/primary): a DSM-IV literature review [Review]. *J Clin Psychiatry* 1993;54:244–255; discussion 256–259.

40. Borbély AA, Tobler I, Loepfe M, et al. All-night spectral analysis of the sleep EEG in untreated depressives and normal controls. *Psychiatry Res* 1984;12:27–33.

41. Kupfer DJ, Frank E, McEachran AB, et al. Delta sleep ratio: a biological correlate of early recurrence in unipolar affective disorder. *Arch Gen Psychiatry* 1990;47:1100–1105.

42. Armitage R, Hoffmann R, Trivedi M, et al. Sleep macro- and microarchitecture in depression: age and gender effects. *Sleep Res* 1997;26:283.

43. Lauer CJ, Riemann D, Wiegand M, et al. From early to late adulthood changes in EEG sleep of depressed patients and healthy volunteers. *Biol Psychiatry* 1991;29:979–993.

44. Dahl RE, Ryan ND, Birmaher B, et al. EEG sleep measures in prepubertal depression. *Psychiatry Res* 1991;38:201–214.

45. Dew MA, Reynolds CF, Buysse DJ, et al. Electroencephalographic sleep profiles during depression. Effects of episode duration and other clinical and psychosocial factors in older adults. *Arch Gen Psychiatry* 1996;53:148–156.

46. Rush AJ, Giles DE, Jarrett RB, et al. Reduced REM latency predicts response to tricyclic medication in depressed outpatients. *Biol Psychiatry* 1989;26:61–72.

47. Giles DE, Jarrett RB, Roffwarg HP, et al. Reduced rapid eye movement latency: a predictor of recurrence in depression. *Neuropsychopharmacology* 1987;1:33–39.

48. Sharpley AL, Walsh AE, Cowen PJ. Nefazodone—a novel antidepressant—may increase REM sleep. *Biol Psychiatry* 1992;31:1070–1073.

49. Nofzinger EA, Reynolds CF, Thase ME, et al. REM sleep enhancement by bupropion in depressed men. *Am J Psychiatry* 1995;152(2):274–276.

50. Gillin JC, Sutton L, Ruiz C, et al. The cholinergic rapid eye movement test with arecoline in depression. *Arch Gen Psychiatry* 1991;48:264–270.

51. Bhatti T, Gillin JC, Seifritz E, et al. Effects of a tryptophan-free amino acid drink challenge on normal human sleep electroencephalogram and mood. *Biol Psychiatry* 1998;43:52–59.

52. Gillin JC, Sohn JW, Stahl SM, et al. Ipsapirone, a 5-HT1A agonist, suppresses REM sleep equally in unmedicated depressed patients and normal controls. *Neuropsychopharmacology* 1996;15:109–115.

53. Ho AP, Gillin JC, Buchsbaum MS, et al. Brain glucose metabolism during non-rapid eye movement sleep in major depression: a positron emission tomography study. *Arch Gen Psychiatry* 1996;53:645–652.

54. Nofzinger EA, Price JC, Meltzer CC, et al. Towards a neurobiology of dysfunctional arousal in depression: the relationship between beta EEG power and regional cerebral glucose metabolism during NREM sleep. *Psychiatry Res Neuroimag* 2000;98:71–91.

55. Wu J, Buchsbaum MS, Gillin JC, et al. Prediction of antidepressant effects of sleep deprivation by metabolic rates in the ventral anterior cingulate and medial prefrontal cortex. *Am J Psychiatry* 1999;156:1149–1158.

56. Clark CP, Gillin JC, Golshan S, et al. Polysomnography and depressive symptoms in primary alcoholics with and without a lifetime diagnosis of secondary depression and in patients with primary major depression. *J Affect Disord* 1999;52:177–185.

57. Drummond SP, Gillin JC, Smith TL, et al. The sleep of abstinent pure primary alcoholic patients: natural course and relationship to relapse. *Alcohol Clin Exp Res* 1998;22:1796–1802.

58. Zarcone VP, Benson KL. *Sleep and schizophrenia. Principles and practice of sleep medicine*. Philadelphia: WB Saunders, 1994:105–214.

59. Benca RM, Obermeyer WH, Thisted RA, et al. Sleep and psychiatric disorders: a meta-analysis. *Arch Gen Psychiatry* 1992;49:651–668.

60. Ganguli R, Reynolds CF, Kupfer DJ. Electroencephalographic sleep in young, never medicated schizophrenics. *Arch Gen Psychiatry* 1987;44:36–44.

61. Tandon R, Shipley JE, Taylor S, et al. Electroencephalographic sleep abnormalities in schizophrenia: relationship to positive/negative symptoms and prior neuroleptic treatment. *Arch Gen Psychiatry* 1992;49:185–194.

62. Keshavan MS, Anderson S, Pettegrew JW. Is schizophrenia due to excessive synaptic pruning in prefrontal cortex? The Feinberg hypothesis revisited. *J Psychiatric Res* 1994;28:239–265.

63. Werth E, Achermann P, Borbely AA. Fronto-occipital EEG power gradients in human sleep. *J Sleep Res* 1997;6:102–112.

64. Keshavan MS, Reynolds CF, Miewald JM, et al. Delta sleep deficits in schizophrenia: evidence from automated analyses of sleep data. *Arch Gen Psychiatry* 1998;55:443–448.

65. Benson KL, Sullivan EV, Lim KO, et al. The effect of total sleep deprivation on slow wave recovery in schizophrenia. *Sleep Res* 1993;22:143.

66. Horne J. Neuroscience. Images of lost sleep [News; Comment]. *Nature* 2000;403:605–606.

67. Keshavan MS, Reynolds CF, Montrose D, et al. Sleep and suicidality in psychotic patients. *Acta Psychiatr Scand* 1994;89:122–125.

68. Lewis C, Tandon R, Shipley JR, et al. Biological predictors of suicidality in schizophrenia. *Acta Psychiatr Scand* 1996;94:416–420.

69. Maixner S, Tandon R, Eiser A, et al. Effects of antipsychotic treatment on polysomnographic measures in schizophrenia: a replication and extension [Report]. *Am J Psychiatry* 1997;155:1600–1602.

70. Keshavan MS, Reynolds CF, Miewald JM, et al. A longitudinal study of EEG sleep in schizophrenia. *Psychiatry Res* 1996;59:203–211.

71. Keshavan MS, Reynolds CF, Miewald J, et al. Slow-wave sleep deficits and outcome in schizophrenia and schizoaffective disorder. *Acta Psychiatr Scand* 1995;91:289–292.

72. Orzack MH, Hartman EL, Kornetsky C. The relationship between attention and slow wave sleep in schizophrenia. *Psychopharmacol Bull* 1977;13:59–61.

73. Smith JR, Karacan I, Yang M. Ontongeny of delta activity during human sleep. *Electroenceph Clin Neurophysiol* 1977;43:229.

74. Feinberg I. Schizophrenia: caused by a default in programmed synaptic elimination during adolescence. *J Psychiatric Res* 1982;17:319–330.

75. Andreasen NC, Arndt S, Swayze V, et al. Thalamic abnormalities in schizophrenia visualized through magnetic resonance image averaging. *Science* 1994;266:294–298.

76. Keshavan MS, Reynolds CF, Ganguli R, et al. Electroencephalographic sleep and cerebral morphology in functional psychoses: a preliminary study with computed tomography. *Psychiatry Res* 1991;39:293–301.

77. Van Kammen DP, van Kammen WB, Peters J, et al. Decreased slow-wave sleep and enlarged lateral ventricles in schizophrenia. *Neuropsychopharmacology* 1988;1:265–271.

78. Keshavan MS, Tandon R. Sleep abnormalities in schizophrenia: pathophysiologic significance. *Psychol Med* 1993;23:831–835.

79. Riemann D. Cholinergic REM induction test: muscarinic supersensitivity underlies polysomnographic findings in both depression and schizophrenia. *J Psychiatric Res* 1994;28:195–210.

80. Griffith JM, O'Neill JE, Petty F, et al. Nicotinic receptor desensitization and sensory gating deficits in schizophrenia. *Biol Psychiatry* 1998;44:98–106.

81. Benson KL, Faull KF, Zarcone VPJ. Evidence for the role of serotonin in the regulation of slow wave sleep in schizophrenia. *Sleep* 1991;14:133–139.

82. Van Kammen DP, Widerlov E, Neylan TC, et al. Delta sleep-inducing-peptide-like immunoreactivity (DSIP-LI) and sleep in schizophrenic volunteers. *Sleep* 1992;15:519–525.

83. Porkka-Heiskanen T, Strecker RE, Thakkar M, et al. Adenosine: a mediator of the sleep-inducing effects of prolonged wakefulness. *Science* 1997;276:1265–1268.

84. Ferre S. Adenosine-dopamine interactions in the ventral striatum. Implications for the treatment of schizophrenia. *Psychopharmacol (Berl)* 1997;133:107–120.

85. Keshavan MS, Pettegrew JW, Reynolds CFJ, et al. Slow wave sleep deficits in schizophrenia: pathophysiologic significance. *Psychiatry Res* 1995;57:91–100.

86. Steriade M. Brain electrical activity and sensory processing during waking and sleep states. In: Kryger MH, Roth T, Dement WC, eds. *Principles and practice of sleep medicine*. Philadelphia: WB Saunders, 1994:105–214.

87. Salin-Pascual RJ, Herrera-Estrella M, Galicia-Polo L, et al. Olanzapine acute administration in schizophrenic patients increases delta sleep and sleep efficiency. *Biol Psychiatry* 1999;46:141–143.

88. Hinze-Selch D, Mullington J, Orth A, et al. Effects of clozapine on sleep: a longitudinal study. *Biol Psychiatry* 1997;42:260–266.

89. Neylan TC, Van Kammen DP, Kelley ME, et al. Sleep in schizophrenic patients on and off haloperidol therapy: clinically stable vs. relapsed patients. *Arch Gen Psychiatry* 1992;49:643–649.

90. Nofzinger EA, vanKammen DP, Gilbertson MW, et al. Electroencephalographic sleep in clinically stable schizophrenic patients: two- vs. six-weeks neuroleptic free. *Biol Psychiatry* 1993;33:829–835.

91. Dahl RE. The development and disorders of sleep. *Adv Pediatr* 1998;45:73–90.

92. Lauer CJ, Schreiber W, Holsboer F, et al. In quest of identifying vulnerability markers for psychiatric disorders by all-night polysomnography. *Arch Gen Psychiatry* 1995;52:145–153.

93. Reynolds CF, Monk TH, Hoch CC, et al. Electroencephalographic sleep in the healthy "old old": A comparison with the "young old" in visually scored and automated measures. *J Gerontol* 1991;46:M39–M46.

94. Monk TH, Reynolds CF, Machen MA, et al. Daily social rhythms in the elderly and their relation to objectively recorded sleep. *Sleep* 1992;15:322–329.

95. Reynolds CF, Hoch CC, Buysse DJ, et al. *Sleep* after spousal bereavement: a study of recovery from stress. *Biol Psychiatry* 1993;34:791–797.

96. Bliwise DL. Dementia. In: Kryger MH, Roth T, Dement WC, eds. *Principles and practice of sleep medicine*, second ed. Philadelphia: WB Saunders, 1994:790.

97. Chenier MC. Review and analysis of caregiver burden and nursing home placement. *Geriatr Nurs* 1997;18:121–126.

98. Bliwise DL, Hughes M, McMahon PM, et al. Observed sleep/wakefulness and severity of dementia in an Alzheimer's disease special care unit. *J Gerontol, Series A, Biol Sci Med Sci* 1995;50:M303–M306.

99. Cohen-Mansfield J, Werner P, Freedman L. Sleep and agitation in agitated nursing home residents: an observational study. *Sleep* 1995;18:674–680.

100. Prinz PN, Peskind ER, Vitaliano PP, et al. Changes in the sleep and waking EEGs of nondemented and demented elderly subjects. *J Am Geriatr Soc* 1982;30:86–93.

101. Ancoli-Israel S, Klauber MR, Butters N, et al. Dementia in institutionalized elderly: relation to sleep apnea. *J Am Geriatr Soc* 1991;39:258–263.

102. Cohen-Mansfield J, Marx MS, Rosenthal AS. Dementia and agitation in nursing home residents: how are they related? *Psychol Aging* 1990;5:3–8.

103. Tandberg E, Larsen JP, Karlsen K. A community-based study of sleep disorders in patients with Parkinson's disease. *Mov Disord* 1998;13:895–899.

104. Factor SA, McAlarney T, Sanchez-Ramos JR, et al. Sleep disorders and sleep effect in Parkinson's disease. *Mov Disord* 1990;5:280–285.

105. Stocchi F, Barbato L, Nordera G, et al. Sleep disorders in Parkinson's disease. *J Neurol* 1998;245:S15–S18.

106. Bernath O, Guilleminault C. Sleep-related violence, injury, and REM sleep behavior disorder in PD. *Neurology* 1999;52:1924.

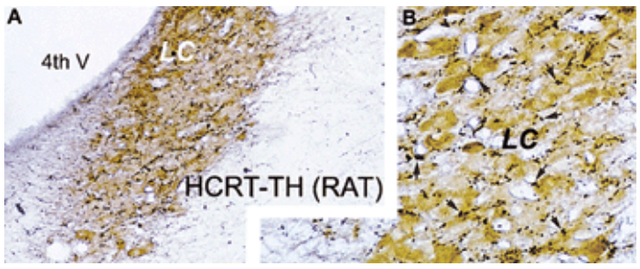
107. Kales A, Ansel RD, Markham CH, et al. Sleep in patients with Parkinson's disease and normal subjects prior to and following levodopa administration. *Clin Pharmacol Therapeut* 1971;12:397–406.

108. Kostic VS, Susic V, Przedborski S, et al. Sleep EEG in depressed and nondepressed patients with Parkinson's disease. *J Neuropsychiatry Clin Neurosci* 1991;3:176–179.

109. Apps MC, Sheaff PC, Ingram DA, et al. Respiration and sleep in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1985;48:1240–1245.

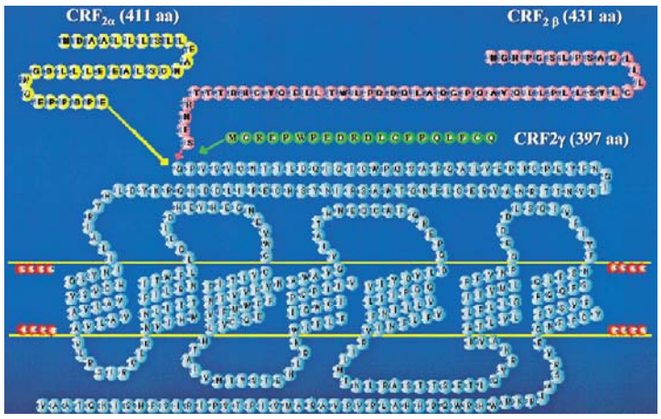
Resources

Color Plates

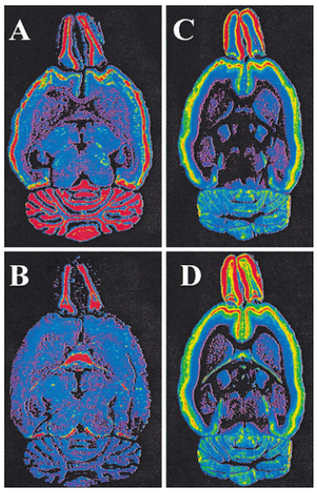


COLOR FIGURE 4.1. (This figure is printed in black and white as Figure 4.1.)

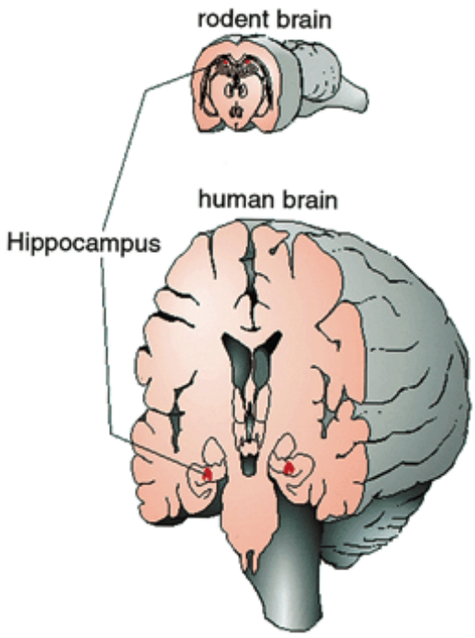
COLOR FIGURE 7.1. (This figure is printed in black and white as Figure 7.1.)



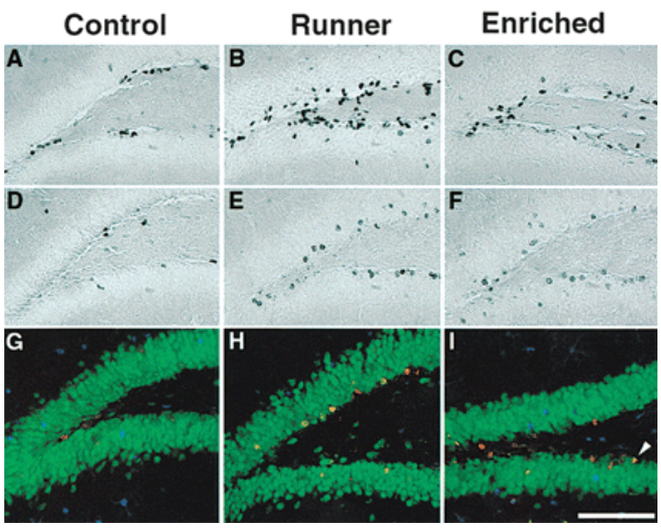
COLOR FIGURE 7.2. (This figure is printed in black and white as Figure 7.2.)



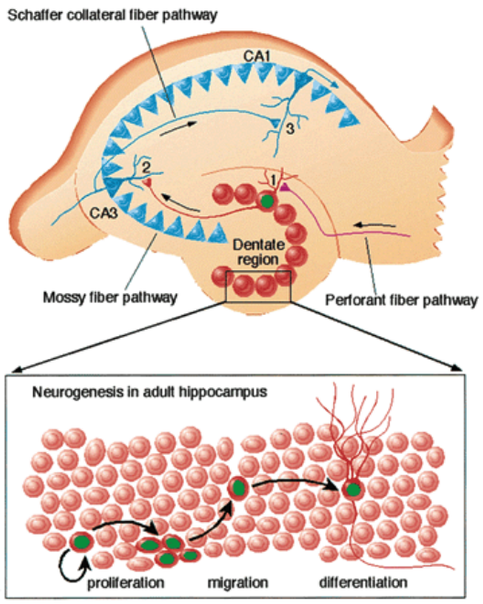
COLOR FIGURE 7.3. (This figure is printed in black and white as Figure 7.3.)



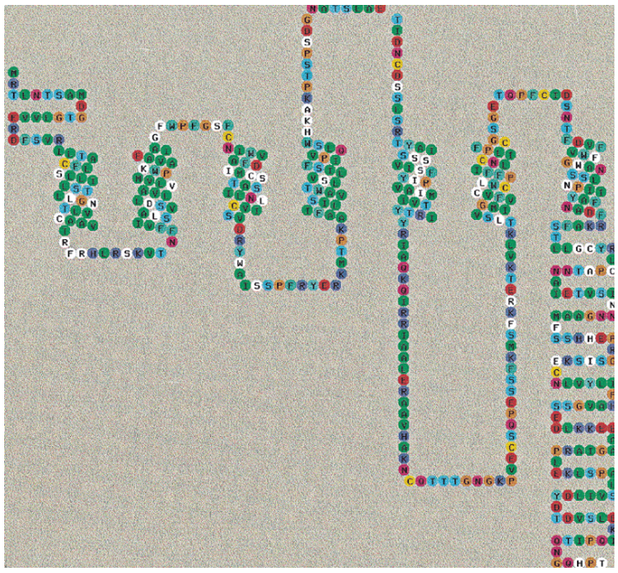
COLOR FIGURE 8.1. (This figure is printed in black and white as Figure 8.1.)



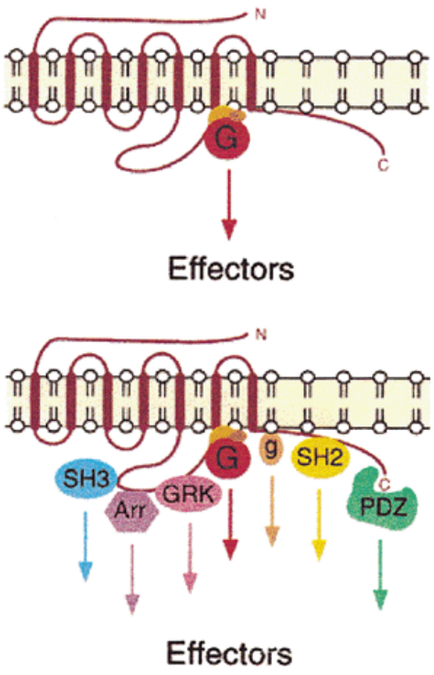
COLOR FIGURE 8.2. (This figure is printed in black and white as Figure 8.2.)



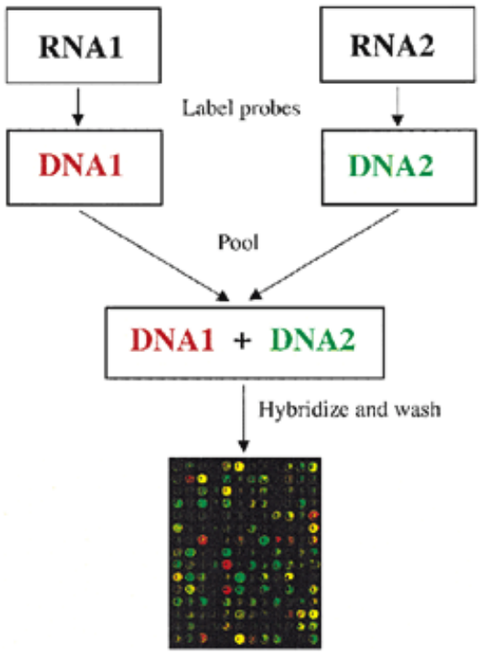
COLOR FIGURE 8.3. (This figure is printed in black and white as Figure 8.3.)



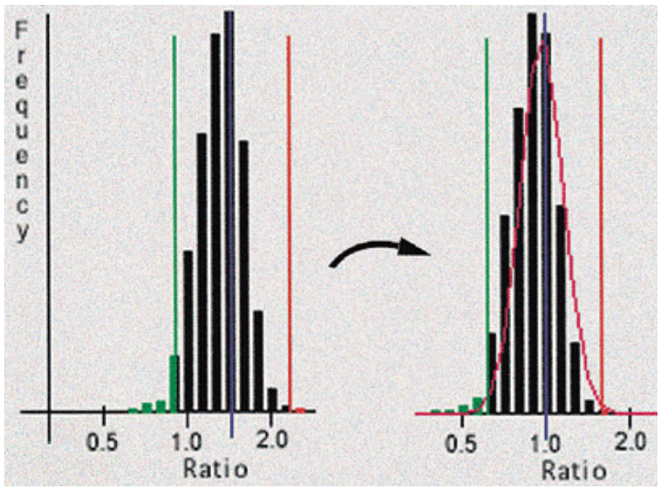
COLOR FIGURE 22.1. (This figure is printed in black and white as Figure 22.1.)



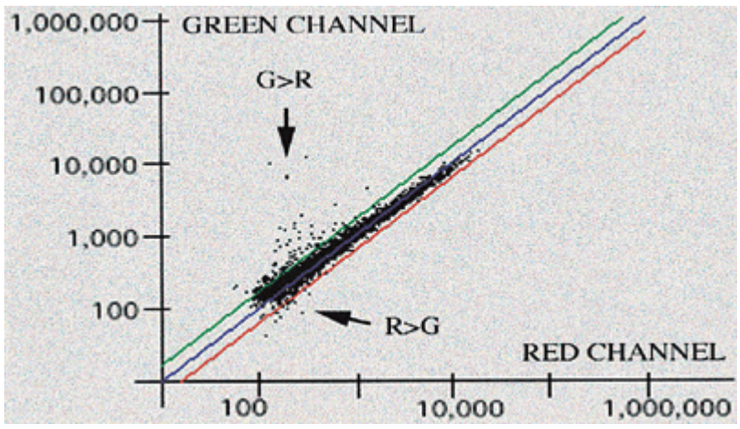
COLOR FIGURE 22.3. (This figure is printed in black and white as Figure 22.3.)



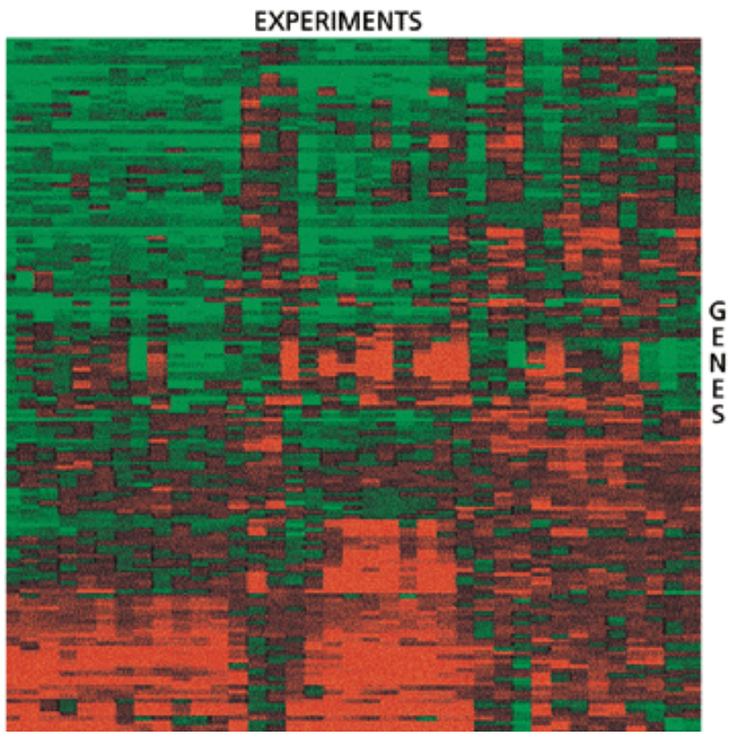
COLOR FIGURE 23.1. (This figure is printed in black and white as Figure 23.1.)



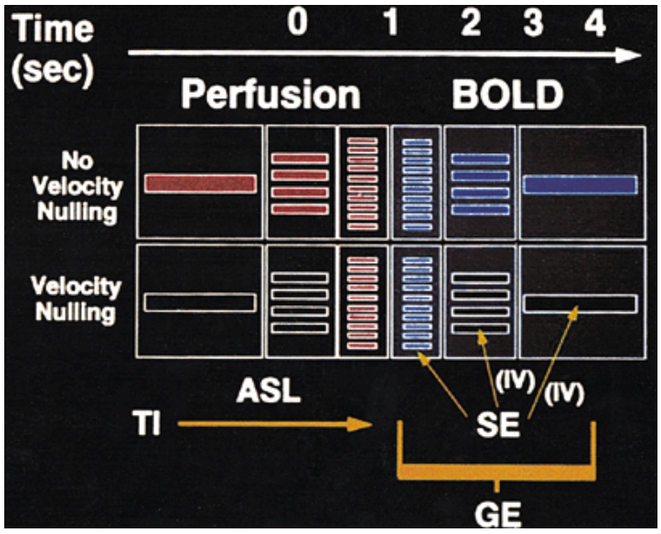
COLOR FIGURE 23.2. (This figure is printed in black and white as Figure 23.2.)



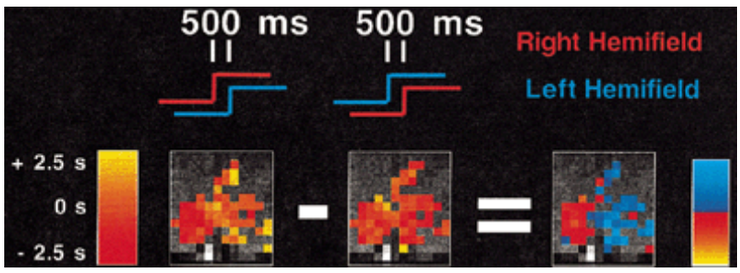
COLOR FIGURE 23.3. (This figure is printed in black and white as Figure 23.3.)



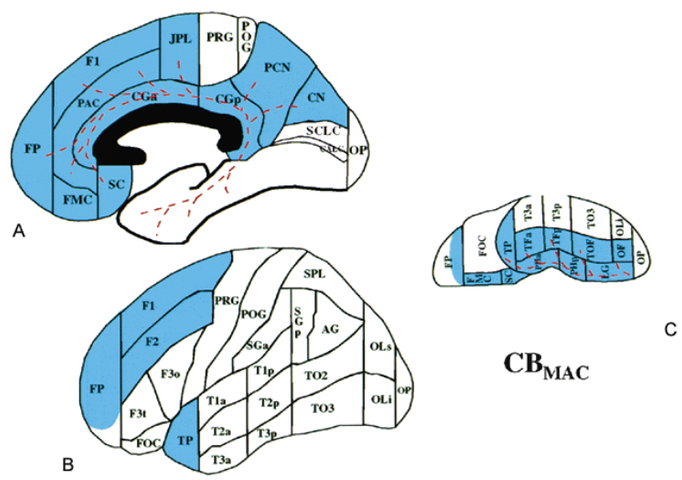
COLOR FIGURE 23.4. (This figure is printed in black and white as Figure 23.4.)



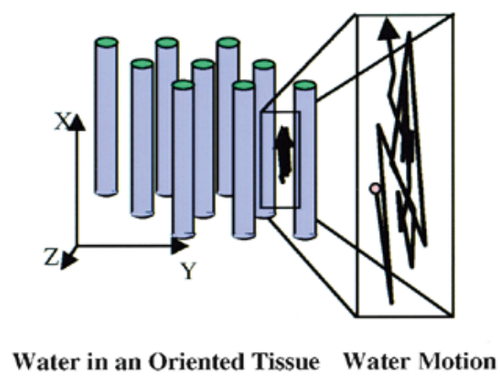
COLOR FIGURE 26.2. (This figure is printed in black and white as Figure 26.2.)



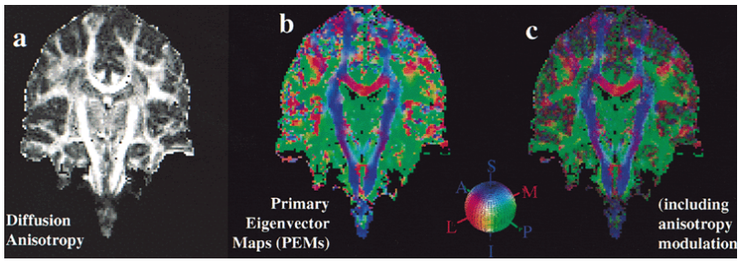
COLOR FIGURE 26.6. (This figure is printed in black and white as Figure 26.6.)



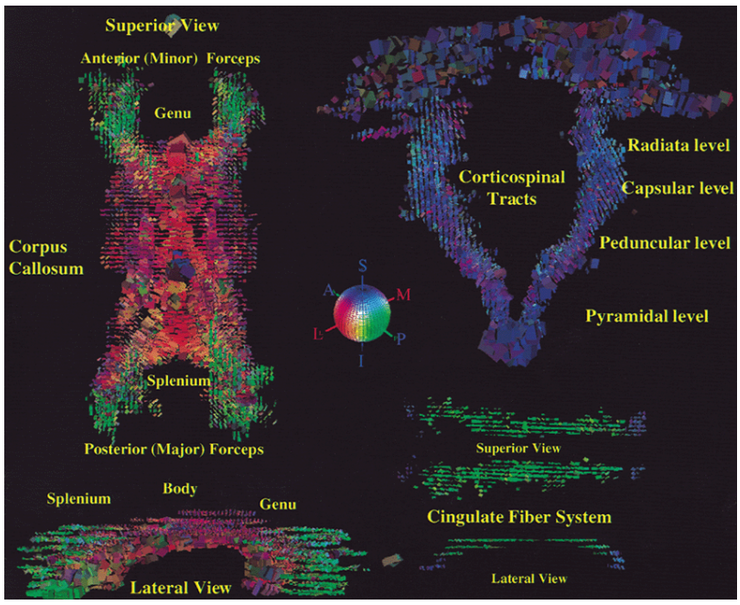
COLOR FIGURE 27.2. (This figure is printed in black and white as Figure 27.2.)



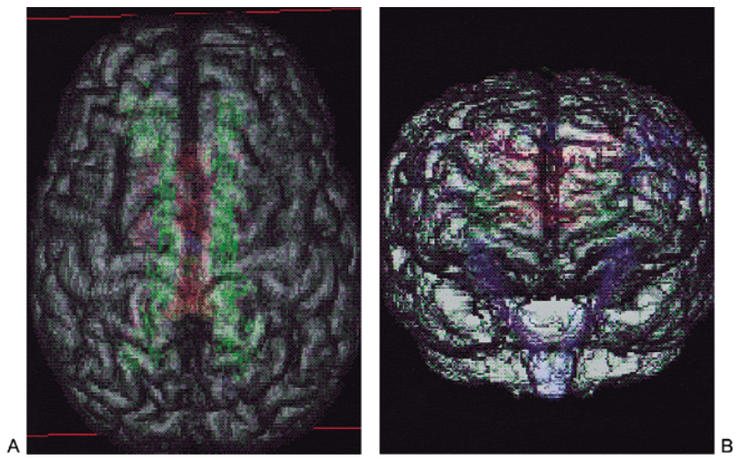
COLOR FIGURE 27.3. (This figure is printed in black and white as Figure 27.3.)



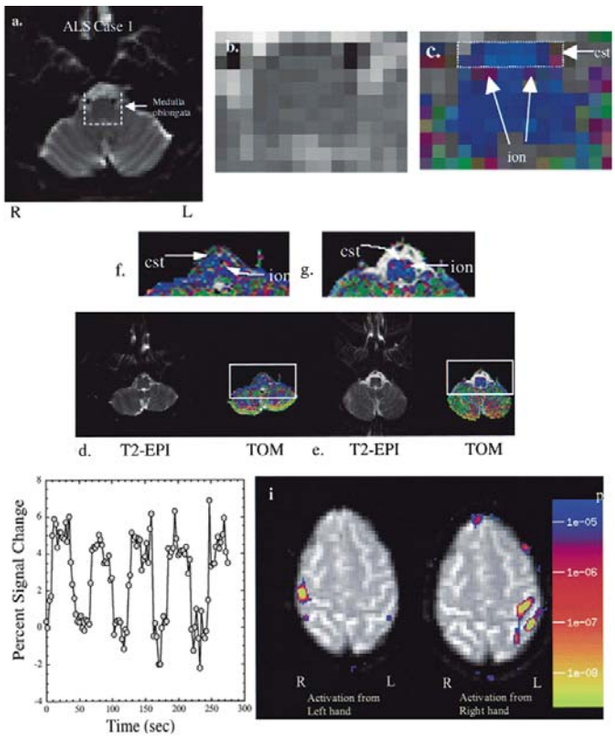
COLOR FIGURE 27.4. (This figure is printed in black and white as Figure 27.4.)



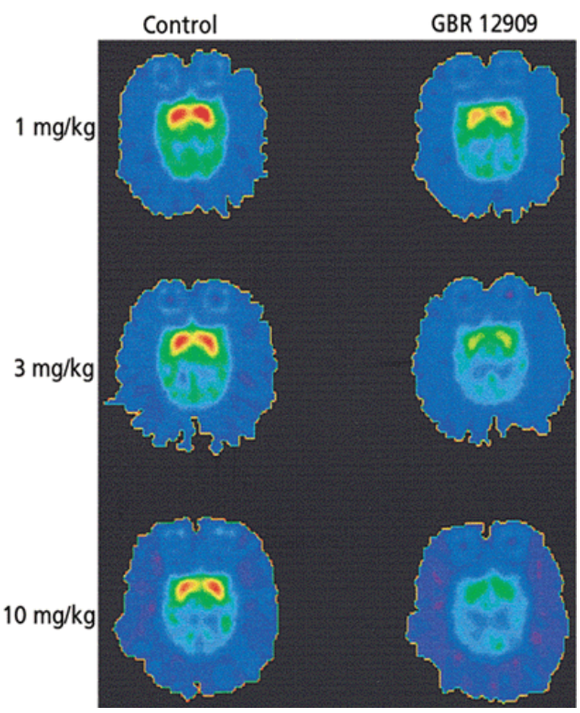
COLOR FIGURE 27.6. (This figure is printed in black and white as Figure 27.6.)



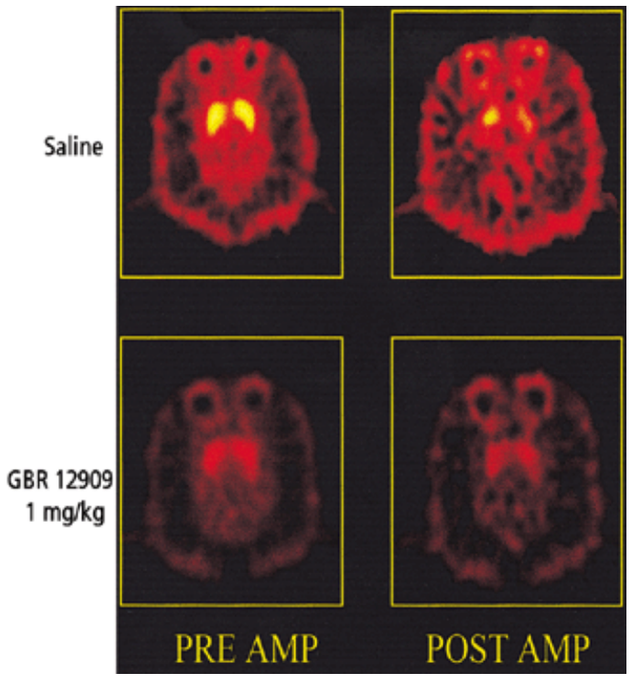
COLOR FIGURE 27.7. (This figure is printed in black and white as Figure 27.7.)



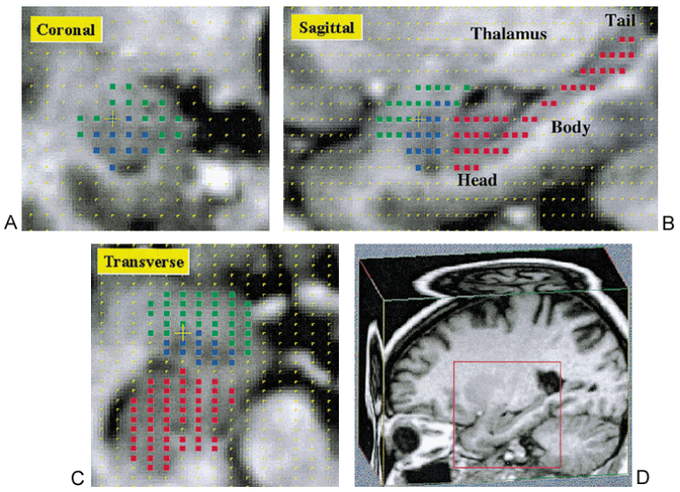
COLOR FIGURE 27.8. (This figure is printed in black and white as Figure 27.8.)



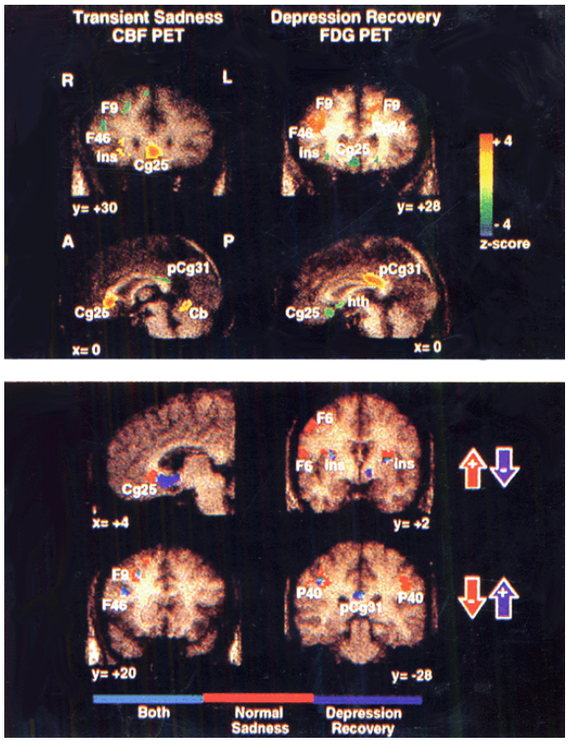
COLOR FIGURE 34.1. (This figure is printed in black and white as Figure 34.1.)



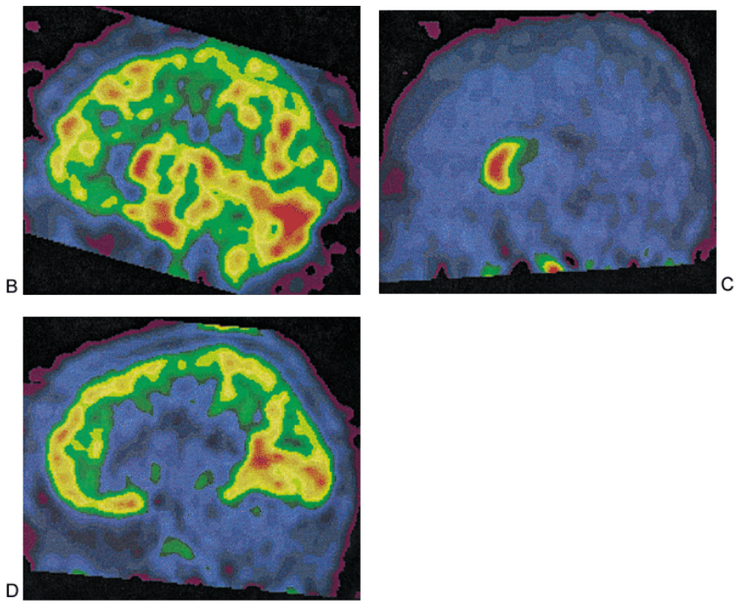
COLOR FIGURE 34.3. (This figure is printed in black and white as Figure 34.3.)



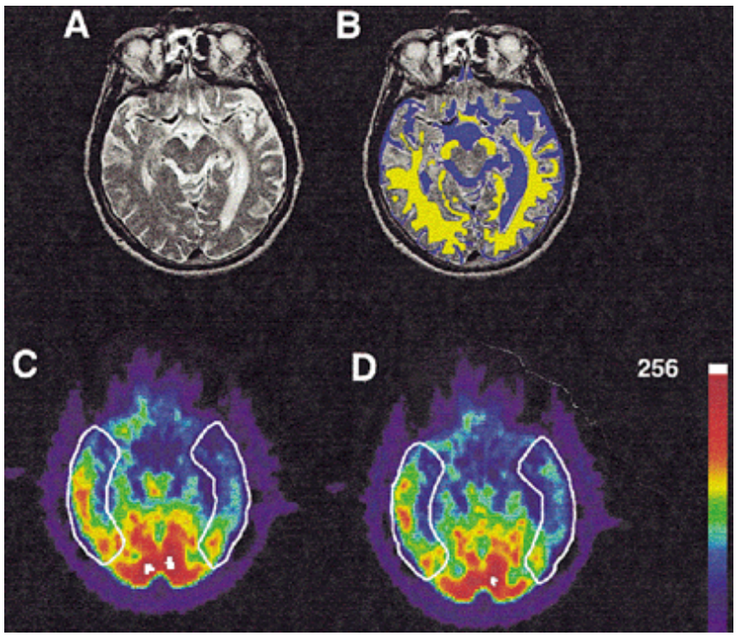
COLOR FIGURE 74.1. (This figure is printed in black and white as Figure 74.1.)



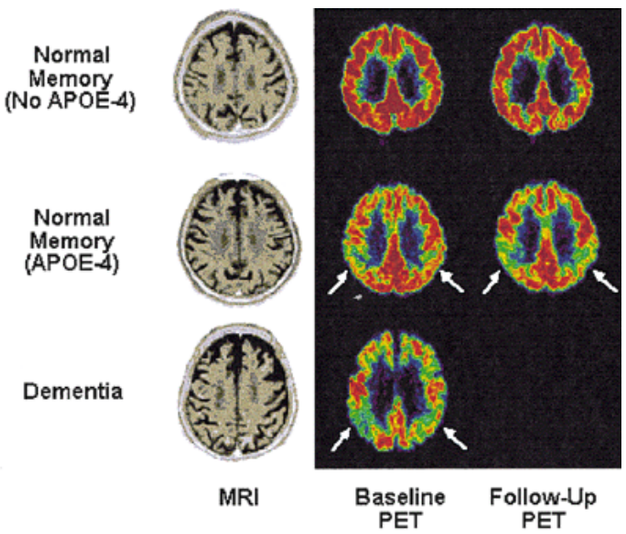
COLOR FIGURE 74.2. (This figure is printed in black and white as Figure 74.2.)



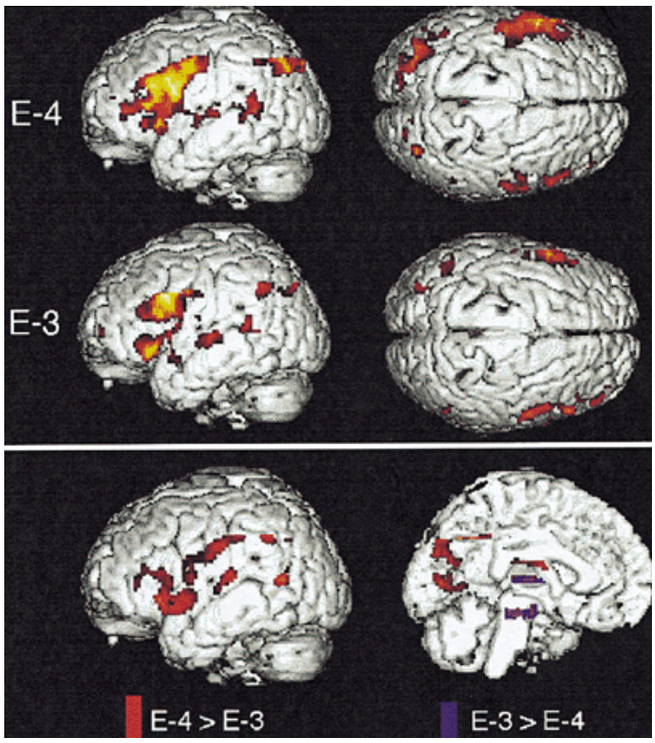
COLOR FIGURE 74.3. (This figure is printed in black and white as Figure 74.3.)



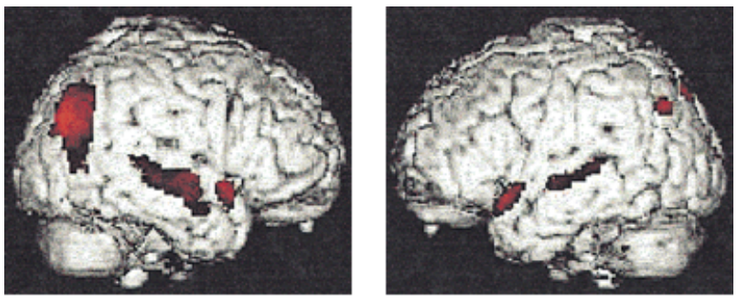
COLOR FIGURE 86.1. (This figure is printed in black and white as Figure 86.1.)



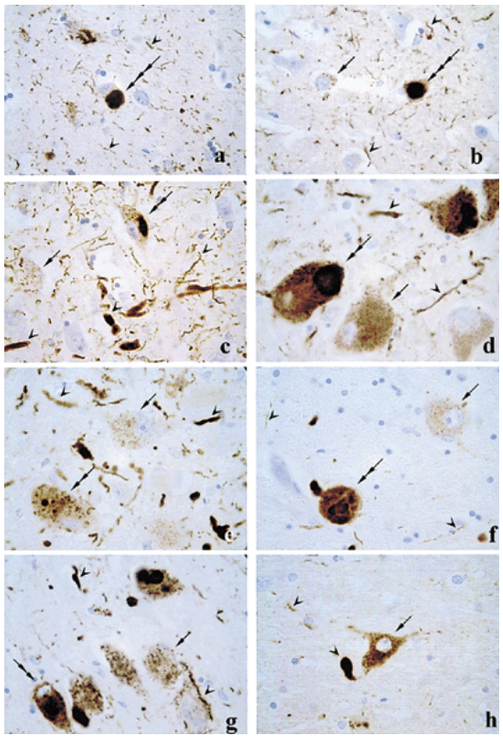
COLOR FIGURE 86.2. (This figure is printed in black and white as Figure 86.2.)



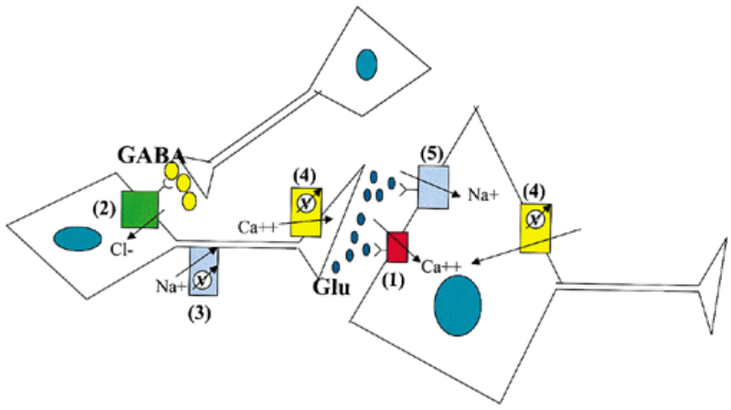
COLOR FIGURE 86.3. (This figure is printed in black and white as Figure 86.3.)



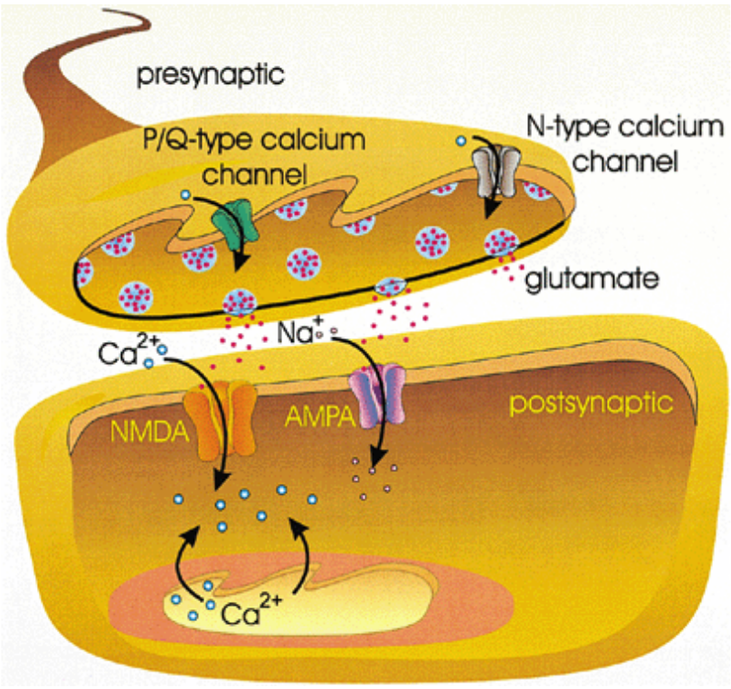
COLOR FIGURE 86.4. (This figure is printed in black and white as Figure 86.4.)



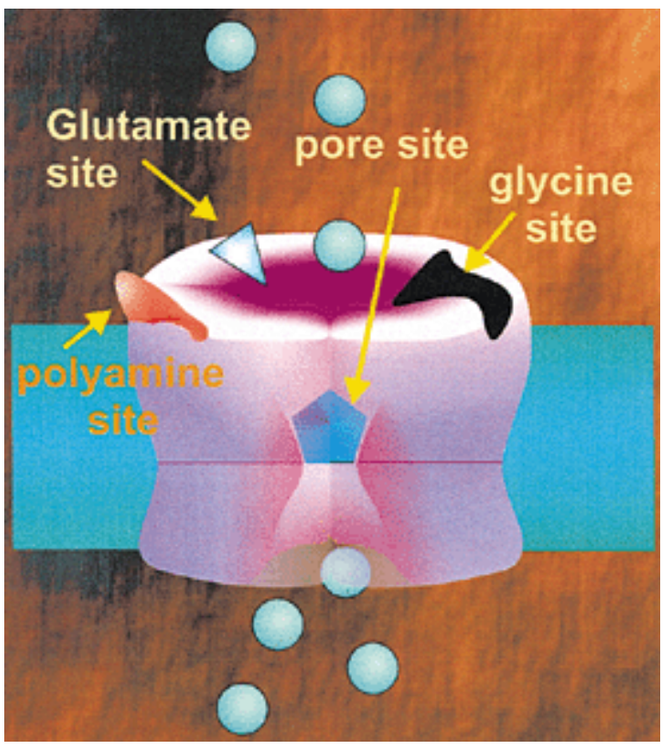
COLOR FIGURE 91.1. (This figure is printed in black and white as Figure 91.1.)



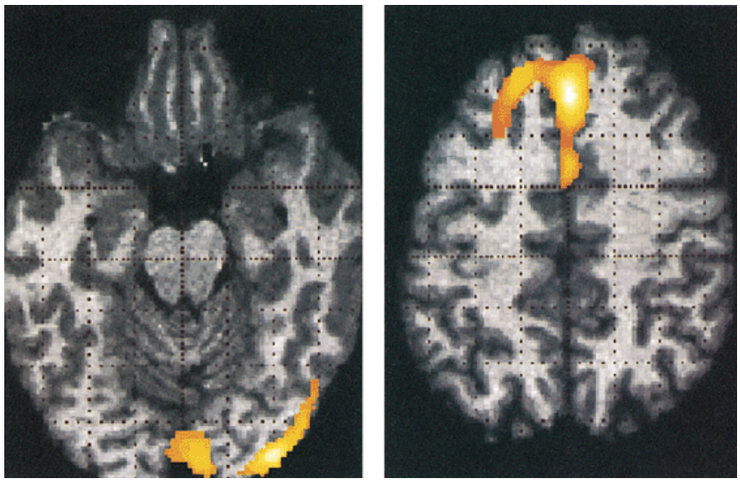
COLOR FIGURE 92.2. (This figure is printed in black and white as Figure 92.2.)



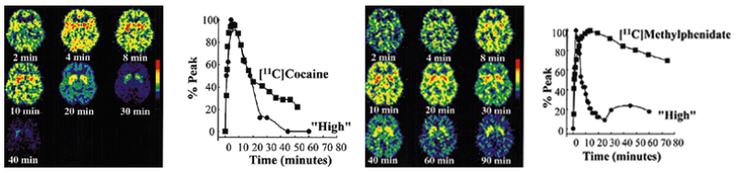
COLOR FIGURE 93.1. (This figure is printed in black and white as Figure 93.1.)



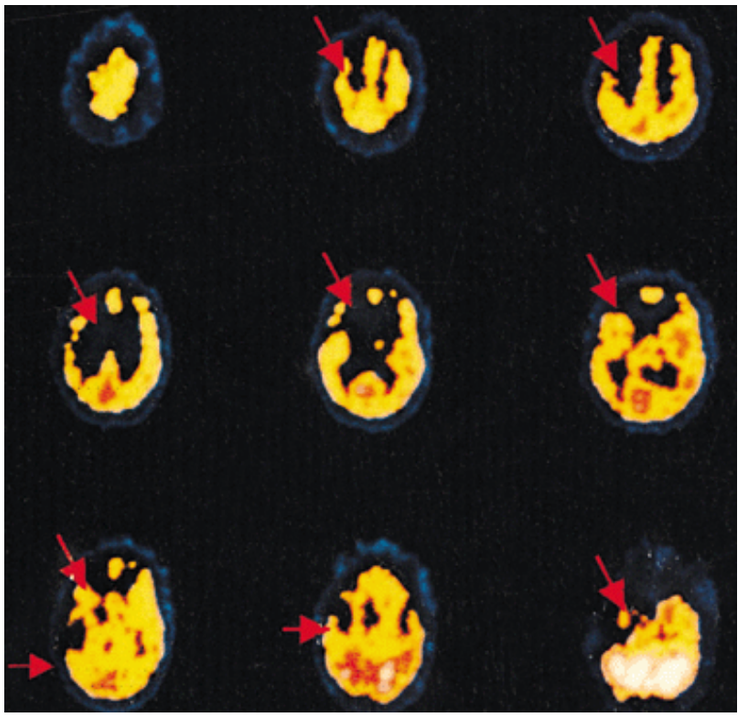
COLOR FIGURE 93.2. (This figure is printed in black and white as Figure 93.2.)



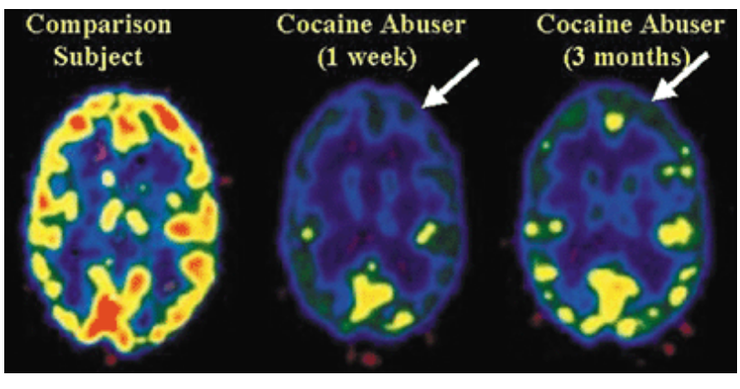
COLOR FIGURE 100.3. (This figure is printed in black and white as Figure 100.3.)



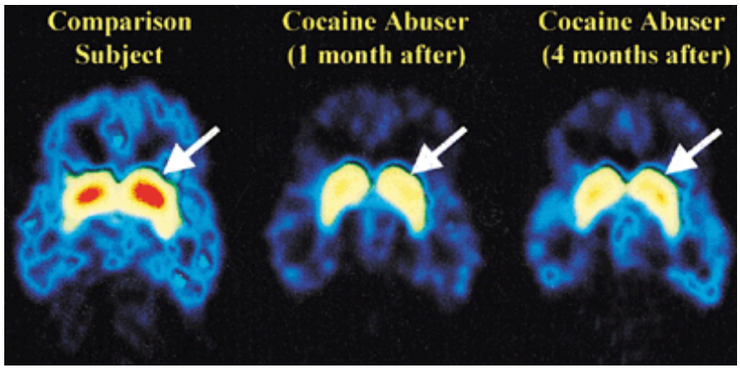
COLOR FIGURE 103.1. (This figure is printed in black and white as Figure 103.1.)



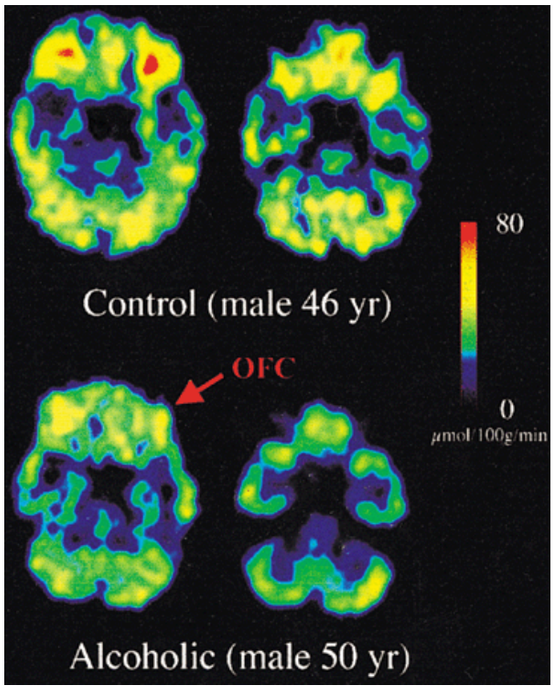
COLOR FIGURE 103.2. (This figure is printed in black and white as Figure 103.2.)



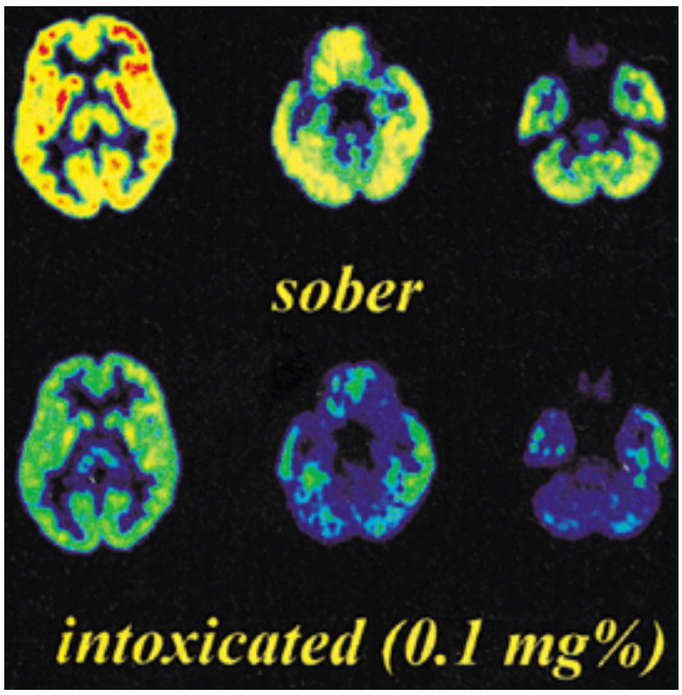
COLOR FIGURE 103.3. (This figure is printed in black and white as Figure 103.3.)



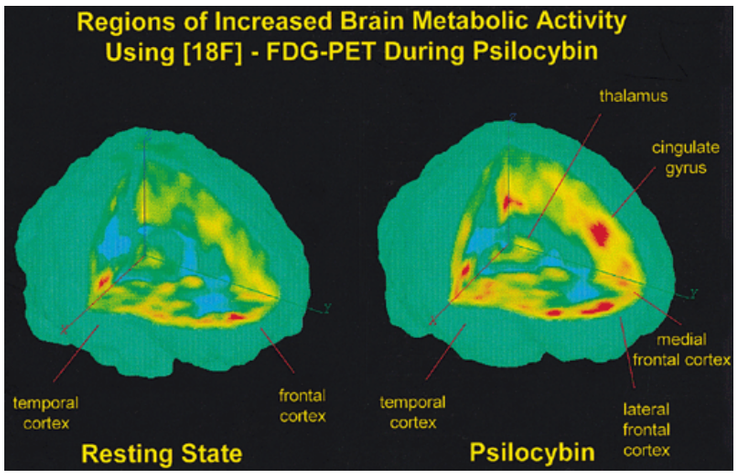
COLOR FIGURE 103.4. (This figure is printed in black and white as Figure 103.4.)



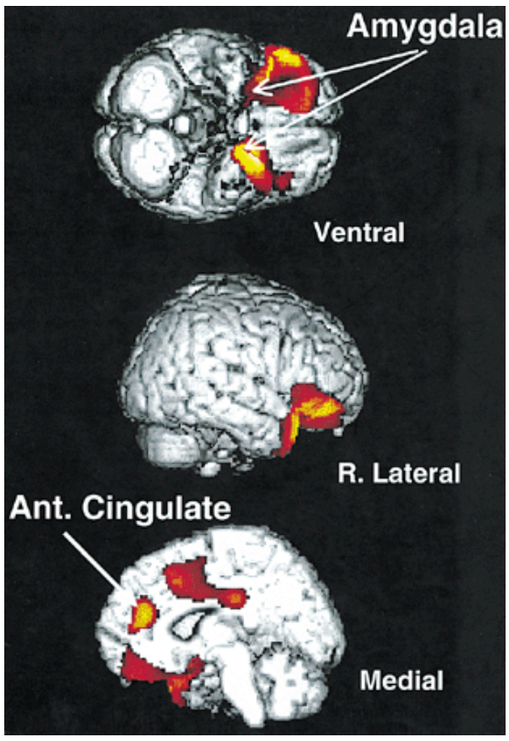
COLOR FIGURE 103.5. (This figure is printed in black and white as Figure 103.5.)



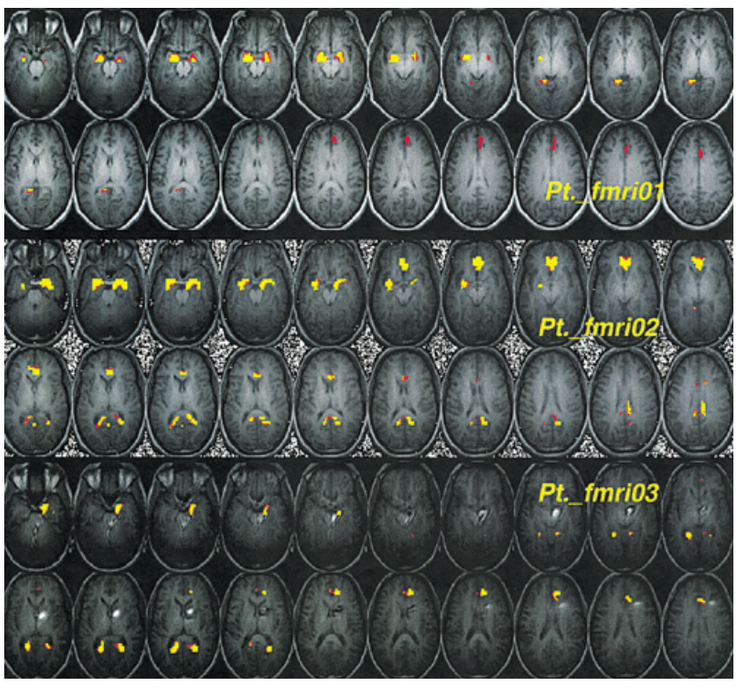
COLOR FIGURE 103.6. (This figure is printed in black and white as Figure 103.6.)



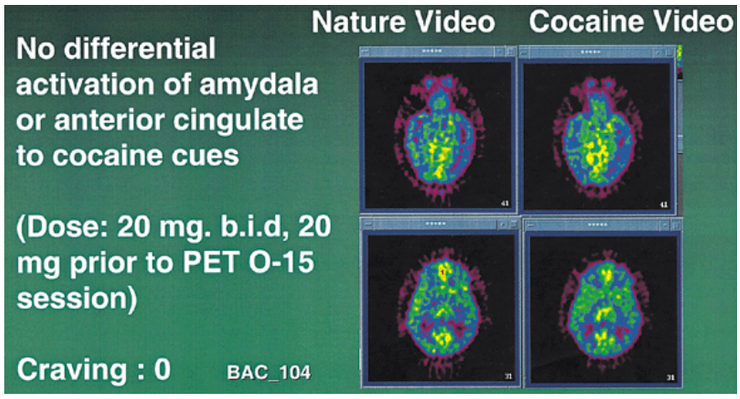
COLOR FIGURE 108.1. (This figure is printed in black and white as Figure 108.1.)



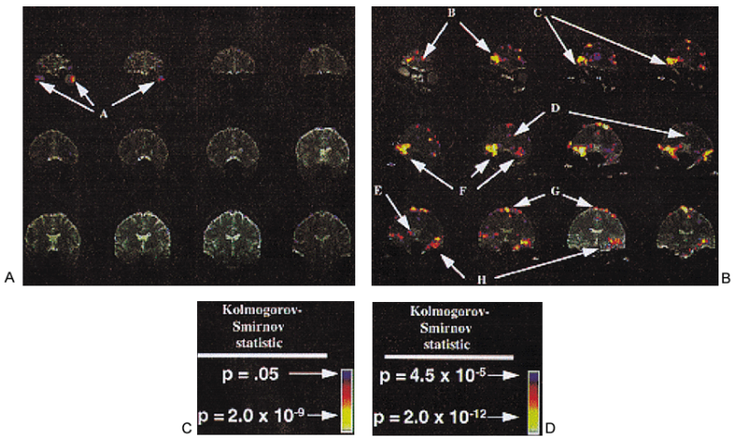
COLOR FIGURE 110.1. (This figure is printed in black and white as Figure 110.1.)



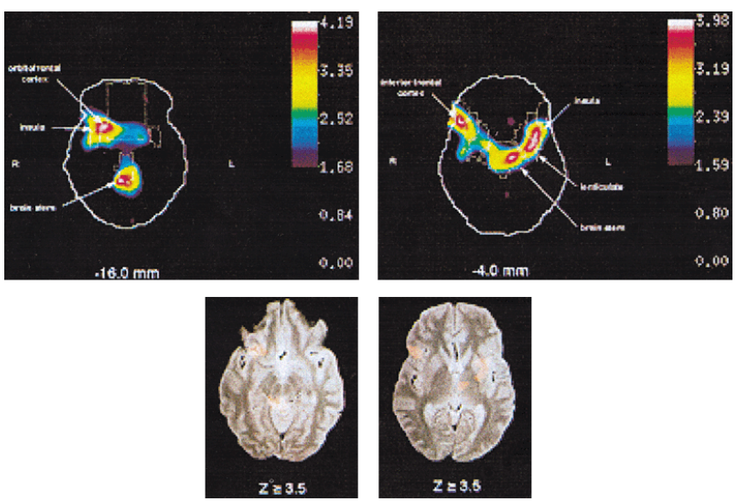
COLOR FIGURE 110.2. (This figure is printed in black and white as Figure 110.2.)



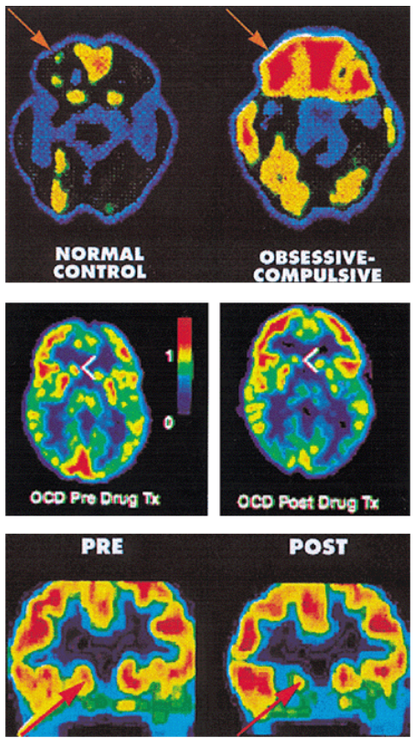
COLOR FIGURE 110.3. (This figure is printed in black and white as Figure 110.3.)



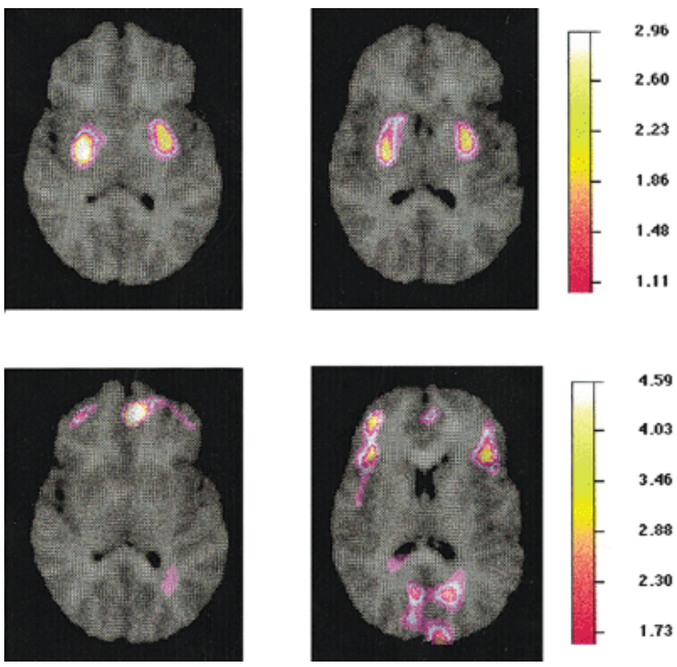
COLOR FIGURE 113.8. (This figure is printed in black and white as Figure 113.8.)



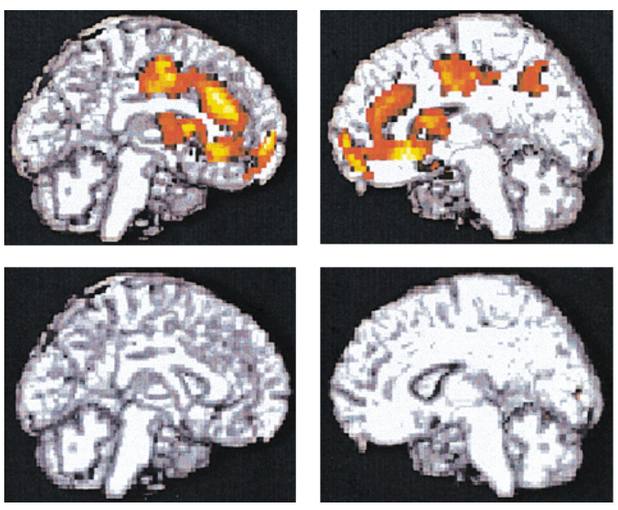
COLOR FIGURE 113.9. (This figure is printed in black and white as Figure 113.9.)



COLOR FIGURE 113.10. (This figure is printed in black and white as Figure 113.10.)



COLOR FIGURE 126.1. (This figure is printed in black and white as Figure 126.1.)



COLOR FIGURE 134.2. (This figure is printed in black and white as Figure 134.2.)

Index

References in *italics* indicate figures; those followed by "t" denote tables

A

B

C

D

E

G

H

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z